5-2012

Technological Aspects of Molecular Diagnosis of Bacterial Infectious Diseases

Christine S. Booth
University of Nebraska-Lincoln, christine.b2@gmail.com

TECHNOLOGICAL ASPECTS OF MOLECULAR DIAGNOSIS OF BACTERIAL INFECTIOUS DISEASES

by

Christine S. Booth

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Chemical Engineering

Under the Supervision of Professor Hendrik J Viljoen

Lincoln, Nebraska

May, 2012

TECHNOLOGICAL ASPECTS OF MOLECULAR DIAGNOSIS OF BACTERIAL INFECTIOUS DISEASES

Christine S. Booth, M.S.

University of Nebraska, 2012

Adviser: Hendrik J. Vijoen

The polymerase chain reaction (PCR) is continually growing in its application, particularly in the field of molecular diagnosis of disease from clinical specimens. The main focus has been in the detection and identification of pathogens. However, quantitative PCR is increasingly utilized to determine initial pathogen load. A well-designed PCR protocol is required in all of these instances. Just as importantly, in the context of disease diagnosis; is the design of the sample processing methodology. The ideal method should concentrate the DNA and effectively isolate a high-quality DNA product, free of PCR inhibitors, while also being simple, reproducible and safe.

The aim of this work is to address the research challenges posed in the preceding paragraph. A previously developed prototype diagnostic system is used to analyze and suggest improvements and an application of the technology is also described. Briefly, the system includes a polystyrene strip that is inserted into a lysis microreactor (LMR) that is fitted with an impeller and temperature control to lyse DNA. The DNA binds noncovalently to the strip and is transferred through a wash step to the thermocycler cuvette for amplification.

The research challenges were addressed by the following:

1. An analytical model was developed to determine the efficiency of each process comprising a PCR cycle. Using this model, reaction conditions can be directly linked to the overall yield and initial template concentration can be determined from real-time PCR data.
2. The flow characteristic of the LMR was solved by computational fluid dynamics to determine the DNA capture efficiency as a function of initial position.
3. Improvements to the use of a non-specific strip for DNA binding were explored by attaching target-complimentary oligonucleotides to a surface.
4. The prototype system was evaluated on a bank of frozen clinical stool samples. Samples were tested for *Clostridium difficile* genomic DNA and the results compared with standard *C. difficile* testing methods used routinely by a hospital clinical laboratory. The prototype system showed 97.5% concordance with standard testing methods.

DEDICATION

To my husband, Johan Prinsloo and my parents, Stephen and Christa Booth. *Niks hiervan sou moontlik wees sonder julle nie! Ek is lief vir julle!*

ACKNOWLEDGEMENTS

I would like to sincerely thank my adviser, Dr. Hendrik Viljoen, for the opportunity, guidance and support to work on this project and to undertake this incomparable learning experience. I would also like to extend a special word of thanks to Dr. Anuradha Subramanian for her continued guidance and encouragement and for going above and beyond in sharing her knowledge and expertise with me. Thank you also, to Dr. Cheryl Bailey, for your suggestions and participation on my examination committee.

Further thanks to past and present group members Dr. Scott Whitney and Dr. Joel TerMaat for provision of instrumentation (with Dr. Viljoen), and together with Mr. Tobi Louw and Dr. Elsje Pienaar for offering ideas and additional support. To other office and lab mates, Dr. Jennifer Calcaterra, Mr. Mostafa Fatemi, Dr. Sanjukta Guha Thakurta, Dr. Nick Whitney, Mrs. MinJeong Scheider and Mr. Michael Taylor, thank you for sharing ideas and experience. To the Chemical and Biomolecular Engineering staff, particularly Mrs. Trish Fenster and Mr. Leonard Akert, a special word of thanks for your continuous helpfulness. To Mr. Tobi Louw, Mr. Hunter Flodman and Dr. Jennifer Kane, thank you for suggestions with my defense presentation. Thank you also to instrumentation specialists, Mr. Christian Elowsky and Mrs. Sara Basiaga.

To my amazing husband, wonderful parents and grandparents (Alfie and Hettie Booth who are no longer with us; and At and Sophie Kok) and the rest of my family; also and my fantastic in-laws (Henry and Annie Prinsloo, the Beijer's and du Plessis'), thank you for your unconditional love and support. And to my friends, thank you for your encouragement, prayers and love.

TABLE OF CONTENTS

List of Tables

List of Figures

PREFACE

Over the last 20 years, PCR has become an indispensable laboratory technique, and is continually growing in its application, particularly in the field of molecular diagnosis of disease from clinical specimens. In addition to detection and identification of pathogens, newer technologies are increasingly utilizing the principles of quantitative PCR (qPCR) to determine initial pathogen load. These methods potentially allow for monitoring of response to antimicrobial therapy or discrimination between microbial colonization and infection. In addition to a well-developed PCR methodology, the sample processing method that is used to isolate the DNA from the clinical specimen is just as important in disease diagnosis. The ideal methodology should concentrate the DNA and effectively isolate a high-quality DNA product, free of PCR inhibitors. Additional requirements are that the sample processing be simple, reproducible and safe.

The aim of this thesis is to address the research challenges posed in the preceding paragraph. The work is divided into four chapters that are described in terms of the prototype diagnostic system shown in Figure P.1. As shown in Figure P.1B, a lysis micro reactor (LMR) has been previously developed to perform clinical sample processing with high efficiency and in processing times of approximately five minutes. The LMR has a capacity of 2 mL and it is fitted with an impeller and temperature control. A polystyrene strip (Figure P.1A), inserted into the LMR at the start of the lysis process, noncovalently binds lysed single stranded DNA on the hydrophobic surface (Figure P.1B). The strip allows the DNA to be effectively concentrated from the clinical sample and provides a

simple transfer method to move the DNA from clinical specimen, through a wash step and to the thermocycler cuvette for amplification (Figure P.1C, D).

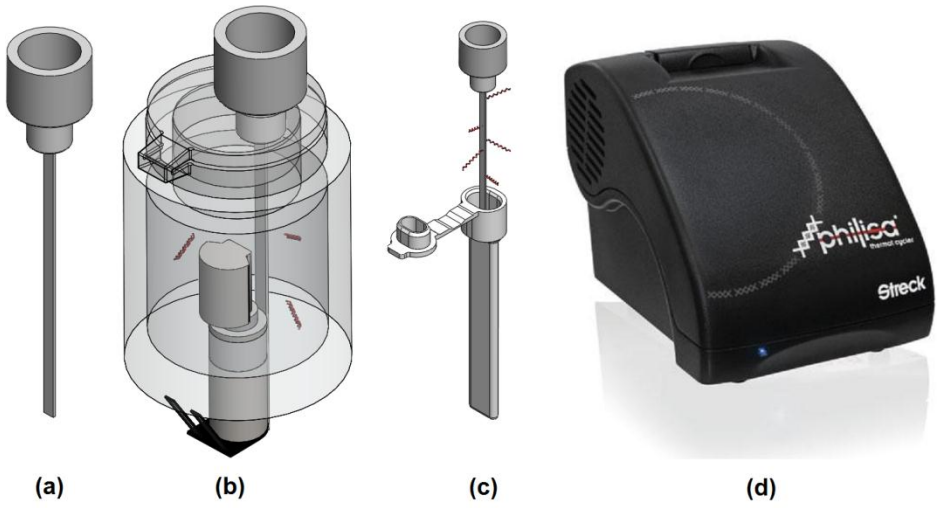

Figure P.1. Prototype diagnostic system. (a) The polystyrene capture strip with cap is inserted into (b) the LMR, which is equipped with an impeller and temperature control. (c) The strip with bound ssDNA is inserted into a PCR cuvette for (d) amplification and detection in a Philisa Thermo Cycler.

In Chapter 1, the PCR yield of each cycle is mathematically analyzed as a function of several processes occurring at each of the steps in the PCR cycle: (1) denaturing (2) annealing (3) polymerase binding, and (4) extension. Explicit expressions are provided for the efficiency of each process and reaction conditions can be directly linked to the overall yield. Consequently, experiments were designed that are specifically controlled by each one of the efficiencies and the results were shown to be consistent with the mathematical model. The experimental data was used to quantify six key parameters of

the theoretical model. An important application of the fully characterized model is to calculate initial template concentration from real-time PCR data.

Chapters 2 and 3 examine the sample processing methodology. In Chapter 2, the flow characteristics of the LMR were solved by computational fluid dynamics and a model was developed for the efficiency of DNA capture as a function of initial position. This analysis may be used to suggest improvements in the strip capture methodology. In Chapter 3, potential improvements to the use of a non-specific polystyrene strip are explored. Selective binding of oligonucleotides to the capture surface is examined through the covalent attachment of complimentary oligonucleotides to glass rods.

And finally, the results from an application of the prototype system (Figure P.1) are presented in Chapter 4. Stool samples are tested for the presence of *Clostridium difficile* (a major gram positive bacterial pathogen of the gastrointestinal tract) and compared with standard *C. difficile* testing methods used routinely by a hospital clinical laboratory.

CHAPTER 1

THEORETICAL AND EXPERIMENTAL ANALYSIS OF THE EFFICIENCY OF THE POLYMERASE CHAIN REACTION

1.1. Introduction

The polymerase chain reaction (PCR) has become a major technology in microbiology, molecular biology and related fields. Whereas PCR still has a lot of qualitative applications, it is increasingly used as a quantitative tool. The sensitivity of PCR permits amplification from a small number of starting templates. However, the exponential increase in product makes the inverse problem difficult – i.e. to infer the starting concentration from a large number of amplicons. Real-time PCR provides a proportional measure of the number of templates at each cycle.

Theoretically, the number of templates should double after each cycle. In practice, the DNA increases by a factor of $(1+\eta)$ where η is the cycle efficiency. Thus an efficiency of $\eta = 1$ would imply a doubling of the DNA concentration. Although the efficiency could change from cycle to cycle, therefore warranting the designation η_j to mark the j^{th} cycle, it is customary to report an overall efficiency (η) for n cycles. Saiki *et al*. (1985) related the overall efficiency (η) and yield (X) as follows: $X = (1+\eta)^n$ and this relation

became the standard way to express the overall efficiency[1] of PCR processes (Keohavong and Thilly, 1989, Li et al., 1988). A small variation in this relation has been proposed by Newton and Graham (1997) if the original DNA is genomic DNA with a length greater than the target DNA length[2]. It has been experimentally observed that yields can vary from cycle to cycle with a general decreasing trend with increasing cycle number, resulting in the characteristic sigmoidal curve (Kainz, 2000; Schnell and Mendoza, 1997a, 1997b; Stolovitzky and Cecchi, 1996). Additional references are listed in Waterfall et al., (2002). Although the use of an overall efficiency is a convenient norm to quantify experiments, it provides no information on cycle-to-cycle changes in efficiency.

The use of $X = (1+\eta)^n$ to infer starting concentrations of DNA has seen application in real-time PCR (rt-PCR) and it has been widely adopted for use in an array of applications including gene expression studies, mutation detection, forensic analysis and pathogen detection with the aim at both clinical diagnostics and food safety (Champe et al., 2008, Logan et al., 2009, Pfaffl, 2004). Two main quantification methods are the standard curve method and the $\Delta\Delta C_T$ method. The $\Delta\Delta C_T$ method is a relative quantification method that assumes 100% efficiency, and uses the differences in crossover threshold (C_T) values

[1] The overall efficiency η in the equation $X = (1+\eta)^n$ has frequently been erroneously reported as the arithmetic average of the individual cycle efficiencies, which it is not.

[2] If the original DNA length is greater than the target length, the first two PCR cycles actually produce sequences of indeterminate lengths and only from the third cycle onwards is the target sequence produced exponentially. Newton and Graham (1997) thus adjust the maximum theoretical DNA amplification factor from 2^n to $2^n - 2n$. However, the original DNA isn't accounted for in their equation and, albeit a minor lacuna, $2^n - 2n - 1$ is more accurate.

between experiment and control reactions to calculate an estimated fold-change in a target gene. The fold-change is defined as (see Livak and Schmittgen, 2001):

$$2^{\Delta\Delta C_T} = 2^{\Delta C_{T,control} - \Delta C_{T,experiment}} \quad (1.1a)$$

where:

$$\Delta C_{T,control} = C_{T,control}^{target} - C_{T,control}^{reference} \quad (1.1b)$$

$$\Delta C_{T,experiment} = C_{T,experiment}^{target} - C_{T,experiment}^{reference} \quad (1.1c)$$

The standard curve method amplifies serial dilutions of known concentrations of both the target and reference gene, along with samples of unknown concentration. The dilution curves are then used to generate a C_T value-concentration curve. When the unknown samples' C_T values are determined, they are correlated to a certain concentration by placement on this curve. The determined concentrations of the reference and target genes are then used to calculate fold-changes between experimental and control reactions.

Pfaffl (2001) proposed a method that combines the standard curve method and $\Delta\Delta C_T$ method. Like the standard curve method it uses dilution methods to calculate the efficiency for a specific reaction. This efficiency (η) is then used in the fold-change equation used by the $\Delta\Delta C_T$ method:

$$(1+\eta)^{\Delta C_{T,control} - \Delta C_{T,experiment}} \quad (1.1d)$$

Liu and Saint (2002a) followed a similar approach but used fluorescence levels at different points in one curve to calculate the efficiency, instead of the dilution curves.

These calculated efficiencies are assumed to be constant throughout the reaction (not varying from cycle to cycle). However, it has been shown that efficiencies are not constant over all cycles and more advanced models have been developed to include the efficiency variations from cycle to cycle (Liu and Saint, 2002b; Platts et al., 2008). However, these models do not provide expressions for the efficiencies of different processes that form part of the overall PCR process and only report a single efficiency per cycle.

Certain models do account for variations in efficiencies of the different stages (denaturing, annealing and elongation) of every cycle (Gevertz et al., 2005; Rubin and Levy, 1996). Gevertz et al. (2005) incorporated annealing and elongation efficiencies into the derivation of a single per-cycle efficiency. The evaluation of the efficiencies required the numerical solution of a set of initial value problems for each cycle. Despite being more rigorous, numerical integration does not lend itself to immediate or convenient implementation by other users. Rubin and Levy (1996) considered the annealing step, but their work was focused on calculating the probabilities for mispriming events in analyzing the effects of different factors on the specificity of PCR.

In this chapter we consider four different efficiencies that each contribute to the overall efficiency. These efficiencies are associated with the denaturing, annealing, ternary complex formation (i.e. polymerase binding to template/primer) and elongation steps. In all cases analytical expressions are provided for the different efficiencies, making it easy

for other users to apply and connect the efficiencies with overall yield and PCR conditions.

Additionally, experimental validation of the mathematical model is presented. Various real-time experiments have been designed to explore reactions that are limited by the annealing-, polymerase binding- and elongation efficiencies. These results have been used to determine the unknown model parameters. Finally, it is shown that this model provides an elegant method to determine initial DNA concentrations, using real-time data and the PCR protocol.

1.2. The Mathematical Model

An analytical model was used to calculate the template concentration S_j for each PCR cycle j. The template is the region of the sample DNA flanked by the sense- and anti-sense primers for replication; thus the initial DNA concentration is equal to the initial template concentration. For a complete derivation of the model, see the Appendix.

The model is based on the following assumptions:

- Symmetry prevails in sense and anti-sense molecules. There are equal numbers of forward and reverse primers and they anneal to equal numbers of sense and anti-sense single stranded DNA strands.
- All of the double-stranded DNA denatures completely to form single-stranded DNA.
- Polymerase damage and DNA damage efficiencies are constant for each PCR cycle.

- The annealing temperature is sufficiently below the primer melting temperature that annealing reactions are irreversible. The same is assumed for the elongation reaction.
- Primer-template annealing does not occur during the elongation phase.
- Partial elongation is not considered. Strands that are not fully extended by the end of the elongation cycle are treated as primers in subsequent cycles.
- The extension rate remains constant, i.e. no slow-down due to pyro-phosphorolysis or dNTP depletion.
- No unwanted side reactions such as primer-dimer formation and mis-priming are considered. Some suggestions are made in the conclusions section on how to include the effects of primer-dimer reactions empirically.

The model calculates an overall per cycle efficiency (η_j), which is the product of three individual efficiencies. The annealing efficiency ($\eta_{j,a}$) is the fraction of available templates that anneal to primers. The polymerase binding efficiency ($\eta_{j,E}$) is the fraction of template-primer (binary) complexes that bind to polymerase to form ternary complexes. Finally, the elongation efficiency ($\eta_{j,e}$) is the fraction of ternary complexes that are fully extended by the end of the elongation step:

$$\eta_j = \eta_{j,a}\eta_{j,E}\eta_{j,e} \tag{1.2}$$

$$\eta_{j,a} = (P_j - P_{j,a})/S_j \tag{1.3}$$

$$\eta_{j,E} = C_{j,e}/(B_{j,a} + C_{j,a}) \tag{1.4}$$

$$\eta_{j,e} = C_{j,c}/C_{j,e} \tag{1.5}$$

Table 1.1. Experimental and model parameters used in analytical model

Experimental parameters	Description	Model parameters	Description
t_a, t_e	Annealing / Elongation phase duration	k_p	Rate of primer annealing
S_0	Initial template concentration	k_c	Rate of polymerase binding at the annealing temperature
P_0	Initial primer concentration	k_c^*	Rate of polymerase binding at the elongation temperature
E_0	Initial polymerase concentration	β	Ratio of template annealing rate to primer annealing rate
V	Polymerase extension rate	η_d	Template denaturing damage
l	Template length	η_{dE}	Polymerase denaturing damage

The variables are defined in Table 1.1. The subscript j identifies the cycle and the subscripts a and e denote values at the end of the annealing and elongation stages respectively. For example, there are S_j templates and P_j primers at the start of cycle j, but at the end of the annealing stage there are $P_{j,a}$ primers left. Thus the number of binary and ternary complexes that have formed during the annealing stage is $(P_j - P_{j,a})$ and the ratio $(P_j - P_{j,a})/S_j$ defines the annealing efficiency. Equations 1.6-1.8 give the primer, ternary and binary complex values at the end of the annealing stage, the number of ternary complexes at the end of the elongation stage is given by Equation 1.9. The ternary complex concentration at the cut-off time $(C_{j,c})$ is the amount of primer-template-polymerase complexes that have formed after $t_c = t_e - l/V$ time has passed in the

elongation phase. The value l/V is the time it takes the polymerase to extend the primer to full length DNA. Thus, $C_{j,C}$ is the concentration of ternary complexes that will fully extend by the end of the elongation phase. This value is calculated using Equation 1.9 with t_e replaced by t_C.

$$P_{j,a} = P_j \left(1 + \gamma(\beta - 1)(1 - \exp\left(-k_p t_a P_j (\gamma(\beta - 1) + 1)^{\frac{1}{1-\beta}}\right)\right)^{\frac{1}{1-\beta}} \quad (1.6)$$

$$C_{j,a} = E_j \left(1 - \frac{(P_j - P_{j,a}) - E_j}{(P_j - P_{j,a}) \exp\left(\left((P_j - P_{j,a}) - E_j\right) k_C t_a\right) - E_j}\right) \quad (1.7)$$

$$B_{j,a} = P_j - P_{j,a} - C_{j,a} \quad (1.8)$$

$$C_{j,e} = \frac{(E_j - C_{j,a})(P_j - P_{j,a}) - B_{j,a} E_j \exp\left(\left((P_j - P_{j,a}) - E_j\right) k_C^* t_e\right)}{(E_j - C_{j,a}) - B_{j,a} \exp\left(\left((P_j - P_{j,a}) - E_j\right) k_C^* t_e\right)} \quad (1.9)$$

The model assumes that the double-stranded DNA strand separate completely (strand separation denaturing efficiency ≈ 1). However, some templates and primers may become damaged during denaturing (Cadet et al., 2002; Hsu et al., 2004; Lindahl and Nyberg, 1972, 1974; Pienaar et al., 2006). The polymerase may also be damaged during this step (Sambrook and Russel, 2000). Taking denaturing damage into account (η_d and η_{dE} for the template and polymerase, respectively), the number of templates, primers and polymerase during each cycle can be calculated from the values at the previous cycle:

$$S_{j+1} = \eta_d \left(1 + \eta_j\right) S_j \quad (1.10)$$

$$P_{j+1} = \eta_d \left(P_j - \eta_j S_j\right) \quad (1.11)$$

$$E_{j+1} = \eta_{dE} E_j \quad (1.12)$$

The variable S_j refers to the template concentration at the beginning of the j^{th} cycle. Therefore, the template concentration at the end of the elongation phase of cycle j is equal to S_{j+1}. This also corresponds to the $(j+1)^{th}$ spectrometer reading, as fluorescence is measured at the end of the elongation phase. To simplify the situation, the first cycle will be counted as cycle 0. Hence, the template concentration at the end of cycle zero is given by S_1, which corresponds to the first spectrometer measurement.

If the values of S_0, P_0 and E_0 are known, then the concentrations of all subsequent cycles can be calculated using Equations 1.2-1.12. First, Equations 1.6-1.9 are used to determine the amount of binary and ternary complexes that have formed after annealing and elongation. These concentrations are then used to determine the cycle efficiencies (Equations 1.2-1.5) and the template, primer and polymerase concentrations at the beginning of the next cycle are calculated (Equations 1.10 – 1.12). The function values S_j can be calculated – clearly quantitative PCR is an inverse problem.

The model parameters are also listed in Table 1.1. The initial conditions and PCR protocol parameters (experimental parameters) are known and fixed before the experiment. The model parameters are unknown and must be determined by matching experimental and theoretical data.

The rate of polymerase binding to form a ternary complex changes as the temperature increases from the annealing temperature to the elongation temperature. The value of

$k_c^* > k_c$ reflects this increase in the polymerase binding rate. The cycle dependent variables are listed and explained in Table 1.2.

Table 1.2. Variables used in analytical model

Variable	Description	Variable	Description
S_j	Template concentration at the beginning of annealing	E_j	Polymerase concentration at the beginning of annealing
P_j;$P_{j,a}$	Primer concentration at the beginning and end of annealing	$B_{j,a}$	Binary complex concentration at the end of annealing
γ_j	Ratio of template to primer concentration	$C_{j,a}$; $C_{j,e}$; $C_{j,c}$	Ternary complex concentration at the end of annealing, elongation and at the cut-off time, respectively
δ_j	Ratio of equilibrium primer concentration after annealing to S_j		

1.3. Materials and Methods

The reference PCR mixture contained 0.5 U KOD Hot Start DNA polymerase (Novagen, Madison, WI). It was estimated that 0.5 U KOD polymerase is equivalent to a concentration of 0.084 µM (Mamedov et al., 2008). The reference mixture also contained 1X polymerase manufacturer's buffer, 200 µM of each dNTP, 3.5 mM $MgSO_4$, 400 µg/ml non-acetylated BSA and 3 µM SYTO13 (Invitrogen, Carlsbad, CA). 0.3 µM of each primer was used to obtain a 1002 bp product. PCR was performed in a PCRJet Thermocycler (Megabase Research Products, Lincoln, NE) in 25 µl reaction volumes containing 1 ng bacteriophage λ genomic DNA. The DNA was ordered from New

England Biolabs (Ipswich, MA) and 1 ng of DNA in 25 µl corresponds to a concentration of 1.27 pM. Thermocycling consisted of a 30 second hot start at 96°C, 90 cycles of 2 s denaturing at 96°C, 3 s annealing at 64°C and 10 s elongation at 72°C. Real-time data was collected at the end of each elongation step.

Seven different experiments were performed to investigate the effects of the key experimental parameters. These parameters are listed in Table 1.3. Each experiment was repeated three times and the average values were calculated. The average values were used to determine the unknown model parameters. The remaining experimental parameters were kept constant ($t_a = 3s,\ V = 300\ bp/s,\ l = 1002\ bp$). The polymerase extension rate V was obtained from Griep et al. (2006).

Table 1.3. Experiments for determining model parameters

Nr	Experiment	S_0 (pM)	P_0 (µM)	E_0 (units)	t_e (s)
1	Reference	1.27	0.30	0.5	10
2	Dilution I	**0.127**	0.30	0.5	10
3	Dilution II	**0.0127**	0.30	0.5	10
4	Reduced primer	1.27	**0.15**	0.5	10
5	Increased primer	1.27	**0.40**	0.5	10
6	Short elongation	1.27	0.30	0.5	**3**
7	Reduced polymerase	1.27	0.30	**0.2**	10

Although a rapid PCR protocol was used, there is still a finite amount of transition time between each of the three phases. To accommodate for ramp-times between the annealing and elongation phase, half a second was added to the elongation time in the mathematical model.

Three additional experiments were conducted using a conservative PCR protocol. This was used to test a method for determining the initial template concentration S_0, as discussed in section 1.5. For these experiments, the annealing- and elongation-time was held constant at $t_a = 10\ s$ and $t_e = 20\ s$. Table 1.4 lists the initial conditions for this additional set of experiments.

Table 1.4. Experiments for determining initial template concentration

Nr	Experiment	S_0 (pM)	P_0 (µM)	E_0 (units)
i	Conservative reference	1.27	0.40	0.5
ii	Conservative dilution I	**0.127**	0.40	0.5
iii	Conservative dilution II	**0.0127**	0.40	0.5

1.4. Results and Discussion

1.4.1 Demonstration of the usefulness of the model

To illustrate the usefulness of this analysis, we first investigate the roles of the different efficiencies on the overall efficiency for different PCR conditions. We do this by assuming certain values for the model parameters prior to experimental verification.

Three different polymerase concentrations will be used and for each choice the elongation period will be varied from Δt_e = 5 s to Δt_e = 10 s and Δt_e = 20 s; where $\Delta t_e = t_e - t_a$. The parameters that do not change are: D_0= 1x10^5 copies, β = 5, k_C = 15 (μM s)$^{-1}$ (Mamedov et al., 2008), η_d = 1, η_{dE} = 0.99, l_{ext} = 400 nt, P_0= 6x10^{12} copies - i.e. 10 picomole, reaction volume is 25 μL (Griep et al., 2006) and the maximum cycle number is 40. We use the simple form of the overall efficiency;

$$\eta_j = \eta_d \left[\eta_{j,a} - \frac{(E_j - \eta_{j,a} S_j)\eta_{j,a} \exp\left((E_j - \eta_{j,a} S_j) k_c (t_c - t_e)\right)}{E_j - \eta_{j,a} S_j \exp\left((E_j - \eta_{j,a} S_j) k_c (t_c - t_e)\right)} \right] \tag{1.13}$$

The value of k_P is not given since k_P cancels out in the product of dimensionless time and only k_C is needed for calculation. In the discussion that follows, we refer to the smallest of η_a, η_E or η_e as the controlling efficiency.

1.4.1.1 Case 1: E_0= 12.6x10^{11} copies

Results for case 1 are presented in Figure 1.1. In Figure 1.1A the different efficiencies are plotted as a function of cycle number. The elongation time is Δt_e = 20 s. The polymerase is in excess and the system is under the control of the annealing efficiency and it tracks the overall efficiency closely. The overall efficiency drops below 90% after cycle 22. The efficiency is less than 10% after 30 cycles and it is expected that increases

in the yield will be exiguous. If the overall yield is calculated, the average value over the first 30 cycles is 81%, but over the 40 cycles it drops to 56%.

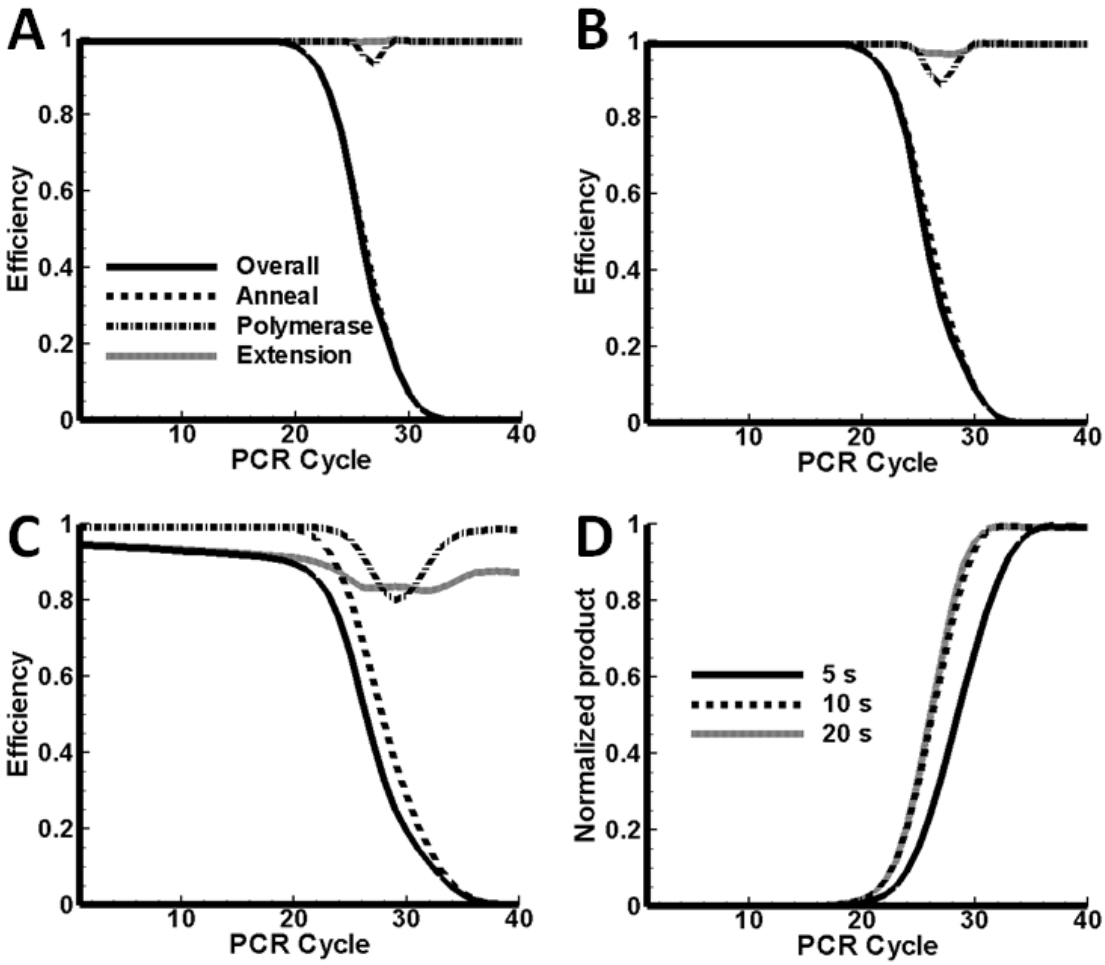

Figure 1.1. Efficiencies as a function of cycle number. D_0= 10^5 copies, E_0= $12.6x10^{11}$ copies, elongation period is 20 s (A), 10s (B) and 5s (C). (D): Normalized DNA product as a function of cycle number. The curves had the same maximum before normalization.

It is expected that the elongation efficiency will lower if the elongation time is shorter. In Figure 1.1B, the efficiencies are shown for the case Δt_e = 10 s (all the other parameters as for Figure 1.1A). The overall efficiency still tracks the annealing efficiency, however a slight decrease is observed in the polymerase and extension efficiencies. There is a brief period between cycle 26 and cycle 28 where the polymerase efficiency drops below 90%. The extension efficiency also lowers during this period, but only down to 96%. In Figure 1.1C the results are shown for an even shorter elongation time, Δt_e = 5 s. Here, the system is under extension control through cycle 24 and under annealing control for the remaining cycles. The localized drop in polymerase efficiency is still present, but the trough spans cycles 26 to 33 and it is deeper. There is even a brief period where the polymerase efficiency is less than the extension efficiency. Whereas $\eta_{j,a}$ is a monotonic decreasing function of cycle number, the polymerase and extension efficiencies exhibit local minima.

Normalized predicted PCR product amounts for the 3 elongation times (20, 10 and 5 seconds shown in Figures 1.1A-C, respectively) are shown in Figure 1.1D. In all three cases the same number of initial copies is amplified to the same final amount. The effect of shorter extension times is to slow template amplification down; more cycles are required to reach the plateau. The mid-points of the curves shift to higher cycle numbers for shorter elongation times, although the copy number remains the same. In Figure 1.1D the two longer extension times give mid-points just beyond cycle 26, but for the shortest time Δt_e = 5 s, the mid-point is at cycle position 28.5.

1.4.1.2 Case 2: E_0= 6.3x10^{11} copies

Results for case 2 are shown in Figure 1.2. The initial polymerase concentration is halved with respect to the amount used in case 1. Results for the three extension times (20, 10 and 5 seconds) are shown in Figures 1.2A-C respectively.

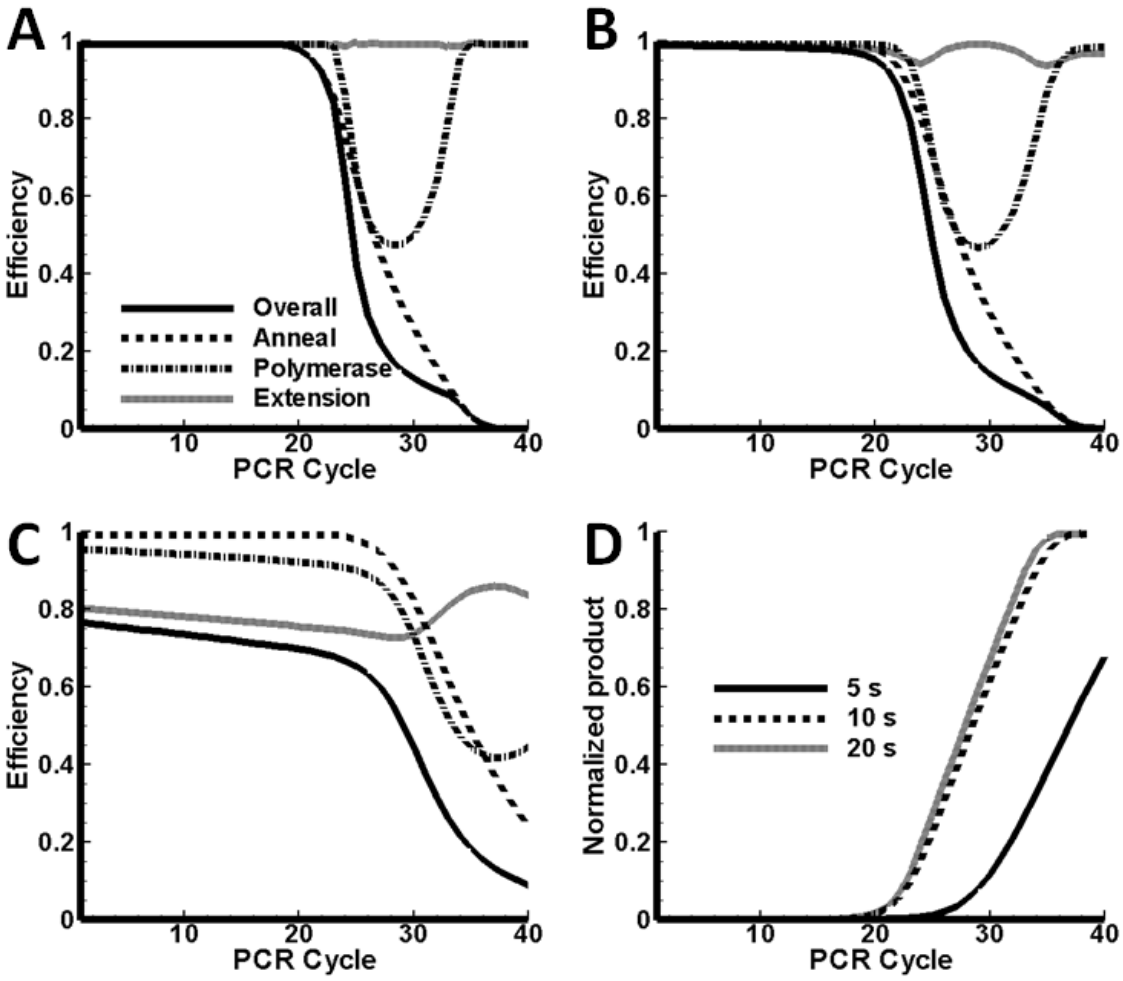

Figure 1.2. Efficiencies as a function of cycle number. D_0= 10^5 copies, E_0= 6.3x10^{11} copies, elongation period is 20 s (A), 10s (B) and 5s (C). (D) Normalized DNA product as a function of cycle number.

In Figure 1.2A the results are shown for Δt_e = 20 s. The reduced polymerase concentration causes a pronounced drop in $\eta_{j,E}$ between cycles 24 and 34 (compare to Figure 1.1A). During this period the number of binary complexes exceeds the number of polymerase molecules, but after cycle 28 this deficit becomes less and the polymerase efficiency begins to increase again - the explanation is a reduction in the number of binary complexes at later cycles, due to increased formation of double stranded DNA during the annealing stage. Compared to the results of case 1, the overall efficiency drops off sooner, and 50% overall efficiency is reached at cycle value 24.5. The extension efficiency remains near unity for the whole PCR reaction, with a subtle double minimum observable.

The results for Δt_e = 10 s are shown in Figure 1.2B. The width of the $\eta_{j,E}$ trough is wider, compared to Figure 1.2A, but the results are qualitatively similar. Also, the reduction in extension time from 20 s to 10 s enhances the double minima in $\eta_{j,e}$; compare $\eta_{j,e}$ in Figure 1.2A with 1.2B.

When the extension time is set to Δt_e = 5 s (Figure 1.2C), the system is under extension control for the first 30 cycles; under polymerase control until cycle 36 and under annealing control for the last four cycles. Here is an example where three different efficiencies controlled the system over the course of 40 cycles. One mechanism overtakes

another as being limiting and the results underscore the nonlinear character of the PCR process.

The plots of normalized DNA product vs. cycle number are shown in Figure 1.2D. The products have been scaled with the same maximum as in Figure 1.1D. The mid-points for Δt_e = 20 s, Δt_e = 10 s are close, at cycle value 27.8 and 28.5 respectively. These values differ from the results for similar extension times in case 1 earlier, and lie close to the midpoint for Δt_e = 5 s (of case 1). The results show that the midpoints shift if the polymerase concentration changes. The product curve does not reach saturation in the case of Δt_e = 5 s (solid curve, Figure 1.2D). If more cycles are added, then the curve continues to increase linearly until it finally plateaus when the primers are depleted. Note that all three curves have different slopes in the linear region. The slope decreases as the extension time is shortened, thus lower extension efficiencies lead to a slow-down of the process.

1.4.1.3 Case 3: $E_0 = 2.1 \times 10^{11}$ copies

Results for case 3 are shown in Figure 1.3. In this case the polymerase concentration is reduced by a factor of 3 with respect to case 2. The results for the three extension times are shown in Figures 1.3A-C.

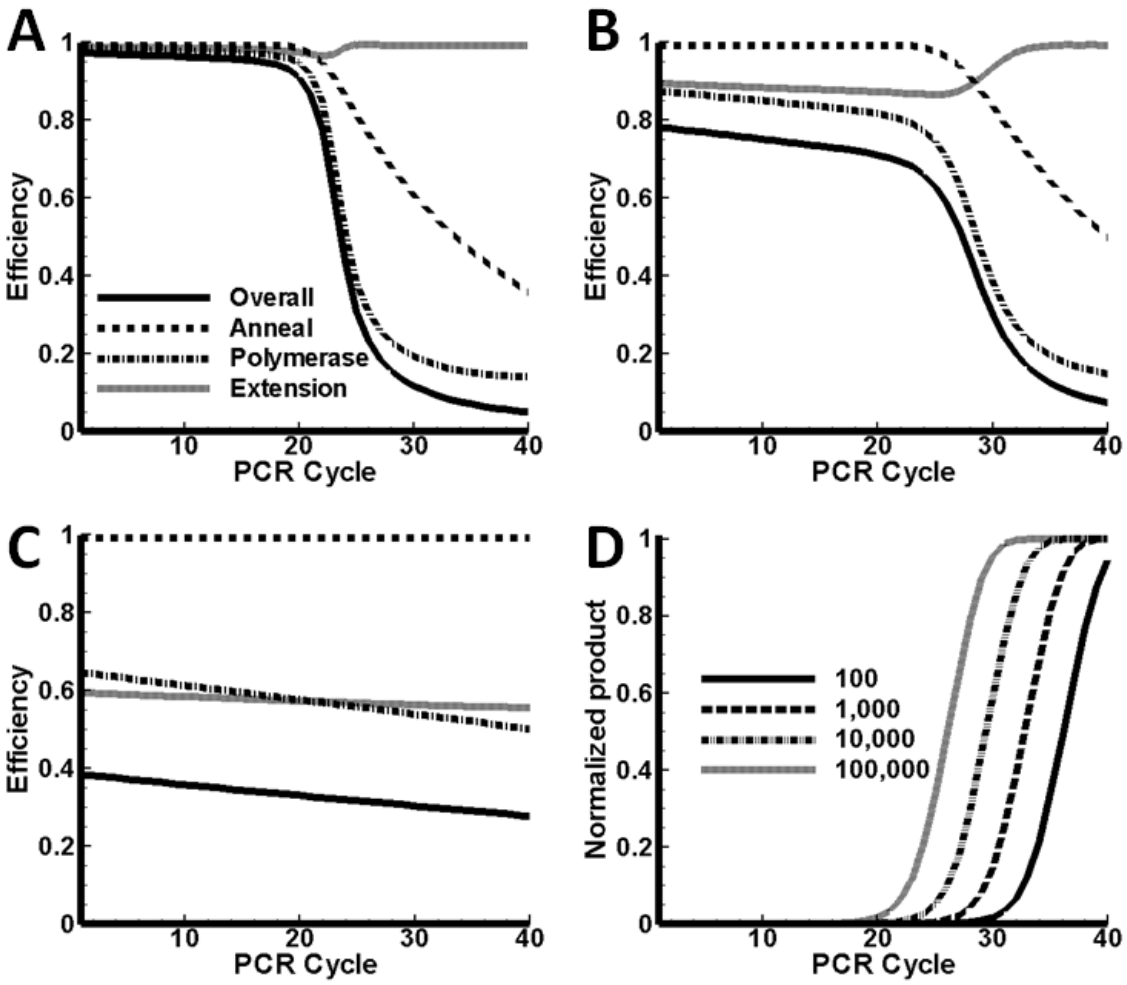

Figure 1.3. Efficiencies as functions of cycle number. D_0= 10^5 copies, E_0= 2.1x10^{11} copies, elongation period is 20 s (A), 10s (B) and 5s (C). (D) Serial dilution study - normalized DNA product as a function of cycle number. D_0= 10^2, 10^3, 10^4 and 10^5 (as indicated in the legend), E_0= 12.6x10^{11} copies, t_E = 20 s.

For Δt_e = 20 s the system remains under polymerase control over all 40 cycles. Both $\eta_{j,E}$ and $\eta_{j,a}$ are monotonically decreasing functions, and $\eta_{j,e}$ exhibits a single minimum. Results for Δt_e = 10 s are shown in Figure 1.3B. The extension efficiency is lower in

Figure 1.3B compared to Figure 1.3A, hence the overall efficiency is lower. However, the system remains under polymerase control. The primers are not depleted at the end of 40 cycles ($\eta_{j,a}$ is still relatively high) and amplification will continue beyond this point, albeit very slowly.

Figure 1.3C presents an example of very poor overall efficiency, where Δt_e = 5 s. For the first 25 cycles the system is controlled by extension, and then by the polymerase concentration. The annealing efficiency remains near unity. The total product formation will be much less than in previous cases.

The final example is a simulation of a serial dilution study. The conditions are the same as for case 1 and the extension time remains constant at Δt_e = 20s. The initial template concentration varies from 10^2 copies to 10^5 copies. The results are shown in Figure 1.3D. The results are as expected for a quantitative PCR experiment. Consecutive midpoints differ by 3.3 cycle values and the slopes are parallel.

1.4.2 Experimental validation of the model

1.4.2.1 Determination of model parameters

The model depends on six parameters (refer to Table 1.1). The parameters are determined by fitting the results of the model to the experimental results. In Table 1.5 the parameters that produced a least square error fit for all experiments are listed. The least square error parameters for each individual experiment were also calculated and used to determine the

standard deviation of each parameter with respect to the best fit for all experiments. This is also shown in Table 1.5. The rate constants are in accordance with Gevertz et al., (2005), who used values of k_P = 1 (μM.s)$^{-1}$ and β = 1.

Table 1.5. Physical parameters determined by matching model predictions to experimental results

k_p = 1.59±0.18 (μM.s)$^{-1}$	$\beta \approx 1$
k_c = 7.08±0.86 (μM.s)$^{-1}$	η_d = 1.00±0.008
k_c^* = 7.08±0.86 (μM.s)$^{-1}$	η_{dE} = 0.947±0.005

In Figure 1.4 the experimental results and the results of the mathematical model are compared for the parameters as listed in Table 1.5.

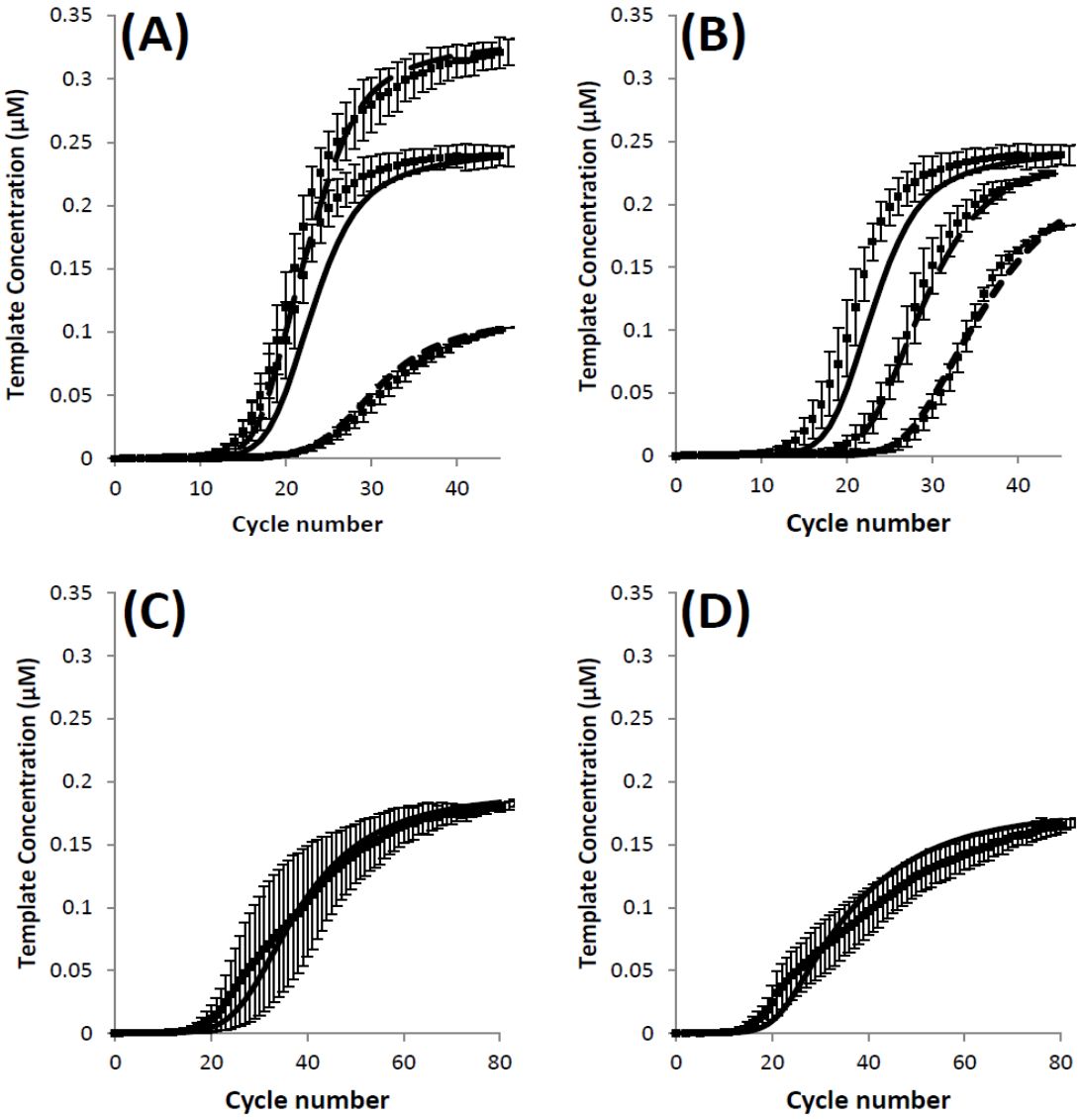

Figure 1.4. Results for experiments 1- 7 (with one standard deviation error bar) and the model predictions (solid lines) for parameter values listed in Table 1.5. (A) Reference (solid line) with increased- (dashes) and reduced-primer (short dashes) experiments. (B) Reference experiment (solid line). Dilution I (dashes) and Dilution II (short dashes). (C) Shortened elongation time experiment. (D) Reduced polymerase concentration.

1.4.2.2. The PCR Efficiencies

Once the parameters of the model have been determined, they can be used to calculate the different efficiencies, as given by Equations 1.2-1.12. The theoretical cycle efficiencies for experiments 1, 3-7 (cf. Table 1.3) are shown in Figure 1.5. The parameter values of Table 1.5 and the concentrations and PCR protocol values (as explained in Section 1.3 and Table 1.3) have been used to model the different experiments.

In Figure 1.5A the efficiencies are shown for the reference experiment. The annealing efficiency is smaller than the polymerase and extension efficiencies; hence the experiment is under annealing control. This is not surprising, since the annealing time is only 3 seconds. However, the polymerase binding efficiency η_E exhibits a local minimum and maximum in the 20 to 30 cycle range. This cycle range is marked by a rapid increase in templates and concomitantly the binary complexes. Therefore the demand on polymerase to form ternary complexes increases. Later, as the plateau phase is approached, fewer binary complexes form (lower demand on polymerase) and the fraction of binary complexes that convert to ternary complexes increases (an increase in polymerase efficiency). The continued decline in the polymerase efficiency during the plateau phase is primarily due to polymerase damage η_{dE}.

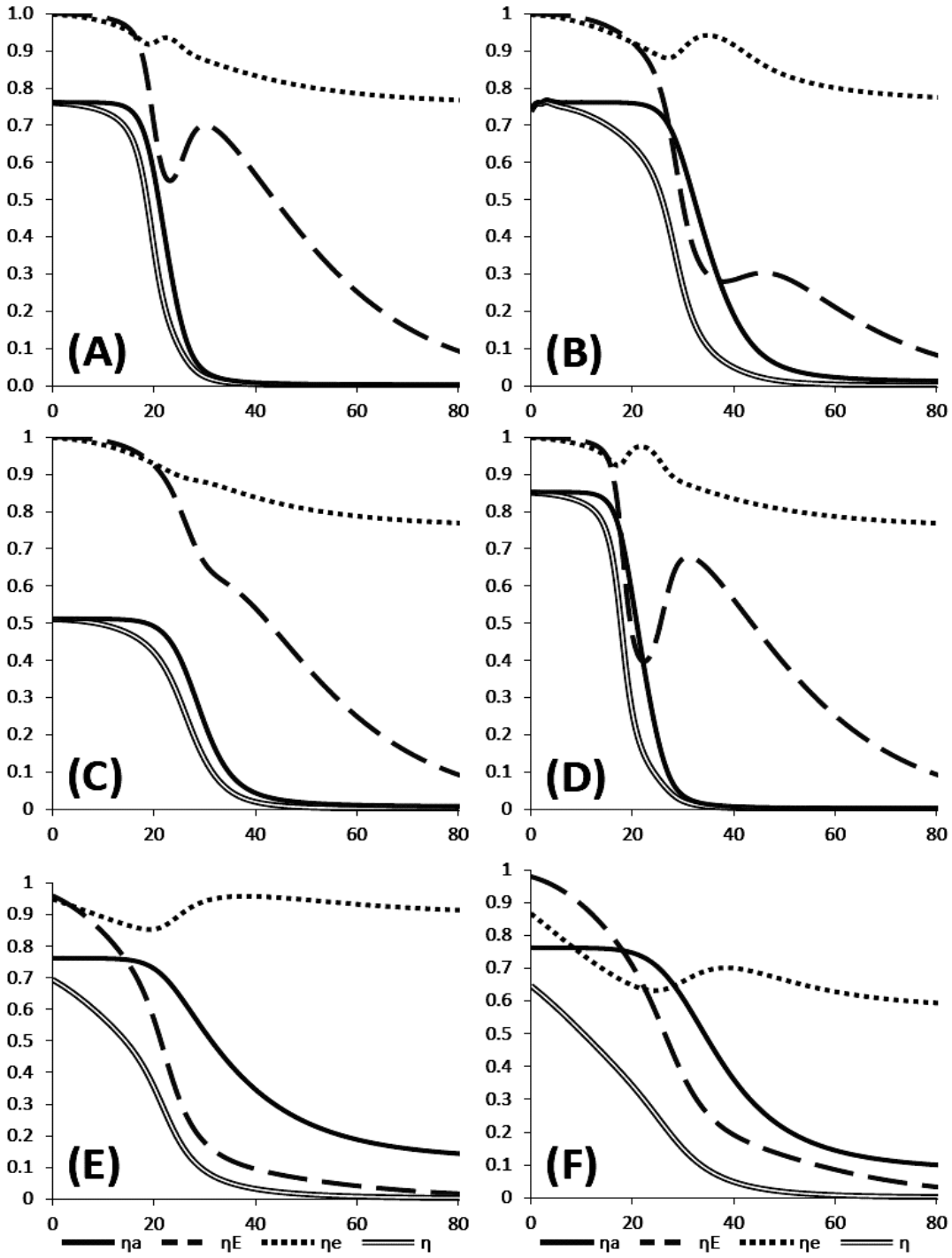

Figure 1.5. The annealing, polymerase binding, elongation and total efficiency for the following experiments: (A) Reference (B) Dilution II (C) Reduced Primer (D) Increased primer (E) and Reduced polymerase (F) Short elongation time.

It can also be noted from Figure 1.5A that the elongation efficiency is the highest of all three, but a small uptick is found in the cycle range that coincides with the local dip in polymerase efficiency. As explained in the previous paragraph, if the fraction of binary complexes that convert to ternary complexes decreases during the period of rapid increase in templates, then the polymerase binding efficiency will decrease ($C_{j,e}$ appears in numerator of Equation 1.4) and the elongation efficiency will increase ($C_{j,e}$ appears in denominator of Equation 1.5).

The efficiency profile is similar for the first dilution experiment (Figure 1.5B). The decrease in annealing efficiency is shifted laterally as a lower initial template concentration is used. The polymerase binding efficiency does play a more significant role – this is due to significant polymerase damage by the time η_E becomes controlling. This leads to a slight overall decrease in efficiency.

The reduced primer experiment (Figure 1.5C) is especially sensitive to the rate of primers annealing (k_P), as this experiment is strongly controlled by annealing efficiency. When the initial primer concentration is increased (Figure 1.5D), the polymerase binding efficiency becomes controlling during the exponential growth period (cycles 20 to 30) as the ratio between available polymerase and binary complexes decreases. The polymerase efficiency plays a much more controlling role when the polymerase concentration is lowered, as shown in Figure 1.5E. Here, the annealing efficiency is only controlling during the initial cycles of the process. After cycle 20, the efficiency is under polymerase

binding control. Figure 1.8F shows the results for an experiment with reduced elongation times. The elongation efficiency is controlling for cycles 10 through 30; then the system is controlled by polymerase binding for the duration of the process.

The overall efficiency in the reduced polymerase and short elongation time experiments decreases gradually, as opposed to the sudden decrease found in the reactions that are purely annealing limited. Compare the overall efficiencies up to cycle 40 in Figures 1.5A, 1.5C and 1.5D with the values in Figures 1.5E and 1.5F. If the system is under polymerase or elongation control, then the template concentration is no longer symmetrical around the inflection point (typical sigmoidal shape), but a slow decrease in the slope after the inflection point occurs (also compare with the respective experimental curves in Figure 1.4C and 1.4D). These experimental results are consistent with the mathematical model.

1.4.2.3 Quantitative PCR Application

In Figure 1.4 the model (Equations 1.2-1.12) has been fitted to the experimental results to determine the parameters. The best fit values are listed in Table 1.5. Of particular importance is $\beta = 1$ (signifying the competition between primer-template and template-template annealing) since it changes Equation 1.6 qualitatively. By taking the limit$\beta \to 1$, Equation 1.6 is written in the simpler form:

$$P_{j,a} = P_j \exp\left(-\gamma_j\left(1 - \exp\left(-k_p t_a P_j \, exp\left(-\gamma_j\right)\right)\right)\right) \quad (1.13)$$

The model, which now comprises of eqns. (1.2-1.5), (1.7-1.13), can be used to solve the inverse problem, i.e. determining the initial template concentration (S_0). If a value for S_0 is guessed, the model can be solved and the resulting curve S_j vs. j can be compared to the experimental curve (on a normalized basis) until a best fit is obtained. This approach is cumbersome.

A simpler procedure is devised by using the midpoint cycle number, which is defined as the cycle that corresponds to half the plateau (or maximum) value: $S_M = \max(S_j)/2$. The mid-point cycle number M is uniquely determined by S_0 and the PCR conditions. The locus of M as a function of S_0 can be determined using the mathematical model and the graph of M vs. $\log_2(P_0/S_0)$ can be constructed. This is shown in Figure 1.6B.

Determining the initial template concentration becomes straightforward: the midpoint cycle number M is determined from the experimental real time results. This value is used to determine $\log_2(P_0/S_0)$ from the graph (constructed using the mathematical model, as above). Finally, this can be used to calculate S_0, as P_0 is known.

1.4.2.3.1 Conservative elongation time

The calculation of the midpoint cycle number locus can be further simplified if the PCR conditions are chosen conservatively. For example, if the elongation times are long with respect to the minimum elongation time l/V, then the effect of the elongation efficiency becomes negligible (i.e. $\eta_e \approx 1$). For our template length and choice of polymerase this

conservative protocol is achieved by setting the elongation time equal to 20 seconds (longer templates/polymerases with slower elongation rates will require longer elongation times). The model reduces to three equations, given by eqns. (1.13-1.15). Note that annealing time and initial polymerase concentration are still present in the model.

$$S_{j+1} = S_j + \min(P_j - P_{j,a},\ \eta_{dE}^{j} E_0) \tag{1.14}$$

$$P_{j+1} = P_j - \min(P_j - P_{j,a},\ \eta_{dE}^{j} E_0) \tag{1.15}$$

Note that Equation 1.14 implies that if the amount of available polymerase ($\mathrm{E_j} = \eta_{\mathrm{dE}}^{\mathrm{j}} \mathrm{E_0}$) is greater than the amount of primer-template complexes ($P_j - P_{j,a}$), then the amount of new templates formed is equal to the amount of binary complexes formed. If $\eta_{dE}^{j} E_0 < P_j - P_{j,a}$, the reaction is limited by the amount of polymerase available.

The locus M vs. $\log_2(P_0/S_0)$ can be calculated using Equations 1.13-1.15 for a conservative elongation time protocol. The application remains the same; the midpoint cycle number is determined from the real time data and used to determine $\log_2(P_0/S_0)$ from the locus. This is shown in Figure 1.6B.

1.4.2.3.2 Conservative elongation time and excess polymerase

If the experiment is setup so that $\eta_{dE}^{j} E_0 > P_j - P_{j,a}$ for all cycles j, then a further simplification can be made:

$$S_{j+1} = S_j + P_j - P_{j,a} \tag{1.16}$$

$$P_{j+1} = P_{j,a} = P_j \exp\left(-\gamma_j\left(1 - \exp\left(-k_p t_a P_j \exp\left(-\gamma_j\right)\right)\right)\right) \tag{1.17}$$

1.4.2.3.3 Conservative elongation and annealing times and excess polymerase

Finally, one can use conservative annealing times to arrive at the model:

$$S_{j+1} = S_j + P_j - P_{j,a} \tag{1.18}$$

$$P_{j+1} = P_{j,a} = P_j \exp\left(-\gamma_j\right) \tag{1.19}$$

In Equations 1.18-1.19, the only factor that limits templates from doubling at each cycle is the competition between single stranded DNA to bind to complementary single stranded DNA instead of primers. It is interesting to note that during the early stages of the experiment, when the primers are in excess and $\gamma_j = S_j/P_j$ is small, the exponential term in Equation 1. 19 is well approximated by a linear expansion. If a linear expansion is used, then Equation 1.18 leads to the following well-known result:

$$S_{j+1} = S_j + P_j - P_j \exp\left(-\frac{S_j}{P_j}\right) \approx S_j + P_j - P_j\left(1 - \frac{S_j}{P_j}\right) = 2S_j \tag{1.20}$$

Equations 1.18-1.19 presents the most ideal case, but it is important to see that all the conservative protocols are only simplifications of the general model. Therefore a quantitative analysis can be done for any set of PCR conditions.

The experiments listed in Table 1.4 correspond to a conservative elongation time protocol, i.e. Equations 1.13-1.15. Figure 1.6A shows the spectrometer readings compared to the simplified model. It is clear that a change in initial template concentration produces a lateral shift in the real-time curve. The mid-point value for each experiment is indicated with a grey cross on Figure 1.6A. In Figure 1.6B we plot the locus of M vs. $\log_2(P_0/S_0)$, shown as the dashed line. The three experimental values of M are also marked on the locus. Suppose the initial concentrations were not known, then the experimentally obtained values of M (Figure 1.6A) would be used to read off $\log_2(P_0/S_0)$ from the dashed line in Figure 1.6B.

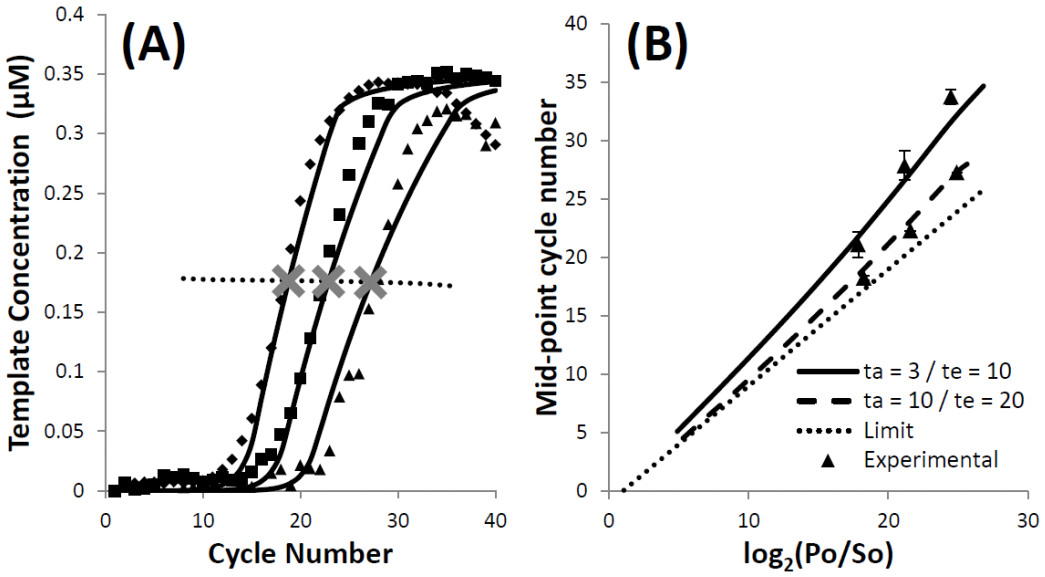

Figure 1.6. (A) Results for experiments i-iii with the simplified model predictions. The dilution curves correspond to an initial template concentration of $S_0 = 1.27\,pM$, $S_0 = 0.127\,pM$ and $S_0 = 0.0127\,pM$. The midpoint cycle number (M) is indicated by an X. The locus of points representing M over a range of S_0 is shown by the dotted line. (B) The midpoint cycle number M as a function of $\log_2(P_0/S_0)$ for the reference (solid line) and conservative reference (dashed line) parameters, over a range of S_0 values. The actual midpoint cycle numbers obtained by fluorescent measurements are shown. As the annealing time is increased, the loci approach a limit function (dotted line).

The fast protocol that was used for experiments listed in Table 1.3 requires that we use the general model (1.2-1.5, 1.7-1.13). The theoretically determined locus for the fast protocol is shown as the solid line in Figure 1.6B. Values of M for experiments 1-3 (Table 1.3) are also plotted on the locus. Finally the locus obtained using

Equations 1.18-1.19 are plotted as the dotted line in Figure 1.6B. Note that the most conservative model forms a lower bound for the other models. It becomes quite clear how PCR conditions impact the template amplification and how to account for protocol changes quantitatively. This analysis becomes especially helpful in a time where rapid PCR is used more in point-of-care diagnosis applications.

1.5. Conclusions

1. If the polymerase is in excess compared to the binary complex and the extension time is long, then the polymerase binding and elongation are not rate-limiting.
2. If the polymerase is in excess compared to the binary complex but the extension time is short, the system is under control of the extension time and the annealing efficiency.
3. If the binary complex is in excess compared to the polymerase, the system is under the control of the polymerase concentration.
4. The efficiency changes from cycle to cycle and different mechanisms may control the system over the course of 30 or 40 cycles.
5. The annealing efficiency is a monotonic decreasing function of cycle number, but η_E^j and η_e^j may not be. A particularly interesting situation arises if the polymerase concentration becomes rate-limiting. Since new templates still form and γ^j continues to increase with each cycle, the annealing efficiency decreases. As a result the binary complexes begin to decrease at some point and the polymerase concentration is no longer deficient – then a notable increase in η_E^j occurs.

6. The model matches experimental results and exposes the underlying factors driving the polymerase chain reaction.
7. Model parameters were determined (Table 1.5) that can be used in future experiments. Some variation is possible for the values of k_c, k_c^* and η_{dE} when different polymerases are used. It is expected that k_p and β will remain constant for many different experiments.
8. Using the model parameters, the full mathematical model was simplified to one that could easily be implemented if a conservative PCR protocol was used.
9. Using model predictions, many PCR reactions can be simulated to find the optimal PCR protocol. This will allow increased throughput of PCR assays.
10. Functions relating the initial DNA concentration to the midpoint cycle number (similar to those first implemented by Higuchi et al., 1993) were created on a fundamental basis, and found to correlate well with experimental data. This can be used to quantify the initial amount of DNA in a sample.

Comment: Two factors may affect the average rate of extension, V. Firstly, the dinucleotide triphosphate (dNTP) concentration may become depleted; in which case the extension rate becomes dependent on the rate of diffusion of dNTPs to the ternary complexes. Secondly, pyrophosphates (PP_i) are produced upon insertion of dNTPS and their concentration builds up in the system. It is possible that a point may be reached where the pyrophosphorolysis reaction could effectively compete with dNTP insertion resulting in slow (if any) net extension. These factors can be accounted for by making V

dependent on PP_i and dNTP transport. Secondly, primer-dimer interactions are often problematic and one will have to resort to numerical solutions to account for the effect. The best alternative, if one wishes to use the analytical results presented here, is to assign a loss factor for primers at each cycle, similar to the polymerase and template losses due to thermal damage.

Though some observations from this model (such as the shift in the curves due to shortened elongation time or reduced polymerase) can be intuitive for scientist familiar with PCR, this model uncovers the underlying efficiencies that are affected by these changes. This unique understanding of the controlling factors of the reaction will aid in optimization and analysis of the reaction.

CHAPTER 2

COMPUTATIONAL FLUID DYNAMICS MODELING OF THE LYSIS MICROREACTOR AND EFFICIENCY OF DNA CAPTURE

2.1 Introduction

As discussed in the preface, when considering the application of PCR to diagnosis of disease by identification of organisms by their nucleic acids, the quality and purity of the DNA greatly influence the outcome of the reaction (Rantakokko-Jalava and Jalava, 2002). Another potential complication is the competition that exists in the PCR mixture between template and primers at low copy numbers, making it imperative to isolate the maximum quantity of DNA possible (Viljoen et al., 2005). This is typically determined by the methods employed during cell lysis and DNA extraction.

Several strategies have been developed to increase the efficiency of the PCR for low copy numbers, including nested PCR, booster PCR, homo-primer PCR, micro-TAS and single molecule water in oil emulsion PCR (Nakano et al., 2003). However, these methods require multiple steps and/or additional reagents when compared to standard PCR; an already multifaceted reaction. In terms of increasing DNA yield obtained from cells in order to overcome such problems, as well as ensuring good quality of DNA containing few inhibitors; the most effective purification strategy has certainly been the use of silica resins to purify DNA (Boom et al, 1990; Tian et al, 2000). These silica-membrane-based resins may be used as is, or with additional sample processing, but it has been shown that

there exist restrictions on applying any single method to all clinical samples (Rantakokko-Jalava and Jalava, 2002).

The challenge therefore still remains to develop an effective cell lysis procedure with optimized DNA recovery for analysis by PCR that is simple, fast and sensitive. There is a need to process clinical samples quickly and effectively, with minimal infrastructure demands and this need goes beyond the diagnosis of any specific infectious disease.

In this chapter, a model of the fluid flow in the LMR is presented and strategies for optimizing DNA capture are explored.

2.2 Computational fluid dynamics modeling of the LMR and DNA capture

The computational fluid dynamics (CFD) modeling of the LMR was performed with the commercial software FLUENT (ANSYS Inc.). LMR lysis harnesses three mechanisms: shear flow (stirrer), thermal action (heating), and chemical action (lysis buffer). The mechanical mixing enhances contact between the lysis buffer and cells, and the shearing action expedites the processing of samples with a variety of constitutive properties (particulate matter or viscoelastic fluids). The impeller rotates at ~70 cycles per second.

The x and y components of the velocity vector are approximately ten times larger than the z component (in absolute values), as shown in Figure 2.1. A periodic steady state is reached after approximately 4,000 iterations. The variability that is superimposed on the

periodic behavior is due to the turbulence; it is more notable in the case of the *z* component (Figure 2.1C).

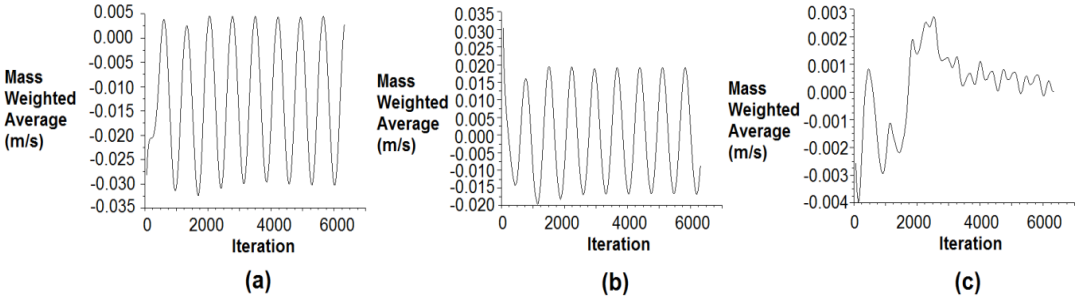

Figure 2.1. Mass weighted average of the (A) x (B) y and (C) z components of the velocity vector with increasing number of iterations of the solution. An appropriate time step size was selected so that the motor position is analyzed for every 10˚ of rotation. The time step size is 0.397 ms since the motor is rotating at 70 RPS. 20 iterations were performed per time step.

In Figure 2.2 side and top views of the trajectory of a DNA molecule is shown at two points in time. The molecule is in a coiled state, and for the purpose of the model it is approximated as a sphere. The DNA molecule flows along streamlines and the proximity to the strip depends on its initial position. The sample trajectory in Figure 2.2 shows a DNA molecule that flows within close proximity of the strip. The polystyrene is hydrophobic and it binds the hydrophobic bases of the DNA molecule when the DNA molecule is in single stranded form. The binding force is assumed to be inversely proportional to the square of the distance. The single stranded DNA molecules are drawn

towards the polystyrene surface, crossing streamlines in the process, when they are sufficiently close to the strip. If DNA is released from a lysed cell, the molecule may swirl around several times in the LMR before the DNA becomes captured – this scenario is shown in Figure 2.2C, D.

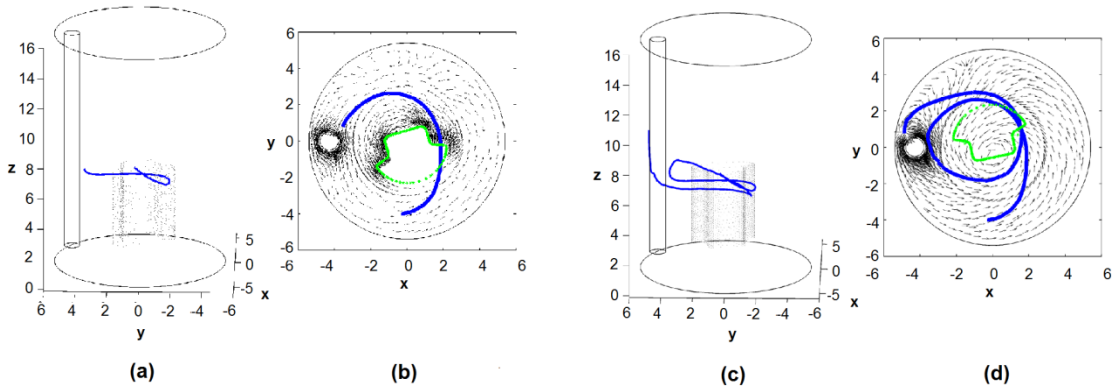

Figure 2.2. (a) Side view of DNA trajectory over 39.7 ms. (b) Top view of LMR with velocity vectors at z = 8.5 mm and t = 11.5 ms. (c) Side view at a later time t = 35.3 ms. (d) Top view at t = 35.3 ms.

It becomes clear from the discussion and the results of Figure 2.2, that competition exists between the attraction of DNA that are close to the strip and the turn speed of the impeller which determines the viscous drag forces on the molecule. At higher impeller speeds the efficiency of binding, which one can express as a probability to bind, is smaller, but more binding opportunities are presented over the same period of time. At the other extreme, if turn speed approaches zero, then the binding process will occur on the time scale of the diffusion process. Therefore one can argue that an optimal impeller speed and an optimal strip location exist that will maximize the binding rate. However,

the CFD results also revealed that molecules can be trapped without ever binding to the strip. Specifically the region directly behind the impeller has lower pressure and molecules tend to be drawn into the wake where they can remain indefinitely.

The fluid flow can be viewed as an operator that maps DNA molecules from an input position into an output position at the end of the simulation period; we defined ten impeller rotations as one mapping period. It must be noted that the flow approaches a mean steady state with noise due to turbulence superimposed on it. One can take the output positions and re-map them over a second simulation period, expecting some small changes due to turbulence. The mapping procedure is constructed as follows. The $Y \times Z$ plane at $X = 0$ is divided into a grid of 20×30 and the center of each block is the initial position of a DNA molecule. If we define the grid positions as $(i, j) \in [1,20] \times [1,30]$, then each molecule can be uniquely identified by its initial condition: $k = 30(i - 1) + j$. Therefore the vectors Y_k^0, Z_k^0 denote the initial y and z positions of the k^{th} molecule. The trajectories of all 600 particles are tracked over a period of ten rotor revolutions (one mapping period). The rotor turns at 70 revolutions per second, therefore the simulation time for one mapping is $t_{map} = 143\ ms$. During this mapping period some particles are captured by the strip and others are still in the fluid phase. Of the molecules in the fluid phase, some may enter the domain of attraction of the strip and get captured, but others may get trapped in the low pressure region behind the impeller. The times and positions of the particles are recorded at the moment they cross the $Y \times Z$ plane at $X = 0$ for the last time in a mapping period (the mapping times can differ from one particle to another

because some may be trapped behind the rotor or get captured on the strip before the end of the period t_{map}, but those in the fluid phase will cross the reference plane at time $t \leq t_{map}$). The new positions and mapping times are calculated for each initial position in vectors $L(k)$ and $\Gamma(k)$ respectively. After the first mapping, the time-to-capture vector is $T(k) = \Gamma(k)$. The new position vector $k_{New} = L(k)$ can serve as input for the next mapping $L(k_{New})$ and the mapping times are updated time: $T(k) = T(k) + \Gamma(k_{New})$. Thus the positions of all the particles and the total time-to-capture are known after n mappings for each particle.

Repeated mappings will converge to a subset of *k* on which the operator *L* is compact. A stochastic component is introduced through the use of changes to the mapping as follows:

$$Y_k^{n+1} = L\,(Y_k^n) + \delta_y \tag{2.1a}$$

$$Z_k^{n+1} = L\,(Z_k^n) + \delta_z \tag{2.1b}$$

The variables δ_y and δ_z are elements of the white noise probability distribution with zero correlation, i.e. $\delta_i \delta_j = 0,\ i \neq j$ which are superimposed on the mean velocities.

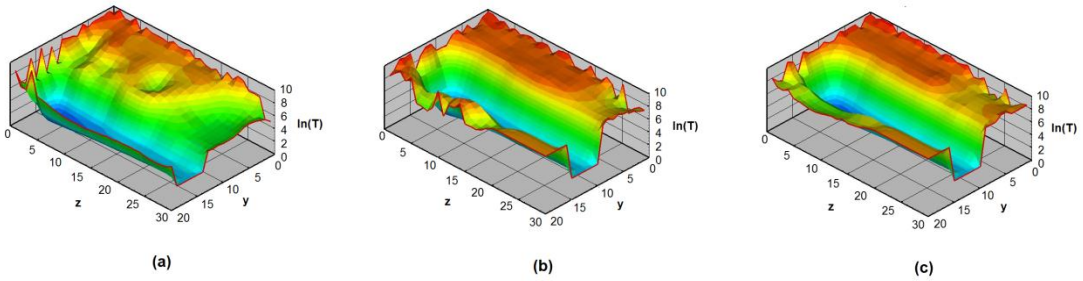

Figure 2.3. Plots of time-to-capture as a function of initial position. (a) Strip is located near perimeter, (b) strip is located near center and (c) strip is located near center, rotor speed is halved.

In Figure 2.3 the natural logarithm of the time-to-capture, $\ln(T(Y,Z))$, is plotted as a function of initial position. In Figure 2.2A the strip is located near the edge of the LMR (inner radius $= 5.4\,mm$), the center of the strip is at $Y = 4.3\,mm$ from center, it is $1\,mm$ from the bottom and it extends beyond the fluid level. Quick capture occurs in the trough area, colored blue in the figure. The capture near the bottom of the trough is slightly faster than near the top. Of the DNA molecules that are released in the center region of the LMR, the ones near the top of the LMR are more effectively captured; notably the green areas. Capture times slow down as one gets closer to the rotor, the red colored areas around the rotor indicate trapped areas from which DNA do not easily escape. Of the 600 initial positions, 245 positions lead to capture, the remainder remains trapped in the LMR. In Figure 2.3B the strip has been moved inwards to the position $Y = 3.3\,mm$. The trough area has now also shifted inwards, and a slow-capture region

appears near the top, close to the perimeter of the LMR. The strip position influences the flow field and the no-capture area that was limited to the rotor proximity before, now extends higher. When the rotor speed is halved, with the strip still at $Y = 3.3\,mm$ (Figure 2.3C), the no-capture areas are smaller. The total number of positions that lead to capture increases from 229 in the case of Figure 2.3B to 249 in the case of Figure 2.3C.

The low pressure region in the rotor's wake traps molecules and if capture can be performed in this area, then the efficiency will improve.

To demonstrate the effectiveness of wake capture, we assumed that the back of the rotor is coated with polystyrene, thus also enabling DNA capture. The strip was located near the perimeter as for Figure 2.3A. The results are shown in Figure 2.4. Not only does the former 'dead zone' become a capture area, but the time-to-capture is reduced for all the other positions. The slowest capture still occurs for positions directly above the rotor, but particles at these positions eventually get captured as well. The efficiency becomes 100%; all initial positions lead to capture events.

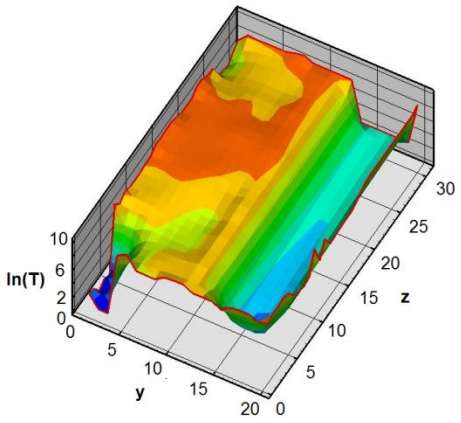

Figure 2.4. Time-to-capture as a function of initial position with capture areas located on the strip and the rear of the rotor.

2.3 Conclusions

The lysis of clinical specimens and the extraction of DNA molecules from the lysate are important steps that precede any nucleic acid amplification test. An analysis of the lysis process and the capture of DNA molecules from the lysate has been presented. The following major conclusions can be drawn from this study.

1. The flow characteristics of a lysis microreactor have been solved by computational fluid dynamics. For the configuration of our LMR, the flow is turbulent at a rotation speed of 70 revolutions per second. The trajectories of DNA molecules have been calculated for different starting positions. The probability to be captured by the inserted strip depends on the initial position. Initial position towards the lower center

of the LMR proves less effective; these molecules tend to get trapped in the rotor's wake.

2. The number of molecules that become captured decreases slightly (< 10%) when the strip is moved from near the perimeters to closer to the center. Slowing down the rotor speeds also improves the capture efficiency.
3. All DNA molecules can potentially be captured from the lysate if the capture area is located in the rotor wake.

CHAPTER 3

PRELIMINARY STUDY ON THE BINDING OF OLIGONUCLEOTIDES TO SURFACES

3.1. Introduction

In the preface we discuss the use of a polystyrene strip to non-covalently bind DNA from a solution. As mentioned previously, the strip serves two purposes (1) it captures lysed DNA from the lysate, thereby concentrating the DNA from the solution and (2) it provides a simple and effective transfer method to move the DNA through a wash step (allowing for the reduction of PCR inhibitors) and then to the thermocycler cuvette.

Although the polystyrene strip does effectively capture DNA from solution in the micromixer (refer to Chapter 4 for additional details), it is a non-specific capture method and may capture many different types of DNA and possibly also proteins and other chemicals present in the clinical specimen. The ideal capture method will be very specific for the molecule of interest, which in this case is the target DNA. Therefore, an improvement of the polystyrene strip method would be to selectively bind only the target DNA from solution. This may be most reliably accomplished by covalent attachment of an oligonucleotide probe to the capture surface (Zammateo et al., 2000). The oligonucleotide will be complimentary to the target DNA, and subsequent hybridization of the target DNA to this oligonucleotide will allow for selective capture.

A multitude of surfaces exist for the attachment of biomolecules and an equivalently large number of chemistries exist to bind oligonucleotides to these surfaces (Pirrung, 2002). In selecting a surface, it was our aim to keep both the current application in mind and also to select a surface that lends itself to future applications and possible technology expansions in our research area of interest. Because of its wide applicability and for the reasons mentioned previously, glass was selected as the substrate of choice. In the context of bioconjugation, the surface hydroxyl groups on glass must be converted to alternative functional groups in order for the surface to be reactive (Hermanson, 2008). The most commonly accepted method of modification is silanization of the glass surface, through which almost any functional group can be introduced (Pirrung, 2002). Aminosilane is widely available and amine functional groups are frequently employed and their behavior generally well understood in the context of DNA and protein chemistry. For these reasons (and others, discussed in detail in section 3.3.2B), aminosilane was selected as the silane of choice for our application.

Much work has already been done on the attachment of oligonucleotides to glass substrates (Walsh et al., 2001; Zammatteo et al., 2000). In the work of Zammatteo et al. (2000) a 255bp (double stranded) DNA fragment was attached using several different chemistries. Phosphorylated and carboxylated DNA was attached to aminated glass slides via aminosilane treated glass surfaces. Aminated DNA was attached to carboxysilane treated glass surfaces and aldehyde-functionalized glass surfaces. The aldehyde functionalized surface was selected as the superior surface due to the high hybridization efficiency, low level of non-specific binding and reproducibility of results. The authors

discuss the immobilization efficiency quantitatively and obtain the maximum hybridization efficiency at an oligonucleotide density of 200 fmol/cm^2 ($1.2x10^{11}$ molecules/cm^2, Zammatteo et al., 2000). Data on the hybridization efficiency is, however, only presented qualitatively and the authors also mention that a loss of aldehyde function (by conversion to carboxyl) is seen upon storage of the slides. The work of Walsh et al. (2001) also deals with the comparison of different immobilization chemistries. A 20 bp amine-functionalized oligonucleotide is attached to glass surfaces. The surfaces are treated with aminosilane, followed by additional chemistry to convert the surface amine group either to a carboxyl (succinylated or PEG-modified) or isothiocyanate group. They found that a one-step EDC procedure on carboxyl-functional surfaces at pH 4.5 resulted in the highest immobilization efficiency (82-89%, or $0.9x10^{13}$ molecules), with a corresponding hybridization efficiency of 58% (Walsh et al., 2001). These works, however, do not address the application of the technology in the context of clinical specimen handling. Also, hybridization efficiencies reported using simple silanization chemistries are typically low and no mention is made of the functional group densities present on the glass before immobilization of oligonucleotides. Hong et al. (2005) were able to produce hybridization efficiencies of between 80 and 100% (refer to Table 3.6) using a cone-shaped dendron molecule that results in mesospacing (around 3 nm separation) between oligonucleotides attached to the surface. The dendron will be synthesized by preparation of various generations, adding an additional level of complexity to surface immobilization.

In this chapter, we describe the binding of oligonucleotides to cylindrical glass rods with the aim at selective capture of a target DNA sequence. We use chitosan beads as a model substrate to explore various immobilization strategies for oligonucleotides. We then use the knowledge gained from these experiments to bind oligonucleotides to glass substrates. We investigate the binding and hybridization efficiencies obtained after performing various pre-treatments of the glass substrates. We also compare the efficiency of using a high cost linker with a low cost linker. Most importantly, we achieve much higher hybridization efficiencies than reported in literature with simple silanization chemistries by addition of Mg^{2+} in both the immobilization and hybridization buffers for glass rod experiments.

3.2. Materials and Methods

3.2.1A Materials for bead production and immobilization

DI water of resistivity 18MΩ·cm was used throughout all experiments. Acid soluble chitosan was purchased from Vansom Inc (Seattle, WA). Carboxymethyl chitosan was purchased from Shanghai Rogone International Trade Co., Ltd. (Shanghai, China). Bis[sulfosuccinimidyl] suberate (BS^3), 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC), N-hydroxysulfosuccinimide (sulfo-NHS) and glacial acetic acid were all purchased from Thermo Fisher Scientific Inc. (Rockford, IL). Anhydrous ethanol (ACS grade, 200 proof) was used. $CaCl_2 \cdot 2H_2O$ (ACS grade), $MgCl_2$ (ACS grade) and hypodermic needles (26G and 25G) were from VWR International LLC. (West Chester, PA). Trizma hydrochloride solution, Tween-20, sodium dodecyl sulfate (SDS, ACS grade), sodium phosphate dibasic (BioXtra, ≥99.0%) and 2-(*N*-

morpholino)ethanesulfonic acid (MES sodium salt) were all from Sigma-Aldrich Co. LLC. (St Louis, MO.). Oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA).

3.2.2A Oligonucleotide design

Oligonucleotides were designed to be complimentary to the *C difficile* tcdB gene. We had previously identified a set of primers that provided reliable amplification of a region of this gene (see Chapter 4 and van den Berg et al., 2006). To allow for maximum binding selectivity, a 100 bp probe was designed from the region spanned by the previously identified primers. Primers were designed with PrimerQuest online oligonucleotide design tool and checked for hairpin formation and self-dimerization using the online tool OligoAnalyzer 3.1 (both tools from http://www.idtdna.com, 2012). The oligo with the lowest level of hairpin formation and self-dimerization was selected. The maximum values for the change in Gibbs free energy for the final probe (Table 3.1) were as follows: ΔG_{dimer} = -10.94 kcal/mole and $\Delta G_{hairpin}$ = -3.05 kcal/mole. The probe was modified (either with an amine or phosphate group) at the 5' end for immobilization. This provides additional flexibility in that the 3' end of the oligonucleotide will be free to participate in a PCR reaction, should this additional flexibility be desired. A complimentary target oligonucleotide (100 bp) was designed with a Cy5 fluorophore on the 3' end to verify hybridization by visualization under the confocal microscope.

Table 3.1. Oligonucleotide sequences used in this study

Name	Sequence
OGN-1-NH_2	5'-NH_2-C6-GA TTA CCT ATA ATT GCA ACT ATT ATA GAT GGT GTA AGT TTA GGT GCA GCA ATC AAA GAG CTA AGT GAA ACG AGT GAC CCA TTA TTA AGA CAA GAA ATA GA-3'
OGN-1-PO_4	5'-PO_4-GA TTA CCT ATA ATT GCA ACT ATT ATA GAT GGT GTA AGT TTA GGT GCA GCA ATC AAA GAG CTA AGT GAA ACG AGT GAC CCA TTA TTA AGA CAA GAA ATA GA-3'
OGN-2	5'-TCT ATT TCT TGT CTT AAT AAT GGG TCA CTC GTT TCA CTT AGC TCT TTG ATT GCT GCA CCT AAA CTT ACA CCA TCT ATA ATA GTT GCA ATT ATA GGT AAT C-Cy5-3'

3.2.3A Bead production

Beads were produced using the experimental setup depicted in Figure 3.1. The chitosan was placed in a 10 mL capacity syringe and a KDS 100 syringe pump (KD Scientific Inc., Holliston, MA) was used to drive the chitosan solution through the needle to form chitosan beads that were then captured in the curing solution. Two syringes were filled and consecutively processed per production run and chitosan beads were collected in 200 mL curing solution contained in a 500 mL glass beaker. Bead sizes below 1 mm were ideally required for binding experiments and the viscosity of the chitosan did not allow for the use of a needle small enough to produce such small beads. Consequently, an XA EF 200 air atomoizing nozzle (fluid cap FC4 and air cap 1003, BETE Fog Nozzle Inc., Greenfield, MA) was used to exploit the principle of Raleigh instabilities thereby producing droplets of smaller sizes than those that could be obtained using only the syringe pump setup. Detailed experimental conditions for each type of bead are listed in sections 3.2.3.1A and 3.2.3.2A.

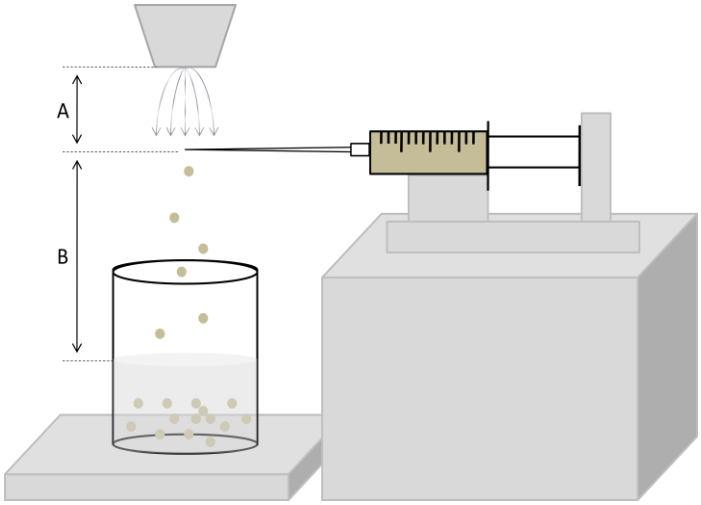

Figure 3.1. Experimental setup for bead production.

3.2.3.1A ASC beads

Acid soluble chitosan was dissolved at 1.75% (w/v) in 1% acetic acid solution. A 26G needle was used. The curing solution comprised 70% of a 0.75 M NaOH solution and 30% anhydrous ethanol. The distances on the experimental setup were as follows: (A) 20 mm and (B) 70 mm. Air flow rate was set to 12 SLM and chitosan flow rate was at 20 mL/hr. Beads were allowed to cure for 1 hour with light stirring and washed 3 times with 10 mMTris·HCl buffer at pH 7.0 before being stored at 4^{o}C.

3.2.3.2A CMC beads

A solution of 10% (w/v) carboxymethyl chitosan was prepared in DI water. A 25G needle was used. The curing solution contained 5% (w/v) $CaCl_2 \cdot 2H_20$ and 30% anhydrous ethanol. The distances on the experimental setup were as follows: (A) 20 mm and (B) 60 mm. Air flow rate was set to 12 SLM and chitosan flow rate was at 10 mL/hr.

Beads were cured for 1 hour, washed 3 times with 10 mM Ca^{2+}-Tris·HCl buffer at pH 7.0 and stored at 4^{o}C.

3.2.4A Oligonucleotide binding to beads

3.2.4.1A Binding of OGN-1-NH_2 to ASC beads and pH study

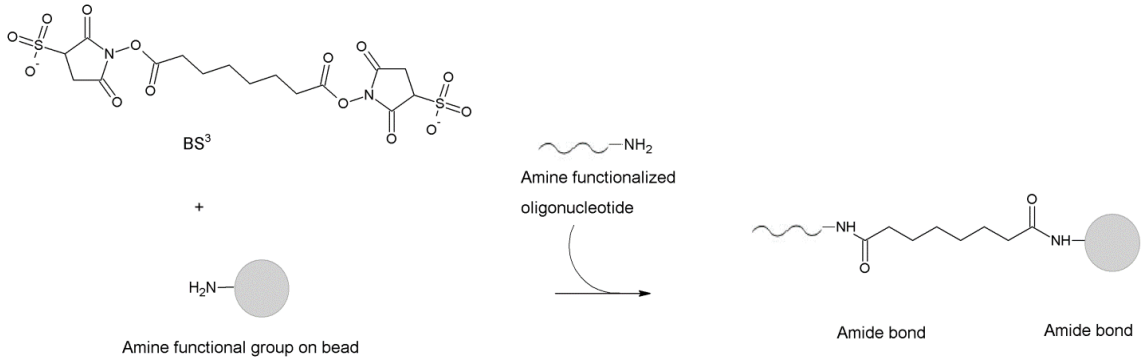

Figure 3.2. Schematic of chemistry used to bind oligonucleotides to ASC beads.

The chemistry used to bind amine-functional oligonucleotide to amine-functional groups on ASC beads is outlined in Figure 3.2.

OGN-1-NH_2 binding via a BS^3crosslinker was investigated at three different pH values (7.0, 7.5 and 8.0) to determine the optimum pH. 160 µL beads were used per reaction tube. Sodium phosphate buffer was prepared at each of the listed pH values. OGN-1-NH_2 was available as a 'lab-ready' mixture in 1X Tris-EDTA buffer. The reaction mixture was prepared in sodium phosphate buffer at of the three pH values and contained 510 µg of OGN-1-NH_2 and BS^3crosslinker at a final concentration of 40 mM. The reaction was allowed to proceed at room temperature in the absence of light. After 30 min, an

additional amount of BS^3 crosslinker was added to bring the final concentration of BS^3 to 80 mM. The reaction was allowed to take place for an additional 30 min. The beads were washed in 0.02% Tween-20, followed by 2 washes of 0.1% SDS and 3 washes of sodium phosphate buffer at the reaction pH. All washes were retained to determine the amount of oligonucleotide bound. The control reaction was performed with all buffers and reagents, but without the addition of OGN-1-NH_2.

Unreacted amine groups were blocked according to the method described in section 3.2.5A.

3.2.4.2A Binding of OGN-1-NH_2 to CMC beads

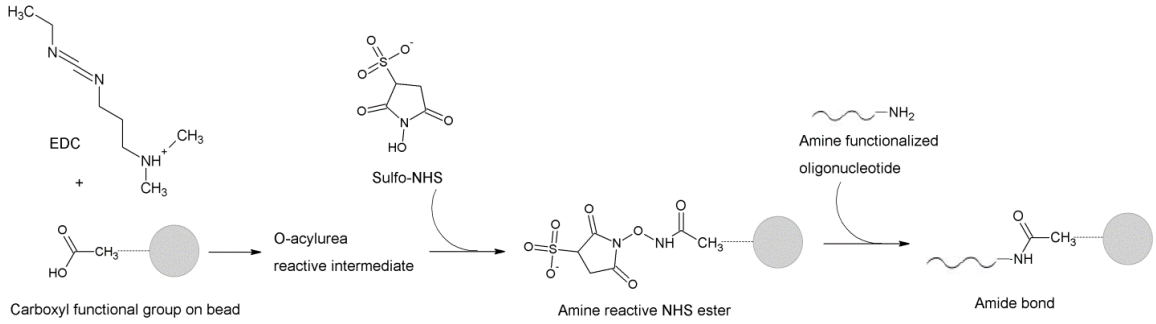

Figure 3.3. Schematic of chemistry used to bind oligonucleotides to CMC beads.

The chemistry used to bind amine-functional oligonucleotide to carboxyl-functional groups on CMC beads is outlined in Figure 3.3. CMC beads (160 µL per tube) were equilibrated for 10 min in Ca^{2+}-MES buffer (100 mM MES, 5% (w/v) $CaCl_2$) at pH 5.0. OGN-1-NH_2 was available as 'lab-ready' mixture. Ca^{2+}-MES buffer was removed and replaced with the reaction mixture containing 510 µg of OGN-1-NH_2 and both EDC and

sulfo-NHS at a final concentration of 100 mM. The reaction mixture was brought to the final volume of 800 µL by addition Ca^{2+}-MES buffer at pH 5.0. The reaction was allowed to proceed at room temperature for 1 hour in the absence of light. Beads were washed once with 0.02% Tween-20, twice with 0.1% SDS and three times with Ca^{2+}-MES buffer. As a control, the reaction was completed in all buffers and reagents, but without the addition of OGN-1-NH_2.

3.2.5A Blocking of unreacted amine groups by acetylation

Acetic anhydride was used to block unreacted amine groups on the surface of chitosan substrates by the method described elsewhere (Hermanson, 2008). A 10 fold molar excess of acetic anhydride was used with respect to the amount of unreacted amines estimated. Typically 3-5 µL of acetic anhydride was added per reaction (20 uL of beads per reaction). Beads were equilibrated in 10 mM acetate containing 5% (w/v) $CaCl_2 \cdot 2H_20$ at pH 5.0 for 10 min. Acetic anhydride was added at 20 minute intervals in aliquots of 33.3% of the total amount. The reaction was allowed to proceed with mixing for 1 hour from the first addition of anhydride. Chitosan beads were stored in 10 mM acetate buffer at pH 5.0.

3.2.6A Hybridization of OGN-2-Cy5

A blocking buffer was prepared containing 10 mM acetate, 0.02% SDS, 1% (w/v) BSA, 5% (w/v) $CaCl_2$ and 2.5 mM $MgCl_2$. The OGN-1 bound beads were incubated for 30 min at 37^oC with the blocking buffer. The blocking buffer was removed and replaced with 50 µg of OGN-2-Cy5 complimentary oligonucleotide and fresh blocking buffer. Tubes

were incubated horizontally in an incubator-shaker at 37^{o}C and 100 rpm for 3 hours in the absence of light. The beads were washed four times with wash buffer (10 mM acetate, 5% (w/v) $CaCl_2$ and 25 mM $MgCl_2$).

3.2.7A Oligonucleotide binding efficiency determination

To determine the amount of OGN-1 bound in each reaction, the amount of OGN-1 that was left unreacted in the subsequent washes was determined and subtracted from the amount of OGN-1 supplied for reaction. Single stranded DNA concentration was estimated by recording absorbance at 260 nm using UV-vis spectroscopy. It was previously determined that the reaction components interfere with this measurement (data not shown), and therefore it was necessary to purify the DNA from the reaction solution.

A 10 kDa Amicon Ultra MWCO membrane was used to perform the required purification. Sufficient purification could not be achieved using only a single spin (data not shown), so each wash was processed multiple times (using a single MWCO membrane multiple times per wash) to obtain a sufficiently clean product. Consequently, it was necessary to quantify the loss at each consecutive spin and also to determine the approximate number of spins required to sufficiently clean the product. This was accomplished by using reaction mixtures containing known initial amounts of OGN-1 mixed with other reagents and also with OGN-1 in DI water only. Measurements were taken before and after each spin to construct a curve of OGN-1 retention vs spin number (data not shown).

3.2.8A Confocal imaging and FTIR analysis

3.2.8.1A Confocal imaging

Following hybridization, beads were viewed on a Nikon A1 confocal laser scanning microscope mounted on a Nikon Eclipse 90I compound microscope at 10X magnification. Cy5 excitation was at 641 nm and emission in the range of 662-737 nm. Both the pseudo-colored red fluorescence image and the transmitted light images were viewed for all experiments.

3.2.8.2A FTIR

Pellets of pressed ASC and CMC with KBr were used in approximately 1:10 ratio. A Thermo Nicolet Avatar 380 FTIR was used to obtain the spectra using 128 scans per sample. For functionalized beads, IR spectra were collected with a Thermo Nicolet Avatar 360 w/ SmartPerformer ATR accessory. The reason for this was the reduced sample volume of functionalized beads available. This latter instrument is equipped with a zinc selenide crystal. Again, 128 scans per sample were used. Both machines provide a 4 nm resolution of data.

FTIR data were baseline corrected and the area under the curve was normalized to 1 over the entire spectrum to correct for concentration differences. Reference spectra (chitosan beads) were subtracted from sample spectra (oligonucleotide immobilized on chitosan beadswith negative peaks indicating loss of features from the reference state and positive peaks indicating gain of features.

3.3.1B Materials for glass immobilization

DI water of resistivity 18MΩ·cm was used throughout all experiments. Borosilicate glass rods were purchased from Fiberoptics Technology Inc. (Pomfret, CT). 1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC), Imidazole, Bis[sulfosuccinimidyl] suberate (BS^3), Fisherbrand Hellmanex II and 3-Aminopropyltriethoxysilane (APTES) were all purchased from Thermo Fisher Scientific Inc. (Rockford, IL). Anhydrous ethanol (200 proof, ASC grade) was used. Acetone (ACS grade), sulfuric acid (ARISTAR®; ACS, FCC Grade), toluene (anhydrous) and methanol (anhydrous) were purchased from VWR International LLC. (West Chester, PA). Hydrogen peroxide (H_2O_2; TraceSELECT® Ultra), Tween-20, sodium dodecyl sulfate (SDS; ACS grade), sodium phosphate dibasic (BioXtra, ≥99.0%), 4-nitrobenzaldehyde (98%) and $MgSO_4$ solution (BioUltra, for molecular biology) were purchased from Sigma-Aldrich Co. LLC. (St Louis, MO). Oligonucleotides were purchased from Integrated DNA Technologies, Inc. (Coralville, IA).

3.2.2B Sizing and cleaning of glass rods

Borosilicate glass rods with numerical aperture (NA) of 0.55 and original dimensions of 0.8 mm by 400 mm were cut to a size of 0.8 mm by 5 mm using a Dremel rotary tool. At least 200 pieces were prepared in this manner to ensure that an adequate number of substrates were available.

A 50 mL solution of 5% of Hellmanex in DI water was prepared in a PTFE beaker and the glass substrates were completely immersed in this solution and sonicated for 5 min at

room temperature. Following sonication, the substrates were rinsed ten times in 50 mL DI water. Samples were subsequently immersed in 50 mL of acetone and sonicated for 5 min at room temperature after which they were rinsed for 10 min on a shaking rotator in DI water. Glass substrates were then sonicated for 5 min in 50 mL ethanol and again rinsed in DI water on a shaking rotator for 10 min. Substrates were finally ultrasonicated in 50 mL of DI water and dried. Cleaned glass rods were used for subsequent reactions.

3.3.3B Pre-activation of glass surface by hydroxylation

Two PTFE beakers were prepared, each containing a different pre-activation solution. One beaker contained 50 mL of piranha solution (3:1 (v/v), H_2SO_4:H_2O_2) and the other 50 mL 65% HNO_3. A single layer of glass rods was placed in each of the beakers and incubated on a shaking rotator for 30 min. The pre-activation solution was removed and the beakers were each filled with 50 mL of DI water and rinsed for 15 min on a shaking rotator. The surfaces were subsequently dried under nitrogen flow for 3 hours to remove all traces of water and stored in a vacuum desiccator until silanization reaction.

3.2.4B Silanization of glass substrates

Silanization was performed under nitrogen purging according to the method previously described (Guha Thakurta, 2010). Briefly, two PTFE beakers each containing 50 mL of 2% (v/v) APTES solution in anhydrous toluene were prepared. Either piranha solution activated- or nitric acid activated glass rods were added to each of the beakers and reacted under light stirring for 30 min. Rods were washed under nitrogen atmosphere and light stirring with (1) toluene (2) toluene:methanol (1:1 v/v) and (3) methanol. Each wash was performed for 5 min and with a 50 mL volume. Rods were cured in a vacuum desiccator overnight before performing oligonucleotide binding reaction. Determination of the amine surface density resulting from silanization of glass rods was determined by the method described in section 3.2.8B.

3.2.5B Oligonucleotide binding to glass substrates

3.2.5.1B BS^3crosslinker

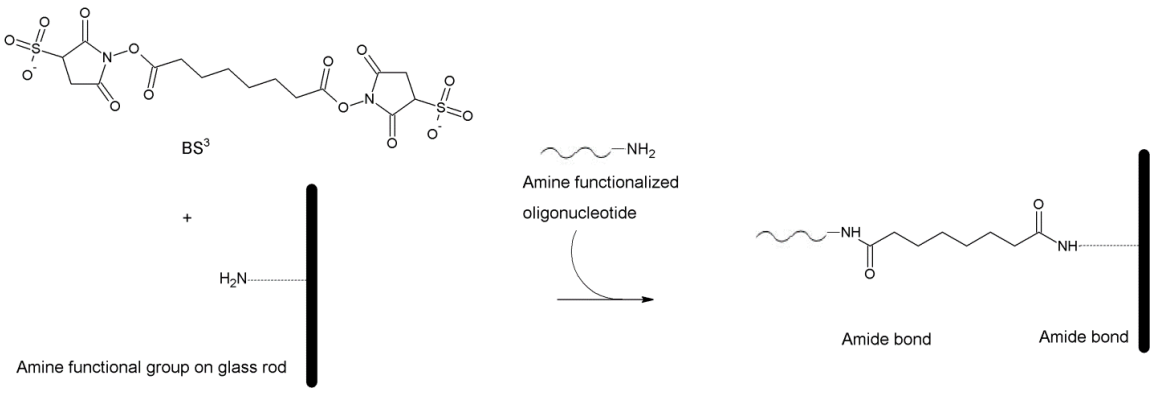

Figure 3.4. Schematic of chemistry used to bind oligonucleotides to aminosilane functionalized glass rods.

The chemistry used to bind amine-functional oligonucleotide to amine-functional groups on glass rods is outlined in Figure 3.4. Glass rods (20 per reaction) were equilibrated in 100 mM sodium phosphate at pH 7.5 for 10 min. The buffer was removed and replaced with the reaction mixture containing 9.2 µg of OGN-1-NH_2, 2.5 mM $MgSO_4$, and BS^3 crosslinker at a final concentration of 40 mM, all prepared in buffer at pH 7.5. The reaction was allowed to proceed with mixing for 30 min at room temperature in the absence of light. An additional amount of BS^3 crosslinker was added to bring the final concentration up to 80 mM, and allowed to react for an additional 30 min as before. Rods were subsequently washed in 0.02% Tween-20, followed by 2 washes of 0.1% SDS and 3 washes of sodium phosphate buffer at pH 7.5. All washes were retained to determine the amount of oligonucleotide bound in the reaction. The control reaction was performed

with all buffers and reagents, but without the addition of OGN-1-NH_2. Unreacted amine groups were blocked by the method described in section 3.2.6B.

3.2.5.2B EDC crosslinker

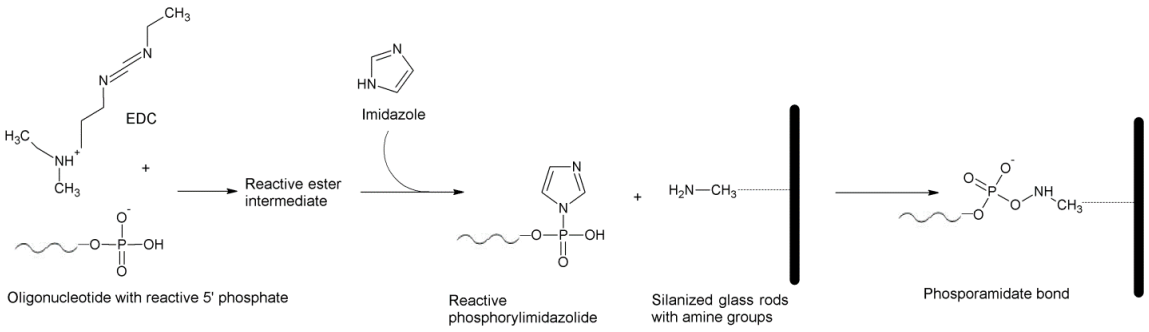

Figure 3.5. Schematic of chemistry used to bind oligonucleotides to to aminosilane functionalized glass rods.

The chemistry used to bind phosphate-functional oligonucleotide to amine-functional groups on glass rods is outlined in Figure 3.5. Glass rods (20 per reaction) were equilibrated in MES buffer at pH 6.5 for 10 min. OGN-1-PO_4 was available as 'lab-ready' mixture in 1x Tris-EDTA buffer. EDC was prepared as a 1M solution in HEPES buffer at pH 7.2 and imidazole was prepared as a 1M solution in MES buffer at pH 6.0 directly before reaction. MES buffer was removed and replaced with the reaction mixture containing 9.2 µg of OGN-1-PO_4, 2.5 mM $MgSO_4$, and both EDC and imidazole at a final concentration of 100 mM. The reaction mixture was brought to the final volume of 400 µL by addition of approximately 320 µL of MES buffer at pH 6.5.The reaction was allowed to proceed on a shaking rotator for 1 hour at room temperature in the absence of light. Rods were then washed in 0.02% Tween-20, followed by 2 washes of 0.1% SDS

and 3 washes of MES buffer at pH 6.5. All washes were retained to determine the amount of oligonucleotide bound. The control reaction was performed with all buffers and reagents, but without the addition of OGN-1-PO_4. Unreacted amine groups were blocked as described in section 3.2.6B.

3.2.6B Blocking of unreacted amine groups by acetylation

Unreacted amine groups were reacted with 10 fold molar excess (typically 4-10 µL per 20 rods) acetic anhydride as describes in section 3.2.5A. 100 mM sodium phosphate buffer at pH 7.5 was used for the reaction and rods were stored in this buffer following the acetylation reaction.

3.2.7B Hybridization of OGN-2-Cy5

The blocking buffer contained 10 mM acetate, 0.02% SDS, 1% (w/v) BSA and 2.5 mM $MgSO_4$. The OGN-1 bound beads were incubated for 30 min at 37^oC with the blocking buffer. The blocking buffer was removed and replaced with 9.2 µg of OGN-2-Cy5 complimentary oligonucleotide and fresh blocking buffer. Tubes were incubated horizontally in an incubator-shaker at 37^oC and 100 rpm for 3 hours in the absence of light. The beads were washed in buffers of increasing stringency and non-hybridized targets were also removed. A single wash was performed with wash buffer A (2X SSC, 0.1% (w/v) SDS, 2.5 mM $MgSO_4$). Two washes were performed with wash buffer B (0.2X SSC, 0.1% (w/v) SDS, 2.5 mM $MgSO_4$). And finally, two washes with only 2.5 mM $MgSO_4$ in DI water. Glass rods were prepared no longer than 12 hours before visualization and kept at 4^oC in 2.5 mM Mg^{2+} solution in the absence of light.

3.2.8B Amine density determination

Amine density from silanization of glass substrates was determined following the method described elsewhere (Guha Thakurta, 2010). The principle is based on the reversible reaction of 4-NBA with the aminosilane covered surface to produce an imine which can then be hydrolyzed to reproduce the 4-NBA. Imine formation was affected under nitrogen atmosphere with constant, light stirring to ensure maximum conversion. The silanized glass rods were submerged in a solution containing 10 mg of 4-NBA, 25 mL of anhydrous ethanol and 20 µL acetic acid (0.08%) and allowed to react for 3 hours at 50^{o}C. The rods were then sonicated three times in anhydrous ethanol, each for 2 min, to remove unreacted 4-NBA. The converted substrates were submerged in 5 mL DI water containing 10 µL acetic acid (0.2%) and allowed to react overnight at 50^{o}C. An aliquot of the reproduced 4-NBA solution was collected and absorbance at 260 nm was recorded and measured against a standard curve to determine the amine density on the surface. The standard curve was constructed for 4-NBA concentrations ranging from 0-50 µg/mL in DI water containing 0.2% acetic acid.

3.2.9B Oligonucleotide immobilization and hybridization efficiency

3.2.9.1B Immobilization efficiency

Immobilization efficiency was determined by the method described in section 3.2.7A.

2.7.9.2B Hybridization efficiency

The amount of OGN-2 hybridized by each substrate was determined by measuring the amount of OGN-2 that was left unhybridized in subsequent washes. This was then subtracted from the amount of OGN-2 supplied to the reaction. DNA was purified from protein components of the hybridization and wash buffers by the Qiagen MinElute PCR purification kit, according to the manufacturer's instructions. Single stranded DNA concentration was estimated by recording absorbance at 260 nm using UV-vis spectroscopy and the initial value was back-calculated by the estimated percentage loss (see section 3.2.7A for details).

3.2.10B Confocal imaging

Following hybridization, glass rods were viewed on a Nikon A1 confocal laser scanning microscope mounted on a Nikon Eclipse 90I compound microscope at 10X magnification. Cy5 excitation was and emission was as listed in section 3.2.8A. The bulk of imaging for the glass rods was performed on the base of the rod. Since the base of the rod was almost always angled (due to the method of cutting), it was section-scanned and the images merged to provide a view of the entire base of the rod.

3.3 Results and Discussion

3.3A Chitosan bead results

3.3.1A Bead production and stability

3.3.1.1A Physical principles of bead formation

It is known that for pre-filming air atomizers, the droplet size is most effectively controlled by adjusting the air flow rate, surface tension and viscosity (Richardson et al., 2002). It has been previously shown by our lab that the chitosan flow is significantly impeded at viscosity values above 800 cP. We kept the surface tension fixed by using a 70% curing solution with 30% ethanol. Our main method of controlling the bead size was by adjusting the air flow rate at the point where the air meets the chitosan. This was done empirically by setting the air flow rate and adjusting the distance A until the desired bead size (100 to 1000 μm) was obtained (refer to Figure 3.1 for details). The distance B and chitosan flow rate are the most important parameters in determining the droplet shape. If the chitosan flow rate is too large or the distance B is too short, the droplets will become disc-shaped. These values were also adjusted empirically to obtain a spherical bead shape.

3.3.1.2A Chemical principles underlying bead formation

Acid soluble chitosan contains reactive amine groups with a pKa of 6.5. Upon dissolution of chitosan in 1% acetic acid (pH 4), essentially all amine groups will be protonated (NH_3^+). A sudden drop in the pH, as is experienced upon contact with the NaOH curing solution, will result in precipitation of the bead from solution by the method of simple coacervation.

Carboxymethyl chitosan contains both reactive amine and carboxyl groups. For the purposes of the preliminary study, oligonucleotides were attached to the carboxyl functionalities. The pKa of the carboxyl groups on CMC is 3.5. At the pH of DI water (around 7); the carboxyl groups will be deprotonated (COO^-). Note that the pH difference is sufficiently large (>> 1 pH point) that the assumption can be made that essentially all carboxyl groups will be deprotonated. The result is that an ionic bond will form between the COO^- on one chitosan molecule and Ca^{2+} in solution. In turn, this Ca^{2+} will crosslink to the carboxyl on another chitosan molecule (ideally most crosslinks will be between carboxyl groups on different strands; clearly some interlinking can occur between the same strands).

ASC beads were not as mechanically robust and perfectly spherical as CMC beads. However, CMC beads are not chemically robust. The Ca^{2+} is easily sequestered from the bead by most buffers, forming a precipitate and care must therefore be taken in the selection of a buffer. Also, addition of Ca^{2+} is required at all times to ensure bead stability.

3.3.2A FTIR

As shown in Figure 3.6, both ASC and CMC FTIR results compare well with the literature (Kumar et al, 2007). The characteristic OH- and amine-stretch at 3450cm^{-1} and C-H stretch at 2900$cm^{-1}$are observed, together with 1650cm^{-1} amide I stretch, which confirms that molecules are not completely de-acetylated. C-O stretches of cyclic alcohol and primary hydroxyl are also confirmed at 1070cm^{-1} and 1030cm^{-1}, respectively.

Presence of amine is confirmed by the band around 1590cm^{-1}. Carboxymethyl chitosan has 2 peaks appearing at around 1570 and 1400 cm^{-1}, indicating the presence of carboxylate groups. Also, the results for ASC and CMC beads are qualitatively the same using either the KBr technique or the ATR accessory (compare Figures 3.6 A and B).

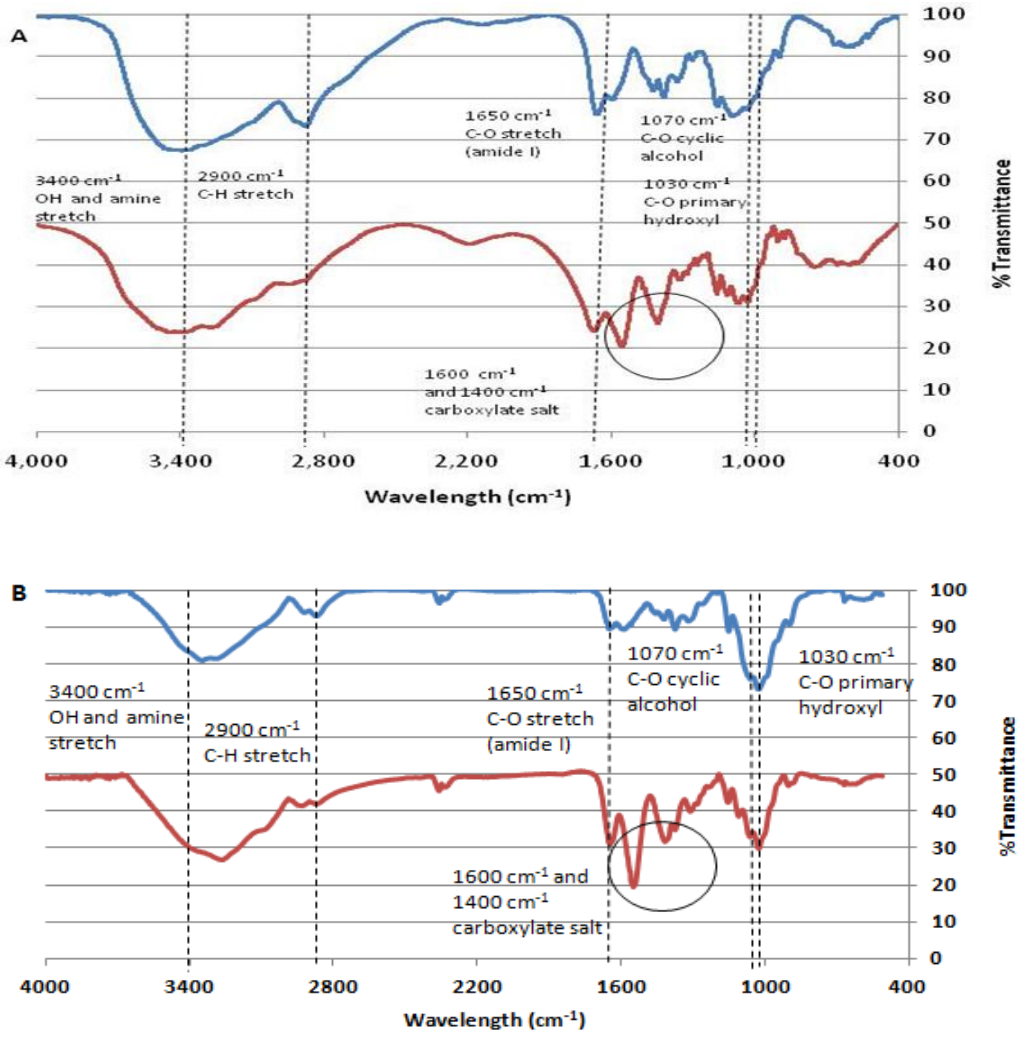

Figure 3.6. FTIR results of (A) ASC (top) and CMC (bottom) powder with KBr pellets vs. (B) ASC (top) and CMC (bottom) beads using ATR accessory

Difference spectra for ASC beads reveal an increase in the amine functionalities, as indicated by the increase of the intensity of the broad peak around 3400 cm^{-1} (Vogel, 1996). The appearance of a peak around 1100-1200 cm^{-1} indicates the presence of phosphate groups (Vogel, 1996). Both these peaks are consistent with the immobilization of oligonucleotide to the bead, since these functionalities are abundant within the DNA structure. The most prominent feature is, however, the large increase seen in the amide I and amide II bands (1650 cm^{-1} and 1550 cm^{-1}). The immobilization chemistry to attach oligonucleotides via BS^3 through ASC amine groups results in the formation of two amide bonds for each immobilized amine group (refer to Figure 3.2 and Figure 3.8B insert).

Difference spectra for CMC beads reveal an increase in the amine functionalities and the appearance of a prominent peak indicating the presence of phosphate groups. As for ASC beads, both the aforementioned peaks are consistent with the immobilization of oligonucleotide to the CMC bead. In the case of the CMC bead, however, we observe only on increase of the amide I band (1650 cm^{-1}). However, if we also consider the large reduction in peaks in the carboxyl reagion (1500-1600 cm^{-1} and 1400 cm^{-1}), we may conclude that the appearance of the amide II peak is masked by the reduction of carboxylate functionalities. This is consistent with the addition of oligonucleotides via EDC/Imidazole through CMC carboxyl groups that are consequently converted to amide groups (refer to Figure 3.3 and Figure 3.9B insert).

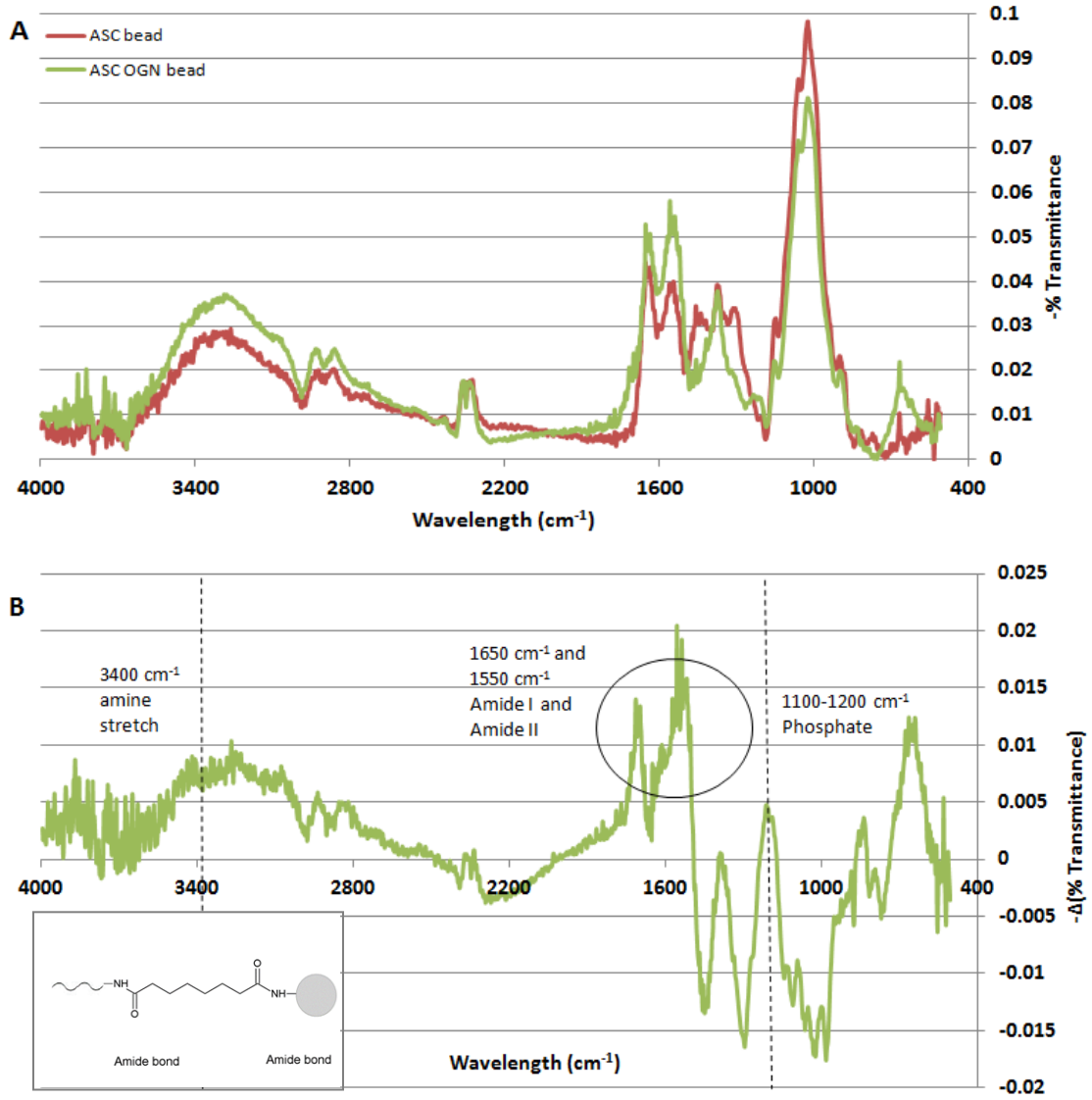

Figure 3.7. (A) Baseline corrected and normalized ATR results of ASC beads before and after oligonucleotide immobilization. (B) Difference spectra of the results shown in (A). Both spectra are inverted for comparison and ease of interpretation of difference spectra.

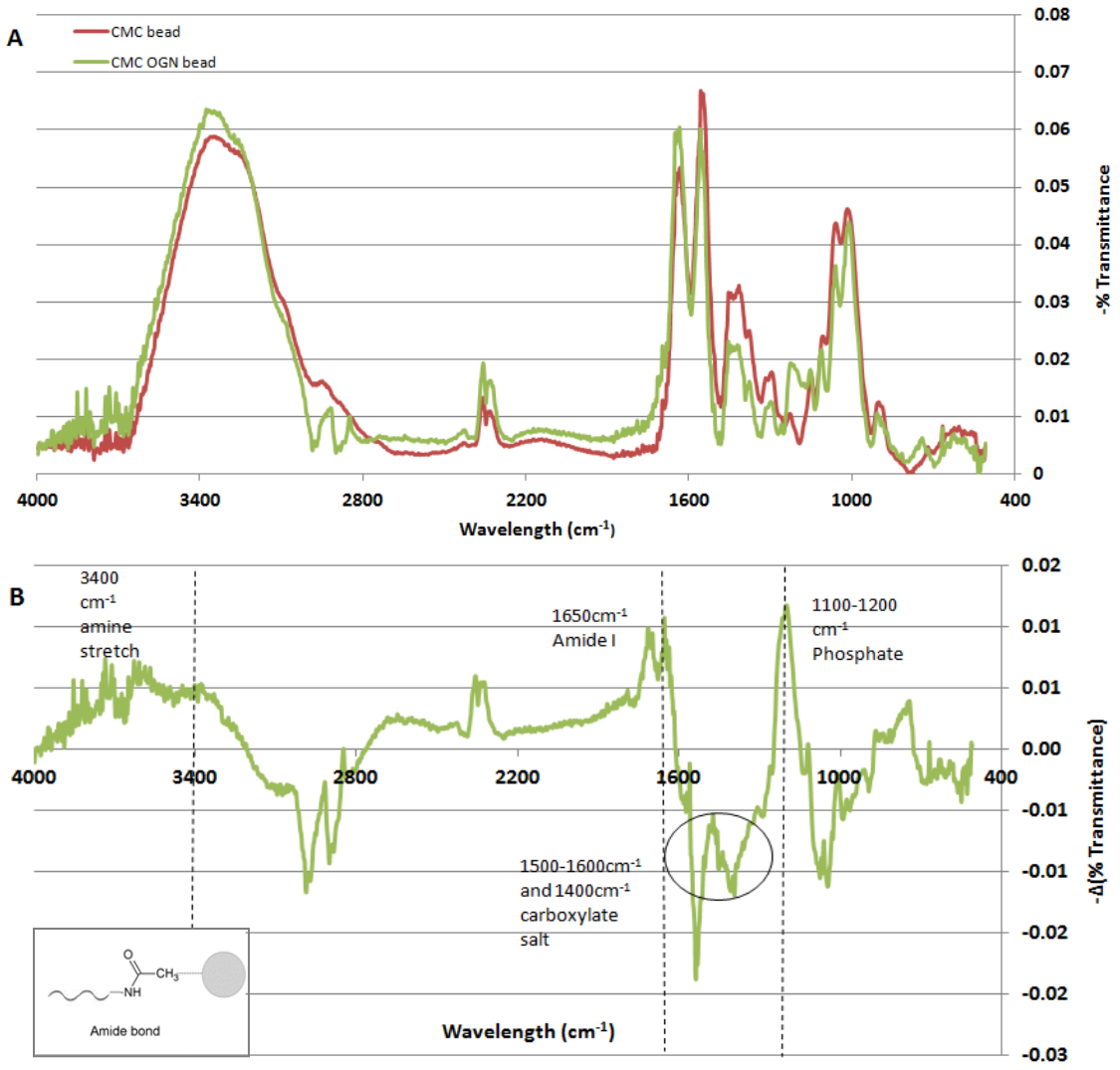

Figure 3.8. (A) Baseline corrected and normalized ATR results of CMC beads before and after oligonucleotide immobilization. (B) Difference spectra of the results shown in (A). Both spectra are inverted for comparison and ease of interpretation of difference spectra.

3.3.3A OGN-1 Immobilization efficiency

General properties of ASC vs. CMC beads are compared in Table 3.2. We observe that the amount of carboxyl molecules available for reaction on CMC beads after Ca^{2+} crosslinking is an order of magnitude less than the amount of amine molecules for reaction on ASC beads. This limiting condition provides a possible explanation for the lower level of immobilization achieved with EDC reaction chemistry as compared to BS^3 chemistry. However, the EDC reaction chemistry is clearly a possible alternative to BS^3 chemistry and was therefore selected for further investigation using glass substrates. It is important to note that although there is theoretically a large excess of reactive groups, it is highly likely that a large percentage of these groups will not practically be available for reaction, largely due to steric effects.

Table 3.2 Properties and calculations of ASC vs CMC beads

Property	ASC beads	CMC beads
Molecular weight	190 kDa	30 kDa
Degree of deacetylation	81.7%	85.5%
Degree of carboxylation	-	80%
Average MW of 1 subunit	186.86 g/mol	218.87 g/mol
Amine units in 1 unit chitosan	1016	116
Carboxyl units in 1 unit chitosan	-	109
Solids content naked beads	3-4%	10-11%
Weight of chitosan in 160 µL beads @ 3%	0.0048g	
(ASC) and 10% (CMC) solids content		
Moles of chitosan	$3.53x10^{-8}$ 1a	$5.3x10^{-7}$ 1b
Moles of amine in 160 µL beads	$2.57x10^{-5}$ 2a	$6.2x10^{-5}$ 2b
Moles of carboxyl	-	$5.8x10^{-5}$ 3
Moles of carboxyl available for reaction	-	$2.6x10^{-6}$ 4

1a,b Moles of chitosan calculation was done by using the solids content to determine the weight of beads and converting to moles by using the MW

2a,b Moles of amine were calculated by using the ratio of chitosan MW and avg MW of a single subunit to determine the amount of amine units per chitosan molecule, this value was multiplied with the moles of chitosan to obtain the moles of amine

3 Moles of carboxyl were calculated the same method as that used for moles of amine

4 Moles of carboxyl available for reaction were determined by measuring the amount of Ca^{2+} remaining in the curing solution by precipitation reaction with H_2SO_4, by comparison to a previously constructed standard curve. This was done in a previous run of bead-making which produced very large beads (data not shown). It was assumed that the ratio of Ca^{2+} crosslinking remained constant, regardless of bead diameter

Table 3.3 Bead immobilization results

Bead type	Reaction pH	Linker/ chemistry used	OGN-1 functionality / Bead functionality	OGN-1 in (µg)	OGN-1 out (µg)	OGN-1 bound (µg)
ASC	7.0	BS3	NH_2/NH_2	510	7±0.8[1]	503±0.8
ASC	7.5	BS3	NH_2/NH_2	510	2±0.3[1]	508±0.3
ASC	8.0	BS3	NH_2/NH_2	510	9±1.4[1]	501±1.4
CMC	5.0	EDC/ Sulfo-NHS	NH_2/COOH	510	217±17[1]	293±17

[1]Sample variance is purely a function of measurement variance; only a single immobilization run was performed for the preliminary study

The results of the pH study of OGN-1 immobilization of ASC beads using BS^3 are summarized in Figure 3.9. It is known that the rate of hydrolysis increases with buffer pH (http://www.piercenet.com, 2012). Conversely, the maximum deprotonation of amine groups will be obtained at higher pH values. The pH study was suggested in order to find the optimum balance between these parameters for our reaction system. The results indicate that there is indeed an optimum pH at a value of 7.5, where the maximum immobilization of OGN-1 to the surface is achieved. If we perform the immobilization of glass rods at this pH value, we predict that we can reliably functionalize a predetermined amount of OGN-1 to a glass substrate.

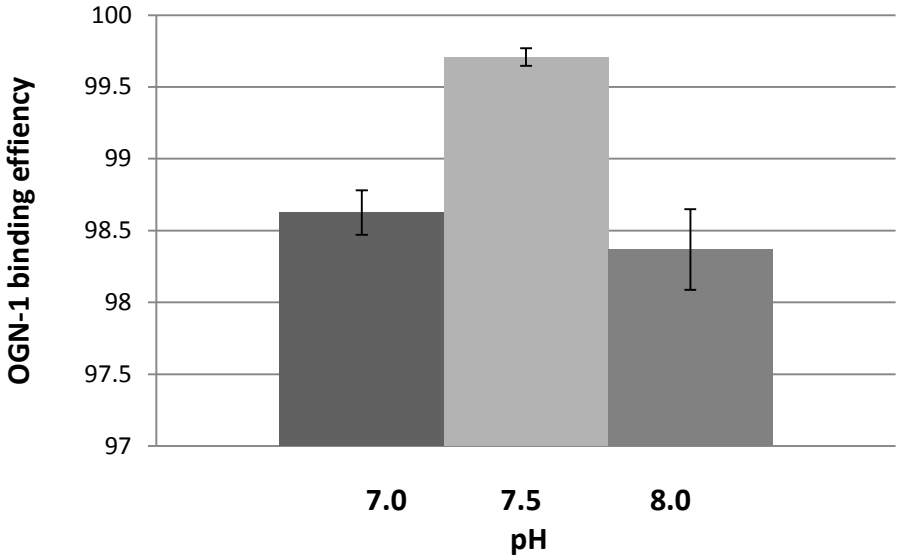

Figure 3.9. OGN-1 binding efficiency with BS^3 on ASC beads as a function of reaction pH. Note: Sample variance is purely a function of measurement variance; only a single immobilization run was performed for the preliminary study.

3.3.4A Solids content

The solids content of the beads were measured before and after OGN-1 immobilization (Table 3.4). The results for both types of beads indicate an increase in solids content, consistent with OGN-1 immobilization.

Table 3.4. Solids content of ASC and CMC beads

Bead type	Chemical treatment (Y/N)	OGN-1 immobilized (Y/N)	Solids content (%)
ASC	Y	N	2.8±0.43
ASC	Y	Y	4.18±0.43
CMC	Y	N	10.75±0.87
CMC	Y	Y	12.75±0.87

3.3.5A Confocal imaging

Confocal imaging results confirm that a high level of OGN-1 hybridization is achieved on both the ASC and CMC beads. It can be seen that the interior of the ASC beads are not accessible for diffusion and covalent attachment of OGN-1, whereas immobilization is apparent throughout the entire CMC bead (Figures 3.9 and 3.10). The perfectly spherical shape and smooth surface of the CMC beads should also be noted. The results indicate that blocking of the unreacted amine groups and replacing them with an acetyl group is an extremely effective strategy in preventing non-specific binding of OGN-2 to the beads.

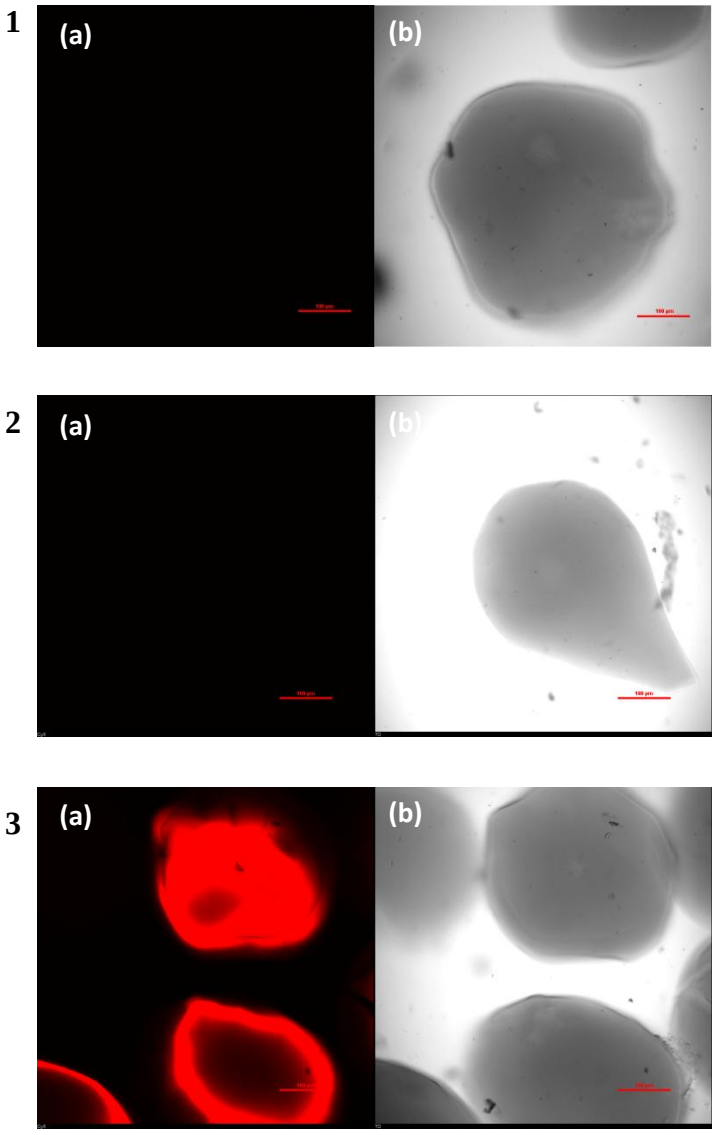

Figure 3.9. ASC beads (1) Negative control (2) Non-specific hybridization (3) Specific hybridization. The channels shown are (a) Cy5 (b) transmitted light for full bead shape visualization.

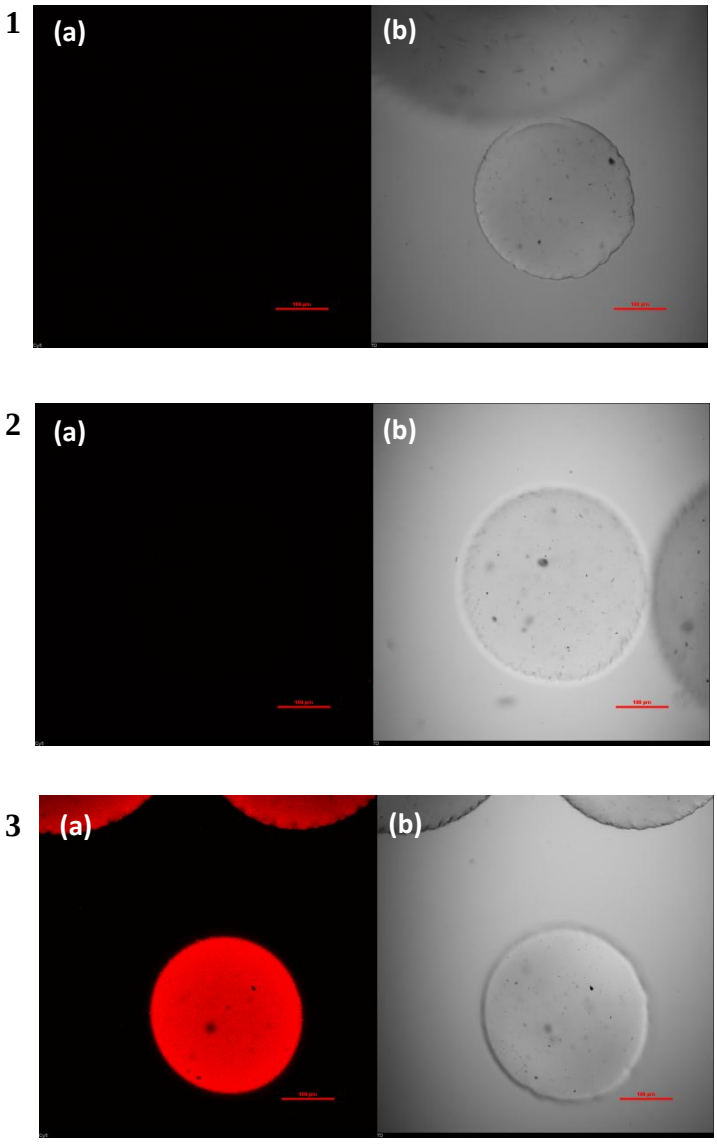

Figure 3.10. CMC beads (1) Negative control (2) Non-specific hybridization (3) Specific hybridization. The channels shown are (a) Cy5 (b) transmitted light for full bead shape visualization.

3.3B Glass rod results

3.3.1B Amine quantification

The results of molecules of amine available per cm^2 are shown in Table 3.5. Using the same technique, a value of $1.2x10^{16}$ was previously obtained on silicon wafers, with a resultant monolayer surface and minimal surface roughness (Guha Thakurta, 2010). The results indicate a 89 fold increase in silane density for nitric acid treated rods and 25 fold increase for piranha solution treated rods. A few possible explanations exist. First, the results could indicate multilayer silane formation, however, this is not easily verifiable on a cylindrical surface. A second possibility is that the method of sizing the rods may result in an increased surface area not accounted for by the theoretical calculation of the surface area. It is additionally possible that the nitric acid treatment on glass rods further increases the surface area by more effectively etching away at the surface. Ideally, the surfaces will be manufactured by a more reproducible method, but this lies outside of the scope of this preliminary study. For this study, the burden was only to prove the presence of amine groups on the surface in order to covalently attach DNA to the surface.

Table 3.5 Amine density after silanization of glass substrates

	Molecules of amine per cm^2 ($x10^{17}$)
Nitric acid treated rods	10.7
Piranha solution treated rods	3.3

3.3.2B OGN-1 Immobilization Efficiency and Blocking of Unreacted Amines

3.3.2.1B Benefits of Mg^{2+} addition for increased OGN-1 immobilization efficiency

It has been well established that the surface density of immobilized oligonucleotides has a large effect on the hybridization efficiency that can be achieved. The amount of oligonucleotide to functionalize on the substrate was therefore an important consideration. Table 3.6 provides a list of hybridization efficiencies corresponding to the optimal oligonucleotide density obtained by various authors in their attempts to functionalize glass substrates. Also included in the table is a calculation for the maximum amount of double stranded DNA theoretically able to provide monolayer surface coverage. This value was selected as the starting point for our experiments.

The absolute value for optimum surface density for each study cannot, however be observed in isolation. There are many factors in the attachment and hybridization steps that influence the final result. As can be seen from the table, a lower surface density does not necessarily translate to a higher hybridization yield. However, the reliable spacing between each individual molecule is an important factor (Hong et al., 2005).

Another effect that has been identified with the aim at increasing the hybridization efficiency is the concentration and type of cations present in the hybridization buffer (Springer et al., 2010; Liu and Tan, 1999). These authors have shown that divalent cations are much more effective at increasing the hybridization efficiency on a surface than monovalent cations since the environment for surface hybridization is much different than that for hybridization in solution. We postulate that the addition of Mg^{2+}

cations to the immobilization buffer will provide an additional benefit in the context of immobilization of a glass substrate, since the addition of Mg^{2+} to the reaction buffer is able to charge titrate the negative phosphate groups of the oligonucleotide backbone. This will reduce electrostatic repulsion between the oligonucleotides and increase the rigidity of the oligonucleotide, allowing for preparation of a densely packed and more importantly- uniform surface.

Table 3.6. Hybridization efficiency and corresponding optimal OGN-1 immobilization density using various methods for oligonucleotide attachment to glass substrates

	Max hybridization efficiency obtained	Chemistry used	Molecules of oligonucleotide bound per cm^2 for max hybridization ($x10^{13}$) / length	Max immobilization efficiency obtained
Hong et al. 2005	80-100%	Cone-shaped dendron w/ 3 nm mesospacing on ethylene glycol modified glass	1.4/15bp (ssDNA)	Not reported
Walsh et al. 2001	58%	One-step EDC with amine OGN attachment to carboxyl – modified aminosilane glass	0.9/20bp (ssDNA)	82-89%
Podyminogen et al. 2001	30%	Benzaldehyde OGN attached to semicarbamide glass	0.05/15bp (ssDNA)	Not reported
Zammateo et al. 200	Not reported	Amine OGN attached to aldehyde-modified glass surfaces	0.012/255bp (dsDNA)	Not reported
Theoretically possible with dsDNA (d= 2 nm)	-	-	3.2/-/-	-

The results of OGN-1 immobilization reveal that both BS^3 and EDC crosslinkers result in high efficiency, reliable OGN-1 attachment to aminosilanized glass rods (Table 3.7).

For BS^3 chemistry, washes were processed 5 times to sufficiently clean the DNA of BS^3, corresponding to only 25% retention of OGN-1 at this point. For EDC chemistry, only 2 processing steps were required, corresponding to 65% OGN-1 retention. The measurements listed in Table 3.6 have been adjusted accordingly (back-calculated). Based on the results obtained with the chitosan beads, the high immobilization efficiency seen with BS^3 is as expected. The results in Table 3.7 indicate no discernible difference between BS^3 and EDC/Imidazole chemistry for attaching OGN-1 to glass rods. The use of BS^3 may become cost-prohibitive on a larger scale and therefore the use of EDC/Imidazole can be recommended for future studies.

Table 3.7. Immobilization efficiency results

Glass treatment	OGN-1 in (µg)	OGN-1 out (µg)	OGN-1 bound (µg)	Immobilization efficiency (%)	Molecules of OGN-1 bound per cm^2 ($x10^{13}$)
BS^3 Nitric	9.2	0.64±0.08	8.56±0.08[1]	93.0±0.8	2.8
BS^3 Piranha	9.2	0.1±0.04	9.1±0.04[1]	98.9±0.4	3.0
EDC/Imidazole Nitric	9.2	0.14±0.04	9.06±0.04[1]	98.5±0.5	3.0
EDC/Imidazole Piranha	9.2	0.55±0.06	8.65±0.06[1]	94.0±0.6	2.9

[1]Sample variance is purely a function of measurement variance; only a single immobilization run was performed for the preliminary study

Comparing our immoblilization efficiency with that reported for Walsh et al. (2001), we see that the efficiency of OGN-1 immobilization using one-step EDC chemistry is increased as much as 10%. However, when we compare the actual oligonucleotide density with that obtained by other authors (refer to Table 3.6); we see an increased

density on the surface to what has been able to be accomplished in the past. We infer that the higher immobilization surface density (due to charge titration) and large immobilization efficiency (due to rigidity of the oligonucleotides) to be achieved by using the method of Mg^{2+} addition. However, the hybridization efficiency will be an important factor in determining whether this approach is a benefit or hindrance for the final application.

3.3.2.2B Benefits of acetylation in blocking unreacted amine groups

Unreacted amine groups and consequent non-specific binding are frequently the main reason that aminosilanized substrates are considered to be inferior to other types of functionalized substrates (Zammatteo, et al., 2000). However, for our application, in which we will immerse our substrate in a clinical specimen (containing a multitude of different proteins and other chemicals), the solution to the problem of non-specific adsorption may not be as simple as merely altering the type of silane used. The groups typically used to functionalize the surface and bind an oligonucleotide (such as carboxyl, aldehyde, epoxide, etc.) are all groups that have a large possibility of interacting with proteins. We therefore used an aminosilane surface and converted the amine groups to acetyl group with acetic anhydride following the immobilization reaction.

3.3.3B OGN-2 Hybridization efficiency

As mentioned in section 3.3.2.1B, hybridization efficiency is a critical parameter in determining the performance of oligonucleotide-functionalized substrates. We have

added Mg^{2+} to the immobilization, blocking and hybridization buffers to investigate the effects of this in the context of immobilization of a glass substrate (Table 3.8).

Table 3.8. Hybridization efficiency results of glass rods

Glass treatment	OGN-2 in (µg)	OGN-2 out (µg)	OGN-2 hybridized (µg)	Hybridization efficiency (%)	Molecules of OGN-1 hybridized per cm^2 ($x10^{13}$)
BS^3 Nitric	9.2	1.42	7.78	90.9	2.5
BS^3 Piranha	9.2	1.42	7.78	85.5	2.5
EDC/Imidazole Nitric	9.2	1.62	7.58	83.7	2.5
EDC/Imidazole Piranha	9.2	1.22	7.98	92.3	2.6

Using the methods outlined above, we observe hybridization efficiencies in the range of 84-92%. This result becomes particularly significant when comparing the EDC reaction results to the results of Walsh et al. (Table 3.6), who employed similar chemistry and conditions, save the addition of $MgSO_4$ to the reaction and hybridization buffers. We see that by making only this minor adjustment, we increase the hybridization efficiency by as much as 34%. We obtain a result similar to that of Hong et al. (Table 3.6), by a much simpler methodology.

3.3.4B Confocal imaging

Confocal imaging was used to supplement the results obtained in section 3.3.2B. Side views of nitric acid treated vs. piranha solution treated glass rods reveal that piranha treatment is more effective at providing a more uniformly distributed immobilization

profile (Figure 3.11). Piranha treatment of glass rods is therefore recommended for future studies.

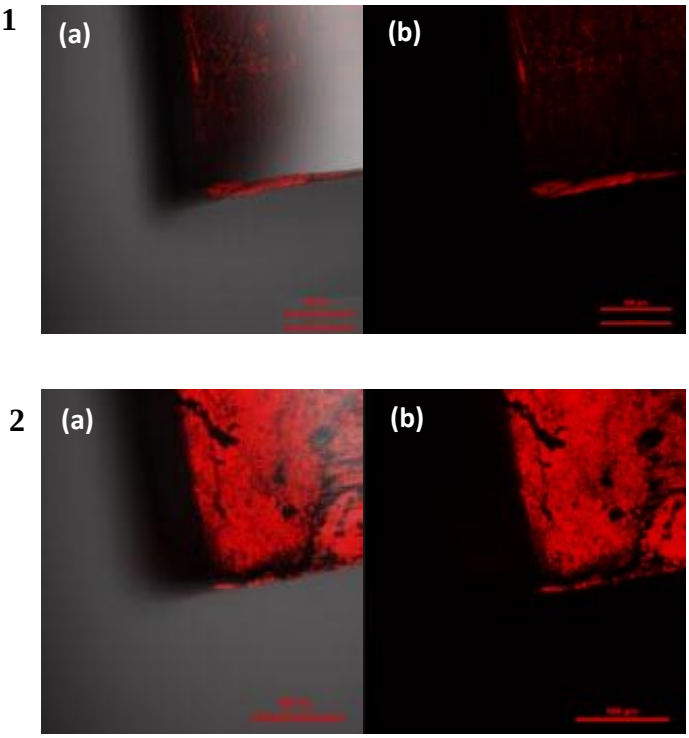

Figure 3.11. Side view of glass rods functionalized following (1) Nitric acid and (2) Piranha solution pre-treatment. The channels shown are (a) combined Cy5 and transmitted light (b) Cy5 for complimentary oligonucleotide visualization.

The bulk of confocal imaging was performed on the base of the cylindrical glass rod (Figures 3.12 – 3.15). In all figures, the first panel is the negative control, which has only been exposed to buffers and reagents, but no OGN-1 or OGN-2. The second panel is the non-specific control that had been exposed only to buffers and reagents, but no OGN-1. The non-specific control was, however, exposed to OGN-2. Some rods contained impurities, typically visualized as distinctive spots on the negative and non-specific controls. The third panel is the positive control. It is clear that acetic anhydride was effective at blocking the unreacted amine groups and essentially eliminating all non-specific binding. The positive control images support the results obtained in sections 3.3.2B.

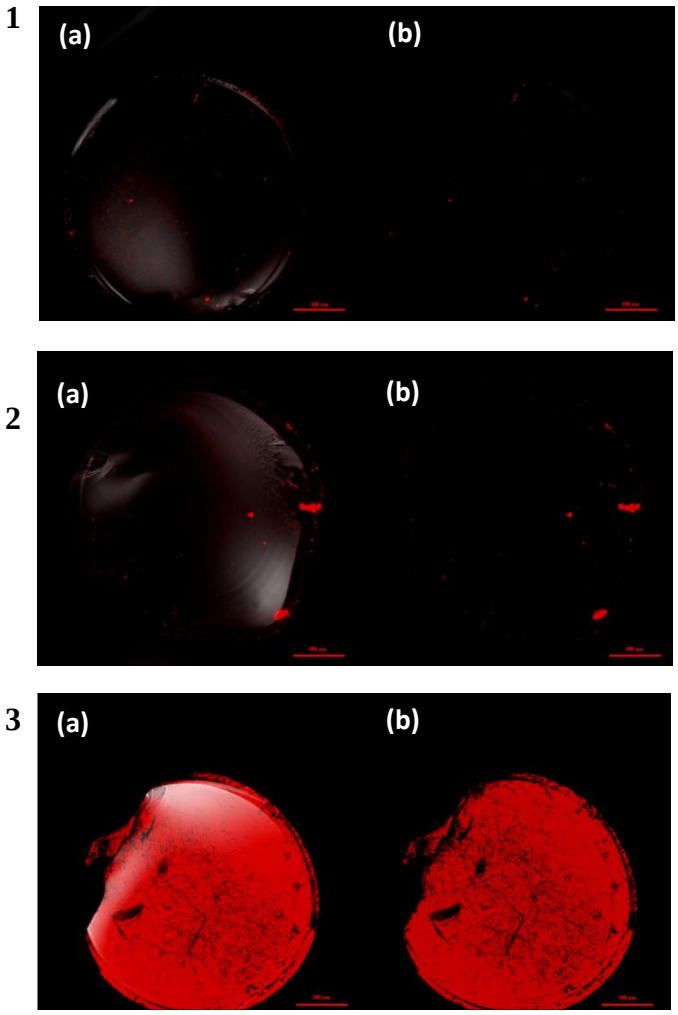

Figure 3.12. Nitric acid pre-treated glass rods. Oligonucleotides were bound using BS^3 crosslinker and free amine groups were blocked by acetylation. (1) Control (2) Non-specific hybridization of complimentary oligonucleotide (3) Specific hybridization of complimentary oligonucleotide. The channels shown are (a) combined Cy5 and transmitted light (b) Cy5 for complimentary oligonucleotide visualization

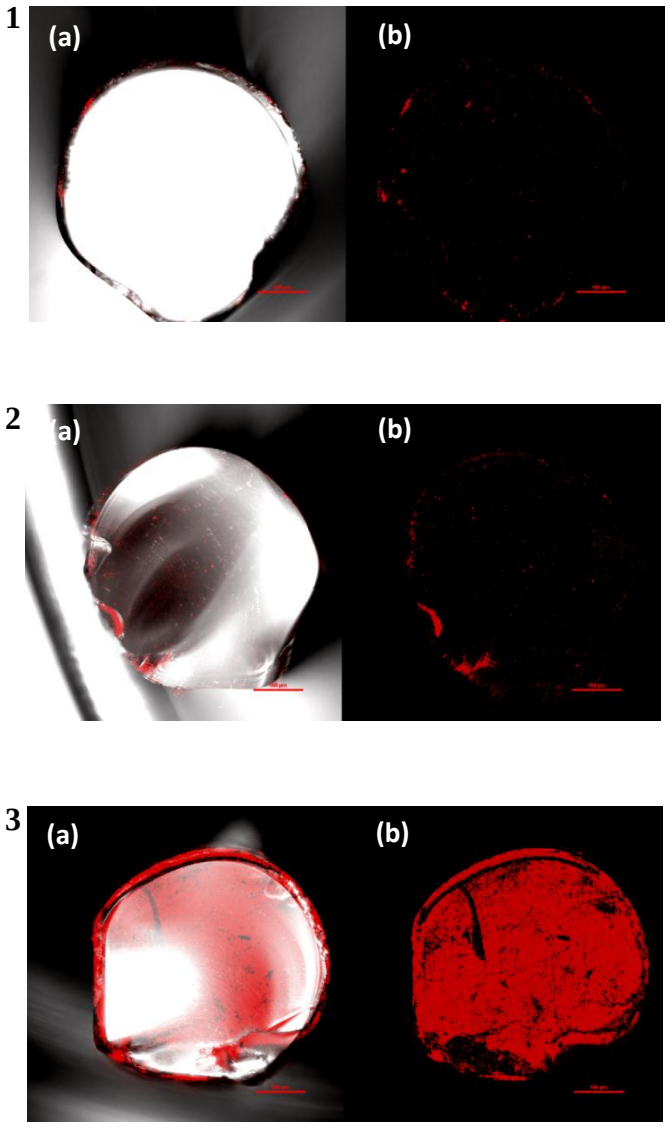

Figure 3.13. Piranha solution pre-treated glass rods. Oligonucleotides were bound using BS^3 crosslinker and free amine groups were blocked by acetylation. (1) Control (2) Non-specific hybridization of complimentary oligonucleotide (3) Specific hybridization of complimentary oligonucleotide. The channels shown are (a) combined Cy5 and transmitted light (b) Cy5 for complimentary oligonucleotide visualization

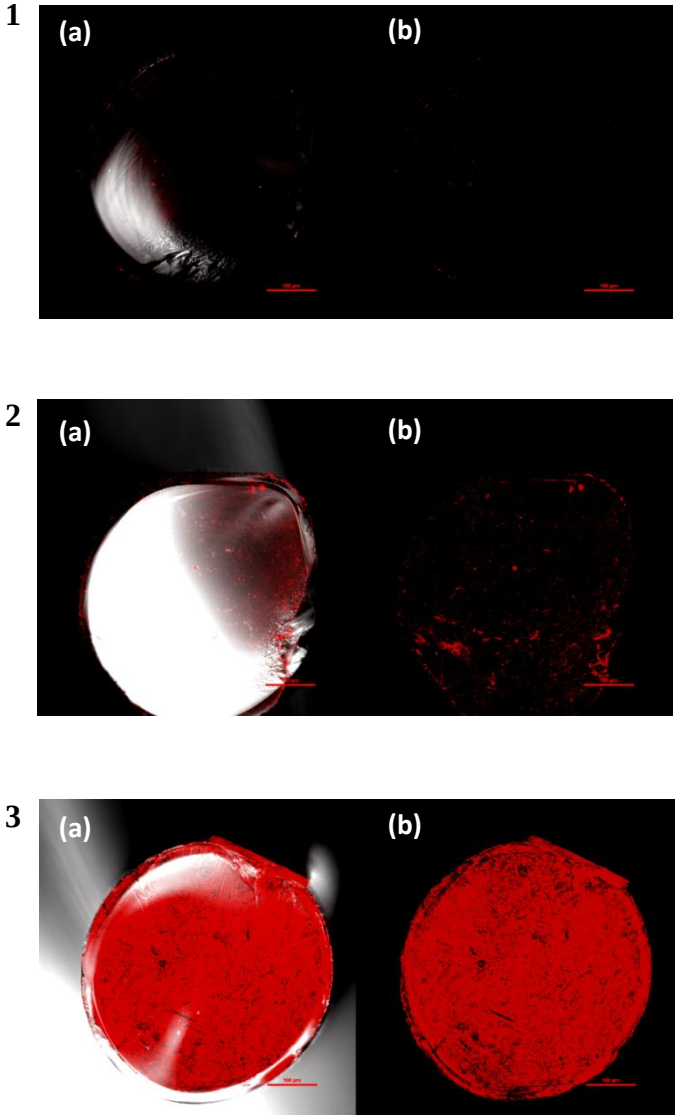

Figure 3.14. Nitric acid pre-treated glass rods. Oligonucleotides were bound using EDC crosslinker and free amine groups were blocked by acetylation. (1) Control (2) Non-specific hybridization of complimentary oligonucleotide (3) Specific hybridization of complimentary oligonucleotide. The channels shown are (a) combined Cy5 and transmitted light (b) Cy5 for complimentary oligonucleotide visualization

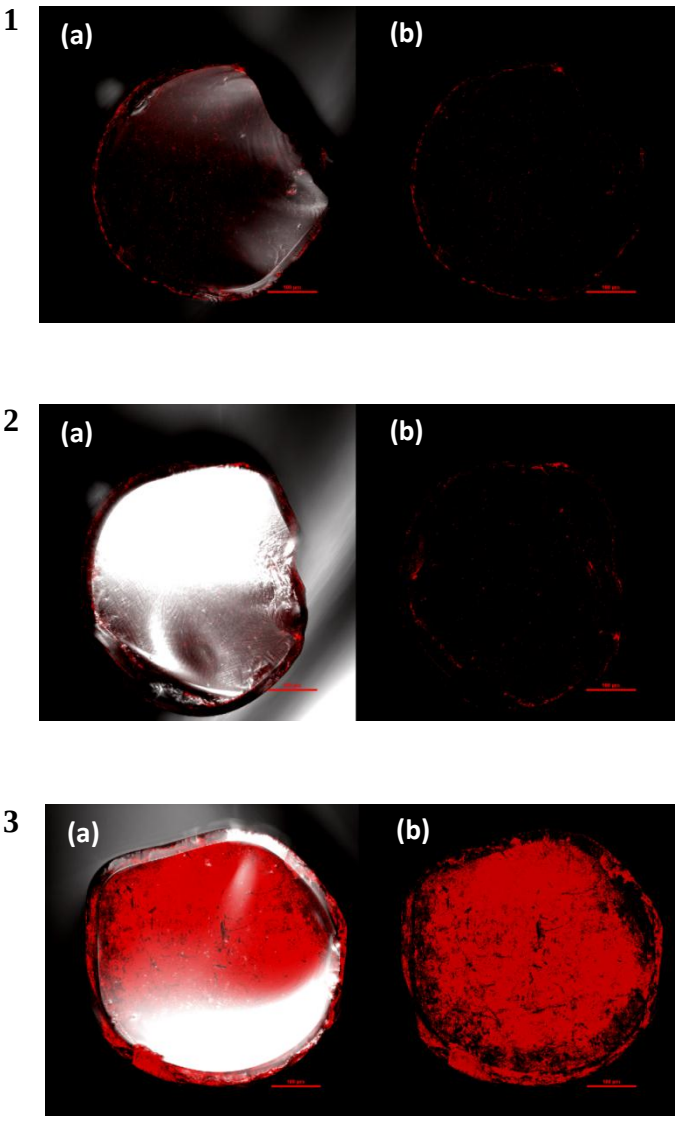

Figure 3.15. Piranha solution pre-treated glass rods. Oligonucleotides were bound using EDC crosslinker and free amine groups were blocked by acetylation. (1) Control (2) Non-specific hybridization of complimentary oligonucleotide (3) Specific hybridization of complimentary oligonucleotide. The channels shown are (a) combined Cy5 and transmitted light (b) Cy5 for complimentary oligonucleotide visualization

3.4. Conclusions

1. Piranha pre-treatment of glass rods results in a more uniform oligonucleotide distribution over the entire surface of the rod than does nitric acid pre-treatment.
2. Using the silanization method of Guha Thakurta (2010), we are able to bind aminosilane to the glass surface at a density of 3-10x10^{17} molecules/cm^2.
3. Both EDC and BS^3 chemistry result in comparable immobilization efficiencies (93-99%). However, the EDC/imidazole is much more cost effective (<5% than the cost of BS^3)
4. Both immobilization efficiency of OGN-1 and hybridization efficiency of OGN-2 are greater than that previously reported in the literature when using simple silanized glass substrates and chemistries. We propose that the increase is attributable to the addition of Mg^{2+} to the reaction and hybridization buffers. Using this simple technique, we obtain a comparable hybridization rate to the use of very complex spacer molecules.

3.5. Future work

1. It is suggested that the specificity of the oligonucleotide be analyzed by using a non-complimentary target tagged with a different fluorophore that can then be hybridized and detected using the methods outlined in this Chapter. In this case, increased temperatures may be considered to increase the stringency of the washes.
2. For samples containing low concentrations of target DNA, it may be worthwhile to investigate the use of a surface containing two oligonucleotides, each specific

for one strand of the target DNA. By this method, twice as much target DNA can potentially be captured at lower concentrations.

3. Investigate the need for a longer spacer between the surface and the bound oligonucleotide to overcome potential steric hindrance problems when binding genomic target DNA molecules.
4. Compare the newly synthesized rods with the polystyrene strip method and conditions outlined in Chapter 4. It may be valuable to add an additional step where non-specific adsorption of proteins to the strip is investigated.

CHAPTER 4

APPLICATION OF THE TECHNOLOGY TO CLINICAL STOOL SAMPLES FOR *CLOSTRIDIUM DIFFICILE* DIAGNOSIS

4.1. Introduction

Clostridium difficile infection (CDI) is the leading cause of nosocomial diarrhea in North America and Europe (Knetsch et al., 2011; Kuehne et al., 2010). Approximately 500,000 CDI cases per year are estimated to occur in hospitals and long-term care facilities in the United States, ranging from asymptomatic cases or mild diarrhea to severe abdominal cramping, fever, and bloody diarrhea signifying active colitis (Rupnik, 2009). Historically, CDI has been associated mainly with antibiotic use in hospitals, but more recently a larger at-risk population has been identified, including hospitalized patients, nursing home residents, and also community dwellers (Ananthakrishnan, 2011; Bartlett, 2010). Epidemiological trends in the last decade have shown marked increases in incidence, severity, persistence, and mortality from CDI, presumably linked to the emergence of the hypertoxigenic NAP1 strain (Bartlett, 2010, Tenover et al., 2010). Unfortunately, the lack of a rapid, accurate, and inexpensive diagnostic test for CDI remains an important barrier to clinical and epidemiological containment of the disease. Laboratory diagnosis of CDI from stool has traditionally been based on detecting the presence of *C. difficile* toxin A and/or toxin B proteins in stool by various methods (Ananthakrishnan, 2011; Bartlett, 2010a; Knetsch et al., 2011; Tenover et al., 2010). EIA tests for the detection of *C. difficile* toxins in stool are widely used because of ease of use

and rapid results reporting. However, existing commercial EIA tests for one or both toxins are relatively insensitive, detecting only 30% to 70% of CDI-related disease (Kvach et al., 2010; Novak-Weekley et al., 2010, Quinn, 2010; Sloan et al., 2008). Biological activity assays (including cytotoxicity and toxigenic culture assays) detect toxin B-mediated cell cytotoxicity in fecal eluates applied to human cell monolayers. Both assays are sensitive tests, but too time-consuming for routine use (Ananthakrishnan, 2011; Knetsch et al., 2011; Tenover et al., 2010). With the development of an EIA for glutamate dehydrogenase (GDH), which is a *C. difficile* cellwall common antigen, the sensitivity for the detection of *C. difficile* approaches 100%; however, because GDH is a ubiquitous protein among both toxigenic and nontoxigenic strains, specificity of the GDH assay is poor (Bartlett, 2010b). Individual EIA tests for either GDH or toxins A or B are therefore viewed as insufficiently specific for diagnosis. A commercially available combined GDH and toxin A/B EIA assay (C. Diff Quik Chek Complete) appears to increase diagnostic sensitivity and specificity (Ananthakrishnan, 2011). When both EIA assays are positive or negative, the interpretation is straightforward and accurate (Bartlett, 2010a; Tenover et al., 2010). If, however, the GDH and toxin results are discordant (typically GDH+/toxin-), then additional reflex or discrepant testing using a highly specific molecular-based assay is required Bartlett, 2010b. Molecular-based assays for the detection of toxigenic *C. difficile* in stool offer increased sensitivity and specificity relative to the EIA tests (Kvach et al., 2010). Genes within the *C. difficile* pathogenicity locus, including those encoding toxins A and/or B (*tcdA* or *tcdB*), are targeted by specific primer sets, usually in a multiplex PCR assay. Identification of these gene sequences in stool samples correlates highly with positive *C. difficile* toxigenic stool assays (Lalande,

2011). A positive PCR assay is therefore thought to directly signify actively toxigenic *C. difficile* strains in stool (Antunes et al., 2010; Rupnik, 2010; Spigaglia et al., 2002). Accordingly, new molecular assays such as the loop-mediated isothermal amplification (LAMP) assay (Illumigene *C. difficile*; Meridian Bioscience, Cincinnati, OH) and the GeneXpert assay (Xpert *C. difficile*; Cepheid, Sunnyvale, CA), which detect pathogenicity locus region sequences, are being increasingly used for confirmatory testing if EIA tests are equivocal. Nonetheless, despite greater diagnostic sensitivity and specificity, current molecular methods are too labor-intensive, complex, and/or costly to be of practical use in many clinical laboratories (Ananthakrishnan, 2011; Rupnik, 2010; Tenover et al., 2010).

In the present chapter, the lysis microreactor (LMR)/PCR assay for *C. difficile* diagnosis is evaluated in a frozen bank of diarrheal stool samples. These results are compared with standard *C. difficile* testing methods used routinely by the hospital clinical laboratory.

4.2. Materials and Methods

4.2.1 Samples and Sample Collection

Diarrheal stool samples (n = 198) were collected from patients for clinical laboratory evaluation of *C. difficile* at the request of the treating physician. Liquid stool samples were processed routinely in the Nebraska Medical Center clinical laboratory as described below. Excess de-identified samples were assigned a study number by laboratory personnel and were transferred to the research laboratory, where they were stored at -20°C until batch testing for study purposes. Samples were thawed once for study

evaluation (data not shown). This study was approved by the Institutional Review Board of the University of Nebraska Medical Center, with a waiver for informed consent.

4.2.2 Standard Clinical Laboratory Testing for *C. difficile*

A dual EIA screen for both GDH and toxin A/B (Wampole C. Diff Quik Chek Complete; Techlab, Blacksburg, VA) was performed according to the manufacturer's instructions. Stool samples testing positive for both GDH and toxin A/B were considered true positive and samples negative for both GDH and toxin A/B were considered true negative; no further testing was considered necessary for these samples (Bartlett, 2010b). For specimens with discordant results (ie., positive GDH and negative toxin A/B result), a LAMP molecular test (Illumigene *C. difficile*; Meridian Bioscience) was routinely performed as a reflex test according to the manufacturer's instructions.

4.2.3 LMR/PCR Assay

As previously discussed in the preface, the LMR/PCR assay consists of three steps: cell lysis, DNA capture via polystyrene strip, and PCR amplification (refer to Figure P.1). This methodology was used to detect the presence of *C difficile* in stool samples for comparison against the standard hospital diagnostic algorithm.

4.2.4 Strip and Lysis Buffer Preparation

Clear polystyrene strips (0.127 mm by 1 mm by 40 mm) were sanded lightly with fine 400-grit sandpaper and incubated overnight in 20 mmol/L EDC HCl ([*N*-(3-dimethylaminopropyl)-*N*-ethyl-carbodiimide hydrochloride]; Sigma-Aldrich,

St. Louis, MO). The strips were washed once with TBST and once with distilled water (pre-lysis wash) and were stored in distilled water until use. The chemically prepared strips were found to be stable for up to 1 month (data not shown). A lysis buffer of 20 mmol/L TCEP [Tris(2-carboxyethyl) phosphine; Sigma-Aldrich] and 20X TE (Tris-EDTA; Sigma-Aldrich, St. Louis, MO) was prepared and stored at -20°C for prolonged use. A wash buffer (TBST) of 10 mmol/L Tris [Tris(hydroxymethyl) aminomethane], 150 mmol/L NaCl, and 0.05% Tween 20 was prepared and stored at 4°C for prolonged use (all from Sigma-Aldrich, St. Louis, MO).

4.2.5 Lysis and DNA Extraction

Stool samples were thawed to room temperature (20°C) before DNA extraction. A specimen of 400 µL of thawed, unformed stool was mixed with 400 µL of lysis buffer and transferred to a 1 mL capacity LMR (Figure 1.1, A and B). Dry chelating resin beads (40 to 80 mg Chelex 100; Bio-Rad Laboratories, Hercules, CA) and a prepared polystyrene capture strip were placed in the LMR. The contents were heated to 92°C and mixed for 5 minutes to lyse the bacterial cells, denature the DNA, and allow the single-stranded forms to bind to the hydrophobic polystyrene strip. The polystyrene strip was removed from the micromixer and washed twice with distilled water (post lysis wash) before being placed in a Roche glass capillary cuvette (Roche Applied Science, Indianapolis, IN) containing 25 µL of previously prepared PCR master mix.

4.2.6 PCR Master Mix Preparation

Amplification of a non-repeat region of the *C. difficile tcdB* gene was performed using previously described 177-bp forward (5'-GAAAGTTCAAGTTTACGCTCAAT-3') and reverse (5'-GCTGCACCTAAACTTACACCA-3') primers (van den Berg et al., 2006). Each 25 µL reaction mixture contained a final concentration of 0.2 mmol/L dNTPs, 4 mmol/L $MgSO_4$, 0.5 U KOD Hot Start DNA polymerase, 1X PCR buffer for KOD Hot Start DNA polymerase (EMD Chemicals, Gibbstown, NJ); 0.4 mg/mL bovine serum albumin (Ambion, Austin, TX); 2 µmol/L SYTO13 (Invitrogen, Carlsbad, CA); and 0.2 µmol/L forward and 0.2 µmol/L reverse primers (UNMC Eppley Molecular Biology Core Lab, Omaha, NE). (The SYTO13 was included for exploration of real-time PCR methodology; data not shown). Glass capillary tubes were used for the PCR reaction.

4.2.7 PCR Amplification

The thermal protocol used for amplifying *tcdB* in a prototype Philisa Thermocycler (Streck, Omaha, NE) included an enzyme activation step at 95°C for 30 s, followed by 30 cycles of 95°C for 3 s and 59°C for 4 s, and then by 15 cycles of 95°C for 3 s, 59°C for 7 s, and 72°C for 10 s. Typical time for amplification was <14 minutes.

4.2.8 PCR Detection

Gel electrophoresis was used to detect PCR amplification products. A 10 µL aliquot of each product was loaded into a 1% agarose gel along with 100-bp DNA ladder (Fisher Scientific, Pittsburgh, PA) and run at 120 V for 30 minutes. Gel images were captured using a Canon PowerShot A650 IS digital camera under the same exposure time for all

gels. Band intensities were measured using ImageJ software (version 1.44; NIH, Bethesda, MD). The average band intensity values correlating with negative (EIA GDH-/toxin-) and positive (EIA GDH+/toxin+) stools were identified as 20.87 ± 5.37 (n = 116) and 68.75 ± 24.46 (n = 82), respectively ($P < 0.001$, Student's t-test). Thus, an intensity of 20.87 ± 5.37 was considered a background value. A positive LMR/PCR result was defined as a minimum band intensity of 31.61, representing background plus 2 standard deviations above the mean.

4.2.9 Analytical Sensitivity Testing

The sensitivity of the LMR/PCR assay was evaluated using stool samples spiked with various concentrations of purified *C. difficile* DNA, which was extracted from *C.difficile* bacterial culture using a NucliSENS EasyMAG automated extractor (bioMerieux, Durham, NC). Stool samples known to be negative for *C. difficile* (GDH-/toxin-) were spiked with DNA at varying concentrations and assayed as described. Two sets of samples with DNA concentrations of 0.5 pg, 0.05 pg, and 0.01 pg/mL were prepared, corresponding to 114, 11, and 2 genomic copies/mL stool, respectively. The number of genomic copies of *C. difficile* was calculated based on the molecular weight of its genome. Analysis of the spiked samples was the same as described for the clinical specimens.

4.3. Results

4.3.1 Analytical Sensitivity with Spiked Stool Samples

Results of analytical sensitivity assays using stool spiked with genomic *C. difficile* DNA are shown in Figure 4.1. Bands of expected product size for *tcdB* sequences defined by the primer set (177 bp) were identified by gel electrophoresis and product was confirmed by sequencing (data not shown). Results of experiments without (Figure 4.1A) and with (Figure 4.1B) Bio-Rad Chelex resin powder added to the LMR lysis mixture are shown. The chelating resin appears to slightly improve the PCR yield, based on increased band intensity, especially at the lower DNA concentrations. For example, at 0.5 pg DNA in 1 mL stool sample, a band intensity reading of 49 was obtained without Chelex resin, compared with a reading of 77.5 with resin. Similarly, at 0.05 pg and 0.01 pg DNA, the average band intensity increased from 34.9 to 42.8 and from 18.1 to 29.9, respectively. Repeated experiments confirmed that DNA concentrations of 11 genomic copies/mL can be detected reliably in stool (data not shown). At 0.01 pg/mL (2 copies), the PCR yield becomes low, and competitive primer-dimer reactions become more pronounced.

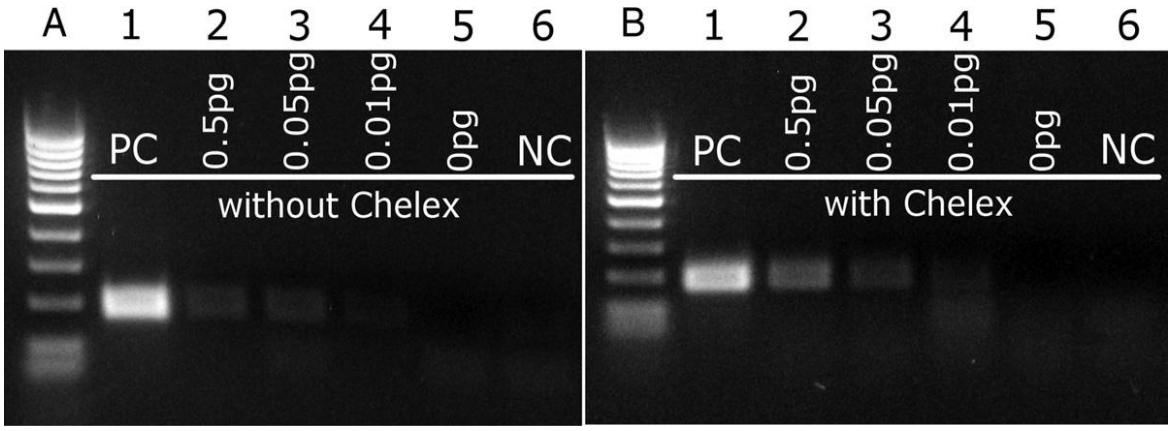

Figure 4.1. Agarose gel electrophoresis results of PCR products after *tcdB* amplification for stool samples spiked with *C. difficile* DNA to concentrations of 0.5 pg/mL (which corresponds to 114 copies/mL; lane 2), 0.05 pg/mL (11 copies/mL; lane 3) and 0.01 pg/mL (2 copies/mL; lane 4). Results are shown without Chelex 100 resin added to the LMR mixture (**A**) or with Chelex 100 resin (**B**). Positive control (PC) indicates GDH+/toxin+ stool; negative control (NC) indicates GDH-/toxin- stool.

4.3.2 Clinical Sample Testing

A total of 198 stool samples from patients with suspected CDI were provided from the Nebraska Medical Center Clinical Microbiology Laboratory for evaluation with the new LMR/PCR method. Of 198 stool EIA results, 129 were EIA concordant: 48 samples were positive for both GDH and toxin A/B and 81 samples were negative for both. The LMR/PCR method had perfect correlation with all 129 EIA concordant results (Table 5.1). Of the 69 EIA discordant samples that were tested with both reflex LAMP PCR and the LMR/PCR method, 64 samples were in agreement (Table 4.1). Of these, 29

samples were positive and 35 were negative by both PCR tests; the remaining 5 samples were positive for *C. difficile* by the LMR/PCR but were negative by LAMP.

Table 4.1 Comparison of C. difficile results with EIA, LAMP and Rapid LMR/PCR for 198 clinical samples

Clinical samples (no.)	Dual EIA: GDH/toxin A, B	Rapid LMR/PCR	LAMP
EIA concordant (n = 129)			
48	+/+	+	Not tested
81	-/-	-	Not tested
EIA discordant (n = 69)			
29	+/-	+	+
35	+/-	-	-
5	+/-	+	-
0	+/-	-	+

4.4. Discussion

This proof-of-concept study showed excellent agreement between the results for *C. difficile* testing by the newly described LMR/PCR technology and the currently used clinical laboratory testing algorithm for *C. difficile*. Overall, the LMR/PCR method was in accord 97.5% of the time with the routine clinical testing pathway, which uses a dual EIA and discrepant LAMP PCR assays. Because there is no universally acknowledged gold standard for toxigenic *C. difficile* diagnosis, the choice of a comparison assay or assays to convincingly demonstrate the accuracy of a newly developed test is not clear. Cell culture-based techniques such as the toxigenic culture and the cytotoxicity assay are notoriously subject to technical and interpretive variability (Novak-Weekley et al., 2010).

Enzyme immunoassays (EIA) are of limited sensitivity and specificity, although they are PCR tests are gaining acceptance as being highly sensitive and specific for diagnosing *C. difficile* infection; however, because they test only for gene sequences and not for the presence of cyto- or enterotoxins, questions remain about whether they actually identify active toxigenic infection (Goldenberg et al., 2011; Kato et al., 2005). In a recent study, LAMP testing for *C. difficile* had a sensitivity and specificity of 98%, positive predictive value of 92%, and negative predictive value of >99%, compared with a composite gold standard consisting of cytotoxin B assay and toxigenic culture (Noren et al., 2011). In fact, molecular assays are often used as a final arbiter test to definitively diagnose or rule out *C. difficile* in cases in which the dual EIA is discordant (ie., GDH+/toxin-). A two-step diagnostic algorithm (initial dual EIA followed by LAMP testing for samples that are EIA positive for glutamate dehydrogenase and negative for toxin) is used routinely at the Nebraska Medical Center's clinical laboratory. Accordingly, we chose to compare the new LMR/PCR to this two-step algorithm as the clinical gold standard diagnostic. This comparison provides a practical, real-life evaluation of the LMR/PCR test performance at an early stage in its development. In the absence of an established gold standard test, it is not possible to calculate sensitivity and specificity based on a true positive or negative. Nonetheless, the favorable comparative results suggest that the LMR/PCR technique is worthy of further development as a point-of-care diagnostic. A significant advantage is the rapid turnaround time for sample preparation and amplification, which took <20 minutes in the experiments reported here. Inclusion of robotic functions to manage sample preparation, as well as a real-time detection system on the PCR thermocycler, would be expected to allow a turnaround time on the order of

15 to 20 minutes. Sensitivity of the LMR/PCR is notably high in spiked stool experiments. The detection level of this method is as low as 0.05 pg/mL of *C. difficile* DNA (corresponding to 11 genomic copies/mL). The detection levels of LAMP and Xpert assays mentioned in their manufacturer package inserts are similar: 64 colony forming units/mL of various *C. difficile* strains, which corresponds to 0.3 pg/mL of DNA (package insert "Illumigene *C. difficile;* Xpert *C. difficile;* 2011 DNA Amplification Assay for the Detection of Cytotoxigenic *C. difficile* in Stool Specimens," Meridian Bioscience, 2011). Clinical sensitivity calculations were precluded in the current study, but it appears that the LMR/PCR test is at least comparable to the LAMP assay, showing agreement in 92.7% (64/69) samples (all GDH+/toxin- by EIA) tested by both methods. In the remaining 7.3% (5/69 samples), the LAMP test was negative and the LMR/PCR test was positive, raising the possibility that the new method could be more sensitive for identifying toxin B gene sequences in stool. In a recent study, the Illumigene LAMP test was found to be 91.8% sensitive, compared with a *C. difficile* toxigenic culture assay, suggesting that it may miss ~8% of presumably true positives (Lalande et al., 2011). As discussed in Chapter 2, the randomness associated with turbulent flow in the LMR assures that lysed DNA could eventually bind to the capture strip, depending on the initial position of the DNA molecule. The forced convection yields hybridization times in minutes, rather than the hours or days that would be expected for DNA diffusion alone without the active mixing. The strip captures lysed DNA from the lysate and it serves as a simple method for transfer of the DNA to the thermocycler cuvette. For this particular application, polystyrene was used for the capture stips since it is known to noncovalently bind to DNA in a solution and is consequently an extremely simple surface to use for

testing the strip functionality hypothesis (Nikiforov et al., 1995). It is suspected that the concentrating effect of the polystyrene capture strip enhances the sensitivity of the LMR/PCR, but there may also be an increase in contamination with that method, and further improvements with the strip capture are expected to greatly enhance the specificity of the capture (also refer to Chapter 3). In essence, the LMR coupled with the capture strip represents a particularly simple, fast, and inexpensive method for isolating genomic and pathogen DNA from stool. The proof-of-concept experiments reported here demonstrate that the LMR/PCR technique may be sensitive and accurate for the detection of *C. difficile*. The system is also fast, with only 20 minutes required for the cell lysis, DNA capture via polystyrene strip, and PCR amplification. The current LMR is an individual temperature-controlled unit for a single sample; however, it would be straightforward to expand it to an eight-well device to match the eight sample positions in the Philisa Thermo Cycle (Streck, Omaha, NE). With the addition of real-time PCR, the throughput of the system could be as high as 192 samples in an 8-hour work day. The lysis apparatus could be injection molded, and none of the reagents are cost-prohibitive. Thus, the goals of an inexpensive, rapid, (sensitive, and accurate) diagnostic test could be met. Based on the limitations of currently available testing modalities for *C. difficile*, efforts to develop new technologies that improve both the diagnostic speed and accuracy for CDI are needed. The LMR/PCR technology described here lends itself to the eventual development of a rapid turnaround point-of-care test not only for *C. difficile* but also for many other infectious pathogens in stool and other body fluids. The simplicity and potentially low cost of the LMR (based on prototype pricing) may help in the development of a reliable, low-cost test for *C. difficile*.

REFERENCES

Ananthakrishnan, A. N., 2011. Clostridium difficile infection: epidemiology, risk factors and management, Nature reviews.Gastroenterology & hepatology 8(1), 17-26.

Antunes, A., Dupuy, B., 2010. Molecular methods to study transcriptional regulation of Clostridium difficile toxin genes, Methods in molecular biology (Clifton, N.J.) 646, 93-115.

Bartlett, J. G., 2010a. Clostridium difficile: progress and challenges, Annals of the New York Academy of Sciences 1213, 62-69.

Bartlett, J. G., 2010b. Detection of Clostridium difficile infection, Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America 31 Suppl 1, S35-7.

Boom, R., Sol, C. J., Salimans, M. M., Jansen, C. L., Wertheim-van Dillen, P. M., van der Noordaa, J., 1990. Rapid and simple method for purification of nucleic acids, Journal of clinical microbiology 28(3), 495-503.

Cadet, J., Bellon, S., Berger, M., Bourdat, A. G., Douki, T., Duarte, V., Frelon, S., Gasparutto, D., Muller, E., Ravanat, J. L., Sauvaigo, S., 2002. Recent aspects of oxidative DNA damage: guanine lesions, measurement and substrate specificity of DNA repair glycosylases, Biological chemistry 383(6), 933-943.

Champe, P. C., Harvey, R. A., Ferrier, D. R. (Eds), 2008. Lippincott's Illustrated Reviews: Biochemistry : Lippincott Williams and Wilkins.

Eaton, J. W., 2010. GNU Octave v. 3.2.3, 2010(5/24/2010).

Gevertz, J. L., Dunn, S. M., Roth, C. M., 2005. Mathematical model of real-time PCR kinetics, Biotechnology and bioengineering 92(3), 346-355.

Goldenberg, S. D., Gumban, M., Hall, A., Patel, A., French, G. L., 2011. Lack of effect of strain type on detection of toxigenic Clostridium difficile by glutamate dehydrogenase and polymerase chain reaction, Diagnostic microbiology and infectious disease 70(3), 417-419.

Griep, M. A., Kotera, C. A., Nelson, R. M., Viljoen, H. J., 2006. Kinetics of the DNA polymerase *pyrococcus kodakaraensis*, Chemical Engineering Science 61(12), 3885.

Guha Thakurta, S., 2010. Design and development of in situ albumin binding surfaces: Evaluation in the paradigm of blood-biomaterial compatibility.

Hermanson, G. T., 2008. Bioconjugate Techniques. : Academic Press.

Higuchi, R., Fockler, C., Dollinger, G., Watson, R., 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions, Nature Biotechnology 11(9), 1026-1030.

Hong, B. J., Oh, S. J., Youn, T. O., Kwon, S. H., Park, J. W., 2005. Nanoscale-controlled spacing provides DNA microarrays with the SNP discrimination efficiency in solution phase, Langmuir 21(10), 4257-4261.

Hsu, G. W., Ober, M., Carell, T., Beese, L. S., 2004. Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase, Nature 431(7005), 217-221.

IDT: Oligo Analyzer, 2012. 2011.

IDT: PrimerQuest, 2012. 2011.

Kainz, P., 2000. The PCR plateau phase - towards an understanding of its limitations, Biochimica et biophysica acta 1494(1-2), 23-27.

Keohavong, P., Thilly, W. G., 1989. Fidelity of DNA polymerases in DNA amplification, Proceedings of the National Academy of Sciences of the United States of America 86(23), 9253-9257.

Kato, H., Yokoyama, T., Kato, H., Arakawa, Y., 2005. Rapid and simple method for detecting the toxin B gene of Clostridium difficile in stool specimens by loop-mediated isothermal amplification, Journal of clinical microbiology 43(12), 6108-6112.

Knetsch, C. W., Bakker, D., de Boer, R. F., Sanders, I., Hofs, S., Kooistra-Smid, A. M., Corver, J., Eastwood, K., Wilcox, M. H., Kuijper, E. J., 2011. Comparison of real-time PCR techniques to cytotoxigenic culture methods for diagnosing Clostridium difficile infection, Journal of clinical microbiology 49(1), 227-231.

Kuehne, S. A., Cartman, S. T., Heap, J. T., Kelly, M. L., Cockayne, A., Minton, N. P., 2010. The role of toxin A and toxin B in Clostridium difficile infection, Nature 467(7316), 711-713.

Kvach, E. J., Ferguson, D., Riska, P. F., Landry, M. L., 2010. Comparison of BD GeneOhm Cdiff real-time PCR assay with a two-step algorithm and a toxin A/B enzyme-linked immunosorbent assay for diagnosis of toxigenic Clostridium difficile infection, Journal of clinical microbiology 48(1), 109-114.

Lalande, V., Barrault, L., Wadel, S., Eckert, C., Petit, J. C., Barbut, F., 2011. Evaluation of a loop-mediated isothermal amplification assay for diagnosis of Clostridium difficile infections, Journal of clinical microbiology 49(7), 2714-2716.

Li, H. H., Gyllensten, U. B., Cui, X. F., Saiki, R. K., Erlich, H. A., Arnheim, N., 1988. Amplification and analysis of DNA sequences in single human sperm and diploid cells, Nature 335(6189), 414-417.

Lindahl, T., Nyberg, B., 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid, Biochemistry 13(16), 3405-3410.

Lindahl, T., Nyberg, B., 1972. Rate of depurination of native deoxyribonucleic acid, Biochemistry 11(19), 3610-3618.

Liu, W., Saint, D. A., 2002a. A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics, Analytical Biochemistry 302(1), 52-59.

Liu, W., Saint, D. A., 2002b. Validation of a quantitative method for real time PCR kinetics, Biochemical and biophysical research communications 294(2), 347-353.

Livak, K. J., Schmittgen, T. D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, Methods (San Diego, Calif.) 25(4), 402-408.

Liu, X., Tan, W., 1999. A fiber-optic evanescent wave DNA biosensor based on novel molecular beacons, Analytical Chemistry 71(22), 5054-5059.

Logan, J., Edwards, K., Saunders, N. (Eds), , 2009. REAL-TIME PCR Current Technology and Applications. : Caister Academic Press.

Mamedov, T. G., Pienaar, E., Whitney, S. E., TerMaat, J. R., Carvill, G., Goliath, R., Subramanian, A., Viljoen, H. J., 2008. A fundamental study of the PCR amplification of GC-rich DNA templates, Computational biology and chemistry 32(6), 452-457.

Nakano, M., Komatsu, J., Matsuura, S., Takashima, K., Katsura, S., Mizuno, A., 2003. Single-molecule PCR using water-in-oil emulsion, Journal of Biotechnology 102(2), 117-124.

Newton, C. R., Graham, A., 1997. PCR (Introduction to Biotechniques Series).Second edition. BIOS Scientific Publishers

Noren, T., Alriksson, I., Andersson, J., Akerlund, T., Unemo, M., 2011. Rapid and sensitive loop-mediated isothermal amplification test for Clostridium difficile detection challenges cytotoxin B cell test and culture as gold standard, Journal of clinical microbiology 49(2), 710-711.

Novak-Weekley, S. M., Marlowe, E. M., Miller, J. M., Cumpio, J., Nomura, J. H., Vance, P. H., Weissfeld, A., 2010. Clostridium difficile testing in the clinical laboratory by use of multiple testing algorithms, Journal of clinical microbiology 48(3), 889-893.

Pfaffl, M. W., 2004. Quantification strategies in real-time PCR, in: Bustin, S. A. (Ed), A-Z of Quantitative PCR. La Jolla, CA, USA, International University Line (IUL), pp. 87-112.

Pfaffl, M. W., 2001. A new mathematical model for relative quantification in real-time RT-PCR, Nucleic acids research 29(9), e45.

Pienaar, E., Theron, M., Nelson, M., Viljoen, H. J., 2006. A quantitative model of error accumulation during PCR amplification, Computational biology and chemistry 30(2), 102-111.

Pirrung, M. C., 2002. How to make a DNA chip, Angewandte Chemie (International ed.in English) 41(8), 1276-1289.

Platts, A. E., Johnson, G. D., Linnemann, A. K., Krawetz, S. A., 2008. Real-time PCR quantification using a variable reaction efficiency model, Analytical Biochemistry 380(2), 315-322.

Quinn, C. D., Sefers, S. E., Babiker, W., He, Y., Alcabasa, R., Stratton, C. W., Carroll, K. C., Tang, Y. W., 2010. C. Diff Quik Chek complete enzyme immunoassay provides a reliable first-line method for detection of Clostridium difficile in stool specimens, Journal of clinical microbiology 48(2), 603-605.

Rantakokko-Jalava, K., Jalava, J., 2002. Optimal DNA isolation method for detection of bacteria in clinical specimens by broad-range PCR, Journal of clinical microbiology 40(11), 4211-4217.

Richardson, J. F., Harker, J. H., Backhurst, J. R., 2002. Coulson and Richardson's Chemical Engineering Volume 2. Elsevier.

Rubin, E., Levy, A. A., 1996. A mathematical model and a computerized simulation of PCR using complex templates, Nucleic acids research 24(18), 3538-3545.

Rupnik, M., 2009. Molecular variability in clostridium difficile large clostridial toxins, in: Brüggemann, H., Gottschalk, G. (Eds), Clostridia: Molecular Biology in the Post-Genomic Era. Norfolk, UK, Caister Academic Press, pp. 131-142.

Rupnik, M., 2010. Clostridium difficile toxinotyping, Methods in molecular biology (Clifton, N.J.) 646, 67-76.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A., Arnheim, N., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, Science (New York, N.Y.) 230(4732), 1350-1354.

Sambrook, J., Russel, D.W., 2000. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Schnell, S., Mendoza, C., 1997a. Enzymological considerations for a theoretical description of the quantitative competitive polymerase chain reaction (QC-PCR), Journal of theoretical biology 184(4), 433-440.

Schnell, S., Mendoza, C., 1997b. Theoretical description of the polymerase chain reaction, Journal of theoretical biology 188(3), 313-318.

Sloan, L. M., Duresko, B. J., Gustafson, D. R., Rosenblatt, J. E., 2008. Comparison of real-time PCR for detection of the tcdC gene with four toxin immunoassays and culture in diagnosis of Clostridium difficile infection, Journal of clinical microbiology 46(6), 1996-2001.

Spigaglia, P., Mastrantonio, P., 2002. Molecular analysis of the pathogenicity locus and polymorphism in the putative negative regulator of toxin production (TcdC) among Clostridium difficile clinical isolates, Journal of clinical microbiology 40(9), 3470-475.

Springer, T., Sipova, H., Vaisocherova, H., Stepanek, J., Homola, J., 2010. Shielding effect of monovalent and divalent cations on solid-phase DNA hybridization: surface plasmon resonance biosensor study, Nucleic acids research 38(20), 7343-7351.

Stolovitzky, G., Cecchi, G., 1996. Efficiency of DNA replication in the polymerase chain reaction, Proceedings of the National Academy of Sciences of the United States of America 93(23), 12947-12952.

Tenover, F. C., Novak-Weekley, S., Woods, C. W., Peterson, L. R., Davis, T., Schreckenberger, P., Fang, F. C., Dascal, A., Gerding, D. N., Nomura, J. H., Goering,

R. V., Akerlund, T., Weissfeld, A. S., Baron, E. J., Wong, E., Marlowe, E. M., Whitmore, J., Persing, D. H., 2010. Impact of strain type on detection of toxigenic Clostridium difficile: comparison of molecular diagnostic and enzyme immunoassay approaches, Journal of clinical microbiology 48(10), 3719-3724.

Tian, H., Huhmer, A. F., Landers, J. P., 2000. Evaluation of silica resins for direct and efficient extraction of DNA from complex biological matrices in a miniaturized format, Analytical Biochemistry 283(2), 175-191.

Thermo Scientific Pierce Protein Research Products: Chemistry of Crosslinking, 2012. 2011.

van den Berg, R. J., Kuijper, E. J., van Coppenraet, L. E., Claas, E. C., 2006. Rapid diagnosis of toxinogenic Clostridium difficile in faecal samples with internally controlled real-time PCR, Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases 12(2), 184-186.

Viljoen, G. J., Nel, L. H., Crowther, J. R., 2005. Molecular Diagnostic PCR Handbook. Dortrecht: Springer.

Vogel, A. I., Tatchell, A. R., Furnis, B. S., Hannaford, A. J., Smith, P. W. G., 1996. Vogel's Textbook of Practical Organic Chemistry. England: Prentice Hall.

Walsh, M. K., Wang, X., Weimer, B. C., 2001. Optimizing the immobilization of single-stranded DNA onto glass beads, Journal of Biochemical and Biophysical Methods 47(3), 221-231.

Waterfall, C. M., Eisenthal, R., Cobb, B. D., 2002. Kinetic characterization of primer mismatches in allele-specific PCR: a quantitative assessment, Biochemical and biophysical research communications 299(5), 715-722.

Zammatteo, N., Jeanmart, L., Hamels, S., Courtois, S., Louette, P., Hevesi, L., Remacle, J., 2000. Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays, Analytical Biochemistry 280(1), 143-150.

APPENDIX

Each PCR cycle consists of three stages: 1) denaturing of the DNA, 2) annealing of primers to single stranded DNA and 3) enzymatic elongation of the complementary strand by the DNA polymerase. The start of the cycle is defined as the beginning of the denaturing step. The overall PCR efficiency of cycle j, η^j, is the product of all of the individual efficiencies for that cycle, i.e. η_d^j, η_a^j, η_E^j, η_e^j for denaturing, annealing, polymerase binding and target elongation respectively. The denaturing damage efficiency of the polymerase, η_{dE}^j, is implicit in η_a^j, η_E^j and η_e^j. A detailed description of notations is provided in the Nomenclature section of the Appendix.

The following sections describe the derivation of expressions for the efficiency for each PCR step. Before continuing, the notation for time must be clarified. Each cycle starts with the denaturing step, but we set $t = 0$ at the beginning of the annealing stage (the denaturing step does not involve integration). Annealing occurs over the span $0 \le t \le t_a$ and the elongation stage is $t_a \le t \le t_e$. The initial number of DNA templates and primers (i.e. before denaturation in the first cycle) are $\overline{D_{init}}$ and $\overline{P_{init}}$. The amount of single stranded DNA available before annealing in the j^{th} cycle is denoted by $\overline{S_0^j}$; the amount available after annealing is denoted by $\overline{S^j}(t_a) = \overline{S_a^j}$ and after elongation, $\overline{S^j}(t_e) = \overline{S_e^j}$. The same is true for all other variables.

A.1. Efficiency of Denaturing

Double-stranded DNA molecules separate into single stranded DNA at the denaturing temperature. DNA is much more susceptible to hydrolytic attack, oxidation and depurination in the single stranded form (Cadet et al. 2002, Hsu et al. 2004, Lindahl and Nyberg 1972, 1974, Pienaar et al. 2006). Therefore a loss of template may occur in this step. An efficiency of denaturing $\eta_d \leq 1$ is defined; such that at the end of the denaturing step, the number of undamaged single stranded DNA that is available for annealing is:

$$\overline{S_0^j} = \eta_d \overline{D_e^{j-1}} \qquad \text{(A.1a)}$$

As mentioned previously, the denaturing efficiency is not an indication of the extent of strand separation, but of thermal damage to DNA. The denaturing temperature is assumed to be high enough to ensure that all the template strands separate. Since $\overline{D_e^{j-1}}$ is the number of double stranded DNA molecules available after the elongation phase at the end of the $(j-1)^{th}$ cycle, there is a one-to-one relationship between $\overline{D_e^{j-1}}$ and $\overline{S_0^j}$.

The polymerase may also incur thermal damage at the denaturing temperature (Sambrook and Russel, 2000). If the initial amount of polymerase is $\overline{E_i}$, then $\overline{E_0^1} = \eta_{dE} \overline{E_i}$ is the amount that is still active at the end of the first denaturing stage. Since the denaturation damage efficiencies, η_d and η_{dE}, depend primarily on temperature and the exposure time (denaturing period), they assumed to be constant from cycle to cycle.

Thus, at the end of the j^{th} cycle, the polymerase amount is

$$\overline{E_0^j} = \eta_{dE}\overline{E_0^{j-1}} = \left(\eta_{dE}\right)^j \overline{E_i} \qquad \text{(A.1b)}$$

For example, even a 1% loss per cycle, leads to a 33% reduction in active polymerase molecules after 40 cycles.

A.2. Annealing Model

The efficiency of the annealing stage depends on competitive binding: 5'-3' single stranded DNA, $\overline{S_0^j}$, could either bind to complementary 3'-5' single stranded DNA strands to form double stranded DNA or to their primers, $\overline{P_0^j}$, to form binary complexes. The double stranded DNA molecules are stable at the annealing temperature and, as mentioned in the list of assumptions, the primer/template products are also considered stable; making an analytical treatise possible. The ratio of templates to primers at the start of the annealing stage in the j^{th} cycle is defined as:

$$\gamma^j = \overline{S_0^j} / \overline{P_0^j} \qquad \text{(A.2)}$$

The ratio is small during earlier cycles ($\gamma^j << 10^{-2}$), but the primers are consumed and the templates increase, therefore γ^j increases with cycle number.

Neglecting non-specific binding or primer-dimer formation, three reactions remain for consideration: (1) two single stranded DNA molecules $\overline{S^j}$ can bind with rate k_S to form double stranded DNA $\overline{D^j}$, (2) a primer can anneal to a single stranded DNA molecule with rate k_P to form a binary complex, $\overline{B^j}$, and (3) a polymerase can anneal to a binary molecule with rate k_C to form a ternary complex, $\overline{C^j}$. Rate constants depend on primer sequences and PCR temperatures, and these constants can be estimated (Mamedov et al., 2008). The annealing reactions are described by the following set of equations (see Figure A.1 for a diagram of the reactions and components):

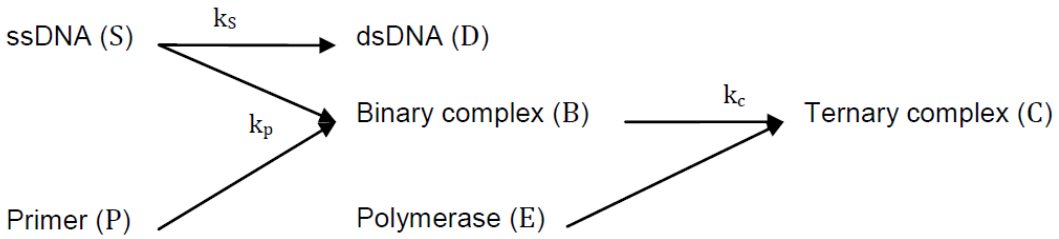

Figure A.1. Schematic diagram of annealing phase reactions showing the formation of double-stranded DNA as well as binary- and ternary-complexes.

$$\frac{d\overline{S^j}}{dt} = -k_S \overline{S^j}\,\overline{S^j} - k_P \overline{P^j}\,\overline{S^j} \tag{A.3a}$$

$$\frac{d\overline{P^j}}{dt} = -k_P \overline{P^j}\,\overline{S^j} \tag{A.3b}$$

$$\frac{d\overline{D^j}}{dt} = k_S \overline{S^j}\,\overline{S^j} \tag{A.3c}$$

$$\frac{d\overline{B^j}}{dt} = k_p \overline{P^j}\,\overline{S^j} - k_c \overline{B^j}\,\overline{E^j} \tag{A.3d}$$

$$\frac{d\overline{C^j}}{dt} = k_C \overline{B^j}\,\overline{E^j} \tag{A.3e}$$

$$\frac{d\overline{E^j}}{dt} = -k_C \overline{B^j}\,\overline{E^j} \tag{A.3f}$$

Converting to dimensionless form simplifies the analysis. All DNA amounts are scaled by the initial amount of single stranded DNA at the start of the annealing step of the j^{th} cycle, $\overline{S_0^j}$. The dimensionless variables are given by $S^j = \overline{S^j} / \overline{S_0^j}$, $P^j = \overline{P^j} / \overline{S_0^j}$, $D^j = \overline{D^j} / \overline{S_0^j}$ etc. Initial values for each cycle in the dimensionless form are: $S_0^j = \overline{S_0^j} \big/ \overline{S_0^j} = 1$, $P_0^j = \overline{P_0^j} \big/ \overline{S_0^j} = 1/\gamma^j$, $E_0^j = \overline{E_0^j} \big/ \overline{S_0^j}$ and $D_0^j = B_0^j = C_0^j = 0$. Time is scaled by the primer/template binding rate constant and the initial DNA quantity, $\tau = t k_P \overline{S_0^j}$. (If k_P has units $1/(\mu M \cdot s)$, then $\overline{S_0^j}$ must be expressed in μM.)

The dimensionless annealing equations are:

$$\frac{dS^j}{d\tau} = -\beta\left(S^j\right)^2 - P^j S^j \tag{A.4a}$$

$$\frac{dP^j}{d\tau} = -P^j S^j \tag{A.4b}$$

$$\frac{dD^j}{d\tau} = \beta\left(S^j\right)^2 \tag{A.4c}$$

$$\frac{dB^j}{d\tau} = P^j S^j - \alpha B^j E^j \tag{A.4d}$$

$$\frac{dC^j}{d\tau} = \alpha B^j E^j \tag{A.4e}$$

$$\frac{dE^j}{d\tau} = -\alpha B^j E^j \tag{A.4f}$$

The parameters $\alpha = k_C / k_P$ and $\beta = k_S / k_P$ are ratios of the reaction rate constants. The symmetry assumption allows the first term on the right hand side of Equation A.5a to be quadratic in S^j, since it is not necessary to distinguish between forward and reverse template concentrations.

Species balance equations for the primers and enzymes are given by:

$$P_0^j = 1/\gamma^j = P^j + B^j + C^j \tag{A.5a}$$

$$E_0^j = E^j + C^j \tag{A.5b}$$

Analytical approximations can be found for Equations A.5a-A.5e. A full derivation of the approximations may be found in sections A.2.1 and A.2.2. The approximations are given by:

$$S^j(\tau) \cong \frac{P^j(\tau)}{\beta - 1}\left[\left(\frac{P^j(\tau)}{\delta^j}\right)^{\beta-1} - 1\right] \tag{A.6a}$$

$$P^j(\tau) \cong \left(\gamma^j\right)^{-1}\left[\left(\gamma^j \delta^j\right)^{1-\beta}\left(1 - e^{-\delta^j \tau}\right) + e^{-\delta^j \tau}\right]^{1/1-\beta} \tag{A.6b}$$

$$B^j(\tau) \cong 1/\gamma^j - P^j(\tau) - C^j(\tau) \tag{A.6c}$$

$$C^j(\tau) \cong E_0^j - E^j(\tau) \tag{A.6d}$$

$$E^j(\tau) \cong \frac{E_0^j\left(1-\gamma^j(P^j(\tau)+E_0^j)\right)}{(1-\gamma^j P^j(\tau))\exp\left(\frac{\alpha(1-\gamma^j(P^j(\tau)+E_0^j))}{\gamma^j}\tau\right)-\gamma^j E_0^j} \qquad \text{(A.6e)}$$

The parameter δ^j is defined as $\delta^j = P^j(\tau \to \infty) = \left(\gamma^j\right)^{-1}\left[\gamma^j\left(\beta-1\right)+1\right]^{1/1-\beta}$. Equations A.6a-A.6e hold for $\beta \neq 1$, and the approximation becomes better for larger values of β, which represent more realistic cases.

The accuracy of the approximations was estimated by comparison to numerical solutions of the model above. Numerical solutions were calculated with GNU Octave v. 3.2.3 (Eaton, 2010), using the Dormand-Prince method. It was found that the approximations are extremely accurate, as illustrated in Figure 1.2. The difference between the approximations and the numerical solutions was calculated for $S^j(\tau)$, $B^j(\tau)$ and $C^j(\tau)$. For $\alpha < 1$ and $\beta = 1+10^{-6}$, it was found that this difference is less than 0.1 for $\gamma = 0.5$ and is less than 0.05 for $\gamma < 0.1$. The error increases with α and is greatest when $\beta \to 1$. The maximum error was 0.16 when $\alpha > 5$, $\beta = 1+10^{-6}$ and $\gamma = 0.5$. However, in all cases, the error tends to zero as $\tau \to 0$ and $\tau \to \tau_a$. Thus, for typical PCR conditions, the error is less than 10% during the initial phase of the reaction and negligible towards the end.

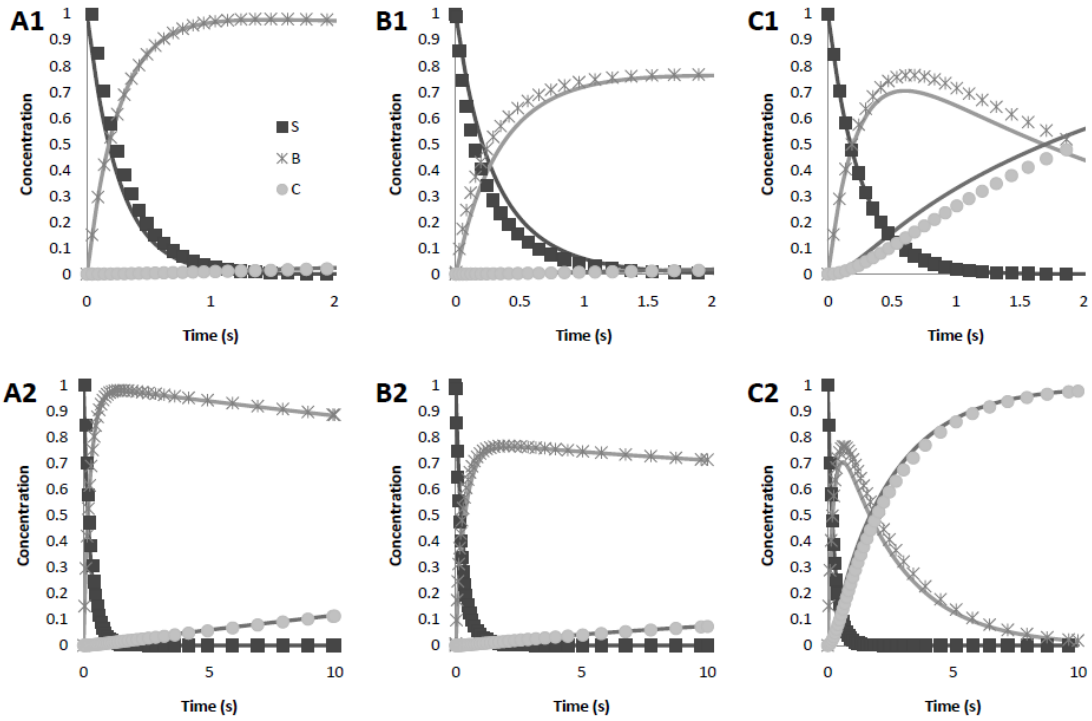

Figure A.2. The analytical approximation (solid line) as well as numerical solutions (markers) for different parameter values. The top row shows the first two seconds of the reaction, while the bottom row shows the first ten seconds. (A1&A2): $\alpha = 0.03, \beta = 1 + 10^{-6}$ and $\gamma = 10^{-3}$. These are the expected values for most PCR experiments. (B1&B2): $\alpha = 0.03, \beta = 1 + 10^{-6}$ and $\gamma = 0.5$. The higher value of γ is characteristic of the last and second to last PCR cycles. (C1&C2): $\alpha = 1, \beta = 5$ and $\gamma = 10^{-3}$. This simulation shows that the approximations hold for different values of α and β.

Two conclusions can be reached from the analytical solutions for typical annealing times $t_a \leq 10s$:

- $P_a^j \approx \delta^j$ and $S_a^j \approx 0$: The primer-DNA annealing reaction approaches completion within $t < t_a$.
- $C_a^j \approx 0$: Using $\alpha << 1$, almost no ternary complex is formed during the annealing stage.

Sections A.2.1 and A.2.2. describe the derivation of the annealing model approximations in detail. For simplicity, the superscript *j* has been dropped, and the equations listed in these sections all apply to a single cycle.

A.2.1 Calculating $\mathrm{P}(\tau)$ and $\mathrm{S}(\tau)$

First, the differential equations describing the primer and single stranded DNA reactions are given by Equations A1.1.1 and A1.1.2:

$$\frac{dP}{d\tau} = -PS \tag{A1.1.1}$$

$$\frac{dS}{d\tau} = -PS - \beta S^2 \tag{A1.1.2}$$

Since $P(\tau) \neq 0 \; \forall \; \tau$, we can divide Equation A1.1.2 by Equation A1.1.1 to get Equation A1.1.3:

$$\therefore \frac{dS}{dP} = 1 + \beta \frac{S}{P} \tag{A1.1.3}$$

This can be solved using an integrating factor to obtain Equation A1.1.4:

$$\therefore \frac{d}{dP}(P^{-\beta}S) = P^{-\beta}$$

$$\therefore S = P^{\beta}\left(\frac{P^{1-\beta}}{1-\beta} - \frac{{P_0}^{1-\beta}}{1-\beta} + {P_0}^{-\beta}S_0\right) \quad \text{(A1.1.4)}$$

If we define $\delta = \frac{(\gamma(\beta-1)+1)^{\frac{1}{1-\beta}}}{\gamma}$, and use the fact that $P_0 = \frac{1}{\gamma}$ and $S_0 = 1$, then Equation A1.1.4 becomes:

$$\therefore S = P\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1} - 1\right)}{\beta-1} \quad \text{(A1.1.5)}$$

Substituting Equation A1.1.5 into Equation A1.1.1 to obtain Equation A1.1.6:

$$\frac{dP}{d\tau} = -PS = -(P)P\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1} - 1\right)}{\beta-1} \approx -aP\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1} - 1\right)}{\beta-1} \quad \text{(A1.1.6)}$$

Here, we approximate P^2 by aP where a is some constant. We will choose the value of a later. This approximation makes it possible to solve the differential equation to obtain Equation A1.1.7. Using separation of variables and partial fractions:

$$\int_{P_0}^{P}\frac{dP}{P\left(\left(\frac{P}{\delta}\right)^{\beta-1} - 1\right)} = \int_{P_0}^{P}\frac{\left(\frac{P}{\delta}\right)^{\beta-2}dP}{\frac{1}{\delta}\left(\left(\frac{P}{\delta}\right)^{\beta-1} - 1\right)} - \int_{P_0}^{P}\frac{dP}{P} = \left(\frac{1}{\beta-1}\right)\ln\left(\frac{\left(\frac{P}{\delta}\right)^{\beta-1} - 1}{\left(\frac{P_0}{\delta}\right)^{\beta-1} - 1}\right) - \ln\left(\frac{P}{P_0}\right) = -\frac{a\tau}{\beta-1}$$

After some manipulation and using $P_0 = 1/\gamma$, we obtain Equation A1.1.7:

$$\therefore P(\tau) = \gamma^{-1}\left[(\gamma\delta)^{1-\beta}(1 - e^{-a\tau}) + e^{-a\tau}\right]^{\frac{1}{1-\beta}} \quad \text{(A1.1.7)}$$

Notice that, $P(0) = 1/\gamma$ and $\lim_{\tau\to\infty} P = \delta$. Since P is monotonically decreasing, we have $\frac{1}{\gamma} > P > \delta$. Let us reinvestigate Equation A1.1.6. If we let $a = P(0) = 1/\gamma$ or $a = \lim_{\tau\to\infty} P = \delta$ then:

$$-\left(\frac{P}{\gamma}\right)\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1}-1\right)}{\beta-1} < -(P^2)\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1}-1\right)}{\beta-1} < -(\delta P)\frac{\left(\left(\frac{P}{\delta}\right)^{\beta-1}-1\right)}{\beta-1}$$

This implies that the approximation with $a = 1/\gamma$ will decrease at a faster rate than the real situation, and the approximation with $a = \delta$ will decrease slower. Hence, letting $a = 1/\gamma$ provides a lower bound (P_l) on $P(\tau)$ and $a = \delta$ provides an upper bound (P_u) on $P(\tau)$. Since one of the goals is the optimization of the annealing time, using the upper bound on $P(\tau)$ will provide a conservative estimate. Hence, we choose $a = \delta$ to get Equation A1.1.8:

$$\therefore P(\tau) = \gamma^{-1}\left[(\gamma\delta)^{1-\beta}\left(1-e^{-\delta\tau}\right)+e^{-\delta\tau}\right]^{\frac{1}{1-\beta}} \qquad \text{(A1.1.8)}$$

The value of S can now be calculated using Equation A1.1.5. To determine the accuracy of this approximation, we calculate the ratio of the upper and lower bounds:

$$R(\tau,\gamma,\beta) = \frac{P_u(\tau)}{P_l(\tau)} = \left[\frac{(\gamma\delta)^{1-\beta}\left(1-e^{-\delta\tau}\right)+e^{-\delta\tau}}{(\gamma\delta)^{1-\beta}\left(1-e^{-\tau/\gamma}\right)+e^{-\tau/\gamma}}\right]^{\frac{1}{1-\beta}}$$

The value of the ratio $R(\tau, \gamma, \beta)$ is plotted for various values of γ and β on $0 < \tau < 4$ on Figure A.3. The higher the ratio, the greater the difference between the upper and lower bounds and the greater the error in the approximation.

From Figure A.3, it is clear that the ratio attains a maximum somewhere on $0 < \tau < 4$. The error increases as $\beta \to 1$ and $\gamma \to 1$. The maximum error attained was less than 1.1 with $\beta = 1 + 10^{-6}$ and $\gamma = 0.5$. This means that, even in the final cycle where the approximation is expected to be poorest, the error is less than 10%. More importantly, however, the ratio decreases to 1 as $\tau \to \infty$, showing that the approximation error tends to zero.

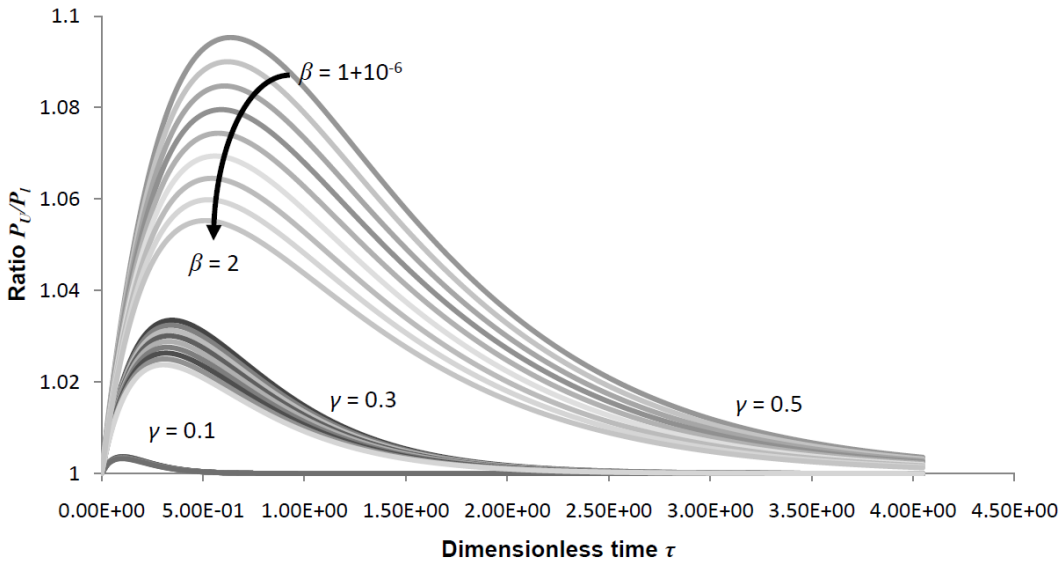

Figure A.3. Three separate bands are seen above, corresponding to $\gamma = 0.1, 0.3$ and 0.5. In each band, the value of $R(\tau, \gamma, \beta)$ increases with β, the top curve in each band corresponding to $\beta = 1 + 10^{-6}$. The maximum ratio achieved is less than 1.1, with $\beta = 1 + 10^{-6}$ and $\gamma = 0.5$.

A.2.2 Calculating $B(\tau)$, $C(\tau)$ and $E(\tau)$

We have the following kinetic equation for E (A1.1.9):

$$\frac{dE}{d\tau} = -\alpha BE \tag{A1.1.9}$$

But the species balances can be rearranged as follows, to obtain Equation A1.1.10:

$$E_0 = E + C \qquad \rightarrow C = E_0 - E$$

$$P_0 = P + B + C \rightarrow B = P_0 - P - C \qquad \rightarrow B = P_0 - P - E_0 + E$$

$$\therefore \frac{dE}{d\tau} = -\alpha E(P_0 - P - E_0) \tag{A1.1.10}$$

We can again separate and use partial fractions to obtain:

$$\therefore \int_{E_0}^{E} \left(\frac{1}{E} - \frac{1}{P_0 - P - E_0 + E}\right) dE = -\alpha \int_0^{\tau} (P_0 - P - E_0) d\tau \tag{A1.1.11}$$

One cannot integrate P with respect to E as this is a non-homogenous term in the differential equation. Furthermore, the integral $\int_0^{\tau} P d\tau$ is a very complex function. Since the value of $P_0 - P$ remains nearly constant for $\gamma \ll 1$, we assume that $P_0 - P \approx constant$. Then Equation A1.1.12 is:

$$\therefore \ln\left(\frac{E}{E_0}\right) - \ln\left(\frac{E + P_0 - P - E_0}{P_0 - P}\right) = \ln\left(\frac{E(P_0 - P)}{E_0(E + P_0 - P - E_0)}\right) = -\alpha(P_0 - P - E_0)\tau \tag{A1.1.12}$$

After some rearranging and recalling that $P_0 = 1/\gamma$, we find Equation A1.1.13:

$$\therefore E(\tau) = \frac{E_0\left(1-\gamma(P(\tau)+E_0)\right)}{\left(1-\gamma P(\tau)\right)e^{\alpha\left(1-\gamma(P(\tau)+E_0)\right)\tau/\gamma-\gamma E_0}} \quad \text{(A1.1.13)}$$

Finally, $B(\tau) = \frac{1}{\gamma} - P(\tau) - E_0 + E(\tau)$ and $C(\tau) = E_0 - E(\tau)$.

A.3. Efficiency of Primer Annealing

The efficiency of primer annealing is defined as:

$$\eta_a^j = \frac{\text{ssDNA bound to primers at } \tau_a}{\text{total available ssDNA}} = P_0^j - P^j(\tau_a) = B^j(\tau_a) + C^j(\tau_a) \quad \text{(A.7)}$$

The right hand side of Equation A.7 is obtained by rearranging Equation A.5a. The annealing efficiency is the sum of the dimensionless binary and ternary complexes at the end of the annealing period $\tau = \tau_a$ as a fraction of total available single stranded DNA. An explicit expression for annealing efficiency is obtained by substituting $P^j(\tau_a) \approx P^j(\tau \to \infty) = \delta^j$ and $P_0^j = 1/\gamma^j$, into Equation A.7:

$$\eta_a^j = \frac{1}{\gamma^j} - \delta^j \quad \text{(A.8)}$$

A.4. Efficiency of Polymerase Binding

The efficiency of polymerase binding is defined as:

$$\eta_E^j = \frac{\text{Total ternary complexes formed at the end of elongation}}{\text{Total binary and ternary complexes formed}} = \frac{C^j(\tau_e)}{B^j(\tau_e) + C^j(\tau_e)} \quad \text{(A.9)}$$

where τ_e is the dimensionless time at the end of the elongation period.

To solve for $C^j(\tau_e)$ in Equation A.9, we use Equations A.5a and A.5b (which are valid for all time), to write Equation A.4d as:

$$\frac{dB^j}{d\tau} = -\alpha B^j E^j = -\alpha B^j \left(E_0^j + B^j - \eta_a^j\right), \quad \text{for} \quad \tau \geq \tau_a \tag{A.10}$$

Note that the term $P^j S^j$ is not present, since $S^j(\tau) \approx 0$ for $\tau \geq \tau_a$. Equations A.10 and A.4e can be solved analytically for the initial data $B^j(\tau_a) = B_a^j; C^j(\tau_a) = C_a^j = \eta_a^j - B_a^j$. The parameter $\alpha = k_C / k_p$ is assumed to only change slightly for the elongation conditions, since it depends on the difference in activation energies of the two rate constants in the quotient and the difference between the annealing and elongation temperatures. The analytical solutions to Equations A.10 and A.4e, valid for $\tau \geq \tau_a$, are:

$$B^j(\tau) = \frac{\left(E_0^j - \eta_a^j\right) B_a^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\left(\tau - \tau_a\right)}}{\left(E_0^j - \eta_a^j + B_a^j\right) - B_a^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\left(\tau - \tau_a\right)}} \tag{A.11a}$$

$$C^j(\tau) = \frac{\left(E_0^j - \eta_a^j + B_a^j\right)\eta_a^j - B_a^j E_0^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\left(\tau - \tau_a\right)}}{\left(E_0^j - \eta_a^j + B_a^j\right) - B_a^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\left(\tau - \tau_a\right)}} \tag{A.11b}$$

Results A.11a and A.11b are used in Equation A.11 to obtain an explicit form for the efficiency of polymerase binding:

$$\eta_E^j = \frac{\left(E_0^j - \eta_a^j + B_a^j\right) + \left(B_a^j / \eta_a^j\right) E_0^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\tau_e}}{\left(E_0^j - \eta_a^j + B_a^j\right) - B_a^j e^{-\alpha\left(E_0^j - \eta_a^j\right)\tau_e}} \tag{A.12}$$

A.5. Efficiency of Elongation

The number of ternary complexes that extend to full-length copies depends on the elongation time. Not all the ternary complexes form at the same time. Those that form early in the elongation step have a better chance to extend fully, compared to complexes that form later in the elongation stage. The efficiency of elongation is framed within this limitation.

Denote the average extension rate at the elongation temperature as V nucleotides per second. If the length that the primer must extend is l_{ext}, then the minimum elongation time that is needed to fully extend a ternary complex is $\Delta t_{\min} = l_{ext}/V$. Therefore a cut-off time t_c exists and ternary complexes that form after the cut-off will not extend completely. The dimensionless form is: $\tau_c = \tau_e - \Delta t_{\min} k_p S_0^j \geq \tau_a$.

The efficiency of elongation is defined as the ratio of the ternary complexes that extend fully, divided by all ternary complexes that have formed.

$$\eta_e^j = \frac{C^j(\tau_c)}{C^j(\tau_e)} \qquad \text{(A.13)}$$

The solution A.11b is substituted in Equation A.13 to arrive at an expression for η_e^j:

$$\eta_e^j = \left[\frac{\left(E_0^j - C_a^j\right)\eta_a^j e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_c-\tau_a)} - \left(\eta_a^j - C_a^j\right)E_0^j}{C_a^j - \eta_a^j + \left(E_0^j - C_a^j\right)e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_c-\tau_a)}}\right]\left[\frac{C_a^j - \eta_a^j + \left(E_0^j - C_a^j\right)e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_e-\tau_a)}}{\left(E_0^j - C_a^j\right)\eta_a^j e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_e-\tau_a)} - \left(\eta_a^j - C_a^j\right)E_0^j}\right] \tag{A.14}$$

Not that the extension begins as soon as a polymerase has bound to a binary complex. Since the ternary complexes may form any time during the elongation stage, a distribution of product lengths may result. For the sake of simplicity, these incomplete products are not carried over to the next cycle in this model. This will have a negligible effect on the accuracy of the model as the partially elongated single stranded DNA will act similarly to a primer in the annealing phase of the next cycle.

A.6. The mathematical model

Four efficiencies have been defined, given by Equations (A.1, A.8, A.13 and A.15). The overall efficiency for the j^{th} cycle is the product of the four individual efficiencies;

$$\eta^j = \eta_d^j \eta_a^j \eta_E^j \eta_e^j, \tag{A.15a}$$

and it takes on the form:

$$\eta^j = \eta_d\left[\frac{\left(E_0^j - C_a^j\right)\eta_a^j e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_c-\tau_a)} - \left(\eta_a^j - C_a^j\right)E_0^j}{C_a^j - \eta_a^j + \left(E_0^j - C_a^j\right)e^{\alpha\left(E_0^j-\eta_a^j\right)(\tau_c-\tau_a)}}\right] \tag{A.15b}$$

The simplicity of Equation A.15b is somewhat misleading. We list the key PCR parameters and where they appear in Equation A.15b:

- Starting composition. The starting polymerase and primer concentrations are scaled with respect to S_0^1. They appear in Equation A.15b, for the 1st cycle as E_0^j and P_0^j. The templates for the next cycle are obtained from the values at the previous cycle: $S_0^{j+1} = S_0^j + \eta^j S_0^j$ and the updated template value is used to find the dimensionless polymerase and primer concentrations at the start of the $j+1^{th}$ cycle. Of course, one must also account for primer consumption each cycle, $P_0^{j+1} = P_0^j - \eta^j S_0^j$.
- The annealing time is implicitly present in C_a^j (cf. Equations A.6c and A.6d). The term τ_c depends on the elongation time, elongation speed and template length.
- The kinetic rate constants appear in dimensionless form as β and α. Temperature settings affect the rate constants. For example, increases in the annealing temperature would reduce binary complex formation.

The annealing efficiency, η_a^j, depends only on β and γ^j and it decreases from cycle to cycle due to primer consumption ($P_0^{j+1} < P_0^j$) and template formation ($S_0^{j+1} > S_0^j$). In Figure A.4 the efficiency of annealing is plotted as a function of γ^j. For a small γ^j (i.e. case of large excess primers), the annealing efficiency is practically 100%, regardless of the β values. If $\gamma^j > 10^{-3}$, then the efficiency starts to drop. The efficiency is more sensitive to larger values of β, because the reaction to form double stranded DNA becomes more competitive (cf. Equation A.4c). The limits of η_a^j are proper; $\lim_{\gamma^j \to 0}\left(\eta_a^j\right) = 1$ and $\lim_{\gamma^j \to \infty}\left(\eta_a^j\right) = 0$.

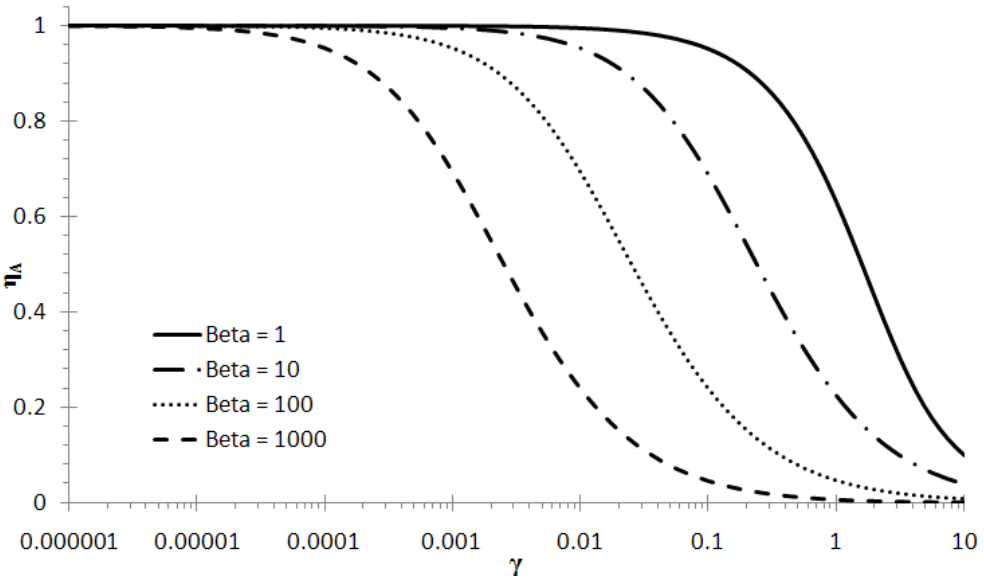

Figure A.4. The annealing efficiency (η_a) as a function of the template:primer ratio (γ).

The parameter α determines the rate of ternary complex formation. The polymerase binding efficiency, η_E^j, will increase with increasing α. However, it is expected that α is small (Mamedov et al. 2008). The expression for overall efficiency Equation A.15b becomes much simpler if no ternary complexes have formed at the end of the annealing stage (i.e. $C_a^j = 0$ in Equation 15b) - this is a good approximation if $\alpha << 1$.

Nomenclature

Symbol	Parameter	Symbol	Parameter
B	Number of binary complexes (primer-single stranded DNA template)	α	Ratio of reaction rate constants, k_C/k_P
C	Number of ternary complexes (primer-single stranded DNA template-polymerase)	β	Ratio of reaction rate constants, k_S/k_P
D	Number of double stranded DNA molecules	δ	Minimum amount of remaining primer after the annealing period
E	Number of polymerase molecules	γ	Ratio of template to primers
k_C	Reaction rate constant for a polymerase binding to a binary complex to form a ternary complex	η	Efficiency
k_P	Reaction rate constant for primer-template annealing to form a binary complex	τ	Dimensionless time
k_S	Reaction rate constant for template-template annealing to form double stranded DNA	**Superscripts**	
l_{ext}	The length that the primer must extend to become another template	j	Cycle number
n	Number of PCR cycles	**Subscripts**	
P	Number of forward/reverse primer molecules	0	Start of annealing period
S	Number of full length top/bottom single stranded DNA template molecules	a	Annealing (end of period when used in reference to time)
t	Dimensional time	c	Cut-off time
$\Delta t_{\min}$	Minimum elongation time	d	Thermal damage to DNA
Δt_e	Elongation hold time	dE	Thermal damage to polymerase
V	Average extension rate of the polymerase at the elongation temperature	e	Elongation (end of period when used in reference to time)
X	PCR yield for n cycles	E	Polymerase binding
Use of overbar indicates dimensional variable.		$init$	Initial, i.e. before denaturation in the first cycle