Advanced microarray technologies for clinical diagnostics

Citation for published version (APA):
Pierik, A. (2011). *Advanced microarray technologies for clinical diagnostics*. [Phd Thesis 2 (Research NOT TU/e / Graduation TU/e), Chemical Engineering and Chemistry]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR724484

DOI:
10.6100/IR724484

Document status and date:
Published: 01/01/2011

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 17. Nov. 2023

Advanced microarray technologies
for clinical diagnostics

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op maandag 19 december 2011 om 16.00 uur

door

Anke Pierik

geboren te Hoensbroek

Dit proefschrift is goedgekeurd door de promotor:

prof.dr. D.J. Broer

The work described in this thesis has been carried out at the Philips Research Laboratories Eindhoven, The Netherlands, as part of the Philips Research program.

Copyright © 2011 by A. Pierik

A catalogue record is available from the Eindhoven University of Technology library

ISBN: 978-94-6191-088-2

Printed by Ipskamp drukkers, Enschede, the Netherlands

Cover design by: Bertina Senders, Harma Feitsma, Anke Pierik

Table of contents

1. **Introduction and scope of thesis**

1.1 General introduction 2
1.1.1 Historic highlights in DNA research 2
1.1.2 Characteristics of DNA 2
1.1.3 Polymerase Chain Reaction 5
1.2 Introduction of microarrays 7
1.2.1 Microarrays: general description and examples of applications 7
1.2.2 Manufacturing of microarrays 8
1.2.3 Protocols of microarray assays 10
1.2.4 Challenges in use of microarrays for clinical diagnostics 12
1.3 Scope and outline of this thesis 14
References 16

2. **Quality control of inkjet technology for DNA microarray fabrication**

Abstract 20
2.1 Introduction 20
2.2 Materials and methods 22
2.2.1 Materials 22
2.2.2 Inkjet printer 22
2.2.3 Droplet inspection and in-line process control 26
2.2.4 Image processing and data analysis 27
2.2.5 Failure mode analysis 29
2.2.6 Quality control methods 31
2.3 Results and discussion 32
2.4 Conclusions 35
References 37

3. **Immobilization of oligonucleotides with homo-oligomer tails onto amine-functionalized solid surfaces and the effects on hybridization**

Abstract 40
3.1 Introduction 40
3.2 Materials and methods 42
3.2.1 Substrate 42
3.2.2 Oligonucleotides 42
3.2.3 Microarray fabrication 43
3.2.4 Microarray processing 44
3.2.5 Fluorescent intensity measurements 44
3.2.6 Immobilization efficiency calculation 46
3.2.7 Hybridization efficiency calculation 47

3.3 Results and discussion 49
3.4 Conclusions 56
3.5 References 60

4. **Rapid genotyping of HPV by post-PCR array-based hybridization techniques**

Abstract 64
4.1 Introduction 64
4.2 Materials and methods 66
4.2.1 Sample selection 66
4.2.2 Sample pretreatment 66
4.2.3 PCR 67
4.2.4 Microarray manufacturing 67
4.2.5 Microarray flow-through optimization 67
4.2.6 Microarray flow-through detection and analysis 68
4.2.7 RLB detection 70
4.2.8 EIA 71
4.3 Results 71
4.4 Discussion 79
4.5 Conclusions 81
References 82

5. **Real time amplification detection on a microarray – real time array PCR –**

Abstract 86
5.1 Introduction real time array PCR 86
5.2 Materials and methods 93
5.2.1 Microarray fabrication 93
5.2.2 Real time array PCR system 93
5.2.3 Real time array PCR protocol 94
5.2.4 Asymmetric real time array PCR 96
5.2.5 Fluorescence measurements and calculations 96
5.2.6 qPCR experiments 97
5.3 Results 97
5.3.1 Proof of concept 97
5.3.2 Optimization of mastermix and PCR protocol 101
5.3.3 Asymmetric real time array PCR 107
5.4 Discussion and conclusions 108
References 111

6. **Mathematical analysis of the real time array PCR process**

Abstract 114
6.1 Introduction 114

6.2 Mathematical modeling 116
6.2.1 Approach 116
6.2.2 Model assumptions 116
6.2.3 Basic reaction equations 117
6.2.4 Rate equations 119
6.3 Results and discussions 122
6.3.1 Geometrical, process and material data 122
6.3.2 Typical results: bulk results 126
6.3.3 Typical results: microarray surface hybridization 127
6.3.4 Optimization of experimental parameters 128
6.4 Comparison with experiments 134
6.5 Parameter sensitivity analysis 137
6.6 Discussion and conclusions 138
References 140
Appendix A: solution of the equations 140
Appendix A1: Solution of real time PCR 141
Appendix A2: Solutions for Real Time Array PCR 144

7. Technology assessment

7.1 Microarray printing 148
7.2 Surface chemistry 149
7.3 Microarray-based tests for clinical diagnostics 149
References 152

Summary 153

Nederlandse samenvatting 157

Dankwoord 163

Curriculum Vitae 167

Chapter 1

Introduction and scope of thesis

1.1 General Introduction

The first part of this chapter gives an overview of historic highlights on DNA research that finally led to the publication of the sequence of the human genome. In the second part an introduction on microarrays, protocols and current challenges in microarray technology is given. The scope of this thesis will be outlined at the end of this chapter.

1.1.1 Historic highlights in DNA research

In 1869, Friedrich Miescher performed experiments on the chemical composition of leukocytes that lead to the discovery of deoxyribonucleic acid (DNA). As this was located in the nuclei of cells, he called it “nuclein”[1]. Almost a century later, in 1944 Ostwald Avery, Colin Macleod and Maclyn McCarty revealed that the molecule DNA, and not proteins as was previously believed, was the carrier of the genetic code[2], which was in 1952 confirmed by Alfred Hershey and Martha Chase who showed that DNA is the genetic material of the T2 phage[3]. A few years later, Erwin Chargaff discovered that the base composition of DNA varies between species. He found that there were always fixed ratios between bases: the same number of adenine as thymine bases, and the same number of guanine and cytosine bases[1,4]. In 1953, the DNA structure was discovered by Rosalind Franklin, Maurice Wilkins, Francis Crick and James Watson, by using X-ray analysis[5]. They proposed a double-helix structure of DNA, for which Wilkins, Crick and Watson obtained in 1962 the Nobel Prize in Physiology or Medicine. In 1956, Kornberg isolated the first DNA polymerase, the enzyme that catalyzes the creation of new DNA strands, after which also the first artificial DNA was synthesized. The genetic code, the way DNA is translated into proteins, was deciphered over a period of five years (1961-1966) in a series of publications[6]. In 1977, Frederick Sanger developed a method to determine the order of nucleotide bases in the genome, and published the first sequenced genome of the bacteriophage ϕ174[7]. The Polymerase Chain Reaction (PCR), which is used to amplify small amounts of DNA, was developed in 1987 by Kary Mullis[8]. The Human Genome Project was started in 1989 and aimed for determining the sequence of DNA bases within the human genome. In 2003 a first complete version of the human genome was published.

1.1.2 Characteristics of DNA

DNA is a polymer consisting of deoxyribonucleotide units. A nucleotide consists of a base, a sugar and one or more phosphate groups. The sugar is a deoxyribose: this sugar lacks an oxygen atom compared to ribose. DNA consists of four bases: adenine (A) and guanine (G), both purines, and thymine (T) and cytosine (C), both pyrimidines[9]. The structure of DNA is given in figure 1.1.

The backbone of DNA, in figure 1.1 indicated in black, consists of deoxyriboses linked by phosphate groups. The 3’-hydroxyl of the sugar is joined to the 5’-end of the adjacent sugar by a phosphodiester bridge. A DNA chain has polarity, which means that one end of the chain has a polar 3'-OH group (3'-end), whereas the other has a 5'-OH group (5'-end). The DNA base sequence which is the variable part of DNA is written in the 5'->3' direction[9].

The three-dimensional structure of DNA consists of two helical chains coiled around the same axis[5]. Both chains follow right-handed helices, but the chains run in the opposite directions due to their specific orientation. The sugars are located on the outside of the chain, whereas the bases are in the inside. The diameter of the helical structure is 20 Å[10], and the adjacent bases in a chain are separated by 3.4 Å. There is an angle of 36 ° between two adjacent bases in the chain, which implies that the orientation of the structure repeats after 10 bases (34 Å)[5]. Along the molecule, alternating larger major grooves and smaller minor grooves are apparent. DNA is a very long molecule; the human genome consists of 2.9 million bases.

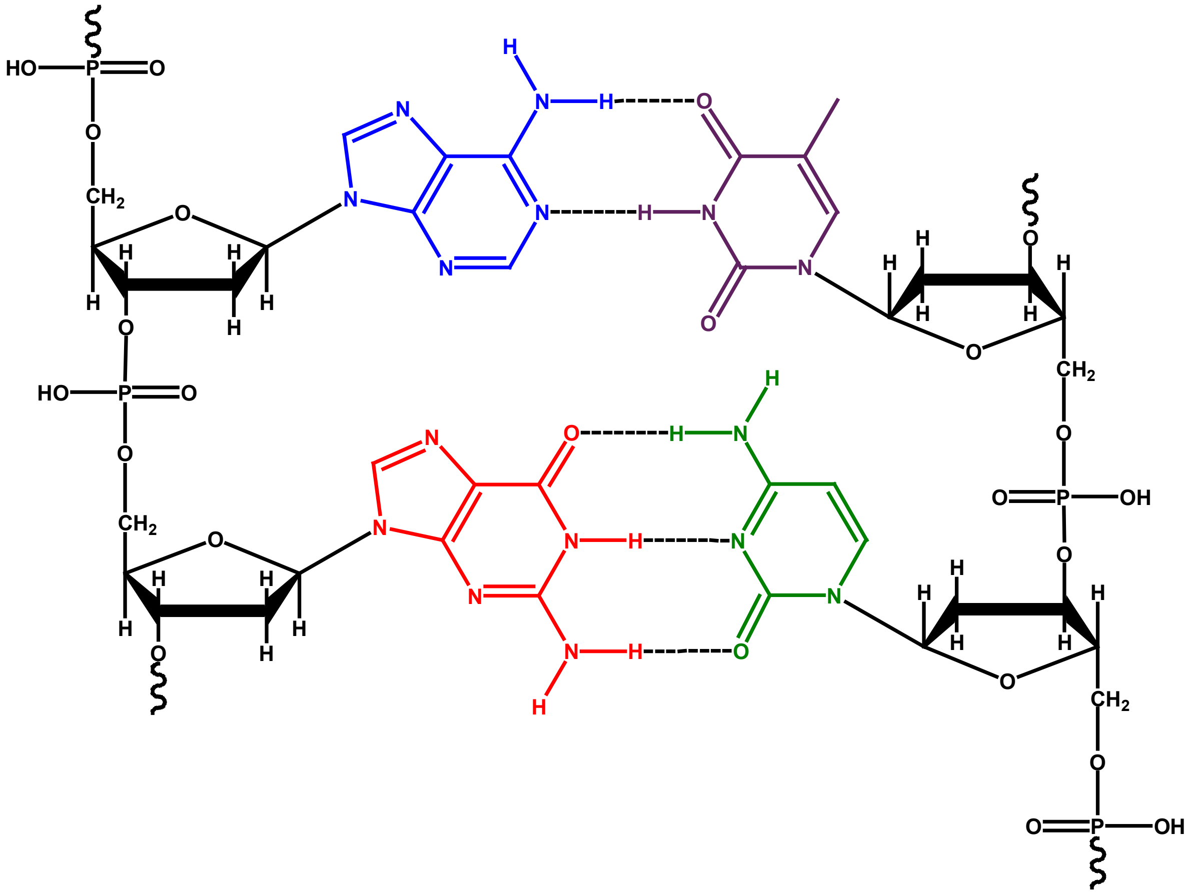


Figure 1.1 Schematic view of the structure of DNA. The DNA backbone is given in black, whereas the four different bases are given in different colors: adenine (A) (blue), thymine (T) (purple), guanine (G) (red) and cytosine (C) (green). The A-T and the G-C form hydrogen bonds, which are visualized as dotted lines. Adenine and thymine are connected by 2 non-covalent hydrogen bonds, guanine and cytosine by 3 hydrogen bonds.

The whole chain is kept together by hydrogen bonds between pairs of bases. One base of the first chain binds to a base of the other chain. The ratio of adenine and thymine, and the ratio between cytosine and guanine in double stranded DNA, is close to unity[4]. This is because adenine is always paired with thymine, and cytosine is always paired with guanine in the complementary strands. Alternative base-pairing is not possible because of sterical hindrance and hydrogen bonding requirements[9]. The planes of the bases are perpendicular to the axis of the chain. The sequence of the bases is not restricted in any way. The sequence of the bases carries the genetic information[5].

DNA denaturation or DNA melting is the reversible process where the double stranded DNA helix unwinds into the two separate single stranded DNA strands. Upon addition of heat or acids or alkali, the hydrogen bonds are disrupted. Denaturation occurs quickly at a certain temperature, which is called the melting temperature. This temperature is defined at the temperature at which half of the double helix structure is lost. This exact melting temperature depends on the base sequence as well as on the presence of salts and their concentration. GC-pairs are more stable than AT-pairs since they are held together by three hydrogen bonds instead of the double hydrogen bond for the AT-pair. Since the GC-regions have a stronger interaction, the AT-rich regions of the DNA will melt first[9,11]. The melting temperature decreases with reduced salt concentrations[12,13]. The melting temperature of longer oligonucleotides can be calculated based on the GC-fraction and the salt concentration by the following empirical correlation:

$$T_m = 16.6 \log M + 0.41(\%GC) + 81.5 \qquad (1)$$

In this equation, T_m is the melting temperature (^oC), M is the molar salt concentration (Na^+), and % GC represents the molar percentage guanine and cytosine[10,13]. Above the T_m, the DNA is largely single stranded, whereas below the T_m, the DNA is for the larger part double stranded.

For shorter DNA molecules (15-30 basepairs) the following rule of thumb can be used to estimate the T_m[11]:

$$T_m = 2(A+T) + 4(G+C) \qquad (2)$$

The opposite of DNA melting is the process of two complementary DNA strands that interact to form one double strand hybrid. This process is called DNA hybridization. DNA annealing is often used to describe the hybridization of an oligonucleotide (a primer or a probe) to a longer polynucleotide. This term is often used in combination with PCR (Polymerase Chain Reaction), see next paragraph. The thermal stability, and thus also the melting temperature, is drastically decreased due to mismatches[14]. A measurement of the melting curves can be used to discriminate a perfect match with a mismatching hybrid. It can also be used to detect Single Nucleotide Polymorphisms (SNP), which are variations in the sequence of a genome of a single nucleotide, and which are associated with disease development and therapy selection in case mismatches are present[14].

1.1.3 Polymerase Chain Reaction

The Polymerase Chain Reaction (PCR) is used to amplify specific DNA target sequences. This is done by repetitive temperature cycling during which the DNA strands are amplified by the enzymatic support of a thermostable polymerase[15,16].

A three-step PCR protocol is performed as follows[9]:

- In case a heat-activated thermostable enzyme is used, the protocol needs an initialization step of 5-10 minute at a temperature of around 95 °C.
- Denaturation phase: during this first part of the temperature cycle, all DNA is denatured at temperatures around 95 °C for 5-30 seconds. As a result only single stranded DNA is obtained.
- Annealing phase: primers specifically designed to amplify certain regions of the template DNA anneal to the now single stranded template DNA. These primers are complementary to a piece of the target sequence and thereby the specificity needed is achieved. This is often done at temperatures around 50-65 °C during 20-30 seconds.
- Elongation phase: during the elongation at around 72 °C , which is the optimal temperature of the DNA polymerase, the primers that have annealed to the template are elongated by the DNA polymerase. Hence, both strands of the target sequence are replicated, to form the so-called amplicons.

These last three steps are repeated 30-60 cycles, and during each cycle the amount of target molecules is maximally doubled. The number of copies in course of a number of cycles is visualized in the upper part of figure 1.2. In the lower part of figure 1.2 a schematic view is given of the processes taking place in within a single PCR cycle. By using labeled primers, at the end of the PCR process, labeled copies (amplicons) of the template containing the target sequence are obtained that can be identified subsequently using a post-PCR detection method, e.g. on a microarray. The annealing and the elongation phase can also be combined in one temperature step, depending on the melting temperature of the hybrid primer-DNA, which is most often done in real time quantitative PCR (qPCR).

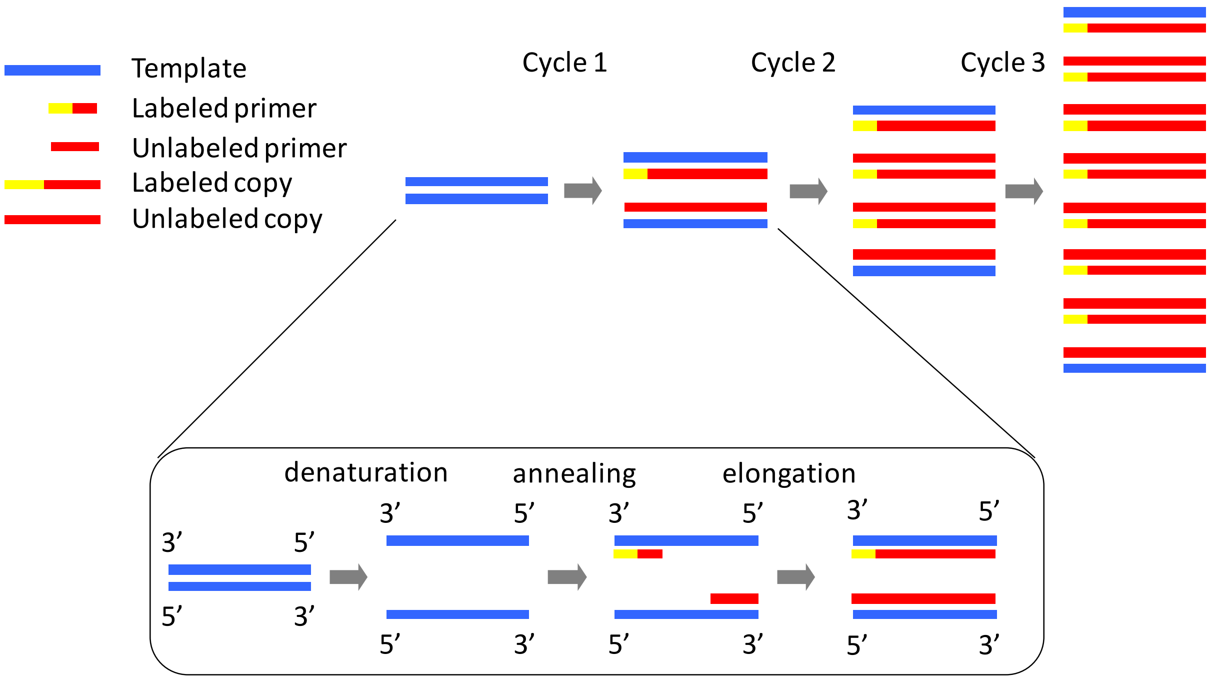


Figure 1.2 Schematic view of the steps in a PCR cycle. The upper part depicts the amount of amplicons at consecutive PCR cycles, whereas the lower part zooms in on the different steps within one amplification cycle. In this case, one of the primers contains a label, thereby producing labeled copies of the target molecule.

In qPCR, the amount of amplified molecules can be detected during each PCR cycle by the use of an intercalating dye or a probe labeled with a fluorophore. An amplification curve is obtained when plotting the obtained fluorescence as a function of cycle number. The cycle number at which the sample fluorescence signal exceeds a chosen threshold value above background signal is called the C_t-value (cycle threshold value)[17]. Since the C_t-value is determined by the amount of target molecules that was present in the sample when the reaction started, qPCR can be used to measure the original number of target molecules present. Multiplexing is achieved by using different probe-target combinations whereas the probes are labeled with different fluorophores. Multiplexing is limited by the number of different dyes that a system can detect, which usually is 6, determined by the spectral overlap.

Figure 1.3 provides a schematic view of two qPCR curves with different target input concentrations. For lower initial target concentrations, more cycles are needed for amplification until the signal is above the threshold value, resulting thereby in a higher C_t-value.

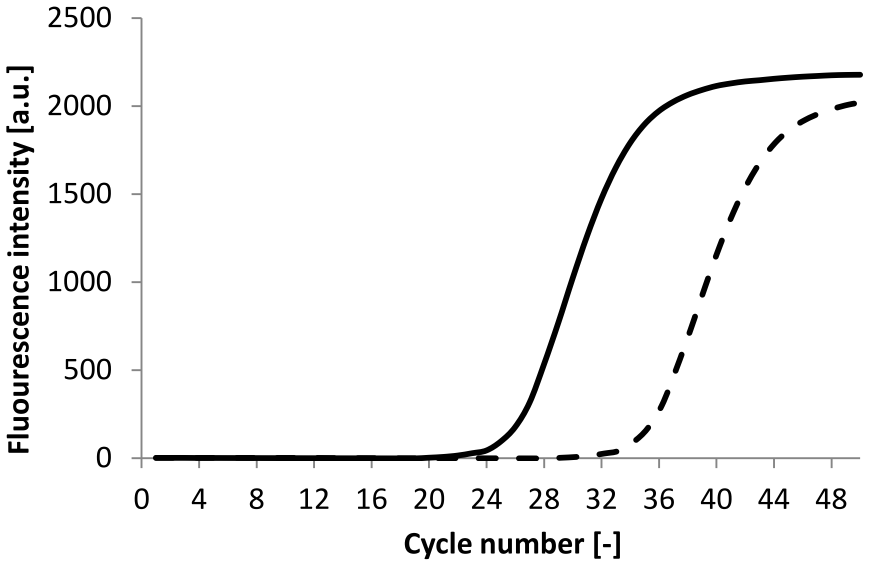


Figure 1.3 Schematic view of a qPCR curve with a relatively high target input concentration (solid line) and lower concentration (dotted line). Since a lower number of cycles are needed to amplify a target present initially in higher amounts, the C_t-value is lower.

1.2 Introduction microarrays

1.2.1 Microarrays: general description and examples of applications

Assays using DNA microarrays are powerful tools for the parallel, high throughput detection and quantification of many nucleic acid sequences[18]. Since their introduction halfway in the nineties of the 20th century[19,20], the DNA technology has evolved rapidly[21]. The use of DNA microarray-based assays have quickly expanded in various fields of molecular biology: e.g. in clinical diagnostics, gene analysis, genotyping applications, DNA mutation detection, as well as for forensic and food applications[22-24].

Microarrays have evolved from Southern blotting[25]. In Southern blotting, DNA fragments separated on an agarose gel by electrophoresis are transferred onto a membrane. The DNA is made single stranded before adding a radio-active labeled probe. The binding of the probe to the blotted DNA through base-pair interaction can be detected by radiography. The radio-labeled probe will form a band on the film at a position corresponding to the complementary sequence on the membrane[10,26].

DNA microarrays are made from non-porous substrates like glass, plastic, silicon, or from porous substrates like porous silicon or glass, or nylon membranes[18]. Onto these substrates, arrayed patterns of many microscopic spots of oligonucleotides are immobilized[24]. Each spot contains multiple copies (around 10^5 molecules/μm^2) of a specific single stranded DNA sequence. These molecules are called the capture probes.

A microarray assay is based on the preferential binding between a DNA target molecule and its complementary probe[23] that is present on the support. Under appropriate conditions, these two single stranded nucleic acid molecules with matching base sequences combine readily and form a hybridized double stranded molecule according to the Watson-Crick pairing rule. When the target molecule is labeled with e.g. a fluorescent dye, this leads in principle to a very sensitive and specific detection method using the mutual specificity between the two complementary strands of nucleic acids: target and capture probe[27]. A microarray hybridization assay consists of using the interaction of a known nucleic acid sequence, -the capture probe- , to identify the presence and the amount of complementary labeled DNA molecules in complex mixtures[28].

The number of different DNA probes varies in accordance with the application[23]. Roughly, discrimination can be made between high density and low density microarrays. In high density microarrays, up to millions of different spots are immobilized onto the substrate. This kind of microarrays is generally used for gene expression profiling, where information on the expression level of thousands of different genes can be obtained within one test. Usually mRNA or cDNA are used as samples. Drawbacks are long data analysis and interpretation times and high cost per test and therefore the use of analyzing large sample numbers is limited[29].

Low density microarrays, where the number of different spots is limited to a few hundred, offer an inexpensive, fast and relatively easy way to analyze a constrained set of DNA sequences and are therefore more suitable for routine applications[29,30].

1.2.2 Manufacturing of microarrays

Two different methods for manufacturing of microarrays exist, namely *in-situ* synthesis (see point 1 below) of the oligonucleotides on the substrate, and deposition of the pre-synthesized oligonucleotides directly onto the substrate (see point 2 below). For the former method, either photolithography or inkjet printing is generally used.

1. *In-situ* and patterned synthesis of the capture probes directly on the support. In this case, the microarray is built up base by base starting at the surface of the substrate. This takes place via the coupling of the 5’-hydroxyl group of the sugar of the last nucleotide to the phosphate group of the next nucleotide. Each nucleotide has a protective group on the 5’-end in order to prevent more than one nucleotide addition during each cycle. Two of the main technologies used in the microarray manufacturing process are photolithography or inkjet printing. Photolithographic masks are used to locally deprotect photolabile groups on the microarray surface only at those specific locations where the next nucleotide will be coupled. Inkjet printing, on the other hand, is used to couple the next nucleotide to the spots on the microarray. Four different print heads are used, each head containing one of the monomers. Both methods of microarray manufacturing are visualized in figure 1.4. Photolithography has a high degree of

position accuracy and reproducibility therein, due to the spatial specificity of the manufacturing process[31]. However, the procedure is inflexible and expensive due to the use of masks. Furthermore, the yield of each synthesized oligonucleotide is subject to wide variation and uncertainty. Using inkjet technology, DNA nucleotides are dispensed from a multi-channel inkjet print head[32,33]. Key challenges here are the required small spot sizes for high density microarrays, as well as the quality control of the process[31]. In-situ manufacturing of microarrays is very suitable for the production of high density microarrays[34]. Since the number of different target sequences can be very high it is virtually impossible to process each capture sequence individually, and therefore the in-situ synthesis is preferred above the deposition of pre-synthesized oligonucleotides.

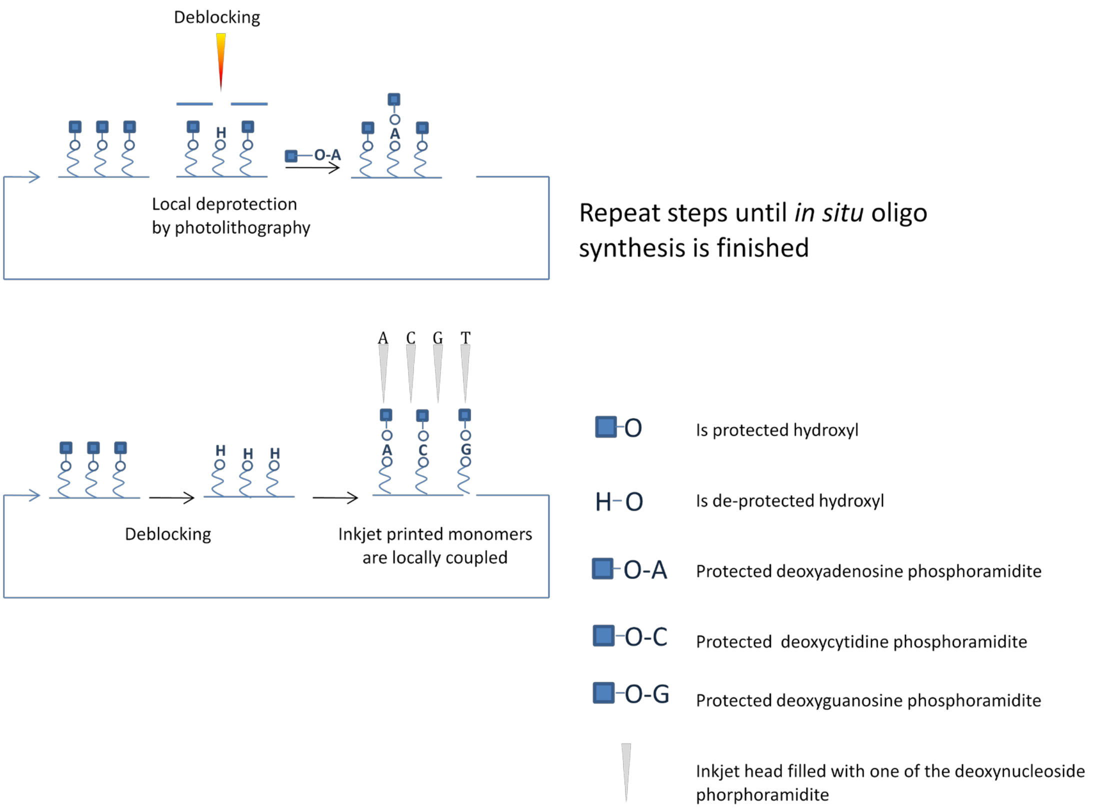

Deblocking
O-A
A
H
Local deprotection
by photolithography
Repeat steps until *in situ* oligo
synthesis is finished
A C G T
A C G
H H H
Deblocking
Inkjet printed monomers
are locally coupled
O
Is protected hydroxyl
H-O
Is de-protected hydroxyl
O-A
Protected deoxyadenosine phosphoramidite
O-C
Protected deoxycytidine phosphoramidite
O-G
Protected deoxyguanosine phosphoramidite
Inkjet head filled with one of the deoxynucleoside
phorphoramidite

Figure 1.4 *In-situ* manufacturing of oligonucleotide microarrays by photolithography (above) or inkjet printing (below).

2. Deposition of pre-synthesized oligonucleotides. In this method, the oligonucleotides are pre-synthesized off-line and subsequently deposited and immobilized onto precise locations on the solid support. Different technologies can be applied to deposit the oligonucleotides onto the substrate. In the past, most often contact spotting (pin-plotting) was used, but today non-contact spotting is the standard[23,27,31]. The advantage of non-contact spotting (e.g. inkjet printing) is the fact that direct surface contact is avoided[32]. For manufacturing of low-density microarrays, deposition of pre-synthesized oligonucleotides is used most.

Generally, two types of diagnostics based on the analysis of nucleic acids on a microarray can be discriminated: genetic (which means determination of the presence or absence of a particular targeted DNA sequence already known) and genomic, which measures gene expression or single nucleotide polymorphisms (SNP analysis)[35,36]. The latter type of microarray tests, which are commonly performed on high-density microarrays, is presently available only in clinical research laboratories because of their high complexity and dependence on expensive microarray laser scanners. The former type of microarray analysis can be performed on low density microarrays.

For (routine) clinical diagnostics it is more cost-efficient to have clinically relevant and reliable information generated by the use of low-density arrays[35]. Therefore, and because our research is directed to the use of genotyping applications for clinical diagnostics, we focus on the use of low density microarrays made by inkjet printing.

1.2.3 Protocols of microarray assays

This paragraph describes the different protocols involving the use of low density microarrays. A typical protocol is visualized in figure 1.5. Before the actual sample analysis can be performed onto the microarray, the microarray has to be manufactured and the sample has to be prepared. The upper left part of this figure describes the preparation steps of the sample, whereas the lower left part describes the manufacturing process of the microarray itself.

DNA is first extracted from the sample and purified. Subsequently, as the target DNA sequences are usually present in far below detectable concentrations, an amplification step is performed in order to increase the concentration. Furthermore, the sequences of interest need to be provided with a label in order to make detection possible afterwards. Multiple detection methods are available, of which fluorescence is most commonly used. Most often for DNA detection, a polymerase chain reaction (PCR) is performed during which the target sequences are amplified and labeled with a fluorophore at the same time. The PCR process is widely used as a molecular biological tool to replicate DNA and can create copies of specific fragments of DNA by cycling through 2 or 3 different temperature steps, as explained in paragraph 1.1.3. At each temperature cycle the amount of DNA is maximally doubled[37].

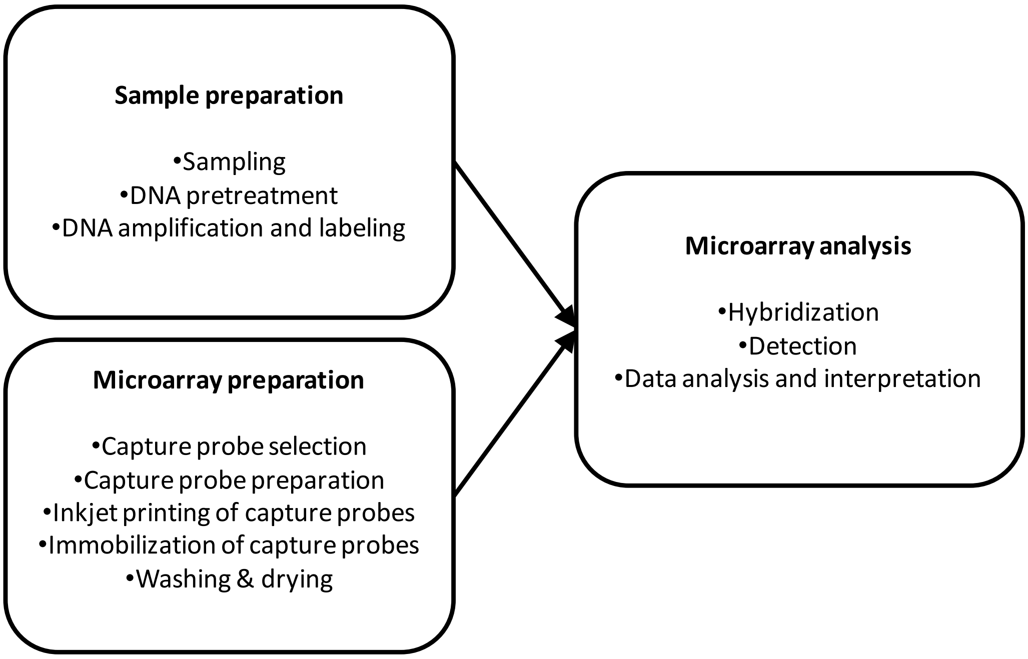

Sample preparation
•Sampling
•DNA pretreatment
•DNA amplification and labeling
Microarray preparation
•Capture probe selection
•Capture probe preparation
•Inkjet printing of capture probes
•Immobilization of capture probes
•Washing & drying
Microarray analysis
•Hybridization
•Detection
•Data analysis and interpretation

Figure 1.5 Process flow for microarray manufacturing and target detection (assay).

Although production capabilities and use of microarrays become increasingly established, significant differences exist with regard to their fabrication techniques and protocols[39]. The microarray process starts with a selection of the sequences for the capture probes. After the synthesis, purification and dry storage, they are dissolved in a printing buffer that contains salts, viscosity enhancers and/or additives for optimizing spot morphology and immobilization efficiency. During printing, between 100 picoliter and 1 nanoliter of liquid is generally applied onto the substrate for each spot, depending on the desired spot dimensions. After printing, the capture molecules are immobilized onto the substrate. The surface chemistry is a crucial step in the manufacturing process since the control of the surface chemistry should ensure a high reactivity, orientation, accessibility and stability of the capture probes[23]. Different surface chemistries exist: amine-functionalized oligonucleotides can be used that covalently bind to aldehyde, epoxy, isothiocyanate[38] or N-hydroxysuccimide activated surfaces. Unmodified oligonucleotides can bind to amine-functionalized substrates like poly-L-lysine, APS, or amine containing membranes. The DNA is often bound to these surfaces by a cross-linking reaction under exposure to 254 nm UV light[38].

The surface area that is not provided with the DNA probes is usually treated (blocked) in order to minimize non-specific binding onto the microarray substrate during hybridization. This is done by blocking the unreacted functional groups of the microarray with chemistries that have a low affinity for DNA. Next to this, this blocking step has the advantage that any unbound capture probes are washed away from the spots on the surface[38]. After drying, the microarray is ready for use.

Before bringing the sample in contact with the microarray, it is often first shortly thermally denatured in order to have single stranded DNA as much as possible. Then the sample is diluted into a high salt concentration buffer since high salt enhances hybridization. After exposure of the sample to the microarray, usually one or more washing steps are applied to wash away unbound molecules[39]. Also stringent washing conditions can be used during which formed cross-reacted hybrids can be removed[40]. After drying the substrate, the microarray is scanned and the fluorescence pattern is measured. The last step concerns the analysis of all the data obtained.

1.2.4 Challenges in use of microarrays for clinical diagnostics

Although microarray technology has led to many discoveries on evaluation of disease susceptibility, diagnoses and prognoses, the translation from research to the clinic has been slower than anticipated, largely due to technical, clinical and marketing challenges[41]. The demands on the quality of the analysis by microarrays for clinical diagnostics are very strict[42]. Microarrays are generally being considered as high complexity products by the Clinical Laboratory Improvement Amendments (CLIA). The Food and Drug Administration (FDA) in general considers microarray products as high risk applications. Therefore specific quality control measures on the manufacturing process of microarrays for clinical diagnostics are highly preferred.

High density microarrays are made by *in situ* synthesis of the probes, which is a process step that is not 100% accurate and does not allow for any confirmation of the fidelity of the synthesis[43]. This means that the microarray probes will contain a significant number of nucleotide chains that are different from the desired design of the sequence[44].

Low density microarrays, on the other hand, allow the use of pre-synthesized oligonucleotides that can be purified by capillary electrophoresis or High-Performance Liquid Chromotography (HPLC) before deposition onto the microarray. Therefore, the starting point is an almost homogeneous population of probes which increases the specificity of the assay[43,44].

However, also in the use of low density DNA microarrays, a number of challenges remain in both the qualitative as well as the quantitative analysis of nucleic acid sequences.

1.2.4.1 Reproducibility and quality of microarray analysis

One important contribution to the lack of reproducible data in microarray analysis lies in the microarray manufacturing process[21]. The transfer of tiny amounts of liquid by non-contact inkjet printing is basically an open-loop process where no feedback is obtained on the actual ejection of a droplet or series of droplets from the dispensing system. Visual inspection of a batch of manufactured microarrays afterwards is possible, though not very straightforward, with transparent biological solutions on glass or porous surfaces. Waiting until a whole batch of microarrays has been processed to see how well the deposition process has performed is not an option in high throughput manufacturing[34].

Also the surface chemistry plays a role: surface parameters like roughness, hydrophobicity, density of chemical moieties, immobilization chemistries, density and orientation of probes, spot geometry and size all have significant effect on the detected spot intensities in microarray experiments[22].

Variability in the microarray manufacturing process is thus a major contribution to the lack of reproducible microarray data.

1.2.4.2 Variations caused by sample preparation

Before the sample can be measured on the microarray, a number of consecutive steps need to be performed. Sample preparation includes a multitude of methods of extraction of nucleic acids from biological samples: shearing, amplification and labeling (or lack thereof). Each step is a potential source of quantitative bias in experimental results[22].

1.2.4.3 Lack of quantitative information

There are a number of reasons why microarrays as commonly used are not yet optimized for quantitative analysis.

First of all, the sampling of fluorescent signals is usually done only once in the course of experiment, which is commonly an end-point read-out measurement. This means that many independent variables are analyzed based on a single sample point for each target[22], (or maximally a limited number of sample points in case the same capture probes are spotted a number of times on different locations on the microarray for redundancy purposes). The measured spot intensity is determined by multiple parameters: next to the unknown target concentration it is also influenced by the capture probe density[45,46] and the hybridization kinetics. Hybridization kinetics are different for each target-capture probe hybrid: not only influenced by the capture probe sequence, but for example also by secondary structures of the target. The intra-molecular base pairs involved in secondary structures stabilize single strand conformation and can thereby slow down hybridization kinetics[47]. Furthermore, microarrays have the fundamental characteristic that they are multiplex systems. This implies that each target can potentially interact with each capture probe spot and with each other target molecule in the bulk solution. One significant problem is the control and estimation of non-specific binding and cross-reactivity to the capture molecules. Cross-reactivity in this case means lower affinity mismatched duplexes involving sequences different than the intended target[44]. Cross-reactivity is generally less probable and has dissimilar binding kinetics compared to specific binding when the target and the interfering species have the same abundance. However, if the concentration of the interfering species is higher, cross-reactivity can play a significant role[48]. A system with severe cross-hybridizations will generally have reduced ability to detect low abundance sequence species and to discriminate closely related species. With an end-point fluorescent measurement only, as commonly applied, discrimination between specific and non-specific binding is basically not possible.

Next to the analysis method itself, also the sample preparation method can play a role in the lack of quantitative data. In case of assays where an amplification step like PCR is used in order to increase the target concentration, the amplification is often performed using a high number of amplification cycles which leads to saturation. This is done to maximize the target concentrations and thus the measured spot intensities. However, it also means that the relation between measured intensity and target input concentration is not available anymore.

The reasons, as mentioned above, all make it difficult to correlate the concentration of the target present in the solution to the measured spot signals. For clinical diagnostics, this quantitative information, which can be obtained by using methods like qPCR, is a drawback in the use of microarrays with only end-point detection of the fluorescent patterns for clinical diagnostics.

1.3 Scope and outline of this thesis

This thesis focuses on low density DNA arrays for the rapid detection of a constrained and specific set of DNA target sequences for clinical diagnostics. As outlined above, a number of performance limitations hamper the translation of microarray-based R&D assays to clinical diagnostic routine. In this thesis, a number of these performance limitations are analyzed and several improvements are investigated. Thereby the adoption of both qualitative and quantitative micro-array based tests in clinical use can be facilitated.

Chapter 2 describes an inkjet system used to manufacture DNA microarrays that are applied to evaluate the microarray manufacturing process. It contains a feedback mechanism that gives direct feedback on print head malfunctioning. This can also be used to study droplet ejection failure mechanisms, since droplet formation and ejection are key for a stable printing process. This extensive study after process stability is considered to be the basis for the reliability of the processes studied in the other experimental chapters in this thesis.

In **Chapter 3**, the immobilization of 5’-end modified oligonucleotides to nylon membranes is investigated in order to improve fundamental understanding of the effect of oligonucleotide sequence on the grafting efficiency. Furthermore, the effect of the use of these modified oligonucleotides on the hybridization is investigated. The surface chemistry studied allows for a cheap and relatively easy process to be implemented for microarray manufacturing.

In **Chapter 4** a DNA detection method based on kinetic hybridization measurements on a porous substrate is benchmarked to the Reverse Line Blot, which is a method currently used in the clinic. This method allows for the measurement of hybridization kinetics as well as the detection of melting curves. A major improvement of this method is the potential identification of cross-hybridizations. For benchmarking, the Human Papilloma Virus is used as a model assay.

Chapter 5 describes a new method to integrate DNA amplification and detection on a microarray. This so-called real time array PCR protocol combines the advantages of high multiplex DNA detection of microarrays with the quantitative properties of qPCR. The fluorescence pattern that develops during cycling on the spotted microarray surface is measured with a confocal scanner built up out of standard components of the electronics industry. In this chapter, feasibility of this method is shown.

Within real time array PCR, a number of simultaneous biochemical processes take place: amplification of DNA in solution and hybridization of the amplicons to the microarray surface. **Chapter 6** describes mathematical models of these biochemical processes. The model can be used to improve the fundamental understanding of this concept and to perform a parameter study to further optimize the real time array PCR process.

In **Chapter 7** the results are summarized and placed in perspective of future research.

References

1. Dahm R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum Genet* 2008, 122:565-581
2. Avery OT, MacLeod CM, McCarthy M. Studies of the chemical nature of the substance inducing transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. *J Exp Med* 1944, 79:137-158
3. Hershey AD, Chase M. Independent functions of viral proteins and nucleic acid in growth of bacteriophage. *J Gen Physiol* 1952, 36:39-56
4. Chargaff E, Lipshitz R, Green C, Hodes ME. The composition of the desoxyrubonucleic acid of salmon sperm. *J Biol Chem* 1951, 191(1):223-230
5. Watson JD, Crick FHC. Molecular structure of nucleic acids. *Nature* 1953, 171:737-738
6. Nirenberg M. Historical review: deciphering the genetic code – a personal account. *Trends in biochem Sci* 2004, 29(1):46-54
7. Sanger F, Air, GM, Barrell BG, Brown NL, Coulson AR, Fiddes JC, Hutchison CA, Slocombe PM, Smith M. Nucleotide sequence of bacteriophage ϕX174 DNA. *Nature* 1977, 265:687-695
8. Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain-reaction, *Method Enzymol* 1987, 155:335-350
9. Stryer L. Biochemistry, 1998. W.H. Freeman and Company, New York, ISBN 0-7167-2009-4
10. Reece RJ. Analysis of genes and genomes. 2005. John Wiley & Sons Ltd, England. ISBN 0-470-84379-9
11. Marmur J, Doty P. Determination of base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J Mol Biol* 1962, 5:109-118
12. Rice SA, Doty P. The thermal denaturation of desoxyribose nucleic acid. *J Amer Chem Soc* 1952, 79:3937-3947
13. Schildkraut C, Lifson S. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* 1965, 3(2):195-208
14. Wallace RB, Shaffer J, Murphy RF, Bonner J, Hirose T, Itakura K. Hybridization of synthetic oligonucleotides to ϕχ174 DNA: the effect of single base pair mismatch. *Nucleic Acids Res* 1979, 6(11):3543-3557
15. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 1985, 230:1350-1354
16. Sambrook J, Russel DW. Molecular Cloning: A Laboratory Manual (3rd ed.) 2001. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 0-87969-576-5
17. Dorak MT. Real-time PCR. 2006. Taylor & Francis Group. ISBN 1-415-37734-X
18. Bodrossy L, Sessitsch A. Oligonucleotide microarrays in microbial diagnostics. *Curr Opin Microbiol* 2004, 7:245-254
19. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray, *Science* 1995, 270:467-470
20. Dufva M, Fabrication of high quality microarrays. *Biomol Eng* 2005, 22:173-184
21. Draghici S, Khatri P, Eklund A, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet* 2006, 22(29):101-109
22. Bischop J, Chagovetz AM, Blair S. Kinetics of multiplex hybridization: mechanisms and implications. *Biophys J* 2008, 94:1726-1734

23. Sassolas A, Leca-Bouvier BD, Blum LJ. DNA biosensors and microarrays. *Chem Rev* 2008, 108:109-139
24. Kumar, RM. The widely used diagnostics “DNA microarray” – a Review. *Am J Inf Dis* 2009, 5(3):207-218
25. Maskos U, Southern EM. Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesized *in sit*. 1992. 20(7):169-1684
26. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975; 98:503-517
27. Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent J. Expression profiling using cDNA microarrays, *Nat Genet suppl* 1999, 21:10-14
28. Mocanu D, Kolesnychenko A, Aarts S, Troost-Dejong A, Pierik A,Vossenaar E, Stapert H. Mass transfer effects on DNA hybridization in a flow-through microarray. *J Biotechnol* 2009, 139:179-185
29. Alvarez P, Saenz P, Arteta D, Martinez A, Pocovi M, Simon L, Giraldo P. Transcriptional profiling of hematologic malignancies with a low-density DNA microarray. *Clin Chem* 2007, 53(2):259-267
30. Gillet J-P, Efferth T, Steinbach D, Hamels J, de Longueville F, Bertholet V, Remacle J. Microarray-based detection of multidrug resistance in human tumor cells by expression profiling of ATP-binding cassette transporter genes. *Cancer Res* 2004, 64:8987-8993
31. Chang TN, Parthasarathy S, Wang T, Ghandi K, Soteropoulos, P. Automated liquid dispensing pin for DNA microarray applications. *IEEE Transactions on Automat Science Eng* 2006, 3(2):187-191
32. Okamoto T, Suzuki T, Yamamoto N. Microarray fabrication with covalent attachment of DNA using bubble jet technology. *Nature Biotech* 2000, 18:438-441
33. Laustedt C, Dahl T, Warren C, King K, Smith K, Johnson M, Saleem R, Aitchison J, Hood L, Lasky SR, POSam: a fast, flexible, open-source, inkjet oligonucleotide synthesizer and microarrayer, *Genome Biol* 2004, 5(8):R58
34. Fisher W, Zhang M. A biochip microarray manufacturing system using inkjet technology. *IEEE Transformations on Automat Science Eng* 2007, 4(4):488-500
35. Mikhailovich V, Gryanudov D, Kolchinsky A, Makarov AA, Zasedatelev A. DNA microarrays in the clinic: infectious diseases. *BioEssays* 2008, 30:674-682
36. Duer R, Lund R, Tanaka R, Christensen DA, Herron JN. In-plane parallel scanning: a microarray technology for Point-of-care testing. *Anal Chem* 2010, 82:8856-8865
37. Zhang C, Xu MJ, Ma W, Zheng W. PCR microfluidic devices for DNA amplification*, Biotech Adv* 2006, 24:243-284
38. Taylor S, Smith S, Windle B, Guiseppi-Elie A. Impact of surface chemistry and blocking strategies on DNA microarrays. *Nucleic Acids Res* 2003, 31(16):e87
39. Lesk AM. Introduction to genomics. 2007. Oxford University Press. ISBN 978 0199296958
40. Dufva M, Petersen J, Poulsen L. Increasing the specificity and function of DNA microarrays by processing arrays at different stringencies. *Anal BioAnal chem.* 2009, 395:669-677
41. Li X, Quigg RJ, Gu W, Rao N, Reed E. Clinical utility of microarrays: current status, existing challenges and future outlook. *Curr Genomics* 2008, 9:466-474
42. Wu Y, de Kievit P, Vahlkamp L, Pijnenburg D, Smit M, Dankers M, Melchers D, Stax M, Boender PJ, Ingham C, Bastiaensen N, de Wijn R, van Alewijk D, van Damme H, Raap AK,

Chan AB, van Beuningen R. Quantitative assessment of a novel flow-through porous microarray for the rapid analysis of gene expression profiles. *Nucleic Acids Res* 2004, 32(15):e123

43. Ramakrishnan R, Dorris D, Lublinksy A, Nguyen A, Domanus M, Prokhorova A, Gieser L, Touma E, Lockner R, Tata M, Zhu X, Patterson M, Shippy R, Sendera TJ, Mazumder A. An assessment of Motorola Codelink™ microarray performance for gene expression profiling applications. *Nucleic Acids Res 2002*, 30(7):e30
44. Dai H, Meyer M, Stepaniants S, Ziman M, Stroughton R. Use of hybridization kinetics for diffentation specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res* 2002, 30(16):e86
45. Peterson AW, Heaton RJ, Georgiadis RM. The effects of surface probe density on DNA hybridization. *Nucleic Acids Res* 2001, 29(24):5163-5168
46. Fuchs J, Fiche J-B, Buhot A, Calemczuk R, Livache T. Salt concentration effects on equilibrium melting curves from DNA microarrays. *Biophys J* 2010, 99:1886-1895
47. Gao Y, Wolf LK, Georgiadis RM. Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res* 2006, 34(11):3370-3377
48. Hassibi A, Vikalo H, Riechmann JL, Hassibi B. Real-time DNA microarray analysis. *Nucleic Acids Res* 2009, 37(20):e132

Chapter 2

Quality control of inkjet technology for DNA microarray fabrication

This chapter is based on:

Pierik A, Dijksman JF, Raaijmakers ATA, Wismans AJJ, Stapert HR. Quality control of inkjet technology for DNA microarray fabrication, *Biotechn J* 2008, 3:1581-1590

Abstract

A robust manufacturing process is essential to make high-quality DNA microarrays, especially for use in diagnostic tests. We investigated different failure modes of an inkjet printing process used to manufacture low-density microarrays. A single nozzle inkjet spotter was provided with two optical imaging systems, monitoring in real time the flight path of every droplet. If a droplet emission failure is detected, the printing process is automatically stopped. We analyzed over 1.3 million droplets. This information was used to investigate the performance of the inkjet system and to obtain detailed insight into the frequency and causes of jetting failures. Of all the substrates investigated, 96.2 % were produced without any system or jetting failures. In 1.6 % of the substrates, droplet emission failed and was correctly identified. Appropriate measures could then be taken to get the process back on track. In 2.2 %, the imaging systems failed while droplet emission occurred correctly. In 0.1 % of the substrates, droplet emission failure occurred that was not timely detected. Thus, the overall yield of the microarray manufacturing process was 99.9 %, which is highly acceptable for prototyping.

2.1 Introduction

Microarray-based assays are developing as powerful high-throughput analysis methods for life sciences and diagnostic applications. As explained in chapter 1, microarrays are substrates covered with multiple microscopic spots consisting of so-called capture probe molecules that are able to bind to specific target molecules in a sample. The number of different spots on the substrate is directly related to the number of different analytes that can be measured in a single measurement. A variety of microarrays already exists for different applications, with various types of capture probes and different manufacturing techniques[1-4].

Existing techniques for dispensing the picoliter to nanoliter droplets containing the oligonucleotides that need to be deposited onto the microarray substrate are contact printing (pin plotting) and non-contact printing (inkjet printing). Inkjet printing is a very robust and more controllable method to dispense the amount of fluid onto the substrate. Furthermore, inkjet printing is flexible in terms of spot size, number of different fluids, and grid layout[5]. Because of the non-contact nature of inkjet printing, this technology is suitable for printing materials onto the substrate without any damage to the print head or substrates, including porous or brittle substrates.

Deposition of presynthesized oligonucleotides is more favorable for low-density arrays compared to *in situ* synthesis of the capture probes on the microarray surface, since only a limited number of fluids need to be handled, and the largest flexibility in the process is obtained. The number of spots and liquids to be processed can be flexibly changed. The work reported here is thus focused on non-contact piezo inkjet printing to manufacture presynthesized high-quality oligonucleotide DNA microarrays.

Inkjet printing is commonly used in the graphical industry and is the leading technology in consumer printers. High-density multi-nozzle print heads, each having an arrangement of a large number of nozzles, deposit relatively large amounts of a small number of different fluids. For graphical applications, the quality of the image is very important. The resolution of a printed image is expressed in dots per inch (dpi), and resolutions of over 4800 dpi can be printed. A considerable number of failing nozzles are acceptable[3] as the printing resolution is higher than the resolution of the human eye.

For microarray manufacturing, the process and the requirements are different. Much smaller amounts of a large number of different fluids are deposited in individual spots onto substrates, consisting of a single droplet per spot or of a burst of multiple droplets fired at a high frequency. One failing nozzle can have important undesired implications: the target molecule may not be identified due to a missing spot or a too low concentration of capture probe molecules in a spot, possibly leading eventually to a wrong diagnosis. There is thus a clear need for a high-quality, robust and reproducible method for manufacturing microarrays[6].

Usually, droplet formation is studied offline before printing, e.g., by direct optical inspection or laser scattering[3,7-9]. Upon applying an electrical pulse to the piezoelectric actuator, a properly functioning nozzle delivers a droplet with a velocity within a given window (defined by a minimum and a maximum value). The straightness (defined as the deviation from the ideal flight path) is within a cone around the ideal flight path with a half apex angle smaller than 10–25 millirad. If a nozzle does not function properly, it can be discarded before the actual printing process starts.

Although the nozzles may emit droplets within the specifications at the start of the process, a nozzle can still fail during the process for several reasons. Apart from electrical and mechanical problems, air entrapment[10], emptying of the print head reservoir, nozzle clogging, fouling of the nozzle front and nozzle front wetting are all phenomena that can cause nozzle dropout during operation. The only way to check the quality of the printing process is by checking the result, as the process itself does not give any direct feed-back. The microarray can be imaged for identification of missing spots or misfirings. This can also be done inline during high-volume production to keep the production speed high[6]. A repair stroke might in some cases be applied to print the missing spots, but at the expense of an increased printing time and it can only be used for correcting missing spots. If seriously misaligned droplets or satellites are visible, the substrates need to be discarded from the batch. This obviously reduces the overall yield of the manufacturing operation. Therefore, in-line and real-time information about the jetting behavior of the print head is necessary for diagnostic arrays for which the presence and correct placement of the spots must be guaranteed at all times. Through accurate and timely identification of nozzle failure as well as the ability to predict upcoming nozzle failings, it is possible to obtain a higher uptime of the print system and an improved overall yield of the process. A quality control method has been described[6], in which a line-scan camera is used to detect the presence of a spot on the substrate by measuring the contrast between salt crystals in the spot and the transparent substrate. This gives valuable information on spot level. However, no information is revealed about the frequency and the underlying causes of missing spots. To find and categorize the causes of missing spots when analyzing substrates is virtually impossible, as in between print head actuation and substrate analysis time has passed. This elapsed time is at least a few orders of magnitude larger than the time it takes to generate a droplet and to form a spot on the substrate (microseconds to milliseconds).

We developed a printing system for microarray prototype production in which the jetting performance is evaluated on droplet formation level, which is as close as possible to the print head where failures can occur. With this system, deviations from proper droplet formation are detected in line and in real time, and immediate action can be taken to stop the printing process and to repair the print head so that it functions correctly again. This closed loop control of the inkjet print head and the process makes the manufacturing process more robust and consequently increases the quality of the microarrays produced. At the same time, such a system can be considered as a measurement tool that allows, on the one hand, precise analysis of the printing process and, on the other hand, the analysis of the precision of the measurement tools used. This analysis validates the use of this type of spotting instrumentation for microarray production.

Over the course of 15 months, a large number of arrays were printed and used for hybridization research[11] as well as for the precise characterization of the spotting process. Through the availability of these process data, statistically relevant data on the inkjet printing process have been gained, and it is possible to investigate different failure mechanisms, to find and to characterize the built-in limitations of the measurement procedures and to identify the areas where the process or print system can still be improved.

2.2 Materials and methods

2.2.1 Materials

All oligonucleotide capture probe molecules were dissolved in MilliQ Ultra-Pure water. As the nature of the experiments performed was diverse, many different oligo sequences at concentrations between several nanomolar and one hundred micromolar were processed. The fluids were stored in polypropylene screw cap vials (Sarstedt, Nümbrecht, Germany) and kept at 4 °C.

2.2.2 Inkjet printer

A microarray spotter was developed based on piezo inkjet technology, see figures 2.1 and 2.2. A precision linear stage with optical ruler with a resolution of 0.25 µm (Newport

Corporation, Irvine CA, USA) was mounted on a granite table. The substrates were placed on a fixture plate attached rigidly to this stage. Vials containing the print fluids to be processed as well as containers for cleaning fluids and waste material were placed on the fixture plate. A bridge was mounted rigidly between this table and a precision linear stage moving perpendicularly to the motion of the first stage. The print head was mounted to this stage with the jets pointing downwards. Any position on the fixture plate could be reached by simultaneously moving the fixture plate and/or the print head. The print head was mounted on a carriage of a vertical precision stage making vertical movements possible. The optics for droplet inspection, a bar code reader and a sensor for detecting fluid levels in the vials and cleaning fluid containers were mounted on the carriage of the print head.

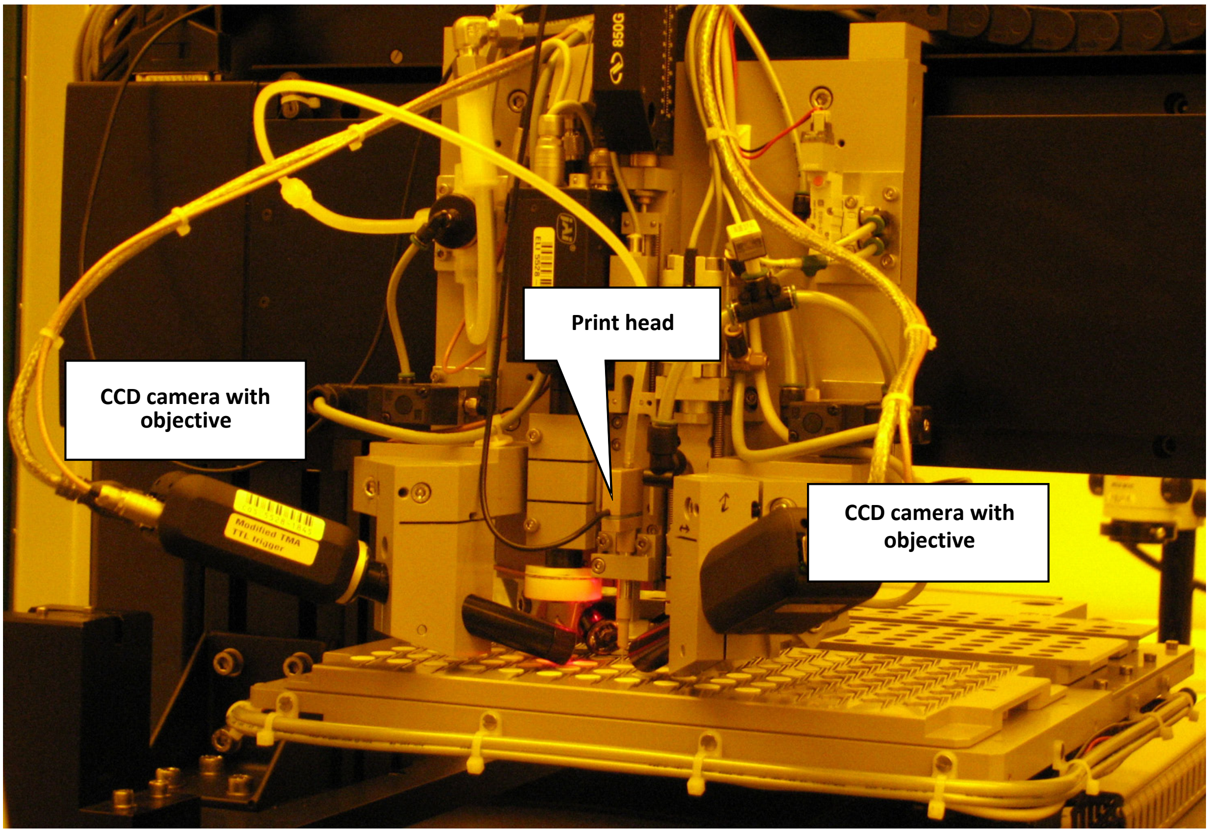


Figure 2.1 Inkjet system used to manufacture DNA microarrays. Onto a first stage, a carrier plate with substrates are mounted. Rigidly to this table, a bridge is mounted with another precision linear stage moving perpendicularly to the motion of the first stage. The print head is mounted to this stage, and jets downwards. Any position on the fixture plate can be reached by simultaneously moving the fixture plate and/or the print head. In the front of the picture, the two cameras are visible that are used to image all droplets jetted by the print head. In figure 2.2, more details of the print head and the imaging systems (CCD cameras with microscope objectives and stroboscopes) are visible.

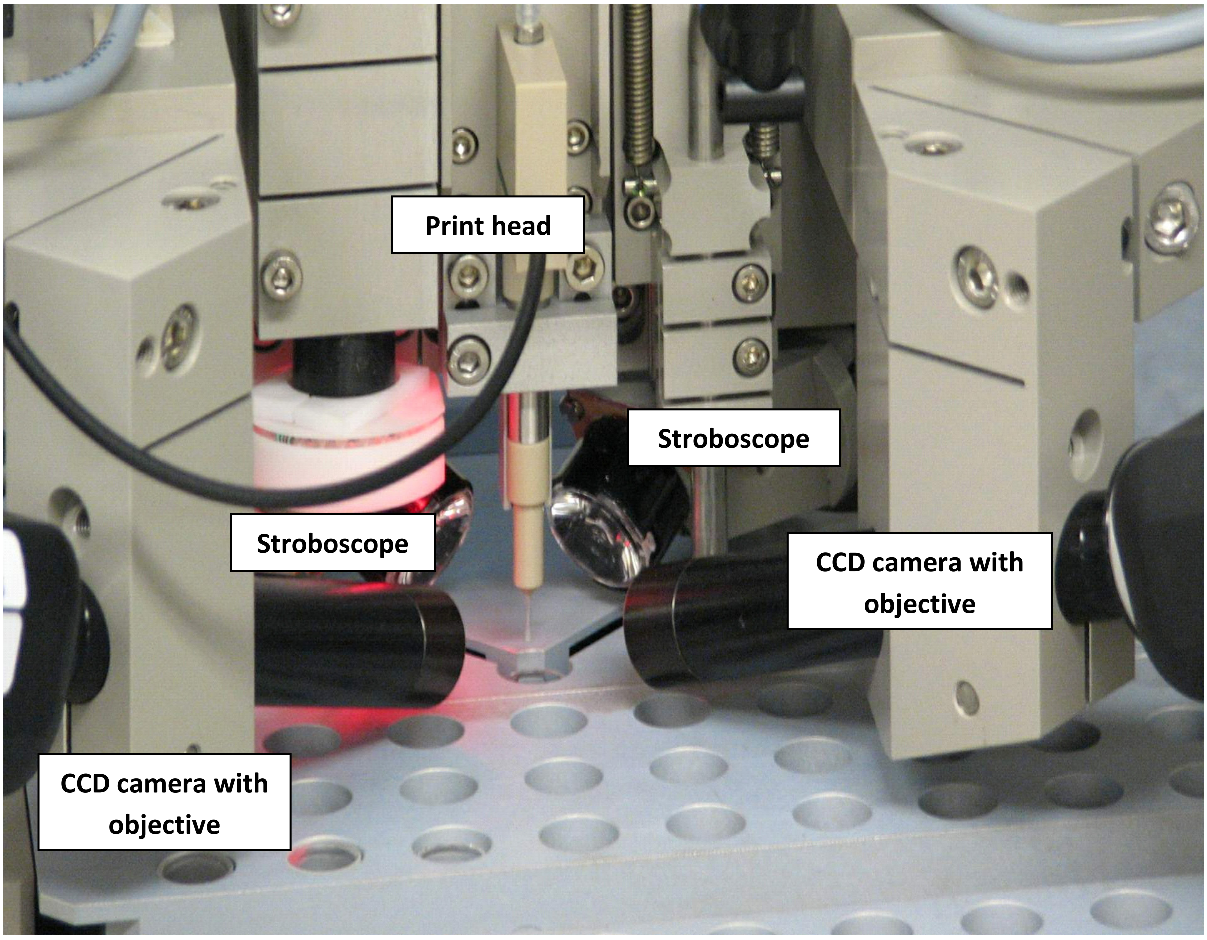


Figure 2.2 Inkjet system zoomed in on print head and the two optical systems. On the front, two CCD cameras with microscope objectives are visible. On the back, two stroboscopes are located. Two other stroboscopes (not visible in this picture) are located underneath the XY-plane.

The print head was a single nozzle Autodrop AD-H-501 pipette (Microdrop Technologies, Norderstedt, Germany), consisting of a glass capillary and an integrated piezo actuator. The nozzle diameter of the print head was 50 µm. A Microdrop AD-E-130-NP pressure control unit maintained a 980-Pa (10-cm water column) under-pressure to the print head to prevent the nozzle from emptying spontaneously. A block pulse, generated by a general-purpose pulse generator, was applied for droplet generation (figure 2.3). When applying this pulse, the meniscus first retracts and then a droplet is generated. The height and the width of the pulse are variable, making it possible to optimize droplet formation for every print head and fluid used. Generally, the voltage applied is between 50 and 70 volts, with a pulse width between 28 and 32 µs. Using these settings, a single droplet can be generated without any additional satellite droplets, with an average volume varying between 100 and 140 pL, depending on the settings, and a droplet velocity of around 2 m/s.

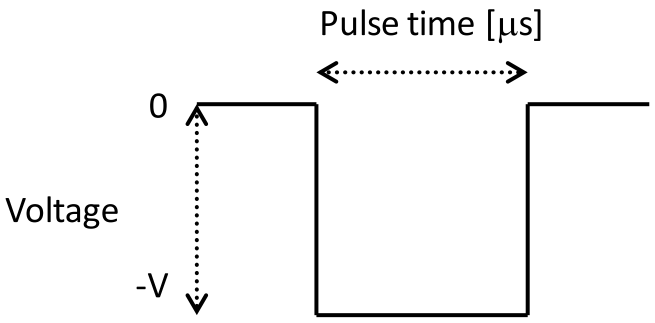


Figure 2.3 Pulse shape as used for inkjet printing of oligonucleotide solutions using the print system described. The pulse is a negative block pulse having a voltage between 50-70 V and a pulse width of 28 – 32 μs.

The software for driving the printer was designed such that in principle any pattern of spots can be deposited. For the experiments reported here, a 14 × 10 rectangular grid of spot locations was used for the microarray pattern. The pitch in x- and y-direction was 400 μm, leading to a spot density of 6.25 spots/mm^2. The average spot volume was 1000 pL. With the printer settings described above, this means that every spot was built up of seven to ten consecutive droplets.

The method for the real-time investigation of the droplet formation using an imaging system is described in section 2.2.3. Even when using identical printer settings, small deviations in droplet volume or straightness occur when the same fluid is printed repeatedly with fluid exchange in between. This depends, among others, on the state and history of the print head or the presence of small particles or air bubbles inside it. On comparing the first and the last few droplets of the same fluid jetted during one run, both the droplet volume and the straightness are generally unchanged. If not, this is usually a forerunner of a complete print head failure, which can be detected with the imaging systems as described below.

The droplets needed to make a single spot are deposited at a frequency of 50 Hz, during which the whole print system remains at exactly the same location. After printing the spot, the print head and substrate plate are moved towards the next spot position with a maximum speed of 10 (print head) and 40 (substrate plate) mm/s. Depending on the distance between two successive spots, this results in an idle time of 1–3 s. Between the printing of two successive fluids, the print head is thoroughly cleaned with cleaning fluid, soft sonic agitation and dry nitrogen gas. All software required to control the functions of the printer has been written in Lab-VIEW.

2.2.3 Droplet inspection and in-line process control

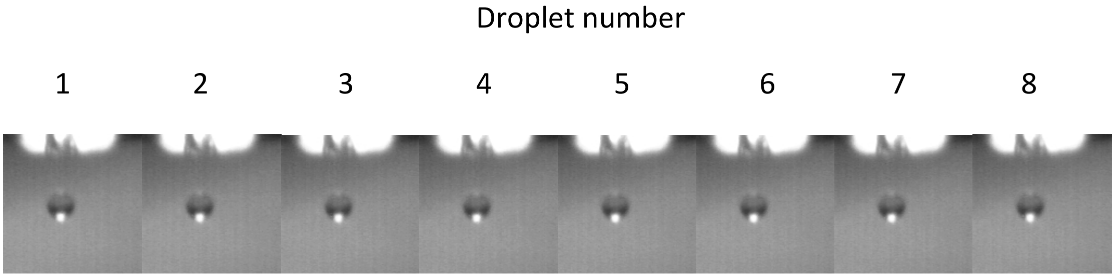
Droplet inspection is always performed before a microarray printing process starts. The tip of the pipette is small (outer diameter at nozzle tip 0.6 mm), making it possible for an optical system to capture an image of a droplet traveling from the nozzle through the air towards the substrate. The droplet is imaged via shadow illumination, resulting in a very sharp image. This image is used for the measurement of the volume of the droplet (converted from measuring the shadow area of the image) and its speed and straightness. Using two visualization systems (projected onto the XY plane) with the optical axes perpendicular to each other, information on the flight path of the droplet in three-dimensional space is obtained. In our case, we used two cameras (Philips Dica 121, Philips Applied Technologies, Eindhoven, The Netherlands) with microscope objectives and two red LED sources (Philips LumiLEDs, San Jose CA, USA). By coupling the stroboscopic flash of the LED directly to the fire pulse, a still image of every single droplet was generated[7]. To obtain the exact flight path of the droplet, a series of ten images was made by gradually increasing the delay time between droplet generation and flash moment. As the distance between the nozzle plate and the substrates was known, the droplet landing position could be calculated from these data. The co-ordinates of the stages were adjusted such that the droplets landed on the desired spot locations on the substrates. This feed-forward procedure for correcting straightness errors turned out to be effective, because during a stable print run the straightness hardly changed.

During inspection, droplet formation can be optimized for volume, speed and straightness by increasing the voltage. To prevent the formation of satellites, the voltage can be slightly decreased to end up with a single droplet. Also the pulse time can be used as a variable for fine-tuning droplet volume and speed. Generally, however, voltage adjustments already suffice to obtain single droplets with a velocity of around 2 m/s. As mentioned before, droplet formation within specifications does not guarantee correct droplet formation during the whole print run, as droplet formation may fail unpredictably. Every droplet jetted during printing is therefore imaged by the same two visualization systems used during droplet inspection. The droplets are illuminated indirectly by reflection via the substrate. This way of illumination results in a lower image quality compared to shadow illumination, but the image captured is still suitable for droplet detection and for the measurement of the position of the droplet in space, (e.g., for detection of changes in straightness). However, this method is not suitable for droplet volume measurements because of its low image resolution. Volume measurements are not needed for process control, as a thorough volume measurement is already done prior to printing during the droplet formation study.

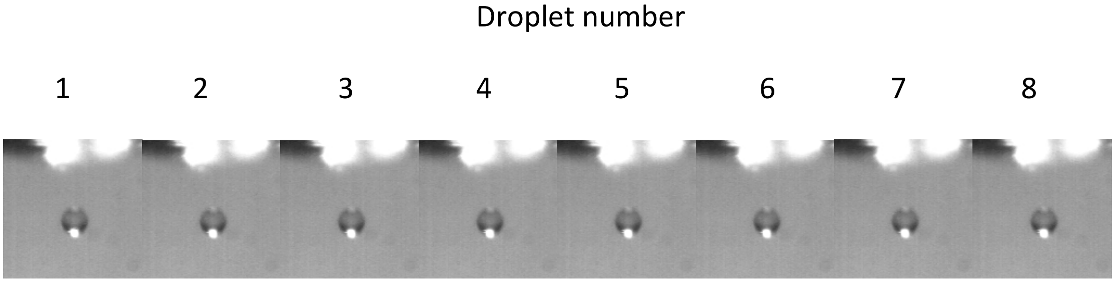
The maximum operating frequency at which these imaging systems are able to image every droplet in a burst of droplets is around 100 Hz. To be on the safe side, the jetting frequency is set on 50 Hz. As the idle time between two spots is longer (1–3 seconds) than the jetting time (0.14–0.2 seconds), using a lower frequency does not result in a significant increase in processing time.

2.2.4 Image processing and data analysis

As explained in the previous sections, the spots are built up of seven to ten droplets, depending on the specific droplet volume. A series of pictures is obtained from both cameras, imaging the situation in between the nozzle and the substrate after a specific delay time from actuation of the piezo element. Subsequent image processing gives information on the presence of the droplet in the air between nozzle and substrate, and furthermore on droplet volume, position and velocity. If no droplet is detected in these pictures, the printing process is interrupted and a sign warns the operator. The images of the last printed spot are shown to the operator to help to identify the cause of failure. This determines the specific actions the operator has to take. Figure 2.4a gives an example of a spot built up out of eight individual droplets, all emitted correctly.

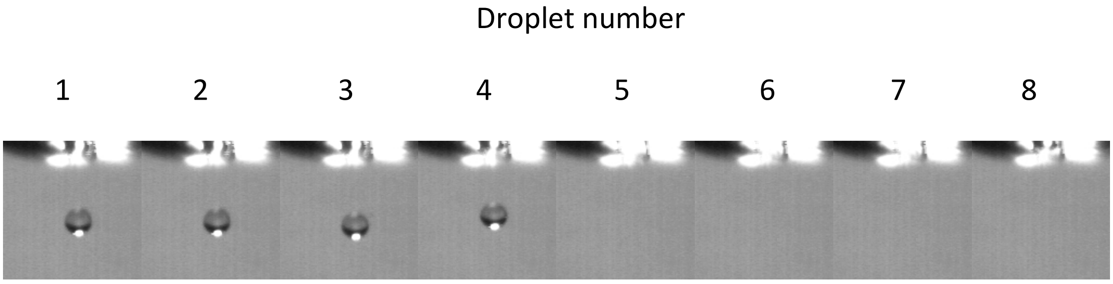


Camera 1

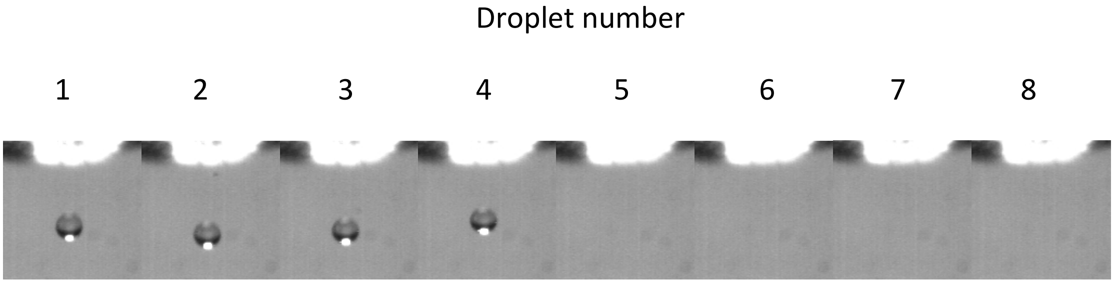


Camera 2

Figure 2.4a Images of a stable droplet formation. Eight droplets that build up one spot on a microarray are made visible by the two imaging systems. It can be seen that the location of the droplet in space does not change, which means that stable droplet formation is stable.



Camera 1



Camera 2

Figure 2.4b Images of an instable droplet formation. Eight droplets that build up one spot on a microarray are made visible by the two imaging systems. It can be seen that the location of the droplet slightly changes in the fourth picture, which means that the droplet velocity decreases and eventually droplet emission stops completely.

The upper and the lower eight pictures are the images of droplet 1–8 by the first and the second camera, respectively. Every image is made after an identical delay time from piezo actuation. As the droplet remains at the same position in the images, a stable droplet formation within one spot is shown. Figure 2.4b shows a slowly decelerating droplet leading eventually to a complete stop of the jetting process. In this case, only 4 droplets are correctly detected and a repair stroke can subsequently be performed to deposit the remaining droplet. All droplet images are stored, sorted by batch number and subsequently classified by substrate number, spot row number and column number.

The strength of this method is the fact that every jetting event is imaged and available for analysis. A viewer in LabVIEW simplifies browsing through this huge amount of data and supports selecting and analyzing events where the process was interrupted. For every process intervention, the corresponding droplet images can be loaded and consulted to identify the failure mode.

2.2.5 Failure mode analysis

From information given by the manufacturer of the print heads, it is highly unlikely that print heads fail due to mechanical and or electrical reasons because the lifetime of the print heads used is expected to be longer than 100 billion print cycles. This number is much higher than the number of jettings used in our process (<500,000 cycles)[12]. All failures identified and analyzed in this study were, on the one hand, caused by the fluids used and the printing method applied, and, on the other hand, by the measurement system built around the print head. A classification of the possible causes leading to a failure and subsequently to a process interrupt is summarized in figure 2.5. By consulting the images and corresponding image-processing data, the cause of a possible interruption of the printing process can be identified.

Four different failure modes are distinguished:

1. Failure of the imaging systems. One or more images are black.
2. Failure of the timing of the embedded software of the imaging systems, i.e. the moment of triggering the imaging systems is incorrectly tuned. In the images, some droplets are missing or not correctly located inside the window.
3. Failure of the image-processing software. The software fails to detect one or more correctly emitted droplets. The droplet recognition algorithm works by subtracting the droplet image from a reference image. A too-strict threshold in the droplet recognition software results in a process interruption. The process is only interrupted if the same droplet is not detected by both cameras.

These three different causes lead to errors that are related to the measurement system (referred to in the figure as situation 1a–1c). They are all situations in which the system fails to correctly detect emitted droplets. The process is therefore unnecessarily interrupted, and these cases are considered as being “false positives”.

4. Failure of the print head. The jetting process deteriorates and eventually completely stops. This is observed as decelerating and missing droplets.

Failure of the print head is the failure mode that leads to a reduced quality of the microarrays. The whole measurement system was built to detect these events and to stop the printing process immediately to prevent further consequences.

Undetected jetting failures (“false negatives”) are to be prevented at any time. Figure 2.6 gives an analysis of the overall process including the highly undesired occasion of undetected jetting failures. As mentioned before, the algorithm for droplet recognition should not be too strict in order to prevent too many unnecessary process interventions (“false positives”). However, a too-soft threshold results in undetected failure events, e.g., other features in the recorded and processed images are identified as droplets. Undetected jetting failures cannot be directly obtained from the image processing data, but need to be extracted manually. As the number of pictures to be analyzed in this study was far too high to check all pictures for droplet presence one by one, all droplets of a selected number of batches were checked manually for droplet presence. In particular, batches were selected where a high number of process interruptions took place. In total, 33,870 pictures were individually analyzed. In the case that droplet emission stops, it would be good to know if such a situation could be predicted based on the information gathered. The droplets jetted directly before the jetting stops often have a decreased velocity, as visible in figure 2.4b, and may also have a different flight path compared to the previous droplets. This might result in a reduced quality of the spot on the substrate. If jetting failure can be predicted, the process could be interrupted or the process parameters (e.g., voltage) adjusted in real time (closed loop droplet emission control). This may result in a longer constant product quality or in a postponement of a process interruption.

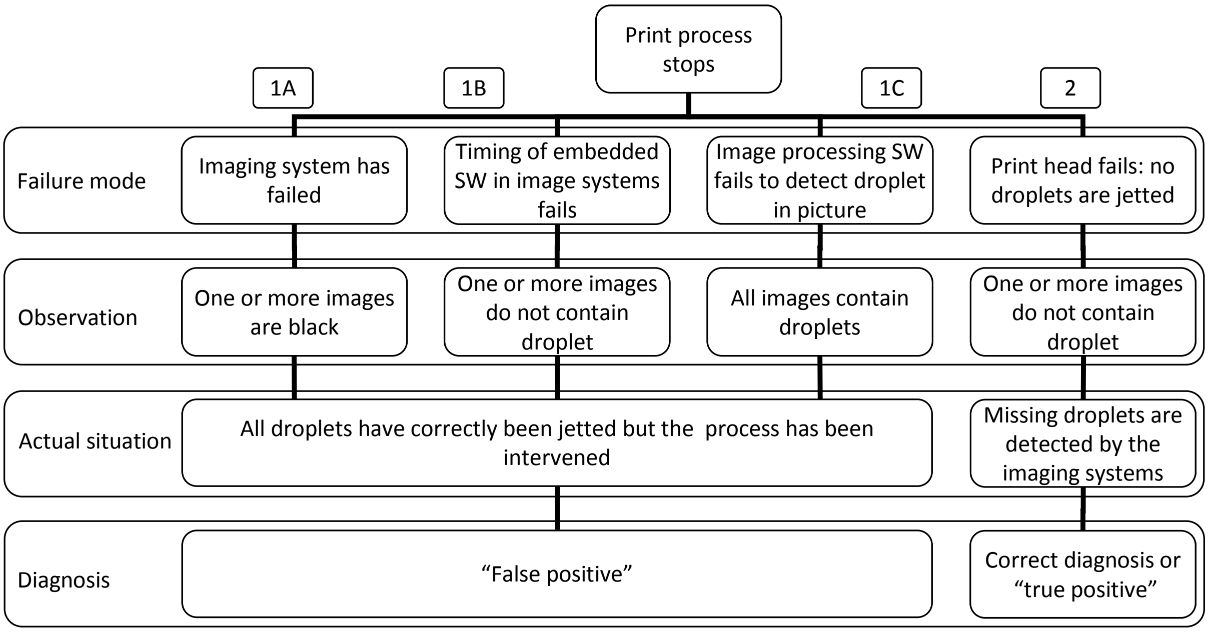
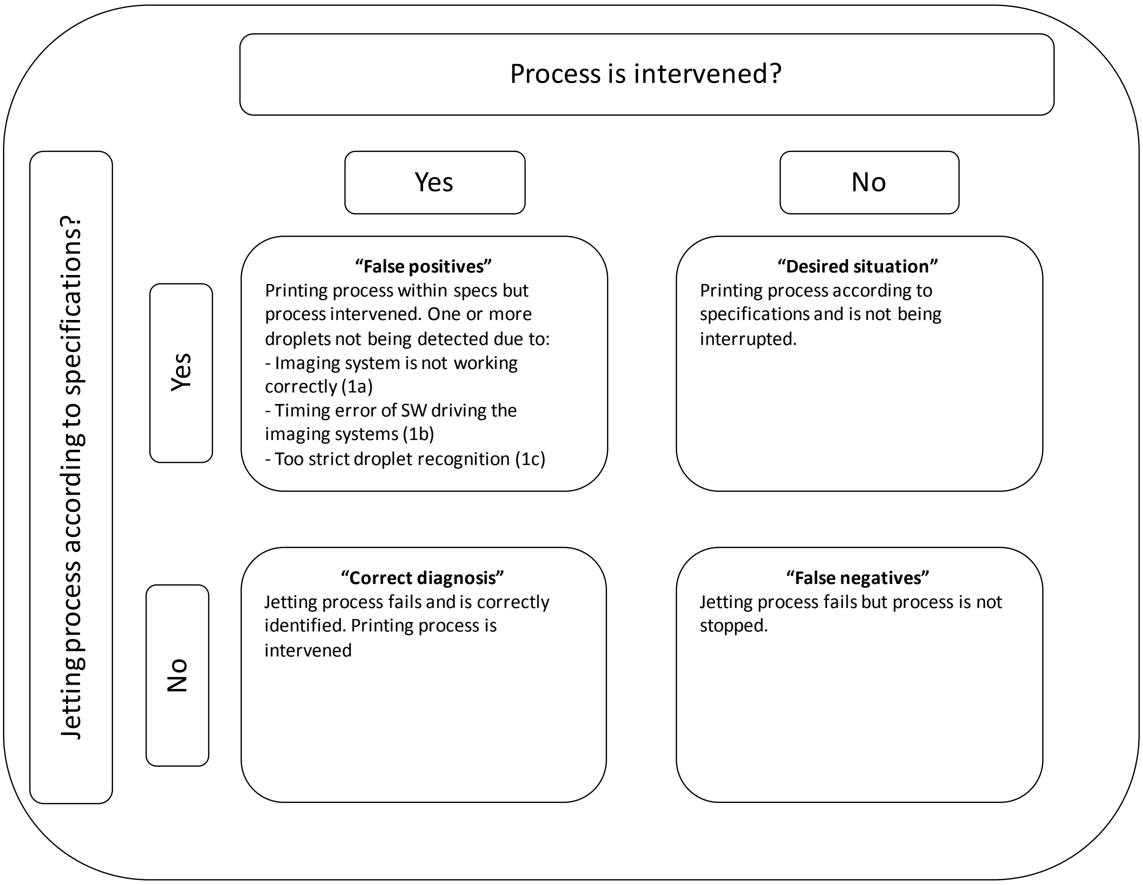

Print process stops
1A
1B
1C
2
Failure mode
Imaging system has failed
Timing of embedded SW in image systems fails
Image processing SW fails to detect droplet in picture
Print head fails: no droplets are jetted
Observation
One or more images are black
One or more images do not contain droplet
All images contain droplets
One or more images do not contain droplet
Actual situation
All droplets have correctly been jetted but the process has been intervened
Missing droplets are detected by the imaging systems
Diagnosis
"False positive"
Correct diagnosis or "true positive"

Figure 2.5 Failure mode classification. a) The terms “false positives”, “correct positives”, and “false negatives” as used in figure 2.5 and 2.6 are related to the printing process. A “false positive” means that the process is unnecessarily interrupted. In case of a “correct positive”, a process intervention is necessary as the process was indeed running out of specifications. A “false negative” means that no indication is given that the process is not running according to specifications. These substrates that miss at least one spot pass the process, while they should have been discarded.



	Process is intervened?	
Jetting process according to specifications?	Yes	No
Yes	**“False positives”** Printing process within specs but process intervened. One or more droplets not being detected due to: - Imaging system is not working correctly (1a) - Timing error of SW driving the imaging systems (1b) - Too strict droplet recognition (1c)	**“Desired situation”** Printing process according to specifications and is not being interrupted.
No	**“Correct diagnosis”** Jetting process fails and is correctly identified. Printing process is intervened	**“False negatives”** Jetting process fails but process is not stopped.

Figure 2.6 Analysis of the jetting failure process.

2.2.6 Quality control methods

The strategy to produce high-quality microarrays contains different levels of quality control, namely:

- Every spot is composed of multiple droplets, reducing the impact of one missing droplet.
- Every fluid is deposited on three to five different locations on the substrate
- In-line and real-time failure detection using optical imaging systems
- Measurement of the droplet flight path during nozzle inspection before printing and subsequently correcting for the landing position
- Fluid height sensing to determine whether enough printing and cleaning fluids are available in the vials
- Bar code reading on the lids of the vials of the fluids to be processed. The print plan is automatically adjusted (configured), avoiding possible operator mistakes.

2.3 Results and discussion

In total, 3,490 substrates in 217 batches with on average 29 spots per substrate were analyzed. All spots had a volume of around 1,000 pL. In total, 1,370,131 droplets were jetted and 2,740,262 corresponding images (two images per droplet) were collected for this analysis. These droplets made up 102,485 spots.

Figure 2.7 gives the results of the process analysis (in the same format as figure 2.6). The process was interrupted 150 times. In 65 of these 150 events, the jetting process itself failed (entry 2). Of these 65 events, 9 (14 %) could be predicted based on the data of the velocity of the droplets of the preceding spots.

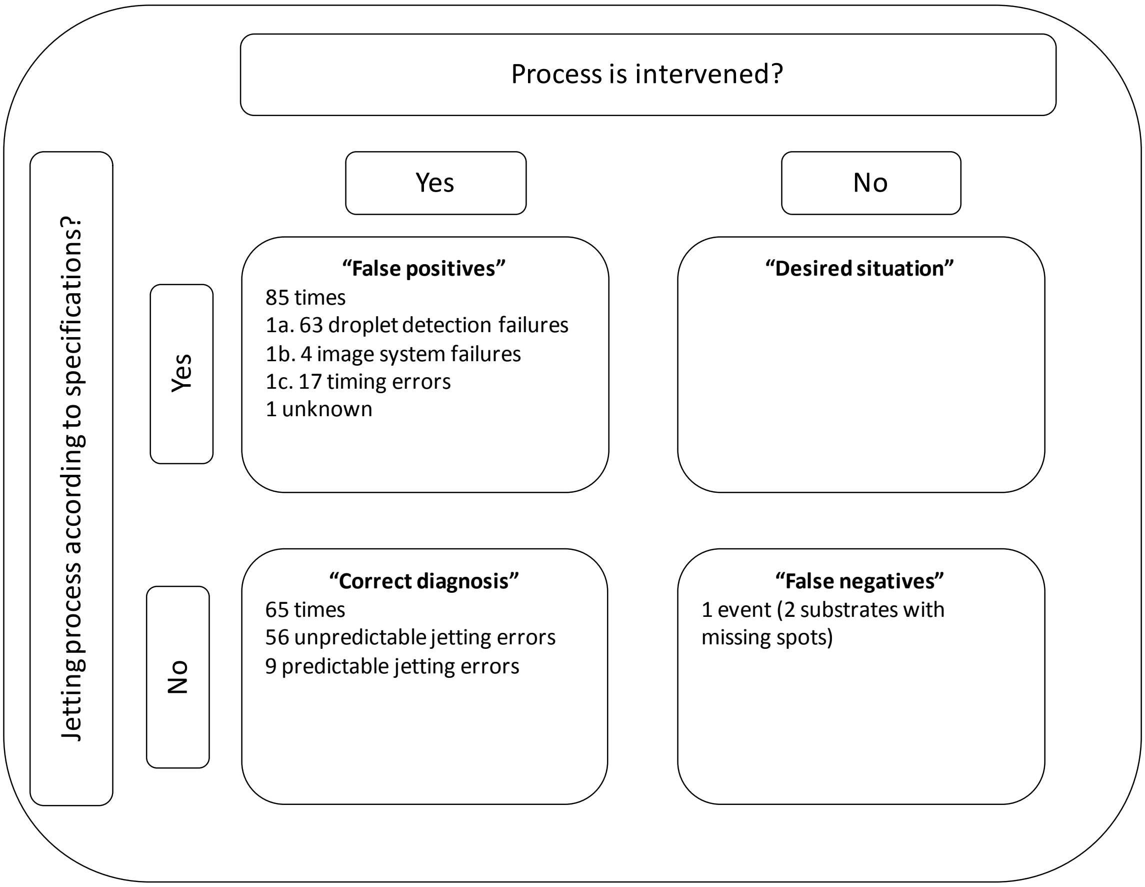

Process is intervened?
Yes
No
Jetting process according to specifications?
Yes
No
"False positives"
85 times
1a. 63 droplet detection failures
1b. 4 image system failures
1c. 17 timing errors
1 unknown
"Desired situation"
"Correct diagnosis"
65 times
56 unpredictable jetting errors
9 predictable jetting errors
"False negatives"
1 event (2 substrates with missing spots)

Figure 2.7 Results of the process analysis performed on spot level with the same categories as indicated in figure 2.6.

The remaining 85 interruptions that were not related to jetting failures can be considered as false positives as they were caused by a failure to detect droplet presence (63), a failure of the imaging system (4), a timing error (17, entry 1c) and unknown cause (1).

In only 1 event was a jetting failure not directly detected, but was determined later during the batch analysis: two substrates were incorrectly marked as correctly produced substrates. This was subsequently attributed to the nozzle being also visible in the image leading to erroneous droplet detection. It is probable that, in the moment prior to this batch, some mechanical changes had been executed to fine-tune the print system, which resulted in a minimal change of the position of the print head in the print head holder. This became apparent at the moment jetting failed as the system could not immediately correctly identify the event. Measures have since been taken to prevent a repetition of this type of error.

As an additional check, 33,870 pictures were manually analyzed. No additional process failures were observed. As these pictures were selected especially from print batches with a high frequency of process failures, this is considered to be a representative sample.

The relationship between the 85 unnecessary and 65 necessary process interruptions and only a single undetected failure event shows that the droplet detection algorithm is finely tuned. With only a single exception, all events were detected without having a significant impact on the up-time of the print system (to put this into perspective: the process was unnecessarily interrupted less than twice a week). The number of unnecessary process interruptions (here 85) can be easily reduced further by setting stricter criteria for a process interruption. As an example, for 85 of the 102,485 spots, the process stopped because of a droplet detection failure (i.e., 0.08 % of all spots). Assuming that the two imaging systems fail at random and making the criteria of a process interruption twice as strict (e.g., only if both cameras fail to detect two subsequent droplets), this would result in process interruptions for 0.00007 % of all spots; for this analysis that would be less than once. It must be noted here that the imaging systems do not fail completely randomly, as some settings might also influence the performance of these imaging systems. From experience, it seems that further refining this criterion greatly reduces the number of unnecessary process interruptions.

Table 2.1 gives the details on batch, microarray, spot and droplet level, showing that 74 % of all batches went without any interruptions. One or more jetting failures occurred in 15 % of all batches. On a microarray level, jetting stopped for 1.6 % of the substrates. The yield increased to 98.8 % when looking to the individual spots. Finally, when investigating the individual droplets jetted, the process was either interrupted or outside the specifications for 656 droplets. This is the worst-case scenario, as it is assumed that all droplets of a spot that are incorrectly deposited (i.e., the sum of the “correct diagnosis” spots and “false negative” spots) are all identified as outside of specifications. For the false positives, only one droplet out of a spot (that comprises seven to ten droplets) is out of specifications. It must be stressed here that these events (with the exception of the false negatives) could all be detected and subsequently corrected. The actual yield of the process is thus 99.9 %.

	Without intervention, without process failure (desired situation)	***With intervention, with process failure (correct diagnosis)***	***With intervention, without process failure (false positives)***	***Without intervention, without process failure (false negatives)***	***Total***
Batch	159	33	24	1	217
	(73.27 %)	(15.21 %)	(11.06 %)	(0.46 %)	(100 %)
Microarray	3,356	57	75	2	3,490
	(96.16 %)	(1.63 %)	(2.15 %)	(0.06 %)	(100 %)
Spot	181,528	64	80	6	181,678
	(99.92 %)	(0.04 %)	(0.04 %)	(0.003 %)	(100 %)
Droplet *	1,369,474	527	80	49	1,370,131
	(99.95 %)	(0.0385 %)	(0.0058 %)	(0.0036 %)	(100 %)

Table 2.1 Yield on batch level, substrate level, droplet level and spot level. *) From the calculations based on droplet level, a worst case scenario is used: it is assumed that all droplets from a spot where jetting failed where considered out of specifications (correct diagnosis or false negatives)

Figure 2.8 gives the total number of process interruptions (solid line) and the number of process interruptions caused by jetting failure (dotted line), as a function of the cumulative number of droplets jetted. A number of stepwise increments are visible, alternated with horizontal lines. Short periods of less efficient printing alternate with longer periods without any problems. This can be explained by, e.g., a slightly mistuned imaging system, leading to many process interruptions one after another in a very short time, or a very difficult jettable fluid (caused by incorrect fluid pre-treatment), leading to a high number of consecutive jetting failures during a single batch. For the data as used in this analysis, 26 % of the jetting failures were caused during three batches that consisted of 132 substrates (3.8 %). By further optimizing protocols, thereby circumventing these events, the printing process could be improved further.

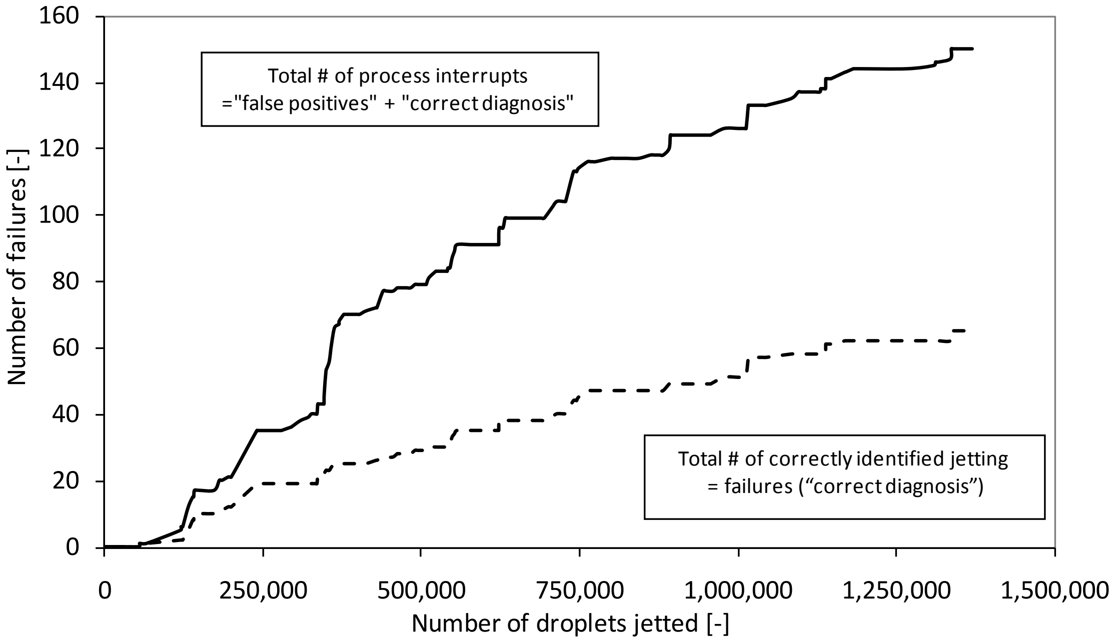


Figure 2.8 Total number of process interruptions (solid line) and the number of process interruptions caused by real jetting failure as a function of time, expressed in number of jets.

As the direct causes of the jetting failures are not yet known, a number of stricter protocols can be applied to investigate the influence of a number of variables on the failure rate. For example, filtration of the fluids, degassing of the fluids, a longer drying time of the nozzle and different initial printer settings (e.g., a higher voltage leads to a higher droplet speed, which may result in less nozzle plate wetting and subsequently lower failure rates): all can have a positive effect on the jetting performance.

2.4 Conclusions

We have described an inkjet system equipped with two camera systems to follow the flight of every droplet in real time and in 3-D space. The recorded images are used to monitor the printing process and to interrupt the process when the optical detection systems do not detect a droplet. We have analyzed a large number of printing sessions (1.4 million droplets) and found 150 failure events of which 65 were related to jetting and 85 to hardware and (embedded) software. From a practical point of view, high-quality printing of microarrays combined with a high uptime of the print system is the desired state. Our results confirm that we have found a very workable compromise. In general, the average number of droplets emitted before a jetting failure is around 21,000. Compared to standard graphical printing, this number is very low. This is due to the fact that little effort was paid to fluid optimization and the different printing strategies (longer idle time of the print head). To place this number into perspective, we need to take into account the low number of droplets jetted for one full substrate, which is around 1,000 (120 spots × 8 droplets). This means that on average the process has to be maintained after every 21 substrates. Going towards large volumes using equipment with multiple nozzles installed and used in parallel, this number of substrates needs to be increased. This inkjet system analysis provides a unique tool for efficiently analyzing future data on jetting failure mechanisms. This will lead to an increased knowledge of jetting failures and to an increasingly robust manufacturing process for microarrays. The system presented in this study worked with operators, who had to intervene in case of a failure. For a final production process, a fully automatic print head maintenance procedure is preferred over operator intervention. Procedures could be implemented that, for example, result in an automatic start of a print head cleaning procedure in case of a jetting failure.

References

1. Dufva M. Fabrication of high-quality microarrays. *Biomol Eng* 2005, 22:173–184
2. Venkatasubbaro S. Microarrays, status and prospects. *Trends Biotechnol* 2004, 22:12 631-637
3. Lausted C, Dahl T, Warren C, King K, Smith K, Johnson M, Saleem R, Aitchison J, Hood L, Laskey SR. POSaM, A fast, flexible, open-source, inkjet oligonucleotide synthesizer and microarrayer. *Genome Biol* 2004, 5, R58.
4. Campas M, Katakis I. DNA biochip arraying, detection and amplification strategies. *Trends Anal Chem* 2004, 23:49–62
5. Fisher W, Zhang M. A biochip microarray fabrication system using inkjet technology. IEEE *Trans Autom Sc. Eng* 2007, 4(4):488-500
6. Sumerel J, Lewis J, Doraiswamy A, Deravi L, Sewell S, Gerdon A, Wright DW, Narayan RJ. Piezoelectric ink jet processing of materials for medicaland biological applications. Biotechnol. J 2006, 1:976–987.
7. Dijksman JF, Duineveld PC, Hack MJJ, Pierik A, Rensen J, Rubingh J-E, Schram I, Vernhout MM. Precision ink jet printing of polymer light emitting displays. *J Mater Chem* 2007, 17:511–522.
8. Hutchings IM, Martin GD, Hoath SD. High speed imaging and analysis of jet and drop formation. *J Imaging Sci Technol* 2007, 51:438–444
9. Van Dam DB, Le Clerc C. Experimental study of the impact of an ink-jet printed droplet on a solid substrate. *Phys Fluids* 2004, 16:3403–3414
10. De Jong J, de Bruin G, Reinten H, van den Berg M, Wijshof, H, Versluis M, Lohse D. Air entrapment in piezo-driven inkjet printheads. *J Acoust Soc Am* 2006, 120:1257–1265
11. Mocanu D, Kolesnychenko A, Aarts S, deJong, A. Pierik A, Coene W, Vossenaar E, Stapert H. Quantitative analysis of DNA hybridization in a flowthrough microarray for molecular testing. *Anal Biochem* 2008, 380:84–90
12. Microdrop Technologies product data (2008). Available at: http://www.microdrop.de/wDeutsch/products/AD_Pipettes_e_0707.pdf

Chapter 3

Immobilization of oligonucleotides with homo-oligomer tails onto amine-functionalized solid surfaces and the effects on hybridization

This chapter is based on:

Pierik A, Dijksman JF, Lub J, Stapert HR, Broer DJ. *Anal Chem* 2010, 82:1191-1199

Abstract

Microarrays have become important tools for the detection and analysis of nucleic acid sequences. Photochemical (254 nm UV) DNA immobilization onto amine-functionalized substrates is often used in microarray fabrication and Southern blots, although details of this process and their effects on DNA functionality are not well understood.

By using Cy5-labeled model oligonucleotides for UV immobilization and Cy3-labeled complementary sequences for hybridization, we measured independently the number of immobilized and hybridized oligonucleotides on the microarray surface. By using a two-color fluorescence LED setup and a novel method to compile the data, a full analysis has been made of the effects of oligonucleotide composition (length and sequence) on both immobilization and hybridization. Short homo-oligomer sequences (tails) of uracils, thymines, and, to a limited extent, guanines attached to a hybridization sequence improve immobilization. We propose a possible mechanism explaining the grafting of these nucleotides to amine-functionalized substrates, and we found evidence that the DNA backbone is possibly involved in the immobilization process. Hybridization, on the other hand, greatly improves as a function of tail length regardless of tail composition. On the basis of statistical arguments, the probes increasingly bind via their tail, with the hybridization sequence becoming more accessible to its complement. We conclude that all tails, sequence independent, improve hybridization signals, which is caused by either improved immobilization (especially thymine and uracil) or improved hybridization (most pronounced with guanine tails).

3.1 Introduction

Microarrays have become important tools for the detection and analysis of nucleic acid sequences[1-4]. Many different applications as well as microarray formats and platforms exist[5] which are characterized by substrate-immobilized capture probes that are able to bind to complementary target molecules. The overall process from microarray manufacturing to hybridization and data analysis is prone to variations, such as unknown deposited active material per spot, uncontrolled substrate binding, or the use of fluorophores that can bleach over time causing reduced intensity. Reducing these variations is a major challenge of microarray technology[5]. Methods have been developed to improve on the reproducibility and reliability of microarray data[4,6-8], and a variety of protocols and surface reactions for immobilization of the capture probes to the solid supports have been investigated in order to optimize hybridization signals[3,9-12].

One specific reactive moiety that is often used to graft DNA to solid surfaces is a primary amine group[12-15]. Unmodified oligonucleotides or polynucleotides are immobilized on amine-functionalized substrates by exposing them to 254 nm UV irradiation or high temperatures[15]. With only a single washing step after the UV exposure, this makes it a fast and cost-effective protocol. This type of substrate is very suitable for microarray production, as grafting of the DNA is induced after deposition onto the substrate. This means that these surfaces will not show a significant decrease in reactivity of the functional groups on the surface over the course of processing time, which might occur with pre-activated substrates which already contain reactive groups[13-14].

Upon application of 254 nm UV or high temperature, it has been shown that especially thymidine and deoxyuridine (modified DNA) residues present in the oligonucleotide are capable to react with the amines attached to the surface[7,18-19]. The grafting efficiency can be further enhanced by attaching tails of multiple oligothymidine phosphates to one end of the oligonucleotide, which leads to a large increase in signal after hybridization with a complementary nucleotide[7]. Variations in immobilization and hybridization efficiencies among different probe-target combinations provided with an identical tail to the probe have been reported. These differences are not yet fully understood. These large variations make it difficult to use microarrays for (semi)- quantitative purposes without the use of extensive calibration tools.

The (photo)chemical reaction mechanism of the immobilization process and the subsequent effect of this reaction on the functionality of the capture probe, i.e., hybridization with a target molecule, is also poorly understood. A possible reaction mechanism between primary amines and thymidines activated by 254 nm UV irradiation has been described[20]. In order to improve hybridization signals and microarray manufacturing methods, a better understanding of the mechanism(s) of the (photo)chemical immobilization processes and their effects on the hybridization on a molecular level are needed.

Currently, most of the reported data on microarray experiments on the photochemical immobilization of DNA on amine-containing substrates correlate a specific parameter to the end result in terms of hybridization signal. Thereby no discrimination has been made between separate effects on immobilization and hybridization, although they can be very different and may even have opposite effects on the end result. For instance, a modification of the oligonucleotide sequence by adding a tail to the capture probe can influence the extent of immobilization. It can also, more indirectly, influence the hybridization efficiency to its complementary part by, for example, a change on the charge of the oligonucleotide[18] or its rotational freedom[21].

In order to obtain a better understanding of the influence of the oligonucleotide composition on both the immobilization and hybridization, we developed an assay and an analysis method that allows us to study these two processes separately. We performed a systematic study on the photochemical immobilization of modified oligonucleotides onto primary amine-containing membranes, as these substrates show high potential because of low cost, safety, easy handling, and simple manufacturing protocol. The hybridization sequences were modified with a tail consisting of one type of nucleotides added to the 5′-end. The deposition into spots on the membrane in a microarray format made it possible to analyze in parallel and at the same time many different surface reactions taking place by different oligonucleotides, varying in tail length or composition, whereby batch-to-batch as well as substrate variations could be removed. By using an assay with two different fluorescent dyes and a two-color fluorescence detection setup, the number of oligonucleotides present in each spot after each process step could be quantified. A novel data analysis method was used to fully discriminate between effects on immobilization and hybridization. We also include in this study the oligonucleotides consisting of oligo deoxyadenosine phosphate, oligo deoxyguanosine phosphate, oligo deoxycytidine phosphate, oligo thymidine phosphate, and oligo deoxyuridine phosphate. For readability, we will further refer to these as adenine, guanine, cytosine, thymine, and uracil, respectively.

3.2 Materials and methods

3.2.1 Substrate

For all experiments we used Nytran Supercharged (SPC) membranes (Whatman, GE Healthcare, UK) with an average pore size of 0.45 µm and thickness of 150 µm. The membranes were punched into 12 mm circular pieces and laser-welded onto nylon rings to facilitate handling and processing.

3.2.2 Oligonucleotides

All HPLC-purified capture oligonucleotides used for immobilization were designed and manufactured according to our specifications and were Cyanine 5 (Cy5) labeled at the 5′-end (Biolegio, Nijmegen, The Netherlands). Before use, all oligonucleotide solutions were diluted in MilliPore UltraPure water to concentrations of 10 µM. All other chemicals were purchased from Sigma-Aldrich. For studying the effect of each oligonucleotide sequence composition on the immobilization onto these substrates, we used oligonucleotides consisting of a stretch of 16 identical nucleotides: $d(N)_{16}$, with N as either adenine (A), cytosine (C), guanine (G), thymine (T), uracil (U), or the oligo(2-deoxyribose phosphate) DNA backbone (i.e., without any bases) (X).

Furthermore, a series of oligonucleotides was designed in order to study separately the effect of specific tails on the immobilization and onto the subsequent hybridization. These capture molecules contained a tail of a oligonucleotide $(dN)n$, with N either as adenine (A), cytosine (C), guanine (G), thymine (T), or uracil (U), having a length n between 0 and 16 bases. All these tails were attached to the 5′-end of an identical 20-mer sequence that was used for hybridization to its complementary 20-mer counterpart. As it is known that thymine is involved in the reaction between DNA and amines present on nylon supports, we designed the identical common sequence without any thymine. This reduced the participation of the hybridization sequence to the immobilization. The hybridization sequence was 5′- ACGACGACACCAGAAAAAGA-3′. Thereby as much as possible the effects of tail composition on immobilization and hybridization can be separately investigated. The Cy5 was always located at the 5′-end. The sequence of the capture nucleotides can thus be summarized as follows: 5’-Cy5-d(N)$_n$- ACGACGACACCAGAAAAAGA-3’, with *N* either A,C,G,T or U, and *n* varying from 0-16. A complete overview of the sequences included are given in table 3.1 below.

Table 3.1 Capture probe sequences included in the analysis of immobilization followed by hybridization. The sequences are given in the 5’-> 3’ direction. ‘Seq’ denotes the sequence 5’-ACGACGACACCAGAAAAAGA’3’. As a reference, the sequence without any modification to the 5’-end is used.

Adenine	**Cytosine**	**Guanine**
AA-Seq	CC-Seq	GG-Seq
AAAA-Seq	CCCC-Seq	GGGG-Seq
AAAAAAAA-Seq	CCCCCCCC-Seq	GGGGGGGG-Seq
AAAAAAAAAAAAAAAA-Seq	CCCCCCCCCCCCCCCC-Seq	GGGGGGGGGGGGGGGG-Seq

Thymine	**Uracil**
TT-Seq	UU-Seq
TTTT-Seq	UUUU-Seq
TTTTTTTT-Seq	UUUUUUUU-Seq
TTTTTTTTTTTTTTTT-Seq	UUUUUUUUUUUUUUUU-Seq

For reference we performed immobilization experiments to verify the effect described in literature[7,15,17] where it was stated that the addition of a tail of thymine improved the immobilization efficiency. We used a slightly different sequence for the hybridization sequence: 5′-TCGACGACTCCAGTAAAAGT-3′.

3.2.3 Microarray fabrication

The Nytran SPC substrates were positioned onto a carrier plate which was mounted in a single nozzle spotter[6]. This spotter is equipped with an Autodrop AD-H-501 piezo driven single nozzle glass capillary print head (Microdrop Technologies, Norderstedt, Germany).The print head generates 100-120 pL droplets upon piezo driven actuation at a frequency of 100

Hz. Droplet volumes were measured before printing, and all spots were built up of 7 to 9 droplets (per spot total deposited volume 1000 pL), which led to a spot diameter of 150-200 µm. Using a print concentration of 10 µM, this resulted in a surface density of approximately $3{\cdot}10^5$ molecules/µm^2. This surface density provided the optimal signals after hybridization. Lower densities resulted in lower signals, whereas higher densities did not lead to higher hybridization signals. As we used membranes, part of the molecules permeate the porous structure of the substrate.

In order to keep the (photo)chemical degradation of the Cyanine 5 dyes caused by ozone[23] and ambient UV irradiation as low as possible, the entire process from printing to subsequent post-processing and all the measurements have been executed inside a clean room with yellow light, thereby reducing UV-light exposure. The ozone concentration was kept below 5 ppb by the use of installed ozone filters.

The microarray layout used in the experiments consisted of a rectangular pattern with a pitch of 400 µm in both directions. Every fluid was deposited at least 3-fold in different spots at different places on the membrane for redundancy purposes. Two substrates were used for each measurement, thereby reducing variations possibly caused by the substrate itself or related to substrate processing, e.g., washing.

3.2.4 Microarray Processing

After printing, the membranes were exposed to varying irradiation exposure times of 254 nm UV. The dose was varied between 0-500 mJ/cm^2 using a Stratalinker 2400 (Stratagene, Agilent, La Jolla, CA, USA) to immobilize the oligonucleotides to the membrane. This equipment had a default power of 10 mW/cm^2; exposure times of the substrate thus varied between 0 and 50 seconds. After UV exposure, the excess material that did not bind to the membrane surface was washed away and the membrane surface was blocked by incubating the samples for 1 hour at 42 °C in a shaking tray with 1 mL of a mixture of 5 × SSC, 0.1% SDS, and 100 µg/mL herring sperm DNA per sample. The membranes were subsequently shortly rinsed with water and dried in an oven for 30 minutes at 40 °C.

During hybridization experiments, the oligonucleotides were hybridized with the complementary oligonucleotide, which was Cyanine 3 (Cy3) labeled at the 5′-end: 5′-Cy3-TCTTTTTCTGGTGTCGTCGT-3′. Hybridization was done during a one-hour incubation at 42 °C in a shaking tray, using per sample 1 mL of a 20 nM concentration of the hybridization oligonucleotide in 5 × SSC, 0.1% SDS, and 100 µg/mL herring sperm DNA. After a short wash using 0.1% SDS in 2 × SSC, the membranes were dried in an oven for 30 minutes at 40 °C.

3.2.5 Fluorescent intensity measurements

After every process step (printing, UV cross-linking, washing/ blocking, drying, and hybridization followed by drying), we measured the intensity of the fluorescent spots on the dry membrane. This was done by using a measurement tool that contained four red LEDs (Philips Lumileds, LXHL-MD1D), optical filters (Omega Optical, XF 1069), and a CCD camera (Sony) for capturing images of the microarray pattern.

Marker spots deposited at the corners of the membrane enabled the image analysis software to put a grid over all spot locations. Each spot has its own square field (26 by 26 pixels) defined by the grid. To measure the fluorescence intensity of the spot, the intensities detected by the pixels in the field are added up and subsequently corrected for background intensity. The background intensity level was calculated by averaging the two lowest values of four small background fields, consisting out of 5 × 5 pixels positioned at the corners of the spot. All intensity measurement results reported are background corrected values. A similar measurement tool equipped with green LEDs (Philips Lumileds, LXHL-LM3C) and suitable filters (Omega Optical, XF1037) was used to measure the intensity of the Cy3-labeled spots after hybridization with the complementary oligonucleotide.

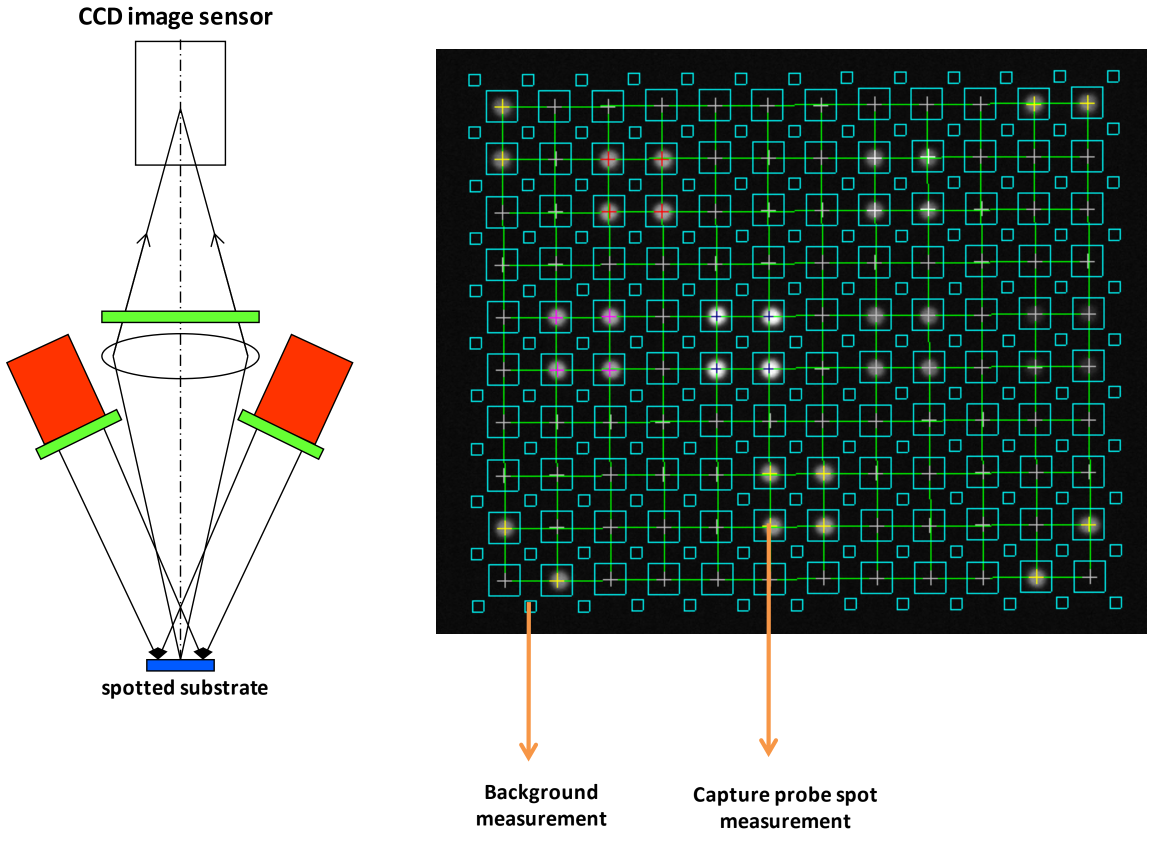


Figure 3.1 Schematic view of the optical measurement tool (left) and a microarray image (right) onto which the grid pattern is projected by the software. This grid pattern consists of 10 x 12 capture probe spots (large squares) of which the fluorescent intensity is calculated. Each spot is surrounded by four areas (small squares) where the local background is calculated and used for background correction of the spots.

3.2.6 Immobilization efficiency calculation

The relation between the spot intensity of the Cy5 fluorophores attached to the oligonucleotides, I^{Cy5}, and the number of deposited molecules N present in a spot can be expressed as:

$$N \propto I^{Cy5} \qquad (1)$$

The proportionality constant varies among experiments, as it depends on, for example, substrate exposure to ambient UV, 245 nm UV, and ozone as well as optical and substrate properties. The exact relation, however, between number of labels and spot intensity is not needed for studying relative differences between grafting of different oligonucleotides deposited on the same substrate at the same time. The immobilization efficiency determines the percentage of the deposited molecules that is able to graft to the substrate and is thus not washed away during the washing/blocking step. During immobilization, part of the Cy5 intensity was lost because of bleaching of the fluorophores. Therefore, the efficiency of immobilization (η_{imm}) was defined by the ratio of the fluorescent Cy5 intensity after washing/blocking ($I_{WB}{}^{Cy5}$) and after UV exposure ($I_{UV}{}^{Cy5}$):

$$\eta_{\mathrm{imm}} = 100\% \cdot \frac{N_{WB}}{N_{UV}} = 100\% \cdot \frac{I_{WB}^{Cy5}}{I_{UV}^{Cy5}} \qquad (2)$$

Furthermore, we found in an earlier optimization study that within a range of (0.12-12) × 10^5 molecules/μm^2 the absolute immobilization efficiency is almost independent on the surface density, see figure 3.2. In this graph it can be seen that the immobilization efficiency marginally increases with increasing surface density. The results are nicely reproducible, as can be seen by the good agreement between the two lines of the different microarray samples. The surface densities we applied in this study ($4 \cdot 10^5$ molecules/μm^2) fall within this range. It can be concluded here that small variations in surface density originating from the microarray manufacturing process do not have influence on the immobilization efficiency.

In order to study the effect of a tail on the immobilization, the efficiencies were normalized to the value of a reference oligonucleotide ($\eta_{imm,no\ tail}$), which is the oligonucleotide without a tail. The gain in immobilization efficiency when using a tail ($gain_{imm,tail}$) obtained at the lowest UV dose, is given by:

$$Gain_{\mathrm{imm,\ tail}} = 100 \cdot \frac{\eta_{imm,\mathrm{tail}}}{\eta_{imm,\mathrm{no\ tail}}} \qquad (3)$$

By normalization to the oligonucleotide without a tail at the lowest UV dose, the influence of the tail composition as well as the UV dose could be separately addressed. Secondly, any effect of degradation of the fluorophores or possible concentration quenching effects were removed, since especially guanine is known to be able to quench fluorophores[24].

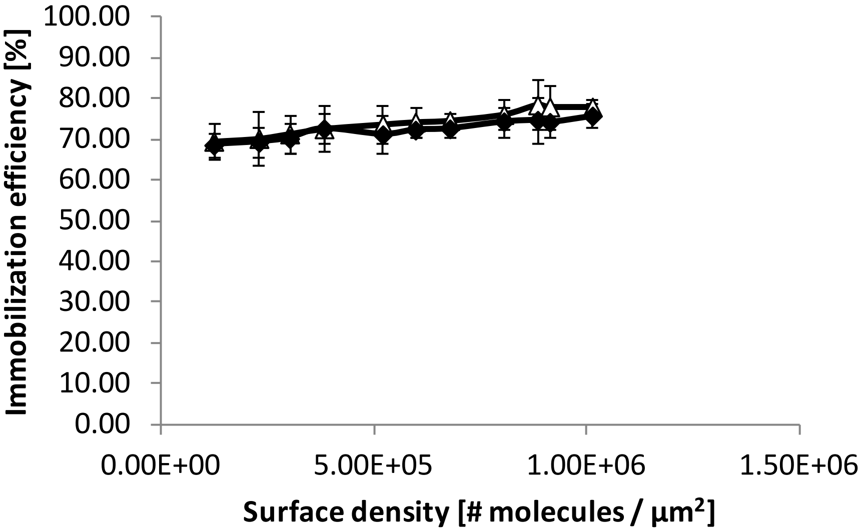


Figure 3.2 Immobilization efficiency as a function of the surface density of deposited capture probes. The error bars represent the standard deviation. The experiment was performed in duplicate, indicated by the two different lines. For this case, 400 mJ/cm^2 UV was used, and the oligonucleotide used was 5'-end labeled Alexa Fluorophore 633 with the following sequence: 5'-GCAAAATGGTGCCGTAACTT-3'

3.2.7 Hybridization efficiency calculation

Hybridization was achieved by using Cy3-labeled oligonucleotides complementary to the Cy5-labeled capture probes. As the surface density of the capture probes was at least 1 order of magnitude below saturation, this is even more the case for the hybridized molecules as hybridization efficiencies are less than 100%. Therefore, the background corrected intensity after hybridization of the Cy3-labeled molecules (I^{Cy3}) was proportional to the number of labeled complementary oligonucleotides hybridized to the capture probes on the substrate (N_{hyb}):

$$N_{hyb} \propto I^{Cy3} \tag{4}$$

For all spots containing capture probes with identical sequences but different tails, hybridization was performed by incubating the complement of these identical sequences with the microarray; therefore, possible bleaching of the Cy3-fluorophore can be considered to be identical for all spots. The total gain when using a tail compared to the reference ($gain_{tot,tail}$) was thus directly related to the ratio between the Cy3-labels measured with and without a tail (I_{tail}^{Cy3} and $I_{no\ tail}^{Cy3}$):

$$Gain_{tot,tail} = 100\% \cdot \frac{I^{Cy3}_{tail}}{I^{Cy3}_{no\,tail}} \tag{5}$$

The efficiency of the overall process (η_{tot}) (the percentage of deposited molecules that eventually hybridize to a complementary molecule) thus depends on both the immobilization efficiency (η_{imm}) and the hybridization efficiency (η_{hyb}):

$$\eta_{tot} = \eta_{imm} \cdot \eta_{hyb} \tag{6}$$

This resulted in an overall efficiency of:

$$\eta_{tot} = 100\% \cdot \frac{N_{WB}}{N_{UV}} \frac{N_{hyb}}{N_{WB}} = 100\% \cdot \frac{N_{hyb}}{N_{UV}} \tag{7}$$

The relative overall total gain of a nucleotide with a tail compared to the reference oligonucleotide without a tail after hybridization (equation 5) could be expressed as follows:

$$Gain_{tot,tail} = 100\% \cdot \frac{\eta_{\mathrm{tot,tail}}}{\eta_{\mathrm{tot,no\,tail}}} = 100\% \cdot \frac{\eta_{\mathrm{imm,tail}} \quad \eta_{\mathrm{hyb,tail}}}{\eta_{\mathrm{imm,no\,tail}} \quad \eta_{\mathrm{hyb,no\,tail}}} = 100\% \cdot \frac{I^{Cy3}_{hyb,\mathrm{tail}}}{I^{Cy3}_{hyb,\mathrm{no\,tail}}} \tag{8}$$

From this expression the gain attributed to the tail for the hybridization ($gain_{hyb,tail}$) could be derived:

$$Gain_{hyb,tail} = 100\% \cdot \frac{\eta_{hyb,\,tail}}{\eta_{hyb,\,no\,tail}} = 100\% \cdot \frac{Gain_{tot,tail}}{Gain_{imm,tail}} \tag{9}$$

The gain in hybridization efficiencies was defined with respect to the reference oligonucleotide, in our case the oligonucleotide without a tail at the lowest UV dose. A gain higher than 100% indicated that there was an effect leading to an improved process, the higher the increase, the larger the effect. A gain lower than 100% indicated there was a negative effect. The use of this method results in a decoupled measurement of immobilization and hybridization efficiency, which means that hybridization signals are corrected for the differences in surface density caused by differences in immobilization, and differences in hybridization efficiency can be completely allocated to phenomena occurring during hybridization. Furthermore, the findings are independent of the absolute values. All experiments were performed in duplicate. For each oligonucleotide, the intensities of three spots were averaged. For each UV measurement, two substrates were used.

3.3 Results and discussion

The background corrected fluorescent spot intensity *I* leveled off only to a limited extent with increasing number of molecules deposited up to 10^6 molecules/µm^2, see figure 3.3.

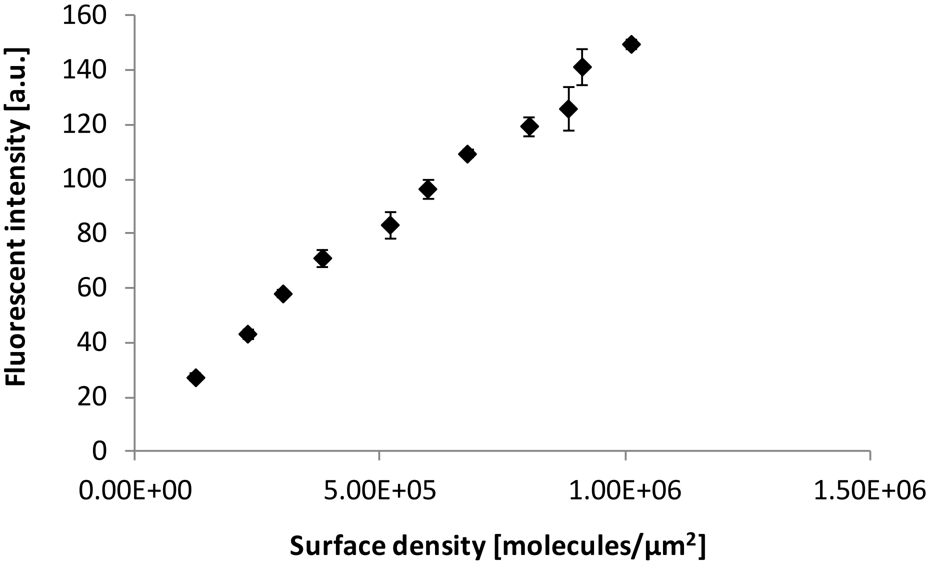
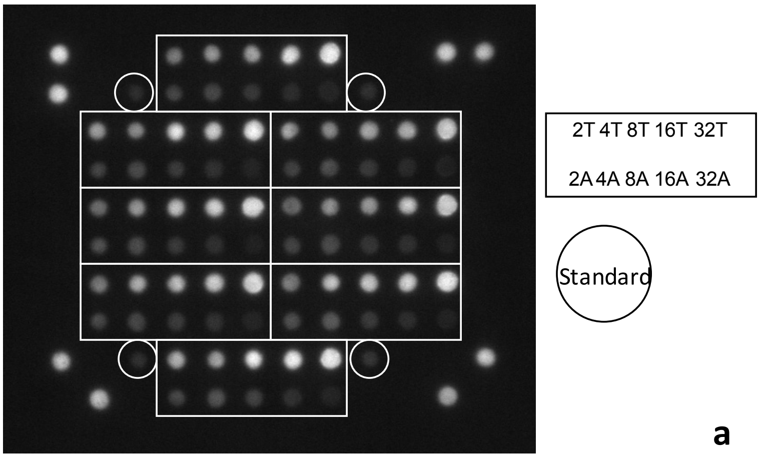


Figure 3.3 Fluorescent spot intensity measured as a function of surface density. The error bars represent the standard deviation.

As we used surface densities in the lower range of (around $4 \cdot 10^5$ molecules/µm^2), we can conclude that fluorescent quenching did not play a significant role. For flat surfaces, an optimal surface density for hybridization of around 10^5 molecules/µm^2 has been reported[2]. At even higher surface densities, the signal will not increase anymore due to larger contributions of surface crowding or concentration quenching[28-30]. This optimal concentration is higher for porous substrates because of the higher surface to projected surface ratio: a few orders of magnitude increase have been reported[25], which means the densities used in this study were far off from the density at which the maximum hybridization signal is reached. Lower densities were purposely used to reduce possible effects of surface crowding. The depth over which the capture probes are present in the membrane is limited. With an estimated porosity of 0.5 and our printing settings, we found that the capture probes are only present in the upper 15-20 µm of the 150 µm[26], which means that the number of oligonucleotides that can be bound to the membrane structure was about a factor of 9 smaller than maximally possible. To be conclusive, we experimentally verified (data not shown) that an increase in surface density of the capture probes led to an increase in hybridization signal and that steric hindrance during hybridization does not play a significant role. In order to make sure that the fluorescent dye is not involved in the immobilization process, we also printed a single thymidine phosphate labeled with Cy5 on the membrane and exposed it to UV. As the pure Cy5 dye itself is not soluble in water and could therefore not be deposited on the substrate, we used a thymine-Cy5. It was found that the immobilization efficiencies were very low (maximum of 3%), and we can thus conclude that the reaction between the labeled oligonucleotide and the surface involves the DNA.

Next it was verified whether our method complies with the literature. In Figure 3.4a-b, the immobilization of a capture probe sequence coupled to a tail of increasing length containing either only thymines or only adenines is shown, at an applied UV dose of 400 mJ/cm^2. Clearly visible is the gain in immobilization efficiency as a function of the number of thymines, as can also be seen in the image of the array itself. Adenine does not exhibit any improved grafting. In that respect, our findings confirmed the results reported by others that thymines react with amine-functionalized surfaces upon photoactivation[7].

To obtain more insight into the reactivity of each nucleotide, a second series of experiments was performed by immobilizing oligomers consisting of 16 identical nucleotides onto the nylon surface, using different UV doses. The absolute immobilization efficiency for each nucleotide is given in figure 3.5. The X16-Cy5 denotes the plain DNA backbone without any bases. Clearly visible is the superior reactivity of thymine, as reported above, and uracil. The other bases also showed increasing immobilization with increasing UV dose. As the oligo(2-deoxyribose phosphate) DNA backbone shows immobilization of the same magnitude, we can conclude that the adenine, cytosine, and guanine bases are not significantly involved in the reaction with the substrate, and that grafting is mainly caused by DNA backbone-substrate interactions.



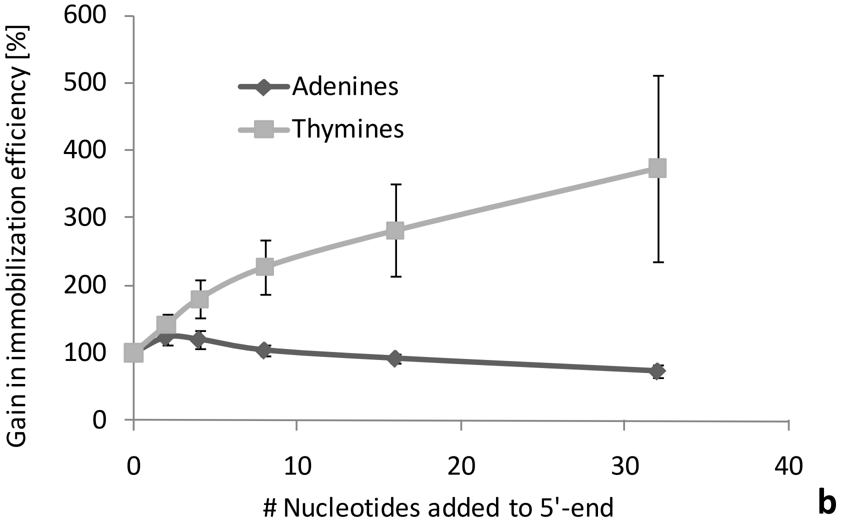


Figure 3.4a Image of a membrane after washing (upper figure). The rectangular box on the right of the image explains the type and order of oligonucleotides as printed in 8-fold on each of the rectangles on the substrate as visible in the image. Every upper row of spots contains five spots of increasing thymine-tail length, whereas the lower row of spots contains oligonucleotides with increasing adenine-tail length. The encircled spots in the image contain the oligonucleotide without any additional tail.

Figure 3.4b Effect of adding adenines (dark grey) and thymines (light grey) to the 5′-end of an oligonucleotide on the immobilization (lower figure). The results are normalized with respect to the standard oligonucleotide (no additional nucleotides). The error bars are the standard deviations. Every measurement point consists of a series of 12 membranes each with 8 spots.

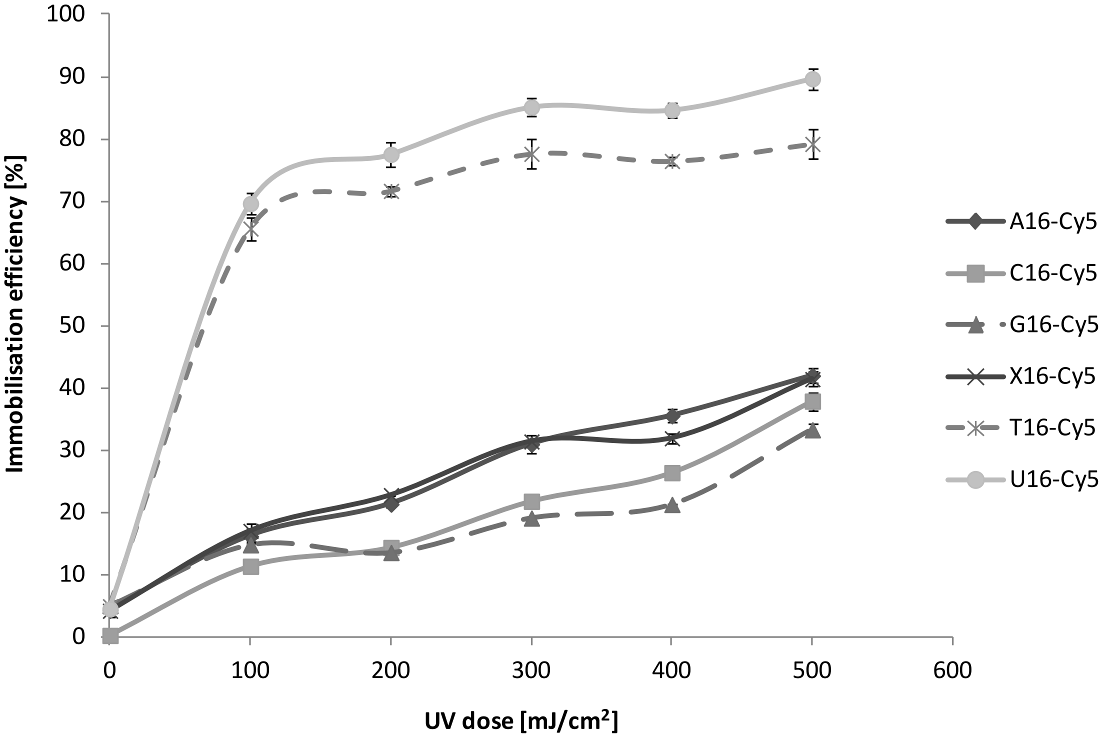


Figure 3.5 The immobilization efficiency of stretches of a 16-mer oligonucleotide built up of one type of nucleotides as a function of the UV dose. X16 denotes a-basic sites. The error bars are in the standard deviation of the intra-membrane variations (spot redundancy of 4).

In the literature, the mechanism of the reaction between primary amines and thymines, or uracils, has been proposed[20]. This reaction mechanism is presented in the upper part of 2.4. It consists of a ring-opening of the primary photoproduct 2 of the thymidine and the primary amine leading to 3, which is followed by a ring-closure leading to a dihydropyrimidine derivative 5. During heating or upon lowering the pH, the dihydropyrimidine ring is cleaved from the deoxyribose moiety 6, liberating the aminosugar moiety. For this reaction mechanism, one would expect that during temperature increase or pH reduction, the intensity of the fluorescent spots would decrease, as the remaining oligonucleotide with the fluorophore detaches from the membrane and is washed away. In our experiments such a decrease in fluorescence after extensive heating of the membrane to 95 °C was not observed, making it unlikely that this mechanism contributes significantly to our results.

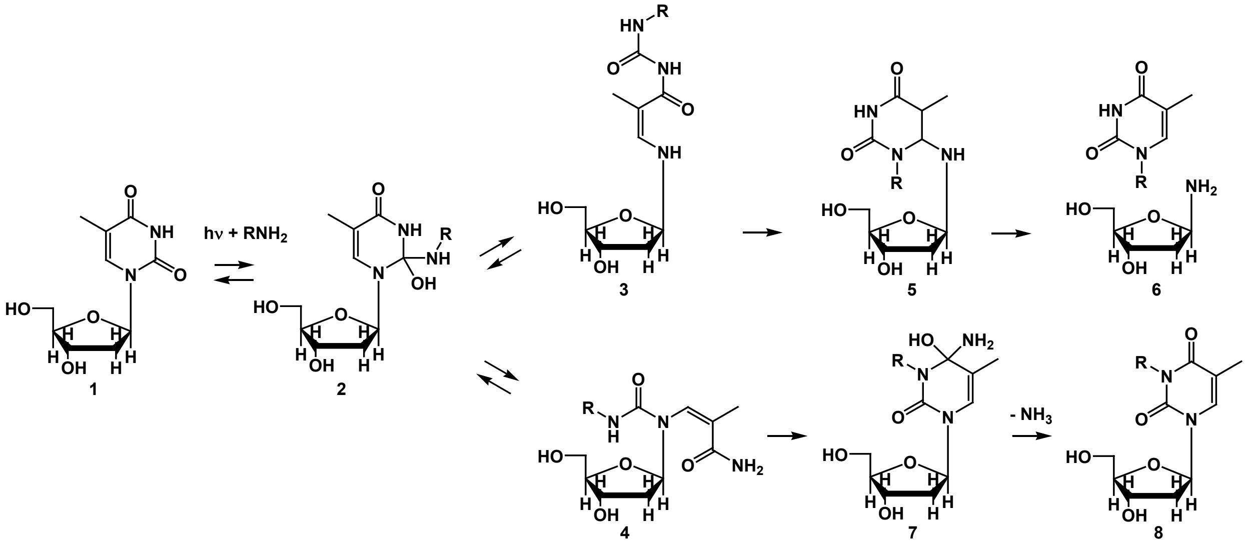


Figure 3.6 Proposed reaction scheme of the reaction between thymines and primary amines. The upper reaction scheme represents the reaction of primary amines with thymines as presented in literature[15]. The lower scheme represents our hypothesis of the reaction between thymines and primary amines which results in a covalent bond between the thymine and the amines and permanent immobilization when the amine is attached to the substrate.

In the lower part of figure 3.6, we propose a hypothetical alternative reaction which can explain our results. Instead of only forming product 3 out of the primary photoproduct 2, we propose that also reaction product 4, an isomer of 3, can be formed, as there is a second available C-N bond which can be broken. Upon ring-closure to 7 and subsequent ammonia release, product 8 is formed. When this route is followed, the nitrogen at position 3 of the dihydropyrimidine ring is substituted by the nitrogen of the primary amino group. This means that the thymidine is not cleaved from the deoxyribose moiety. Consequently, the oligonucleotide with the fluorophore cannot be detached from the membrane and be washed away and is thus in line with the observed results. A similar reaction scheme can be made for uracil bases, which only lacks the methyl group at the 5-position of the pyrimidine ring. The higher reactivity of uracil compared to thymine can be explained by the more electrophilic carbonyl group because of the lack of the electron-donating methyl group. The lower reactivity of adenosine, guanosine, and cytidine under these conditions might be due to the higher aromaticity and/or electron density compared to that of thymine and uracil. Nonetheless, more work is required for experimental verification of this mechanism.

Subsequently, we investigated the effect of attaching tails consisting of a single type of nucleotide to a 20-mer hybridization sequence. The method of this study was comparable to the experiments described previously but far more extensive, as it included all the remaining nucleotides, an improved hybridization sequence (removing the reactive thymines from the hybridization sequence), a UV window between 100 and 500 mJ/cm^2, and the inclusion of the subsequent hybridization step. We limited the UV window to 500 mJ/cm^2. At higher UV doses up to 800 mJ/cm^2, the immobilization efficiencies leveled off (data not shown), whereas there is also an increased risk of DNA degradation at higher UV doses. We included the thymine and uracil tails as being the most reactive units toward the membrane and used for comparison the guanine and cytosine as the nonreactive tails. In case no UV was applied, very low spot intensities after washing were obtained, meaning hardly any probes were bound to the surface, and therefore it was not possible to quantitatively measure these spot signals.

Figure 3.7a presents the gain in immobilization efficiencies for an increasing number of nucleotides added to the hybridization sequence, for the highest UV dose (500 mJ/cm^2). At 500 mJ/cm^2 UV dose, binding of DNA is higher than at 100 mJ/cm^2 (which is the reference case to which all values are normalized), and this leads to gains in immobilization efficiencies of higher than 100%, independent of the oligoncucleotide sequence. The immobilization is slightly improved when a guanine tail is added, which means that guanine shows, to a limited extent, reactivity toward the nylon membrane. This is an unexpected result, as from figure 3.5 it appears that the guanine nucleotides are the least reactive. This effect was visible for all UV doses (data not shown). Cytosine again shows grafting to the surface, which could previously be explained by DNA backbone reactivity. It is remarkable that there is no effect of the length of the cytosine tail on the immobilization efficiency, as an increased length of reactive DNA backbone is expected to lead to a higher degree of immobilization. One hypothesis is that immobilization of the DNA backbone is not randomly distributed over the length of the oligonucleotide, but reactivity is preferably via one of the end groups. However, the exact mechanism of such an end-group immobilization is not clear. Most pronounced are again the effects of the thymine and uracil bases with strong and increasing immobilization efficiency, which leveled off at longer tail length.

Figure 3.7b shows the same data but now presented as a function of the UV dose. The number of oligonucleotides bound to the surface was enhanced at higher UV doses, even though the cytosine tails are not reactive. It is clearly visible that there was no effect of the length of the cytosine tail on the number of oligonucleotides attached to the surface, as all curves coincide, which is contradictory to the results with a much longer (400) cytosine tail found by Saiki et al[27]. The increased binding of the oligonucleotides with the nonreactive cytosine tails can be explained by DNA backbone binding, which is also visible in figure 3.5. In this figure, also a clear increase in immobilization efficiency with increasing UV dose was observed for the plain DNA backbone.

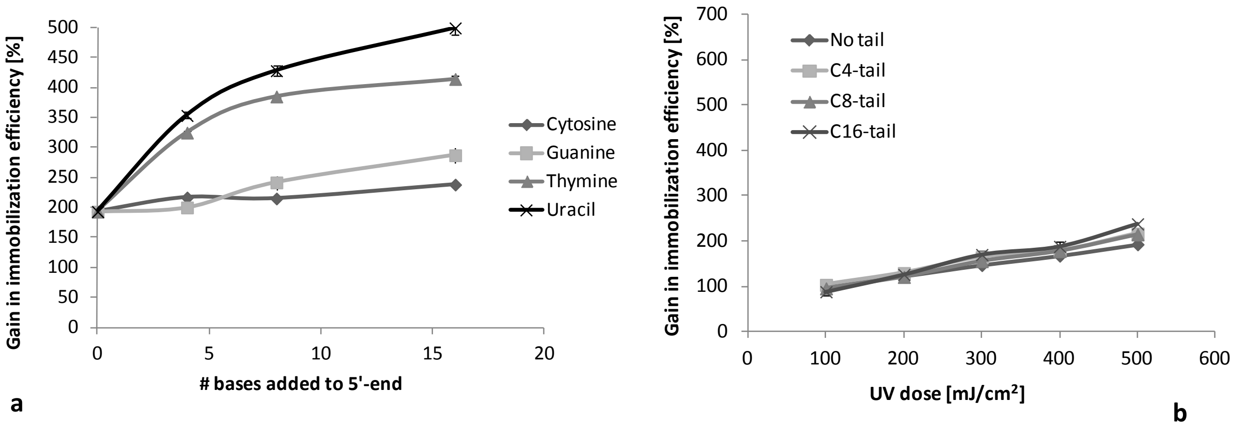


Figure 3.7a Gain in immobilization efficiency as a function of the number of deoxyribonucleotides added to the 5′-end for the four different bases included in this study, measured at 500 mJ/cm^2 UV dose. Without any tail, the gain in immobilization efficiency is around 200%. Twice as many oligonucleotides without a tail bind to the membrane at 500 mJ/cm^2 compared to 100 mJ/cm^2 (which is the reference). This is also visible in part b, where the gain in immobilization efficiency is plotted against the UV dose. **Figure 3.7b** Gain in immobilization efficiency as a function of UV dose for different lengths of cytosine tails added to the 5'-end.

Figures 3.8a-d demonstrate the effect of the tail on the hybridization efficiencies of the oligonucleotides with different tails for all UV doses, calculated according to equation 9. Hybridization is thus corrected for different surface densities because of differences in immobilization, and the effect displayed here is only the improvements in the hybridization itself.

Tail lengths of eight nucleotides or higher significantly improved hybridization. Regardless of the composition of the tails, the hybridization efficiency improves with increased length, the most pronounced with guanine, the least with uracil. This effect of improved hybridization cannot be explained fully by a spacer effect, as it is also visible in case of nonreactive tails, which means that these oligonucleotides are not preferably attached to the surface by its tail. As DNA backbone reactivity plays a role, we propose a more statistical explanation where all nucleotides have an equal chance of reaction. With increasing tail length, statistically the chances that the probe becomes attached to the membrane via one or more nucleotides from its tail sequence increase, and thus the probability that the hybridization sequence is not affected by the photochemical immobilization increases as well. This results in a more accessible hybridization sequence, leading to higher hybridization efficiencies. A second possible explanation is that short rigid oligomers become more flexible (indicated by the persistence length) with increasing tail length, thereby improving their rotational freedom and thus ease of the hybridization process. This effect is difficult to quantify, since not much information is reported about the persistence length of single stranded DNA[22].

The UV dose does not influence the hybridization efficiency, as all lines overlap well in Figure 3.8a-d. This would imply that within the UV dose window applied in these experiments, there is no damage to the DNA, which is confirmed by Kimura et al. for poly-carbo-diimide substrates[11].

Figure 3.8e-h indicates the overall effect of the tail on the signal after hybridization, which was defined as the product of immobilization and hybridization efficiency. The use of thymine and uracil tails leads to the best overall results due to both improved immobilization and hybridization. The overall performance from the use of guanine tails is also reasonable, for the most part attributed to an improved hybridization. Cytosine (and also adenine; data now shown) has only a small positive effect due to increased hybridization performance only.

3.4 Conclusions

We have studied the immobilization and hybridization of modified oligonucleotides onto amine-containing surfaces. We have used nylon membranes as a model system for their low cost and easy handling. For reference, we also performed experiments on non-porous amine-functionalized microarray glass slides (data not shown) which show comparable results. As with all these amine-functionalized substrates, the same surface chemistry is used, and the results of this study and the surface chemistry method are therefore not limited to nylon membranes but are applicable to other amine-functionalized substrates such as microarray glass slides.

We have used a two-color fluorescence measurement technique and image data analysis to study the effect of the photochemical immobilization of oligonucleotides on nylon membranes and the effects thereof on its functionality. A microarray format was used that made it possible to study the many different reactions taking place on the spots on the membrane surface. Instead of focusing on the overall signals after hybridization, we were able to quantify the number of molecules present during each individual process step, making the causes of changes in the overall signal visible.

Model capture probes were used to study the influence of the nucleotide sequence on the immobilization as well as the hybridization separately. As the involvement of the thymines in the immobilization was already known, we have deliberately removed thymines from the hybridization sequence for reduction of bias.

It was found that there are two possible mechanisms explaining the immobilization of oligonucleotides onto the amine-functionalized membranes, the first one being grafting of the DNA backbone which explains binding of the DNA regardless of its composition. One possible explanation of this reaction can be radical reactions, induced by the UV exposure of the DNA. As the UV exposure was done under ambient conditions, the formation of oxygen radicals which can react with the DNA is therefore a likely explanation for nonspecific DNA reactions. A second mechanism is a nucleotide-specific mechanism involving thymine, uracil, and to a limited extent also guanine. The combination of these two mentioned mechanisms explains why all nucleotides, independent of the oligonucleotide sequence, bind to the membrane, and why this is enhanced by the addition of especially thymine and uracil.

In practice, capture probes containing thymines in the hybridization sequence are used. This means that thymines inside the hybridization sequence are able to react with the membrane surface as well, and as a result parts of the capture probes might lose their functionality. By adding a tail of thymines or uracils to the hybridization sequence, this effect can be reduced, as statistically more thymines of the tail will react with the membrane and the hybridization sequence is less affected, leading to sequences with better hybridization performance. Also, tails built up of other nucleotides were found to improve hybridization, as statistically more end-group immobilization takes place, leading to higher hybridization efficiencies. This explains why tails that do not directly contribute to immobilization can still improve hybridization. The results from this study can be used to improve microarray manufacturing using amine-functionalized substrates, by using, for example, tails consisting of different types of nucleotides for both immobilization and hybridization improvement.

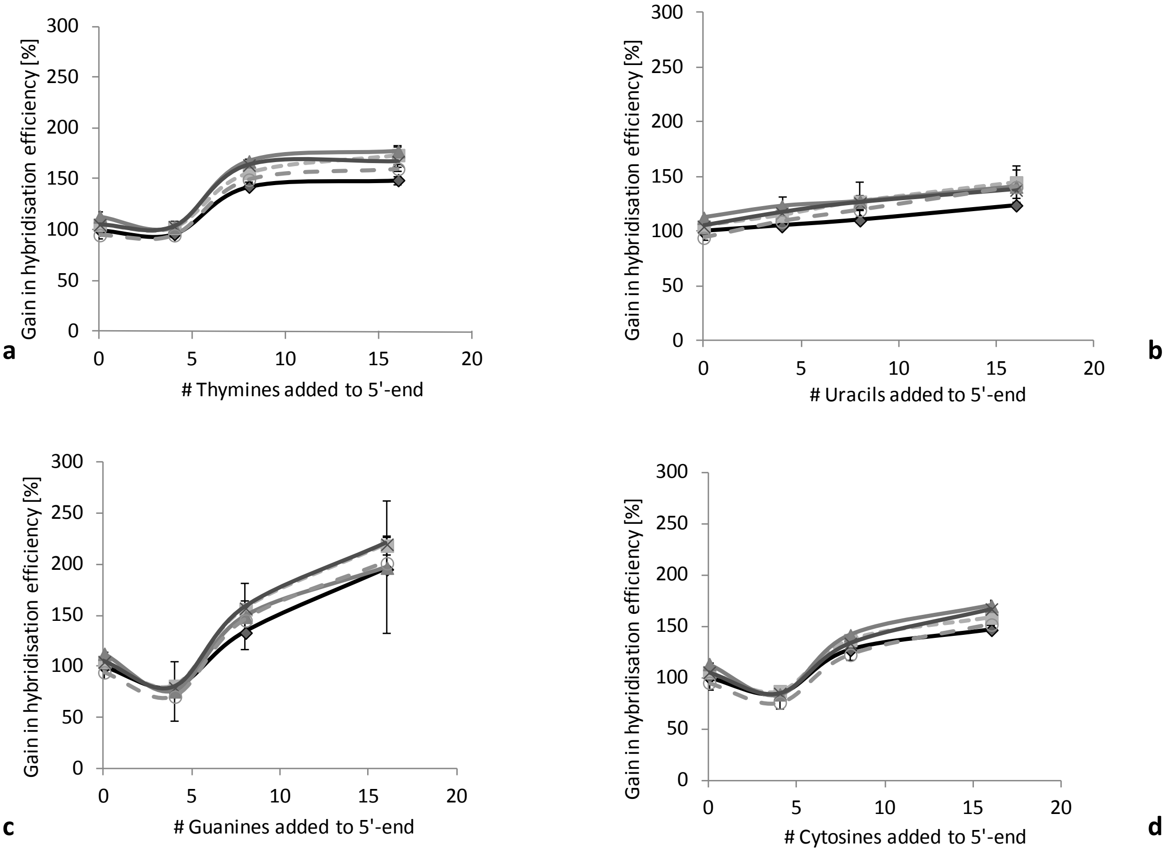


Figure 3.8a-d Gain in hybridization efficiency as a function of the number of deoxyribonucleotides added to the 5′-end for thymine (a), uracil (b), guanine(c), and cytosine (d). The lines in each graph represent different UV doses utilized during the immobilization step: diamonds = 100 mJ/cm^2, squares = 200 mJ/cm^2, triangles = 300 mJ/cm^2, cross = 400 mJ/cm^2, circles = 500 mJ/cm^2. The error bars represent the minimum and maximum value of the measurement points. For the guanine measurements, only 300 and 500 mJ/cm^2 gave large errors.

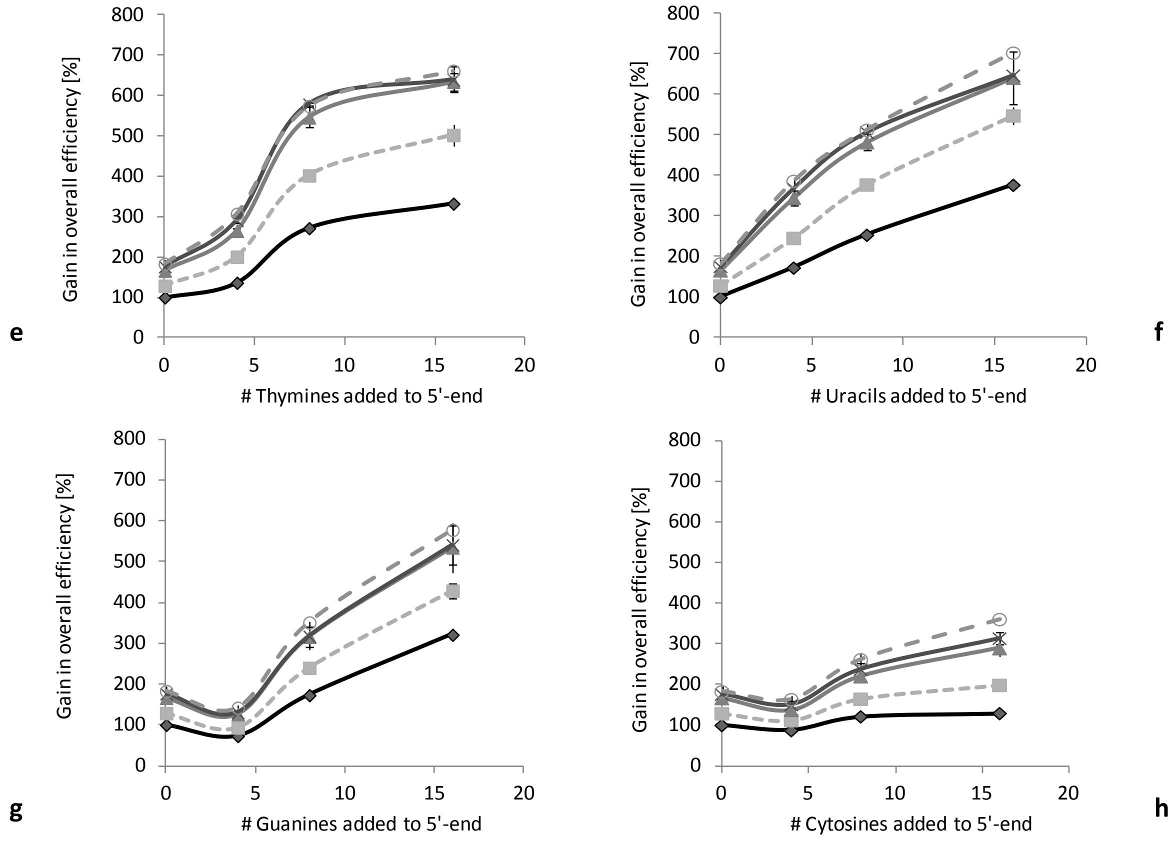


Figure 3.8e-h Gain in total efficiency as a function of the number of deoxyribonucleotides added to the 5′-end for thymine (e), uracil (f), guanine (g), and cytosine (h). This gain is the product of the relative immobilization efficiency and the relative hybridization efficiency. The lines in each graph represent different UV doses utilized during the immobilization step (diamonds = 100 mJ/cm^2, squares = 200 mJ/cm^2, triangles = 300 mJ/cm^2, cross = 400 mJ/cm^2, circles = 500 mJ/cm^2. The error bars represent the minimum and maximum value of the measurements.

References

1. Dufva M. Fabrication of high quality microarrays. *Biomol Eng* 2005, 22:173-184
2. Guo Z, Guilfoyle RA, Thiel AJ, Wang RF, Smith LM. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res* 1994, 22:5456-5465
3. Zammatteo N, Jeanmart L, Hamels S, Courtois S, Louette P, Hevesi L, Remacle J. Comparison of different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. *Anal Biochem* 2000, 280:143-150
4. Auburn RP, Kreil DP, Meadows LA, Fischer B, Matilla SS, Russell S. Robotic spotting of cDNA and oligonucleotide microarrays. *Trends Biotechnol* 2005, 23:374-379
5. Heller MJ. DNA microarray technology: devices, systems and applications. *Annual Rev Biomed Eng* 2002, 4:129-153
6. Pierik A, Dijksman JF, Raaijmakers A, Wismans A, Stapert H. Quality control of inkjet technology for DNA microarray fabrication. *Biotech J* 2008, 3:1581-1590
7. Brown TJ, Anthony RMJ. The addition of low numbers of 3'thymine can be used to improve the hybridization signal of oligonucleotides for use within arrays on nylon supports. *J Microbiol Meth* 2000, 42:203-207
8. Shi LM, Reid LH, Jones WD, Shippy R, Warrington JA, Baker S et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 2006, 24:1151-1161
9. Joos B, Kuster H, Cone R. Covalent attachment of hybridizable oligonucleotides to glass supports. *Anal Biochem* 1997, 247:96-101
10. Huang E, Zhou FM, Deng L. Studies of surface coverage and orientation of DNA molecules immobilized onto preformed alkanethiol self-assembled monolayers. *Langmuir* 2000, 16:3272-3280
11. Kimura N, Oda R, Inaki Y, Suzuki O. Attachment of oligonucleotide probes to poly carbodiimide-coated glass for microarray applications. *Nucleic Acids Res* 2004, 32: e68
12. Wu P, Castner DG, Grainger DW. Diagnostic devices as biomaterials: a review of nucleic acid and protein microarray surface performance issues. *J Biomat Sci-Polym E* 2008, 19(6): 725-753
13. Wang H-Y, Malek RL, Kwitek AE, Greene AS, Luu TV, Behbahani B, Frank B, Quackenbush J, Lee NH. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biology* 2003, 4:R5
14. Taylor S, Smith S, Windle B, Giuseppe-Elie A. Impact of surface chemistry and blocking strategies on DNA microarrays. *Nucleic Acids Res* 2003, 31(16):e87.
15. Nierzwickibauer SA, Gebhardt JS, Linkkila L, Walsh K. A comparison of UV cross-linking and vacuum baking for nucleic acid immobilization and retention. *Biotechniques* 1990, 9:472-478.

16. Gong P, Grainer DW. Comparison of DNA immobilization efficiency on new and regenerated commercial amine-reactive polymer microarray surfaces. *Surf Sci* 2004, 570:67-77
17. Cheng F, Gamble LR, Grainger DW, Castner DG. X-ray photoelectron spectroscopy, time-of-flight secondary ion mass spectrometry, and principal component analysis of the hydrolysis, regeneration, and reactivity of N-hydroxysuccinimide-containing organic thin films. *Anal Chem* 2007, 7:8781-8788
18. Church GM, Gilbert W. Genomic sequencing. *P Natl Acad Sci Biol* 1984, 81:1991-1995
19. Calevro F, Charles H, Reymond N, Dugas V, Cloarec JP, Bernillon J, Rahbe Y, Febvay G, Fayard JM. Assessment of 35mer amino-modified oligonucleotide based microarray with bacterial samples. *J Microbiol Methods* 2004, 57:207-218
20. Saito I, Sugiyama H, Furukawa N, Matsuura T. Photochemical ring opening of thymidine and thymine in the presence of primary amines. *Tetrahedron Lett* 1981, 22:3265-3268
21. Zhang WT, Ni HH, Capp MW, Anderson CF, Lohman TM, Record MT. The importance of coulombic end effects: Experimental characterization of the effects of oligonucleotide flanking charges on the strength and salt dependence of oligocation (L^{8+}) binding to single-stranded DNA oligomers. *Biophys J* 1999, 76:1008-1017
22. Tinland B, Pluen A, Sturm J, Weill G. Persistence length of single-stranded DNA *Macromolecules* 1997, 30:5763-5765
23. Fare TL, Coffey EM, Dai HY, He YDD, Kessler DA, Kilian KA, Koch JE, LeProust E, Marton MJ, Meyer MR, Stoughton RB, Tokiwa GY, Wang YQ. Effects of atmospheric ozone on microarray data quality. *Anal Chem* 2003, 75:4672-4675
24. Marras SAE, Kramer FR, Tyagi S. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res* 2002, 30 (21):e122
25. Wu Y, de Kievit P, Vahlkamp L, Pijnenburg D, Smit M, Dankers M, Melchers D, Stax M, Boender PJ, Ingham C, Bastiaensen N, de Wijn R, van Alewijk D, van Damme H, Raap AK, Chan AB, van Beuningen R. Quantitative assessment of a novel flow-through porous microarray for the rapid analysis of gene expression profiles. *Nucleic Acids Res* 2004, 32 (15): e123
26. Dijksman JF, Pierik A. Fluid dynamical analysis of the distribution of ink jet printed biomolecules in microarray substrates for genotyping applications. *Biomicrofluidics* 2008, 2, 44101
27. Saiki RK, Walsh PS, Levenson CH, Erlich HA. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotices probes. *P Natl Acad Sci USA* 1989, 86:6230-6234
28. Steel AB, Herne TM, Tarlov MJ. Electrochemical quantitation of DNA immobilized on gold. *Anal Chem* 1998, 70:4670-4677

29. Lee CY, Nguyen PCT, Grainger DW, Gamble LJ, Castner DG. Structure and DNA hybridization properties of mixed nucleic acid/maleimide-ethylene glycol monolayers. *Anal Chem* 2007, 79:4390-4400
30. Peterson AW, Heaton RJ, Georgiadis RM. The effect of surface probe density on DNA hybridization. *Nucleic Acids Res* (2001), 29:5163-5168

Chapter 4

Rapid genotyping of Human Papilloma Virus by post-PCR array-based hybridization techniques

This chapter is based on:

Pierik A, Zwanenburg C, Moerland E, Broer DR, Stapert HR, Adriaan van den Brule JC. *J Clin Microbiol* 2011, 49(4):1395-1402

Abstract

Kinetic hybridization measurements on a microarray are expected to become a valuable tool for genotyping applications. A method has been developed that enables kinetic hybridization measurements of PCR products on a low-density microarray. This is accomplished by pumping a solution containing PCR products up and down through a porous microarray substrate. After every pumping cycle, the fluorescently labeled PCR products hybridized to capture probes immobilized on the solid surface of the porous microarray substrate are measured. By this method, both binding curves and high-resolution melting curves are obtained instead of the single endpoint hybridization intensities as with commonly used post-PCR array-based hybridization techniques. We used 20 subtypes of the human papillomavirus (HPV) as a model system to test our detection method and blindly analyzed 216 clinical samples. We compared our microarray flow-through method with a reference method, PCR followed by a reverse line blot (RLB). Real-time hybridization measurements followed by high-resolution melting curves of low concentrations of fluorescently labeled HPV targets on a microarray were successfully carried out without any additional chemical signal amplification. The results of our new method were in good agreement (93 %, with a kappa coefficient κ of 0.88 [95 % CI, 0.81 to 0.94]) with the RLB results. All discrepant samples were analyzed by a third method, enzyme immunoassay (EIA). Furthermore, in a number of cases, we were able to identify false-positive samples by making use of the information contained in the kinetic binding and melting curves. This clearly demonstrates the added value of the use of kinetic measurements and high-resolution melting curves, especially for highly homologous targets.

4.1 Introduction

PCR-based methods allow for the rapid, cost-effective, and automated diagnosis of pathogens and have become a powerful tool in microbiology and virology during the last decade[1]. Real-time PCR makes the simultaneous detection of a number of targets possible. In practice, this number is limited to around six targets because of the spectral overlap of the different fluorophores that have to be discriminated in a single reaction mixture volume. One way of detecting more targets in parallel is by splitting the sample into several separate PCR volumes and performing different multiplex PCRs on each volume. However, especially in the case of low target concentration, splitting sample volumes increases the risk of target copy numbers being below the limit of detection. Therefore, the number of times that a sample can be split should be limited. For the parallel detection of a larger number of different PCR products, a separate, post-PCR detection step is a suitable alternative. Several post-PCR techniques exist to detect amplified PCR products, generally based on hybridization assays. They have the advantage that the multiplex assays have higher target capabilities than real-time PCR. Genotyping is especially necessary for highly homologous targets, such as the human papillomavirus (HPV), for which over 100 different types have been identified. The types can be divided, based on their phylogenetic relationship and their presence in benign or malignant cervical lesions, into high-risk and low-risk types[2]. Different molecular techniques exist to identify HPV DNA or RNA, and a number of post-PCR tests are commercially available. With a broad-spectrum PCR, many different HPV types can be amplified in one sample using a single primer pair. Costs and labor are therefore reduced compared to those necessary to run specific amplifications in parallel. Different broad-spectrum HPV PCRs are available, including GP5$^+$/GP6$^+$ [3,4], MY09/11 [6], and CPI/IIG [7].

Post-PCR genotyping tests for the determination of a specific virus type are important because molecular epidemiology clearly indicates that certain human papillomavirus types are the principle cause of invasive cervical cancer, which is the second most common cancer in women worldwide [3,8]. This kind of test is frequently performed by a hybridization assay. One established technique is the reverse line blot (RLB), which captures and detects specific subtypes on a membrane support. This method has been proven to be suitable for the identification of specific HPV types and the detection of multiple infections[4,9] but has the disadvantage of being laborious and time consuming. Other commercially available tests include the Linear Array HPV genotyping test (Roche Diagnostics), the INNO-LiPA HPV test (Innogenetics NV), and the PapilloCheckTest (Greiner Bio-One GmbH), all using a general primer PCR followed by a hybridization detection. Furthermore, a test has been developed based on a general primer PCR followed by hybridization with type-specific oligonucleotide probes coupled to fluorescently labeled Luminex polystyrene beads, making the differentiation of to up to 100 different targets possible[10].

Other tests aim to identify a high-risk infection only by using either enzyme immunoassay (EIA) tests (Digene; Qiagen) or cocktails of group-specific detection probes (Amplicor; Roche Diagnostics). RNA-based tests are also available, e.g., bioMérieux and Norchip offer a nucleic acid sequence-based amplification (NASBA) test that detects the expression of the E6/E7 viral oncogenes.

One commonality of all available DNA detection tests is a sequential hybridization and detection step. This results in a single endpoint intensity measurement. Discrimination between a-specific and specific binding is difficult. Despite stringent washing conditions, cross-hybridization is still known to occur and is one of the main causes of false-positive results[11,12]. An advantage of real-time PCR is that some of the detection chemistries used (e.g., intercalating dyes or detection probes like molecular beacons) allow for a post-PCR high resolution melting curve analysis, which gives additional information on specificity[13,14]. Especially for highly homologous targets like HPV, additional information on melting temperatures, like that obtained in real-time PCR, is valuable to reduce false positives, but it cannot be obtained with the currently available methods.

In this chapter, we describe a method that combines the dynamic measurements of real-time PCR with the high multiplex detection capability of the post-PCR hybridization arrays.

We have developed a method that integrates microarray hybridization and detection and evaluated this kinetic hybridization method for genotyping applications. With this system, PCR products are pumped up and down through a porous microarray substrate (membrane) at a controlled temperature. During binding (hybridization) and the subsequent unbinding (melting), the fluorescence pattern is imaged after every pumping cycle. By concatenating these images, the kinetics of hybridization and melting can be followed in detail over time. In reference 15, a microarray flow-through platform is disclosed that can also detect PCR products. In this platform, special aluminum-oxide porous substrates are used that have an increased specific surface area. In this chapter, standard commercially available porous nylon membranes are used as the microarray surface. We have investigated the use of this method for genotyping applications and have chosen HPV as a model system, due to the method’s relevance for discriminating different HPV variants. As a benchmark for this genotyping analysis, we used the reverse line blot (RLB). This system is technically the most identical because the same substrate material and the same primer sets are used. Both methods also allow the processing of sample types such as cervical tissue and could therefore be used in a feasibility study that included this sample type.

4.2 Materials and methods

4.2.1 Sample selection

For optimization of the microarray performance, HPV PCR products obtained from clinical samples with known HPV types, as well as HPV 16 PCR products obtained from human cervical SiHa cell lines, were used. Subsequently, clinical samples with unknown HPV infections were first analyzed by using the EIA and the RLB methods. Based on the outcomes of these tests, 180 samples were selected to be analyzed by the microarray flow-through method for the evaluation of the kinetic measurements. The 180 samples were divided into three sample groups consisting of 60 samples each. The first group had yielded contradictory results between EIA and RLB or had shown the presence of multiple infections. The second group consisted of cytological cervical cell suspensions that showed borderline atypical squamous cells of undetermined significance or that mild dyskaryosis. A third group of samples included DNA from formalin-fixed, paraffin-embedded cervical intraepithelial neoplasia grade I (CIN I) and CIN II tissues. Each group also included a number of samples that were HPV negative based on the outcome of the EIA or RLB. Furthermore, each sample group contained a total of 12 sample processing controls and PCR blanks for quality control. In total, 216 samples were included in this study.

4.2.2 Sample pretreatment

Cytological cervical cell suspensions in Preserv-Cyt fixative (ThinPrep; Cytyc, Marlborough, MA, USA) were centrifuged, and after removal of the supernatant, the remaining pellet was put into 180 µl lysis mix (containing 3.3 mg/ml proteinase K from a proteinase K high pure PCR template preparation kit [Roche Diagnostics, Basel, Switzerland]) in proteinase K buffer

(10 mM Tris-HCl [pH 7.5] and 0.5 % Tween 20 [Sigma-Aldrich, St. Louis, MO, USA]). After mixing, the samples were incubated in a block heater for 1 h at 56 °C. The samples were boiled and centrifuged for 1 min at 8,000 x g. The formalin-fixed, paraffin-embedded tissue samples were pretreated as described previously[4], using an overnight incubation at 37 °C with the lysis mixture as described above.

4.2.3 PCR

All pretreated samples were subjected to a DNA quality check by amplification of β-globin by using the primers PCO_3 and PCO_5 for PCR, previously described in reference 2. For HPV detection, 10 µl of the pretreated sample was subjected to PCR in a 50-µl reaction mixture volume using the general primers $GP5^+/GP6^+$ in a 9700 thermal cycler (Applied Biosystems, CA, USA) with the PCR program described in reference 4. These primers anneal to a homologous region of the HPV types[4]. For the microarray analysis, the $GP6^+$ primer was labeled with cyanine-5 dye (Cy5) (Biolegio, Nijmegen, Netherlands), whereas for the RLB, the $GP6^+$ primer contained biotin.

4.2.4 Microarray manufacturing

Biodyne C membranes (Pall Biosupport, East Hills, NY, USA) were punched into 12-mm circular pieces and laser welded onto nylon rings to facilitate handling and processing. The membranes were activated by incubation for 10 min at room temperature with a 16 % 1-ethyl-3-dimethylaminopropyl carbodiimide hydrochloride (EDC) solution (Sigma-Aldrich, St-Louis, MO, USA). After a short rinse with Milli-Q water, the membranes were dried and then positioned on a single-nozzle inkjet printer[16]. 5'-Amine-functionalized oligonucleotide capture probes (all high-performance liquid chromatography purified; Biolegio, Nijmegen, the Netherlands) were dissolved in print buffer (60 mM $NaHCO_3$ buffer, pH 8.5) to a concentration of 40 µM. Each spot of the microarray was built up with 400-picoliter print fluid. The final spots, containing the capture probes attached to active groups on the inner surface area of the membrane, were formed by evaporation of the solvent. Each fluid was printed in triplicate for purposes of redundancy. After printing and drying, the remaining active groups on the membrane were deactivated by 8-minute incubation with 0.1 M NaOH. The final step of the array manufacturing procedure was blocking the surface for 5 minutes at 50°C using a 2 x SSPE–0.1 % SDS blocking solution (1 x SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4,and 1 mM EDTA, pH 7.7) and then incubating it for 1 min with a 20 mM EDTA solution at pH 8.0.

4.2.5 Microarray flow-through optimization

The microarray flow-through method was optimized for the concentrations of the printed capture probe fluids, capture probe sequences, and hybridization conditions. The capture probe sequences for the Reverse Line Blot, as originally published[4], were adapted for the microarray analysis to make use of the “position-of-label” effect[17]. The hybridization sequences were chosen such that the position of the capture probe was close to the 5'-end of the amplicon containing the fluorophore, which can account for a considerable increase in hybridization signals, especially for short capture probes as used in this study[18]. All designs were evaluated by using the database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) and checked for homology by using BLAST. They were subsequently evaluated for analytical sensitivity and specificity by hybridization of 2 nM PCR product of each HPV type on an individual microarray that contained spots of both the RLB and the position-of-label capture probes. Sensitivity was addressed by comparing the spot intensities of the RLB and the new microarray oligonucleotide sequences of each HPV type. All final sequences obtained during this optimization were used in further experiments and are given in table 4.1. In total, 20 different HPV types were considered in this study; all of the high-risk HPV types were used, including a number of suspected high-risk types which we categorize as high risk (HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82). We also used a number of low-risk types (HPV 6, 11, 26, 53, and 66) that are the most prominent low-risk HPV types known today. The overall sensitivities were assessed based on hybridizations of dilution series of PCR products and were compared with those of RLB.

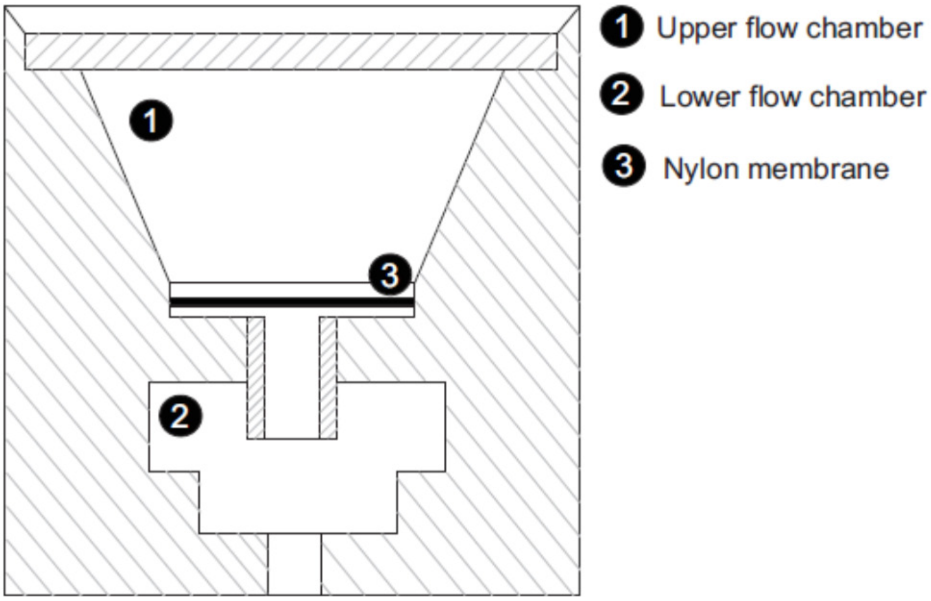


Figure 4.1 Experimental setup[20] showing the membrane that contains the capture probes and the upper and lower flow chamber.

4.2.6 Microarray flow-through detection and analysis

The volume containing the PCR products was diluted in Milli-Q water to 100 µl. The PCR products were denatured by heating the samples up to 95 °C for 5 minutes and placing them on ice. The samples were diluted in 400 µl of hybridization buffer, with a final buffer composition of 1.9 x SSPE and 0.08 % SDS.

Figure 4.1 shows a schematic of the flow-through device[20]. The porous microarray substrate was positioned between a so-called upper flow chamber and a lower flow chamber. Not shown are the means to apply pressure to the contents of either the upper chamber or the lower chamber to force the solution through the porous microarray substrate. Prior to hybridization, the solution was preheated to the hybridization temperature of 42 °C.

During hybridization, the solution was forced to flow through the membrane up and down from one chamber to another about 30 times at a rate of 1 cycle per minute. During the time the solution was flowing through the porous microarray substrate, the fluid containing the fluorescently labeled PCR products was in direct contact with the capture probes on the membrane and hybridization could occur. The fluorescence intensities of the spots were measured from above using a charge-coupled-device (CCD) camera (Sony) and an illumination system with four red light-emitting diodes (LEDs) (LXHL-MD1D Lumileds; Philips) and optical filters (XF 1069; Omega Optical) suitable for the excitation and detection of the Cy5 dye. After each pumping cycle, at the moment when all the fluid was below the membrane in the lower flow chamber and no fluorescent liquid was on top of the membrane (minimizing background signal), the microarray spot pattern was imaged. For each HPV spot type, the fluorescence intensity was measured once during each pumping cycle, and thus, the cumulative fluorescence intensity during the whole hybridization process could be followed over time. This procedure allowed for (semi) real-time hybridization kinetics measurements without washing. After the hybridization step, melting curves were obtained by slowly increasing the temperature from the hybridization temperature of 42 °C to 95 °C at a rate of 1 °C/minute, allowing for chemical and thermal equilibration[21]. During melting, the mixture was pumped up and down through the membrane at one cycle per minute. The assay conditions were optimized for the hybridization buffer composition, temperature, and pressure levels. Kinetic hybridization curves and melting curves were obtained by plotting the spot intensities as a function of time and temperature, respectively. Melting temperatures were determined at the temperature where the spot intensity equaled 50 % of the original signal.

The total analysis time for a 27-cycle hybridization curve and a subsequent melting curve measurement of 53 measurement points thus takes 80 minutes, much faster than the RLB, which needs several hours for hybridization and detection. The background-corrected intensities (following the procedure as in reference 16) of three corresponding spots were averaged. A cutoff value of three times the standard deviation was used. The identification of the HPV type was initially based on the level of the hybridization signals at the endpoint.

Table 4.1 Capture probe and primer sequences including the length, used in the Reverse Line Blot[4] and the microarray flow-though method.

HPV type	Sequence 5'-3' RLB	Length (nt)	Sequence 5'-3' flow-through microarray	Length (nt)
HPV 6	tccgtaactacatcttcca	19	tataaagagtacatgcgtcat	21
HPV 11	tctgtgtctaaatctgctac	20	gaatacatgcgccatgtg	18
HPV 16	cattatgtgctgccatatc	19	gagtacctacgacatggg	18
HPV 18	tgcttctacacagtctcct	19	gcagtatagcagacatgttg	20
HPV 26	gtacattatctgcagcatc	19	aaacaatttataagacatggcgaa	24
HPV 31	gcaattgcaaacagtgatac	20	agagtatttaagacatggtgag	22
HPV 33	tgcacacaagtaactagtga	20	aaagaatatataagacatgttgaag	25
HPV 35	ctgctgtgtcttctagtga	19	ggaatatttaaggcatggtgaa	22
HPV 39	atagagtcttccataccttc	20	ggaatataccaggcacgtg	19
HPV 45	taatttaacattatgcctc	19	agtacatatgaccctactaagt	22
HPV 51	tgctgcggtttccccaa	17	tccaagtaactttaagcaatatat	24
HPV 52	gaataccttcgtcatggc	18	ggaataccttcgtcatggc	19
HPV 53	tgtctacatataattcaaagc	21	aacagtatgttagacatgcag	21
HPV 56	cagaacagttaagtaaatatg	21	tcagtaccttagacatgtgg	20
HPV 58	tatgcactgaagtaactaag	20	ggaatatgtacgtcatgttgaa	22
HPV 59	tctactactgcttctattcc	20	agaatatgccagacatgtgg	20
HPV 66	agctaaaagcacattaactaa	21	tcaataccttcgccatgtg	19
HPV 68	ctgaatcagctgtaccaaat	20	ggaatatattaggcatgttgag	22
HPV 73	acaggctagtagctctact	19	aaggaatatttaagacatgcagaa	24
HPV 82	actgctgttactcaatctg	19	gcagtacattaggcatggg	19
B-glob			aaactgggcatgtggaga	18
GP5+	tttgttactgtggtagatactac	23	tttgttactgtggtagatactac	23
GP6+	gaaaaataaactgtaaatcatattc (Biotin at 5'-end)	25	gaaaaataaactgtaaatcatattc (Cy5 at 5'-end)	25

Melting curves were available for all hybridizations and were used to investigate discrepancies between the RLB and the microarray flow-through method. The outcomes of the flow-through measurements were compared to the results of the RLB. Cohen's kappa coefficient was used as a tool to analyze the reliability between the two methods.

4.2.7 RLB detection

The reverse line blot (RLB) protocol was adapted from reference 4. RLB makes use of a Biodyne C membrane (Pall Biosupport, East Hills, NY) onto which many different capture oligonucleotides were immobilized in rows by using a miniblotter MN45 (Immunetics, Cambridge, MA) in house. Each oligoprobe had a sequence complementary to one particular HPV type that was being detected. Perpendicular to the rows of deposited probes, up to 45 different biotin-labeled PCR products, as well as additional negative and positive controls and internal quality controls, were pipetted into parallel channels. After hybridization of the PCR products with the capture probes, the hybridization products were made visible by incubation with antibiotin conjugates and measured with enhanced chemiluminescence detection (ECL).

4.2.8 EIA

The EIA was performed as previously described [2,4]. In the EIA, checking for infection with high-risk or low-risk HPV is done by incubation with a cocktail of group-specific detection probes. The low-risk cocktail contained probes for low-risk types in addition to those in the microarray or the RLB, namely, HPV 34, 40, 42, 43, 44, 54, 55, 57, 61, 70, 71, 72, 81, 83, and 84. Furthermore, HPV 73 and 82 were part of the low-risk cocktail while being considered high risk throughout the analysis described in this paper.

4.3 Results

The position-of-label effect, as described previously[17] explains the increase of the hybridization intensity when the capture probe is located closer to the fluorophore that is attached to the amplicon. We evaluated this effect in our study. The data collected clearly show that positioning the hybridization site of the amplicons and, thus, the capture probe closer to the fluorophore indeed increases signal intensity. The results in figure 4.2 show this effect, with the design of different capture probes of the HPV 16 amplicon as an example. In order to rule out any specific involvement of the Cy5 dye, additional experiments were performed with Alexa Fluor 633-labeled PCR products, which resulted in the same trend.

To make optimal use of this effect, we have accordingly redesigned all 20 capture probe designs originally described in reference 4. Sixteen of the 20 capture probe sequences indeed showed improved hybridization signals. A few of these new sequences, however, showed cross-hybridization, and therefore, the capture probe sequences of HPV 6, 45, and 51 were redesigned. For the other sequences showing cross-hybridization, specificity could be increased by increasing the stringency of the hybridization conditions (optimization of buffer conditions). The optimized capture probe sequences used in our experiments are listed in table 4.1.

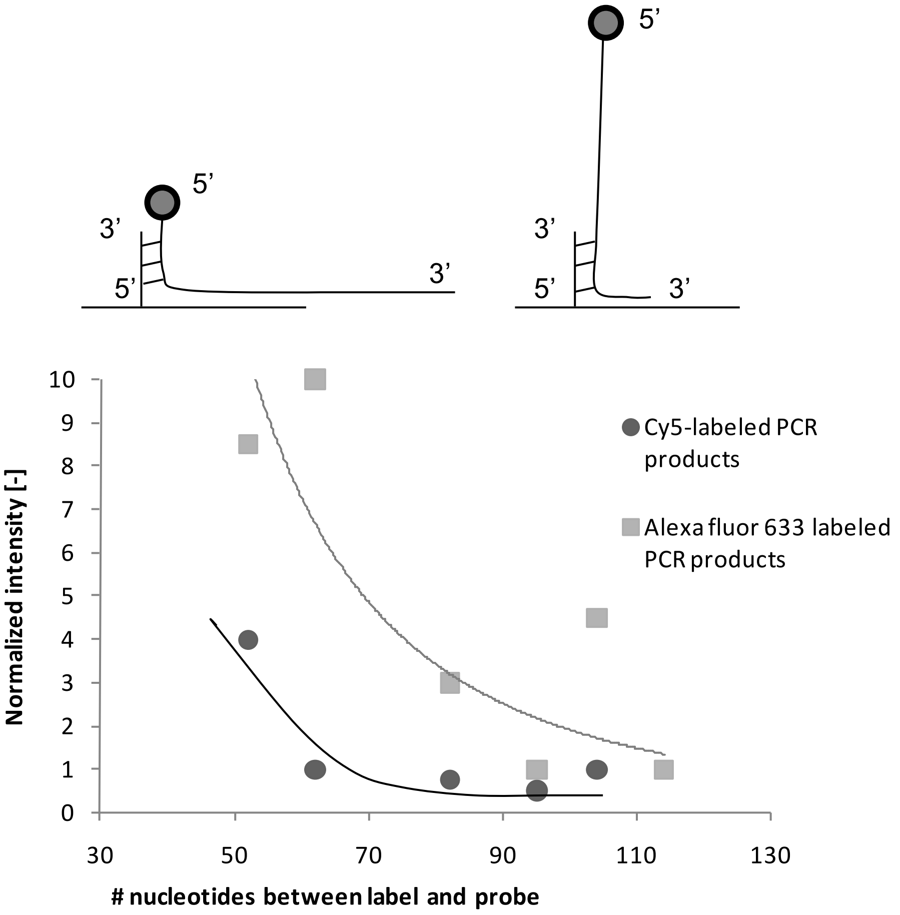


Figure 4.2 Influence of the position-of-label effect for HPV 16. (a) Schematic representation of a label (circle) near the hybridization probe (upper left) and at a distance (upper right). The closer the label to the position of the capture probe, the higher the hybridization signal. (b) The gain in hybridization intensity is plotted against the number of nucleotides that separate the label from the site where the amplicon is attached to the capture probe. The intensity for the capture probe that is the furthest away from the fluorophore is set at a value of 1. Comparison of the data for the Cy5-labeled and the Alexa Fluor 633-labeled PCR products supports (dye-independent) improvement in hybridization intensity when using capture probes complementary to the region close to the fluorophore. The relatively large scatter of the data is partly caused by small differences in melting temperature.

Figure 4.3 shows both the binding curves and the high-resolution melting curves obtained when 2 nM PCR product of a mixture of HPV 68 and 73 was incubated on the microarray. The left graph shows kinetic hybridization curves. It is clear that the spots containing capture probes for HPV 33, 35, and 51 also give positive signals because of cross-hybridization. During the first few cycles, the cross-hybridization signals increase, while during later cycles, the spot intensities decrease, which is a characteristic feature for competitive hybridization[18,19]. Based purely on the fluorescence levels obtained at the endpoint, after cycle 27, discrimination between matching and non-matching oligonucleotides is not possible. The right-hand graph shows the different melting curves: the perfect matches clearly have a higher melting temperature than the nonmatching oligonucleotides. Only after evaluation of the melting curves can it be concluded that cross-hybridization to HPV 35 and HPV 51 is the cause of false-positive hybridization signals. Furthermore, the kinetic hybridization measurements show that HPV 33 cross-hybridization disappears over time.

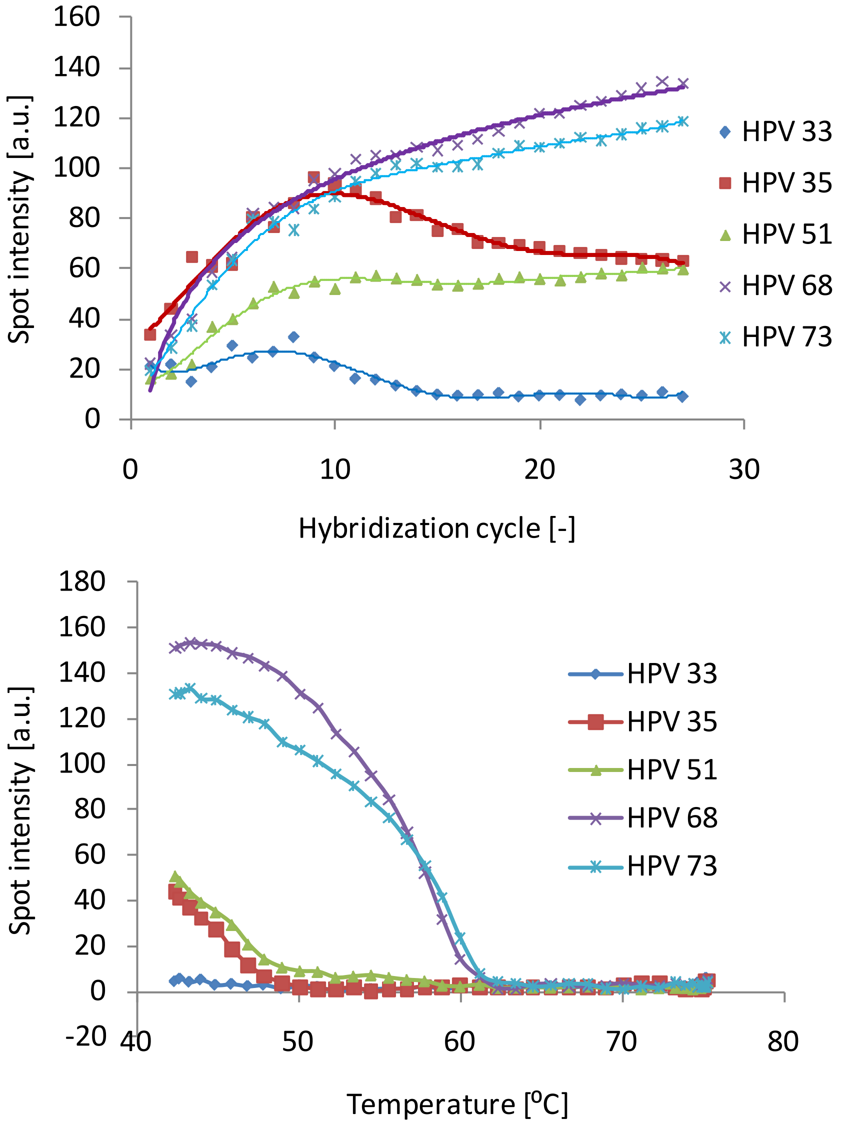


Figure 4.3 Kinetic hybridization curves (upper) and melting curves (lower) of a sample containing a 2 nM mixture of HPV 68 and HPV 73. Cross-hybridization of HPV 33, 35, and 51 is clearly observed. HPV 33, 35, and 51 show degrees of homology of 82 % with HPV 68 and 83, 88, and 75 %, respectively, with HPV 73. A distinction can be made between the perfect matches and the mismatches by consulting the melting curves, where clear differences in melting temperatures are observed. BG: background; a.u.: arbitrary units.

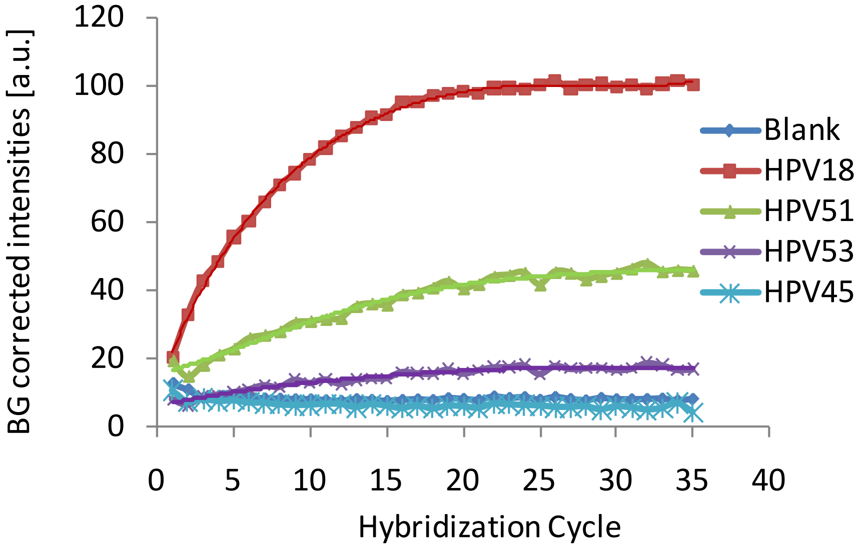
The kinetic hybridization curves can be used to estimate binding kinetics. Based on the Langmuir model for adsorption, the apparent association constant can be calculated, assuming that the dissociation constant is much lower than the product of the association constant and the target concentration. This value is equal to 7.10^5 $l \cdot mol^{-1}$ s^{-1}, which is nicely in line with values found by others[18,20]. This shows that these kinetic curves give a reasonable understanding of the hybridization process and can be used for optimization thereof.

The analytical sensitivity of the flow-through method, depending slightly on HPV type, was found to be around 100 pM, which is comparable to the sensitivity of the RLB. This means that the flow-through method allows for endpoint PCR without chemical signal amplification as needed with the RLB. Therefore, washing and additional chemical processes are eliminated, reducing time and costs.

Figure 4.4 shows the hybridization curves and melting curves of a typical clinical sample. The flow-through analysis confirmed the triple HPV infection found by reverse line blot analysis (HPV 18, 51, and 53). As HPV 45 has a high homology to these three HPV types, this specific hybridization curve was added to show the high specificity of the method. The melting curve analysis afterwards delivered melting temperatures of 52, 54, and 51 °C for HPV 18, 51, and 53, respectively, which is fully in agreement with the theoretical values that were calculated based on the hybridization sequences.

Table 4.2 shows the overall results of the analysis of the 216 samples, as well as the results for each of the three subgroups, for both the RLB and the microarray method. Group 1 contained the samples with many HPV infections, including a total of 15 double or triple infections. This was confirmed by the microarray flow-through method. Group 2 and 3 contained more HPV-negative samples, again in agreement with the results of the microarray. The results for these groups are representative of the numbers of positive samples found in practice with these sample types.

Overall, the outcomes of the two different methods do match: 93 % agreement is found between the two methods, with a Cohen's kappa coefficient (κ) of 0.88 (95 % CI, 0.81 to 0.94), showing a very high level of agreement (table 4.3). In each group, the results of the two measurement methods were not fully in line with each other in a number of cases. This number was the highest in group 1, which can be explained by the high number of multiple infections in this group: in almost half of these discordant cases, part of the diagnosis was still correct, but not all the HPV subtypes were detected by both methods.



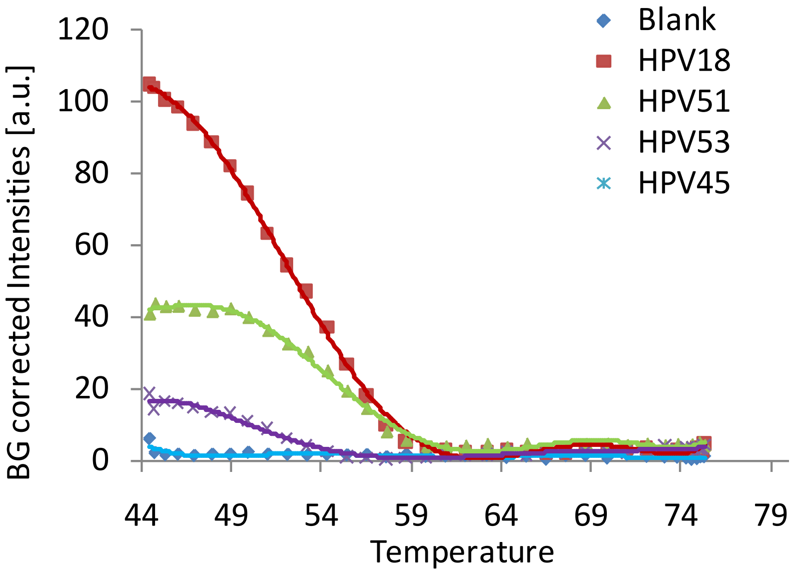


Figure 4.4 Kinetic hybridization curve of amplified DNA extracted from a clinical sample with a triple HPV infection (HPV 18, 51, and 53) (upper graph), and the corresponding melting curve analysis (lower graph). HPV 45, which is the HPV type with the highest degree of homology with HPV 18, 51, and 53 (90, 80, and 75 %, respectively), is added to show the high selectivity, as it clearly does not give any hybridization signal or melting curve. BG, background; a.u., arbitrary units.

Table 4.2 Comparison of the three different groups of sample types and overall comparison for the two post-PCR detection methods, reverse line blot and microarray flow-through.

Group	Number of samples with result in[a]			
	RLB	Microarray		
		Neg	LR	HR
1	Neg	6	0	0
	LR	1	5	0
	HR	4	1	43
2	Neg	36	1	0
	LR	0	2	0
	HR	1	0	20
3	Neg	19	0	2
	LR	0	4	0
	HR	2	0	33
Overall	Neg	61	1	2
	LR	1	11	0
	HR	7	1	96

[a] *Neg: negative; LR: low-risk HPV type; HR: high-risk HPV type*

Table 4.3 Statistical analysis of the comparison of the two post-PCR detection methods.

	Overall	Group 1	Group 2	Group 3
% agreement	93	90	97	93
Kappa factor (95 % CI)	0.88 (0.81 – 0.94)	0.75 (0.64 – 0.86)	0.93 (0.88 – 0.99)	0.87 (0.81 – 0.94)

As all data on melting temperatures were available for every microarray analysis, it is possible to extract a melting temperature for each hybridization. We investigated whether a potential relationship exists between the concentration of a specific PCR product present in the sample and the associated melting temperature. We did not find such a correlation within the range of concentrations as obtained after the amplification step. Next, we compared the melting temperature with the average melting temperatures of the same HPV type for each sample and investigated the samples where there was a difference larger than 2 °C. We identified 2 cases (10 %) where cross-hybridizations occurred on the flow-through microarray, and these cases could thus be identified correctly. As no melting temperatures can be obtained from RLB analyses, it is not possible to identify whether there are any cross-hybridizations on the RLB leading to any discrepancies.

Other reasons for discrepancies can be explained by differences in detection method and threshold value. In the flow-through method, the cutoff was set to three times the standard deviation. In three (15 %) of the discordant cases, the flow-through method gave signals just above the cutoff value. Overall, in 12 of the 180 cases, the differences between the microarray and the RLB readings could lead to a difference in diagnosis.

A further comparison with the EIA method was made with the 12 samples that gave contradictory results in the microarray flow-through method and the RLB. The raw data for each sample for the three different methods are presented in table 4.4. In 7 samples, the RLB method was in full agreement with the EIA. In 4 of these 7 cases, the signals obtained in the EIA were just above the cutoff value. In the remaining 3 samples, a high intensity was obtained in the EIA, whereas the microarray did not detect the infection. In 3 of the samples, the RLB gave a negative result that was in contrast to the results for the flow-through method and the EIA, where a high-intensity HPV infection was identified. Also, in 2 samples, the EIA found additional low-risk HPV infections with a weak intensity.

Table 4.4 Raw data of discrepancies between the results of the RLB and those of the microarray flow-through[a]

		Microarray		**RLB**		**EIA**[b]		
Group	**Sample no.**	**Result**	**HPV type found**	**Result**	**HPV type(s) found**	**Result**	**Intensity HR**	**Intensity LR**
1	7	Neg	-	HR/LR	16, 58	HR/LR	0.943	1.631
	8	Neg	-	LR	6	LR	-	2.773
	12	Neg	-	HR	16, 18	HR	0.342	-
	21	Neg	-	HR	18	HR/LR	9.999	0.354
	58	Neg	-	HR	45	HR	9.999	-
	70	LR	11	HR/LR	11, 16	HR/LR	0.429	9.999
2	131	LR	1	Neg	-	LR	-	9.999
	135	Neg	-	HR	82	HR/LR	9.999	0.401
3	148	HR	52	Neg	-	HR	1.300	-
	158	Neg	-	HR	33	HR	0.281	-
	165	HR	16	Neg	-	HR	9.999	-
	209	Neg	-	HR	31	HR	0.367	-

[a] For the two genotyping methods, the specific HPV type(s) found is specified. In cases where a high-risk and a low-risk infection were found (HR/LR), the result was classified as high risk. Neg, negative; HR, high-risk HPV type; LR, low-risk HPV type. [b] For the EIA, the fluorescence intensities in the HR and the LR cocktails are given in arbitrary units. The value 9.999 means that the spots were at saturation.

In order to compare the reverse line blot and our new flow-through principle, it was investigated which method was more efficient in detecting an HPV infection in the same sample set. In table 4.5, each HPV infection is classified as being detected either by both methods or by only one. Most HPV types were detected by both methods. There are no indications that one detection method is more sensitive than the other for the detection of a specific HPV type, and failed detections are randomly distributed rather than systematically found for specific HPV types.

Table 4.5 Detection of HPV infections in the different assays. In total, 159 HPV infections were detected in 216 clinical samples. GP5$^+$/GP6$^+$ PCR was followed by both reverse line blot detection and flow-through microarray detection.

HPV type	No. of HPV infections detected by:			Total no. detected (n = 159
	Both methods (n = 129)	RLB only (n = 17)	Microarray only (n = 13	
1			1	1
6	6	1		7
11	3			3
16	27	3	1	31
18	13	1		14
26	2			2
31	13	1	2	16
33	9	1		10
35	4	1	1	6
39	2	1		3
45	3	1		4
51	10		1	11
52	5		2	7
53	6			6
56	7	2	2	11
58	5	1	1	7
59	1			1
66	7			7
68	1			1
71	1			1
73	6	1	1	8
82	1	1		2

4.4 Discussion

In this study, we have investigated the added value of measuring the binding kinetics and melting curves of PCR products to and from capture probes immobilized on a microarray. As a model assay, we have used genotyping of the human papillomavirus, of which over 100 different types are known. Molecular epidemiology clearly indicates that infection with certain human papillomavirus types is the main factor involved in the development of cervical cancer, which is the second most common cancer in women worldwide [3,8,22]. The detection of HPV is hampered by a high number of homologous targets, which are known to cross-hybridize on a microarray. The method described in this paper combines, on the one hand, the advantages of microarrays and, on the other hand, the high-resolution melting of real-time PCR[13,14] as originally developed by Idaho Technology, Inc. (Salt Lake City, UT), and now commonly applied in an increasing number of real-time PCR machines.

Our method enables real-time fluorescence detection on a microarray during hybridization and melting. It allows for increased multiplex detection compared to the number of targets possible in real-time PCR, as the detection of the different targets is spatially separated on a solid surface. Our method is especially suitable for highly homologous targets, like HPV, or single nucleotide polymorphism (SNP) detection, since a high-resolution melting curve is part of the detection protocol. First, we have optimized the sequences of the hybridization probes in order to take advantage of the position-of-label effect[17]. This led to considerably increased hybridization signals, and we were thereby able to obtain sensitivities comparable to the analytical sensitivity of the reverse line blot (100 pM) without the need for additional signal amplification methods, such as incubation with antibiotin conjugates and enhanced chemiluminescence detection, as used in the RLB.

The kinetic curves, including the high-resolution melting curves, give additional understanding of the hybridization process. First, they provide more information on specificity than an endpoint hybridization measurement, which is standard in the tests available on the market today (e.g., INNO-LiPA HPV test or PapiloCheckTest and the Linear Array HPV test). References 10 and 18 stress the problem of cross-hybridization, which can lead to undesired false-positive results. It has been shown that cross-hybridizations can have a large influence on the hybridization kinetics, as well as the expected spot intensity levels. Furthermore, the ratio between perfect matching spots and mismatching spots changes over incubation time, both with hybridization and with subsequent washing steps.

The kinetic measurements of the flow-through method enable better discrimination between perfect match and mismatch hybridizations. In addition to this, dynamic curves can also be used for process optimization in terms of binding kinetics and minimization of potential cross-hybridization. In this study, we have presented the analysis of 180 clinical samples with unknown HPV infections; we were able to identify a number of false-positive results solely on the basis of the melting curves. The outcomes of these measurements were compared with the results of the reverse line blot. Overall, there was an agreement of 93 %. The lowest numbers were found in the sample group which contained a high number of multiple infections and known difficult cases. This group, group 1, contained samples that were already seen to be difficult to diagnose, as the EIA and the RLB gave contradictory results. When only the samples that were more representative in terms of what is normally found in the field (group 2 and 3) were taken into account, an overall sensitivity and specificity of 95 % were found. A Cohen’s kappa coefficient of 0.88 was found (95 % CI, 0.81 to 0.94), indicating a very high level of agreement between the two detection methods.

A number of discrepancies between the two methods can be explained by signals just below or above the cutoff values. For the microarray flow-through analysis as presented here, we have used an analytical threshold value based on three times the standard deviation of the noise levels. This should be further optimized, also taking into account the clinical relevance. The flow-through method has the additional possibility of using a dynamic threshold value, depending on the specific application, which might exclude clinically non-relevant false-positive results.

We have compared all the results of the discordant samples with the results obtained by a third measurement method, which was the EIA. In 25 % of these discordant samples, the flow-through method obtained results similar to those of the EIA, whereas in 58 % of the discordant samples, the EIA confirmed the results of the RLB. In the remaining 2 samples, the EIA found additional low-risk types next to a high-risk type, as was also found in the RLB. From the analysis given in table 4.4, it seems that the flow-through method is slightly less sensitive than the RLB, as negative results were found more often. It should be mentioned, however, that in a number of cases (4 out of 7) where the microarray did not detect an infection, the EIA intensities were very low (slightly above the cutoff value of 0.2), which might indicate a low viral load in these samples. In addition to this, in a number of the cases of discrepancies, multiple infections were present. It is not possible to use the same PCR products for the microarray method as are used with the RLB or EIA, as the RLB needs a biotin labeled primer and the flow-through method a Cy5-labeled primer. Therefore, a separate PCR was needed for the microarray-based assay. The mutation-tolerant GP5$^+$/GP6$^+$ PCR generally gives low concentrations of PCR products because this PCR is optimized to amplify all HPV types in a single run, thereby sacrificing endpoint concentration. As the GP5$^+$/GP6$^+$ PCR is complex, especially when multiple virus types are present, the concentrations of different HPV types at the end of the PCR can vary. A low viral load or the presence of multiple infections combined with a different PCR can lead to slightly different PCR product concentrations, which can lead to detection signals just below or above the cutoff values. This might explain the negative outcomes of the microarray flow-through method where the EIA/RLB results were positive.

For the RLB, the same PCR products were used as for the EIA, and therefore, discrepancies between the results for the RLB and EIA cannot be explained by this reasoning. The microarray method can be further optimized by an investigation of the clinical sensitivity. Furthermore, when comparing this method to the RLB or the EIA, especially in the case of low viral loads, the effect of using different PCRs, as well as differently labeled primers, should be further investigated. When the results for each genotype are assessed individually, as is done in table 4.4, there are only 4 cases out of 159 where an HPV infection was detected by the RLB while the microarray-based method failed to detect it. We are therefore tempted to conclude that the sensitivities of the RLB and the microarray flow-through method are comparable.

An advantage of the method presented in this study is that through miniaturization and the use of fluorescence instead of chemiluminescence, the amount of chemicals needed is much lower. No toxic chemicals are used, unlike other detection methods. There is no need for signal amplification steps, which can cause additional variations. All post-PCR steps in the process are integrated into one final automatic hybridization and detection step which lasts just over 1.5 hours. The real-time binding and unbinding kinetics can lead to significantly increased understanding and improvement of the assay, as kinetic curves are measured instead of hybridization endpoints. The system can be further improved by a full implementation of the melting curve analysis, thereby automatically identifying false-positive results caused by cross-hybridization.

4.5 Conclusions

We conclude that we have shown the added value of the kinetic hybridization measurements and high-resolution melting for genotyping applications. This principle combines the advantages of real-time PCR (multiple measurements instead of an endpoint measurement, and in addition, a high-resolution melting curve) with the high, multiplex detection capabilities of a microarray. As high-resolution melting analysis is becoming increasingly relevant and state of the art in real-time PCR, we would like to stress the added value of kinetic hybridization measurements and melting curves in microarray analysis. Combining the potential of this principle and the good initial results with clinical samples, we believe that this valuable principle enables the improvement of clinical PCR-based diagnostics.

References

1. Malorny B, PaccassoniE, Fach P, Bunge C, Martin A, Helmuth R. Diagnostic real-time PCR for detection of Salmonella in food. *Appl Environ Microbiol* 2004 70:7046–7052
2. Jacobs MV, Snijders PJF, van den Brule AJC, Helmerhorst TJM, Meijer CJLM, Walboomers JMM. A general GP5$^+$/GP6$^+$-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *J Clin Microbiol* 1997, 35:791–795
3. Muñoz NM, Bosch FX, de Sanjosé S, Herrero R, Castellsagué X, Shah KV, Snijders PJF, Meijer CJLM. Epidemiologic classification of human papillomavirus types associated with cervical cancer. *New Engl J Med* 2003, 348:518–527
4. Van den Brule AJC, Pol R, Fransen-Daalmeijer N, Schouls LM, Meijer CJLM, Snijders PJF. GP5+/6+ PCR followed by reverse line blot analysis enables rapid and high through-put identification of human papillomavirus genotypes. *J Clin Microbiol* 2002, 40:779–787
5. De Roda Husman A-M, Walboomers JMM, van den Brule AJC, Meijer CLJM, Snijders PJF. Processing of long-stored archival cervical smears for human papillomavirus detection by the polymerase chain reaction. *Brit J Cancer* 1994, 72:412–417
6. Manos MM, Ting Y, Wright D, Lewis AJ, Broker TR, Wolinsky SM. Use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. *Cancer Cells* 1989, 7:209–214
7. Tieben LM, Ter Schegget J, Minnaar RP, Bouwes Bavinck JN, Berkhout RJM, Vermeer BJ, Jebbink MF, Smits HL. Detection of cutaneous and genital HPV types in clinical samples by PCR using consensus primers. *J Virol Methods* 1993, 42:265-280
8. Kjaer S, Høgdall E, Frederiksen K, Munk C, van den Brule AJC, Svare E, Meijer E, Lorincz A, Iftner T. The absolute risk of cervical abnormalities in high-risk human papillomavirus-positive, cytologically normal women over a 10-year period. *Res Cancer* 2006, 66:10630–10636
9. Molijn A, Kleter B, Quint W, van Doorn L-J. Molecular diagnosis of human papilloma virus (HPV) infections. *J Clin Virol* 2005 32S:S43–S51
10. Schmitt M, Bravo IG, Snijders PJF, Gissmann L, Pawlita M, Waterboer T. Bead-based multiplex genotyping of human papillomaviruses. *J Clin Microbiol* 2006, 44:504–512
11. Jones J, Powell NG, Tristram A, Fiander AN, Hibbitts S. Comparison of the PapilloCheck DNA micro-array human papillomavirus detection assay with Hybrid Capture II and PCR-enzyme immunoassay using the GP5/6$^+$ primer set. *J Clin Virol* 2009, 45:100–104
12. Poljak M, Marin IJ, Seme K, Vince A. Hybrid Capture II HPV test detects at least 15 human papillomavirus genotypes not included in its current high-risk probe cocktail. *J Clin Virol* 2002, 25:S89–S97
13. Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, and Pryor RJ. High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin Chem* 2003, 49:853–860

14. Zhou L, Wang L, Palais R, Pryor R, Wittwer CT. High-resolution melting curve analysis for simultaneous mutation scanning and genotyping in solution. *Clin Chem* 2005, 51:1770–1777
15. Wu Y, de Kievit P, Vahlkamp L, Pijnenburg D, Smit M, Dankers M, Melchers D, Stax M, Boender PJ, Ingham C, Bastiaensen N, de Wijn R, van Alewijk D, van Damme H, Raap AK, Chan AB, van Beuningen R. *Nucleic Acid Res* 2004 32(15):e123
16. Pierik A, Dijksman JF, Raaijmakers ATA, Wismans AJJ, Stapert HR. Quality control of inkjet technology for DNA microarray manufacturing. *Biotech J* 2008, 3:1581–1590
17. Zhang L, Hurek T, Reinhold-Hurek B. Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridizations. *Nucleic Acids Res* 2007, 33:e166
18. Bishop J, Wilson C, Chagovetz AM, Blair S. Competitive displacement of DNA during surface hybridization. *Biophys J* 2007 92:L10–L12
19. Zhang Y, Hammer DA, Graves DJ. Competitive hybridization kinetics reveals unexpected behavior patterns. *Biophys J* 2005, 89:2950–2959
20. Mocanu D, Kolesnychenko A, Aarts S, Troost-deJong, A, Pierik A, Vossenaar E, Stapert H. Mass transfer effects on DNA hybridization in a flow-through microarray. *J Biotechnol* 2009, 139:179–185
21. Mocanu D, Kolesnychenko A, Aarts S, Troost-deJong, A, Pierik A, Coene W, Vossenaar E, Stapert H. Quantitative analysis of DNA hybridization in a flow-through microarray for molecular testing. *Anal Biochem* 2008, 380:84–90
22. Walboomers JMM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJF, Peto J, Meijer CJLM, Muñoz N. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999, 189:12–19

Chapter 5

Real time amplification detection on a microarray

- Real time array PCR -

This chapter is based on:

Pierik A, Boamfa MI, van Zelst M, Clout DEW, Stapert HR, Dijksman JF, Broer DJ, Wimberger-Friedl, R, *in review*

Abstract

Real time quantitative Polymerase Chain Reaction (qPCR) is widely used as a research and diagnostic tool for amplification and detection of specific DNA sequences. Notwithstanding its many powerful features, the method is limited in the degree of multiplexing to about 6 due to spectral overlap of the available fluorophores. When a higher degree of multiplexing is needed, microarray-based methods are commonly used. However, in clinical practice transferring post-PCR liquids to a microarray is not preferred due to contamination, hands-on time and potential human errors. The increasing availability of genetic information (e.g. genome sequence data of micro-organisms of infectious diseases, or Single Nucleotide Polymorphism (SNP) detection in the human genome contributing disease development) also results in a trend towards increasing quantitative multiplex detection. Hence there is a need for integrated high-multiplex qPCR systems. A method is presented that allows quantitative, multiplexed amplification by combining volume PCR amplification and real-time surface-based monitoring by confocal scanning, based on microarray hybridizations. This method does not require any manipulation of PCR product and runs in one single and closed container. A compact confocal scanner was employed, based on miniaturized optics from DVD technology and combined with a flat thermocycler for simultaneous scanning and heating. Feasibility of this method has been proven in a single-plex assay, and quantification of the target input concentration has been shown. The method is especially promising for the quantitative detection of multiple homologous sequences, e.g. viral (such as HCV, Human Papilloma Virus subtyping, or multiple Single Nucleotide Polymophism (SNP) detection.

5.1 Introduction real time array PCR

PCR is a powerful method to increase the number of target DNA molecules that are initially present in very low concentrations. In real time quantitative PCR (qPCR), as explained in chapter 1, amplification curves are detected. Since a complete amplification curve is obtained, the cycle number at which the signal becomes above the background can be used to quantify the target input concentration. Multiplexing is achieved by parallel detection of different probes each labeled with a different fluorophore. Typically up to 6 different dyes and thus targets can be discriminated. Further multiplexing is not possible because of the spectral overlap of the fluorophores[1].

For the detection of more DNA sequences an alternative technology needs to be chosen. Different ways of detecting multiple DNA targets are given in table 5.1, scored on a number of aspects: the capability for quantification (defined as the capability to measure the initial number of target molecules present in a sample), the multiplexing capability, and sensitivity. Additionally also the biomolecular control is included in this table, which describes the degree of complexity regarding the biochemical processes that take place on a molecular scale. A more complex system (indicated with a '-' in the table) may possibly lead to less control over the biochemical processes taking place.

Splitting the sample into multiple reactions wells and performing different PCR reactions in each well is a straightforward option to increase multiplexing. The disadvantage, however, is a reduced sensitivity since the sample needs to be diluted, split and divided over multiple wells. The lower limit of detection increases proportionally to the number of wells.

The biomolecular complexity is considered to be higher for the higher multiplex techniques since more potential reactions can run in parallel in a single reaction volume, and thereby the complexity is intrinsically higher.

Hybridization of PCR products onto a microarray is one of the most applied methods for high multiplex detection since it allows for the parallel detection of a lot of different targets in one single run. Specificity and multiplexing are obtained by spatial resolution on the microarray surface rather than the use of different fluorescently labeled probes, like in real time PCR. The disadvantage is that the target molecules have been amplified using for example PCR in order to obtain reasonable concentrations, and this step is often performed until saturation is reached. Stopping the PCR reaction in the logarithmic phase followed by microarray hybridization is an option to prevent saturation and loss of quantitative information[2]. For samples containing an unknown number of targets, it is difficult to determine the cycle number during which to stop the PCR reaction such that a reasonable concentration of target molecules is achieved without reaching saturation. The dynamic range of this method is limited to typically 2-3 orders of magnitude in concentration. Such a dynamic range does not satisfy the requirements of many high-performance applications[3]. Furthermore, for this method, a number of intermediate steps need to be performed in between amplification and hybridization, which results in higher contamination risks which is especially for clinical applications an important factor to take into account.

Table 5.1 Characteristics of different methodologies for multiplex DNA detection.

	qPCR	**Parallel qPCRs**	**Post-PCR hybridization**	**Real time array PCR**
Quantification	+	+	-	+
Multiplexing	-	+	+	+
Sensitivity	+	-	+	+
Biomolecular control	+	+	-	-

The method referred to as real time array PCR[4] combines bulk amplification by PCR with surface specific detection on a DNA microarray, constructively building on the advantages of both powerful and well-established techniques. The integration of different DNA amplification methods with DNA detection has previously been reported in literature, though in many cases the processes were still performed consecutively[5-9], with the disadvantage that although contamination risks are reduced, quantitative information is not obtained. Basically it is still about end-point PCR followed by hybridization. With the method as investigated in this chapter, instead of measuring after completion of the amplification steps only, the amplified target molecules are detected by microarray hybridization during each individual PCR cycle[4]**.** In this way, the PCR is monitored on the microarray and thereby quantification of target input concentration is possible. This requires a number of technical issues that need to be addressed. When surface hybridization is measured during PCR, the solution containing the labeled PCR products, PCR ingredients as well as labeled primers are in contact with the microarray, generating a significant background signal. Furthermore, hybridization and detection thereof should be done rapidly, ideally within one or a few minutes in order to prevent long analysis times and possible degradation of reagents. One of the main challenges of real time array PCR is the rapid measurement of weak hybridization signals on the microarray surface in high background fluorescence: maximization of the signal to background ratio is needed. Furthermore, since PCR and hybridization are executed in parallel in the same volume, both processes automatically experience the same conditions. The individual processes, however, do not have the same optimal circumstances for best performance: hybridization is favored in high-salt concentration, whereas PCR requires low-salt conditions. One additional challenge thus lies in finding optimal matrix conditions where both processes can proceed under the same circumstances.

One way to reduce background signals is by using advanced detection technologies. Surface specific detection methods enabling a substantial background reduction, like evanescent wave methods based on total internal reflection, are known[4]. These methods allow for the detection of hybridization signals throughout each PCR cycle. Next to using advanced technologies that make use of background suppression, the issue can also be addressed by designing a method that does not create high background to begin with. In reference 10 a real time multiplex PCR system has been shown that makes use of immobilized primers onto hydrogel pads. Subjecting the system to amplification conditions results in elongation of the immobilized primers. Detection of hybridized targets to the elongated primers is done by measuring the fluorescence of an intercalating dye (SYBR green) by a (CCD) camera. Background is initially low since it scales with the amplicon concentration. A different approach of real time array PCR detection is disclosed in reference 11. Here, the microarray itself is rotated out of the solution after each PCR cycle and the target molecules are incubated onto the microarray. After washing to remove background fluorescence, the array can be imaged directly without the need for additional background reduction.

However, this requires a complicated rotating mechanism, extra handling and is prone to contamination.

Another aspect to consider is the amplification method itself, for which different possibilities exist. Probes can be used to capture the labeled targets that are elongated in the solution on the microarray[4], as depicted in figure 5.1, or primers can be elongated and labeled on the surface [8,10,12], see figure 5.2. These figures schematically represent situations in the bulk and at the microarray surface during the real time array PCR process, at the end of the denaturation (left), annealing/hybridization (middle) and elongation (right).

In figure 5.1, the amplification reaction is driven by primers in the bulk. The use of immobilized capture probes follows the standard way of microarray hybridization and detection which means amplification in the bulk and target hybridization on the microarray surface. In this case, either the reverse or the forward primers present in the bulk carries a fluorescent label needed for detection.

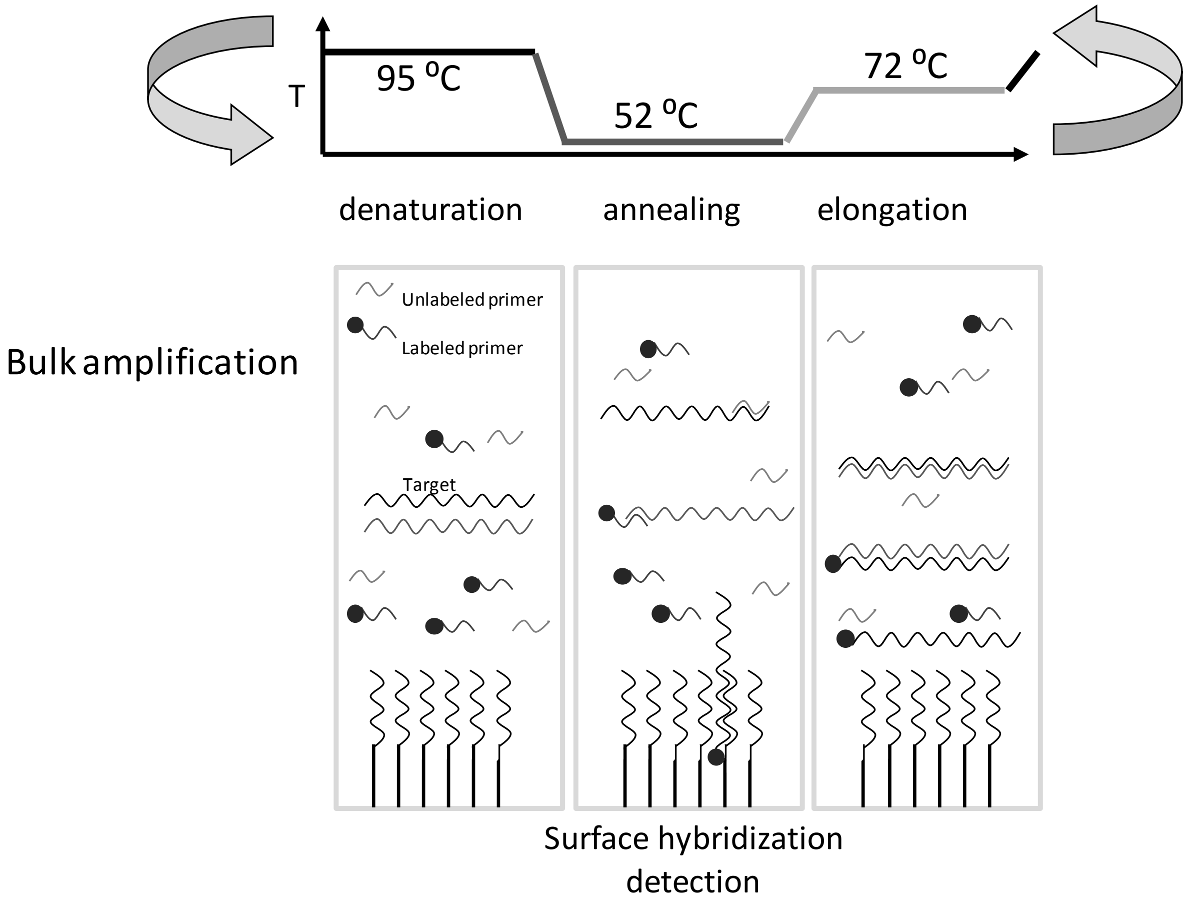


Figure 5.1 Schematic view of real time array PCR process with capture probes on the surface and amplification performed in the bulk.

The formed labeled amplicons can diffuse towards the microarray surface and hybridize to its complementary capture probe during the annealing/hybridization step. During amplification, the target concentration increases. As a result, in the successive PCR cycle the fraction of occupied capture probes during hybridization will therefore be higher. By using a surface-specific detection technology the targets captured by the probes can be detected throughout the amplification reaction. By repeating this procedure during a number of consecutive cycles, a PCR curve can be obtained by plotting the surface fluorescence as a function of cycle number. After finishing the amplification, an (optional) additional hybridization step at a prolonged time at the hybridization temperature allows for a more accurate hybridization measurement. Additionally, a high resolution melting curve can be applied to further address the specificity of the formed hybrids[13].

Figure 5.2 represents the case where one of the two primers is immobilized on the surface, whereas the second primer is present in the bulk. After denaturation, all double stranded DNA becomes single stranded. During the following annealing step, one strand of the target molecule hybridizes to the primers on the surface, while the other strand of the target molecule anneals to a primer in the bulk. Upon increasing the temperature, elongation of the primers onto which target molecules are hybridized occurs on the surface. In parallel the elongation of the second primer hybridized with a target strand happens in the bulk. After denaturation, all double stranded DNA denatures, the targets go back into solution, and the elongated primers remain immobilized on the microarray surface. Since target strands are amplified in the bulk, more targets become available for surface hybridization for the next cycles.

Detection during the amplification by the use of a surface specific detection technology can be done either by the use of an intercalating dye, the incorporation of labeled nucleotides (as included in figure 5.2), or hybridization with a complementary labeled primer or probe. In case labeled nucleotides are used, fluorescent signals on the microarray surface are built up permanently since the elongated primers are attached to the microarray surface. The advantage of the method as presented in figure 5.2 is that in case of a multiplex PCR, each individual PCR reaction is done only at the spot where the specific primer is immobilized, and therefore the reactions are spatially separated on the surface. There are less undesired side-reactions between primers and amplicons, and lower chances on formation of primer-dimers in the bulk. However, in practice, compared to traditional bulk PCR the amplification efficiency is often much lower[14] and therefore the immobilized primers are usually also added in solution in order to enhance the amplification reaction. The use of primers that are elongated provides fewer options for obtaining reasonable melting curves after amplification. This method thus has a drawback in addressing the specificity of the target-capture probe hybrid (see chapter 4), and additionally the degree of specificity is lower compared to the case as visualized in figure 5.1. An additional problem might be the fact that the enzyme and the nucleotides that are needed for elongation of the immobilized primers need to diffuse towards the surface. An advantage of this method is the possibility to build up a permanent signal on the microarray surface when incorporating labeled nucleotides in the elongated primer during elongation. This means that hybridization signals can be higher since the signals are cumulative. Additionally, hybridization measurements can be de-coupled from the bulk amplification process, providing possibilities to scan during the elongation or denaturation phase or even continuously during amplification. When using capture probes on the microarray surface, as indicated in figure 5.1, hybridization signals need to be completely built up from the start during each PCR cycle, which may result in lower signals (or alternatively longer hybridization times needed).

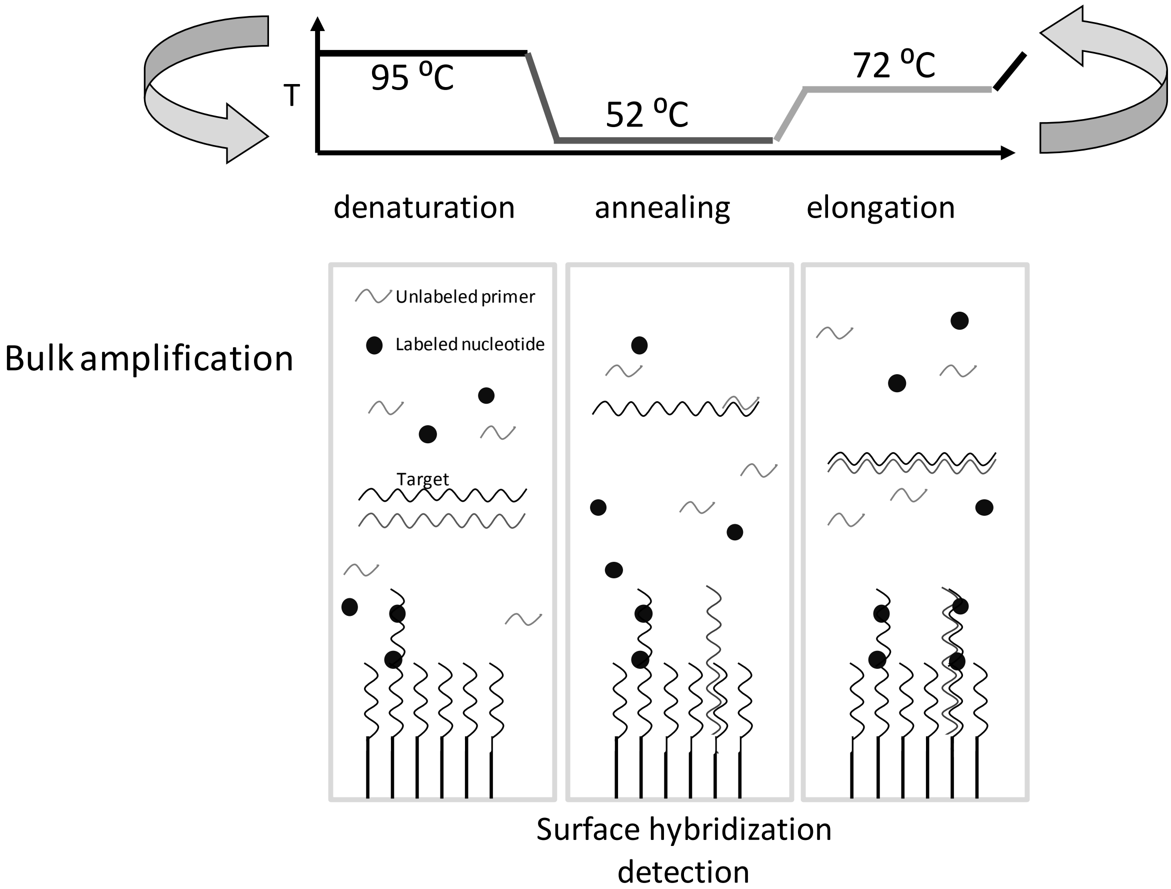


Figure 5.2 Schematic view of real time array PCR process with primers on the surface that are elongated during the amplification reaction. In this schematic view, already one primer that is elongated during a preceding cycle is visualized in the left part of this figure. In this specific example, labeled nucleotides are incorporated in the immobilized primer during elongation.

As much as possible on proven and robust technologies has been built upon in the approach as presented in this chapter. Two specific primers in solution, one of them labeled, and capture probes immobilized on the microarray are used to serve an optimal degree of specificity and freedom in probe design. As a consequence, strong background suppression is needed to deal with the background signals generated by the labeled amplicons and primers in solution. Confocal scanning is a well-known microarray detection technique. The confocal detection eliminates contributions from auto-fluorescence of the optical components in the light path by selecting the fluorescence signal coming from the focal area only. Thereby, also the contribution from fluorescence of the supernatant solution is suppressed. At the same time, fluorescence coming from the spots is enhanced by the high intensity of a focused beam that is used for excitation of the fluorophores, as well as the high collection efficiency of an objective lens. Using an epi-fluorescence arrangement where excitation and detection is done from the same side of the cartridge, a heater can be integrated at the opposite side of the scanning optics.

The real time array PCR instrument was designed and built leveraging on Philips' know-how in optical storage technology. This resulted in a compact prototype onto which additionally a thermocycler was integrated. In figure 5.3a-d, the cartridge, the integrated instrument, the thermocycler, and the confocal concept are depicted, respectively.

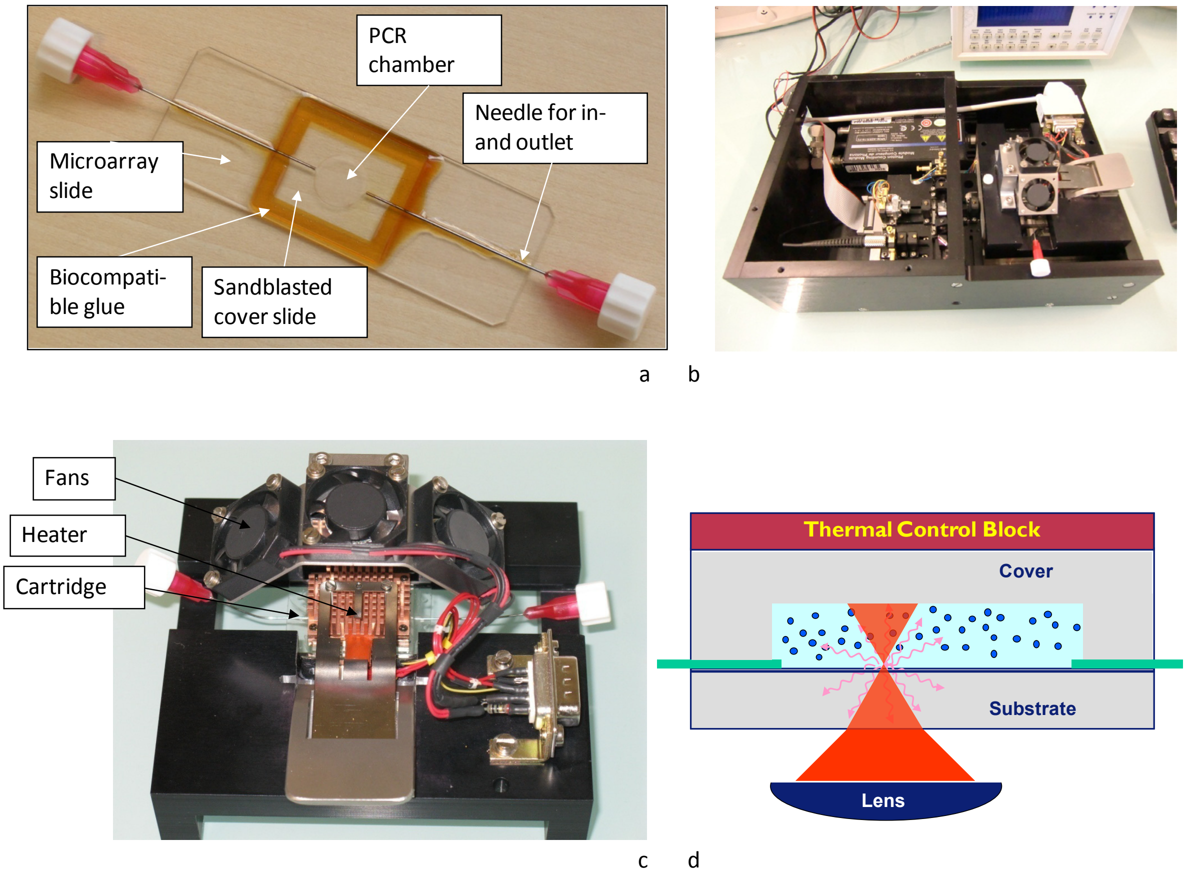

PCR chamber
Needle for in- and outlet
Microarray slide
Biocompatible glue
Sandblasted cover slide
a
b
Fans
Heater
Cartridge
Thermal Control Block
Cover
Substrate
Lens
c
d

Figure 5.3 Cartridge (a) and instrument, (b) prototype complete integrated fluorescent scanner and (c) thermocycler for the real time array PCR. Figure d shows schematically the principle of confocal scanning.

In this chapter, proof of concept of monitoring real time DNA amplification on a microarray by using confocal scanning technology is shown. qPCR is included as a benchmark. Dilution series of target concentrations are included to study if quantification of target input concentrations is possible. These data are also used for calculation of the amplification efficiency. Routes for further improvement are investigated, based on the optimization of the temperature protocol and the PCR mix. Additionally, the added value of an asymmetric PCR instead of the use of equimolar primer concentrations is investigated.

5.2 Materials and methods

5.2.1 Microarray fabrication

All capture probes (Biolegio, Nijmegen, The Netherlands) were dissolved to concentrations of 3 µM in a 75 mM sodium phosphate buffer of pH 8.5 (all ingredients from Sigma-Aldrich, St-Louis, MO, USA). Superamine2 microarray glass slides (ArrayIt Corporation, Sunnyvale CA, USA) were used. Each capture probe solution was deposited by an inkjet printer[15] in four-fold onto the substrates using 1 nanoliter spot volume.

All capture probes were 5’-end modified with 16 thymines for optimal immobilization during the subsequent photoactivated crosslinking step[16] by exposure to 300 mJ/cm^2 of 254 nm UV light. Additionally, a spacer sequence (5’-CACACACACACACACA-3’) was included between the point of attachment and the hybridization sequence to increase accessibility. UV-exposure was followed by a combined washing and blocking step for 1 hour at 42 ^{o}C in 5 x SSC, 0.1% SDS and 0.1 mg/ml herring sperm DNA (all ingredients from Sigma-Aldrich, St-Louis, MO, USA). Fluorescently labeled oligonucleotide spots with no sequence homology to the target templates were used as control spots.

The cartridge was manufactured by gluing a cover glass onto a microarray substrate with a sand-blasted reaction chamber in it, using a biocompatible two component epoxy (Epotek – 353ND from Epoxy Technologies). In the adhesive layer, two syringe dispense tips are integrated (5125-1.5-B type from EFD) for in- and outlets (prior to epoxy application the needles were fixed with Dymax 142-M, a biocompatible UV curable glue). In this way, a reaction chamber of 25 µl with means for inlet and outlet was created. All PCR ingredients were mixed externally and then injected into the PCR chamber via standard syringe connections. After filling, the needles were closed by pressing them together with clamps.

5.2.2 Real time array PCR system

The experimental device combines confocal fluorescence detection with thermal cycling in a compact device, see figure 5.3b and c for pictures of the scanner and figure 5.3d for the concept of confocal scanning. The scanner is built around a standard DVD pick-up unit (objective lens NA=0.65) with auto focus feedback loop, locked to the micro-array surface inside the chamber. The bi-aspheric lens is corrected for the substrate thickness. A diffraction-limited spot (HWFM 0.6 micron, 660 nm wavelength) is scanned over the microarray surface at typically 50 mm/s. An image is created by a meander-like movement driven by a xy-stage with motors from the tray drive unit of a DVD player. The forward resolution is determined by the integration time (min 1 µm), the lateral step size can be chosen. For the array-PCR, scans were built out of 100 lines, 5 mm long, taken at a lateral step of 25 µm, to cover the 5x4 mm area of the microarray, in about 40 seconds. The sample holder includes a thermal unit with clamps that hold the cartridge in place. The heating is achieved by a thin film heater and cooling is done by means of three fans from a PC motherboard. The temperature of the heater that is in direct contact with the cartridge surface is actively controlled (PID).

5.2.3 Real time array PCR protocol

Different PCR mixes have been used: initially experiments were done to show feasibility with a ready-to-use mastermix kit: all PCR ingredients were diluted into Geneamp Fast PCR Mastermix (Applied Biosystems, CA, USA). Since the exact content of this kit, especially the type and the amount of Taq polymerase, is not known, additional optimization experiments were performed using mixes consisting of the individual and known PCR ingredients (all from Applied Biosystems, CA, USA).

For all experiments 0.1 % Bovine Serum Albumin (BSA) was added for reducing non-specific binding of PCR ingredients to the walls of the cartridge (like the glass, needles etc). Different target sequences were used as model systems: the femA gene of methicillin-susceptible *Staphylococcus epidermidis* (femA SE) and the same gene of methicillin-resistant *Staphylococcus Aureus* (femA SA) as well as the mecA gene of methicillin-resistant *Staphylococcus Aureus (*mecA SA)[17]. These targets were added as plasmid DNA, except where stated otherwise. Primers were used in 300 nM concentrations. All reverse primers were 5'-end Cyanine-5 (Cy5) labeled. All sequences of the nucleotides included in the study are given in table 5.2.

Two different real time array PCR programs were run: a two-step and a three-step protocol. The latter program started with a 5-minute hot-start at 98^{o}C followed by 60 cycles of each 40 s at 72 ^{o}C, 20 s at 95 ^{o}C and 120 s at 52 ^{o}C. These times include heating (2 ^{o}C/s) and cooling (-1 ^{o}C/s) times. The actual heater settings were 3 ^{o}C higher to compensate for an offset between heater and PCR mixture. The microarray surface was scanned during the last 30 seconds of the annealing step. The two-step protocol consisted of an equal hot-start protocol and a cycling program of 40 s at 95 ^{o}C followed by 140 s at 60 ^{o}C. Scanning of the microarray is again done during the last 40 seconds of the annealing phase. In this protocol, the annealing and elongation steps are combined (both processes occur simultaneously) into a single step at 60^{o}C , which was possible with the design of the primers used. This protocol has the advantage of having more time available for hybridization without giving in on total assay time. The elongation speed is lower since the activity of the enzyme is optimal around 72 ^{o}C step.

The real time array PCR protocol can be followed by an extended hybridization during a predefined time, typically at lower temperatures to increase hybridization signals, e.g. 50 ^{o}C. Hybridization build-up at the microarray surface is measured in the same way as described above, the time in between two successive scans is 40 seconds. A typical result of a real time array PCR curve followed by hybridization is given in figure 5.4. Hybridization signals observed after finalizing the real time array PCR protocol are usually much higher since during the amplification process, less than 1 minute of hybridization time is available during each PCR cycle.

Based on this type of curves, a rough estimation of the kinetics can be made. For the real time array PCR process, time scales are short and the number of target molecules captured is thus small. This means that the concentration of target molecules in the bulk can be assumed not to be significantly influenced by the hybridization reaction. Furthermore, due to the low fraction of occupied capture probes, the dissociation reaction (target molecule renaturating from the capture probe back into the solution again) can be neglected compared to the hybridization reaction and also potential spot saturation does not play a significant role. For the initial cycles where the renaturation of double stranded DNA is not significant (at least beyond the C_t-value), it can be assumed that hybridization signals scale with the concentration of the target molecule. Comparable curve shapes like in qPCR therefore can be expected since also in qPCR the measured signals are proportional to the concentration of the target molecule. It has experimentally been verified that hybridization signals at the lower time scales correlate with the concentration of the target molecule (data not shown). A more detailed analysis of the reaction occurring in the bulk and at the microarray surface is given in chapter 6.

Capture probes of the original length as in reference 17 and increased length were designed based on the Taqman sequences as found in literature[17]. For the experiments performed in the Geneamp mastermix, the longer capture probes performed slightly better and therefore the hybridization data of this probe are used in the experimental results section. For the home-made mix, however, the shorter probes were used in the analyses since these showed slightly higher signals.

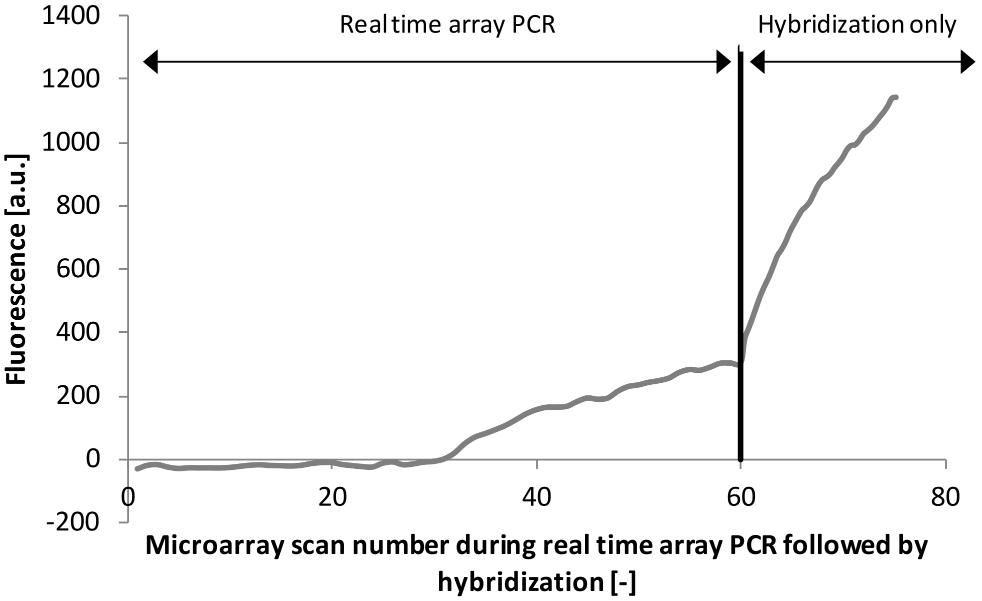


Figure 5.4 Typical result of a fluorescence scan of a microarray during the real time array PCR followed by hybridization. During the first 60 cycles, the real time amplification of $2.5.10^5$ copies of the femA SE target is monitored. A signal above background can be observed starting around cycle number 31. The real time array PCR process is followed by an extended period of time at 50 ^oC. Note that the timing before cycle number 60 is related to the PCR process (120 seconds annealing/hybridization, 40 seconds elongation and 20 seconds denaturation). The timing after cycle number 60 is different, each 40 seconds the fluorescence of the microarray is scanned while the microarray is kept at a constant temperature of 50 ^oC.

5.2.4 Asymmetric real time array PCR

Asymmetric real time array PCR experiments were performed using non-equimolar concentrations of both primers. The reverse, labeled, primer was added in 300 nM concentration, whereas the forward unlabeled primer in 50 nM concentration. The asymmetric real time array PCR experiments were all done using the Geneamp mastermix and the 3-step protocol.

5.2.5 Fluorescence measurements and calculations

For each hybridization measurement, the fluorescent intensities of two spots of the same capture probe sequence were averaged. Background correction was done by subtracting average background fluorescence in the vicinity of the spots. For the determination of the C_t-value, two different methods were available. For absolute comparison with qPCR curves, first a four-parameter Boltzmann curve fit, which is a common method for PCR curve fitting[18], was carried out to fit the real time array PCR curves. An absolute threshold of 25 kCounts/s for the fluorescence was then taken as the cut-off value for determination of the C_t-value. This method closely resembles the way of working in commercial qPCR instruments. For comparison with real time array PCR experiments, an alternative method was employed which takes into account the effect of scattering on the measurement data caused by the noise level on the measurements. C_t-values were calculated by determining the cycle number at which the average (background corrected) fluorescence intensity exceeded three times the average standard deviation on the background. The background is taken as the average fluorescence of cycle 1-15.

5.2.6 qPCR experiments

For reference experiments, real time PCR experiments were performed on a Biorad CFX96 system (Hercules, CA, USA). For detection, a Taqman probe was used, see table 5.2, with a sequence identical to the short capture probe. This probe also contained a tail of 16 thymines, which did not influence the amplification (data not shown). The real time PCR protocol was optimized for having the fairest benchmark, which was, after an initial 5-minute hotstart 60 cycles of 15 seconds at 95 ^oC followed by 60 seconds at 60 ^oC. For all experiments, the same MasterMix was used as in the real time array PCR experiments.

Table 5.2 Oligonucleotide sequences for the primers and probes used for real time array PCR

Oligonucleotide type	Oligonucleotide sequence
mecA SA forward primer	5’- CAT TGA TCG CAA CGT TCA ATT T -3’
mecA SA reverse primer	5’- TGG TCT TTC TGC ATT CCT GGA -3’
mecA SA capture probe	5’- AAG ATG GTA TGT GGA AGT TAG ATT GGG ATC ATA GCG TCA T -3’
mecA SA Taqman probe	5’- TGG AAG TTA GAT TGG GAT CAT AGC GTC AT -3’
femA SA forward primer	5’- TGC CTT TAC AGA TAG CAT GCC A -3’
femA SA reverse primer	5’- AGT AAG TAA GCA AGC TGC AAT GAC C-3’
femA SA capture probe (short)	5’- TCA TTT CAC GCA AAC TGT TGG CCA CTA TG-3’
femA SE forward primer	5’ - CAA CTC GAT GCA AAT CAG CAA -3’
femA SE reverse primer	5’ - GAA CCG CAT AGC TCC CTG C -3’
femA SE capture probe (short)	5’- TAC TAC GCT GGT GGA CTT CAA ATC GTT ATC G- 3’
femA SE capture probe (long)	5’- AAG TAG TTT ACT ACG CTG GTG GAA CTT CAA ATC GTT ATC G-3’
femA SE Taqman probe	5’- TAC TAC GCT GGT GGA CTT CAA ATC GTT ATC G- 3’

5.3 Results

5.3.1 Proof of concept

To investigate the amplification characteristics of the real time array PCR concept compared to conventional real time PCR, 2.5×10^5 copies of the femA gene of methicillin-susceptible *Staphylococcus epidermidis* (femA SE) were amplified, using genomic DNA, in GeneAmp mastermix using both conventional qPCR and real time array PCR (figure 5.5). The same comparison was done for the mecA gene of *Staphylococcus aureus (mecA SA)*, figure 5.6.

The dashed line represents the values obtained during bulk qPCR whereas the solid line is found by fitting the real time array PCR data data (four parameter Boltzman fit), gaining good quality fits (mecA SA: R^2 = 0.99, femA SE: R^2 = 0.98). The diamonds are the experimental values. For comparison, the fluorescent intensities are normalized to the maximum value of each measurement.

In both figures, the two PCR curves are almost similar, though the steepness of the real time array PCR curve is in both cases lower, which can be attributed to a reduced amplification efficiency in the bulk. The C_t-values of both methods are comparable.

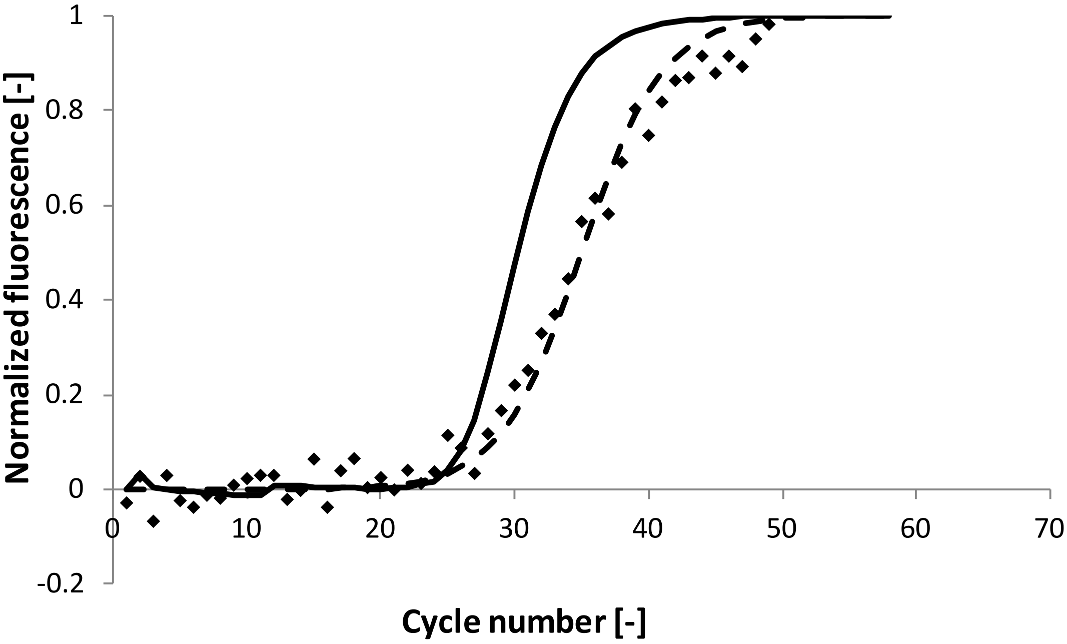


Figure 5.5 Amplification curves for 2.5.10^5 plasmid copies of the femA SE gene, as obtained with conventional qPCR (solid line) and real time array PCR (dashed line). The diamonds represent the experimental data, whereas the dashed line represents a four parameter Boltzmann curve fit used to guide the eye[1].

---

[1] The Boltzmann fit used is $y = A_2 + \frac{(A_1 - A_2)}{1 + \exp\frac{x - x_0}{dx}}$. In this equation, A_1 gives the ground asymptote, whereas A_2 is the maximum asymptote, x_0 is the inflection point and dx is related to the steepness of the curve.

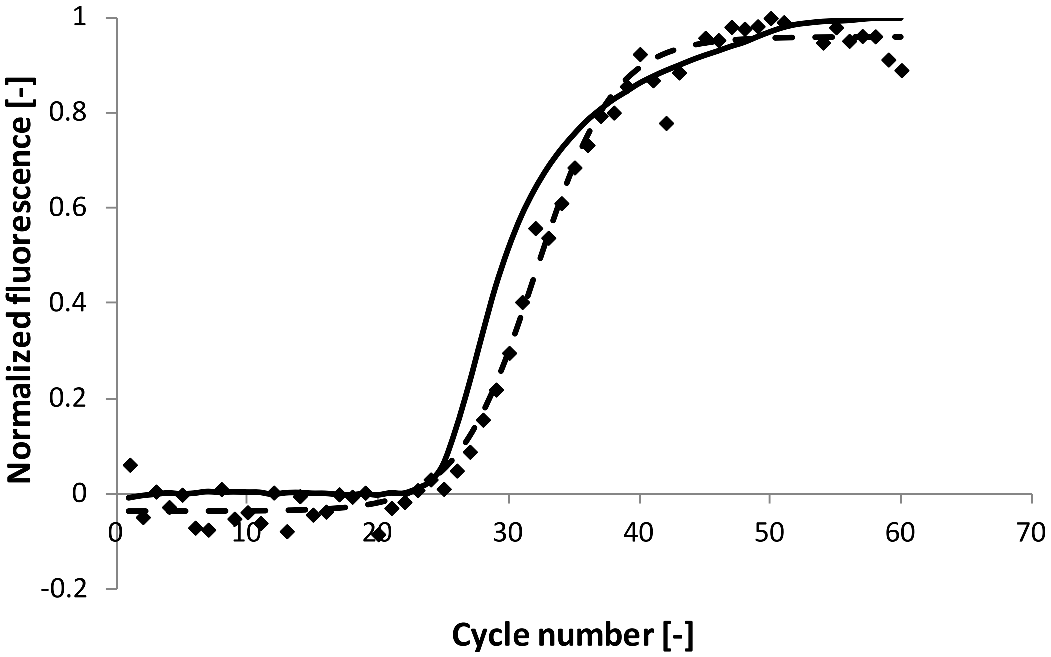


Figure 5.6 Amplification curves for 2.5.10^5 plasmid copies of the mecA SA gene, as obtained with conventional qPCR (solid line) and real time array PCR (dashed line). The diamonds represent the experimental data, whereas the dashed line represents a four parameter Boltzmann curve fit used to guide the eye. For details on this fit: see footnote of figure 5.5.

Next, a dilution series ranging from 2.5x10^5 to 25 copies per reaction of the methicillin-susceptible *Staphylococcus epidermidis* (MSSE) genome was performed. Figure 5.7 shows the C_t-values versus target input concentrations obtained with qPCR and real time array PCR. Indeed also for the real time array PCR data a clear relation between the target input concentration and the C_t-value is achieved, confirming that this method can be employed for quantitative analysis. Secondly, the lowest copy number could still be detected using both methods. A no template control (NTC) was negative as expected. From these data, a semi-log fit between input concentration and C_t-value was obtained, which was used to evaluate the array PCR efficiency. An efficiency of 97 % was found in qPCR, the array PCR efficiency appeared to be 58 %.

The absolute C_t-values at the higher target input concentrations do not differ much. The absolute C_t-value is determined by a number of factors including the amplification efficiency, quenching efficiency (for qPCR) and detection limit of the measurement instrument. Since part of these factors is related to the instrument used, it is therefore not evident that both methods should give equal absolute C_t-values. For the higher target input concentrations, a similar C_t-value was found for both qPCR and real time array PCR even with a reduced amplification efficiency of real time array PCR compared to qPCR. This means that intrinsically the detection sensitivity of real time array PCR is at least as good as of qPCR[19].

The efficiencies of the amplification reactions are calculated, based upon a limited set of data, using the derivative of the line when plotting the C_t-value versus the logarithmic target input concentration. In figure 5.7, it seems that the two lines of the qPCR and the real time array PCR cross at high input concentrations. However, this is not to be expected: more likely is that at high input concentrations, the lines will run in parallel (the offset between the two lines being determined by the detection properties of the two systems) and that for higher cycle numbers during the amplification, the efficiency of the real time array PCR will be lower (e.g. because of decreasing activity of the Taq polymerase). This becomes visible in a steeper line for the lower input concentrations.

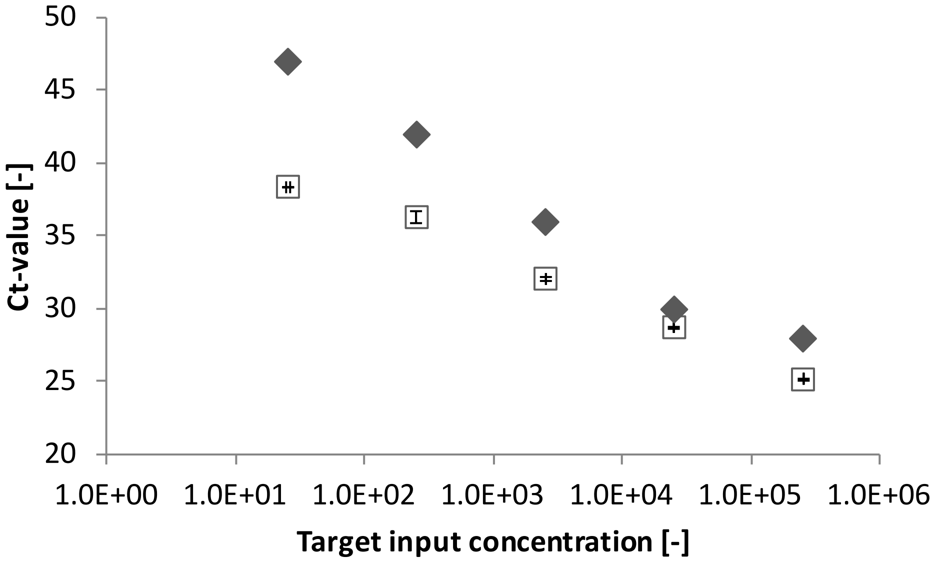


Figure 5.7 C_t-values as a function of target input concentrations for qPCR and real time array PCR for the MSSE target amplification. The curve fits were based on the average value of two spots. These measurements were all performed using the Geneamp mastermix. C_t-values were calculated based on the curve-fit data. The error bars for the qPCR data show the minimum and maximum range of the two measurements. The real time array PCR result consists of a single set of measurements.

Down to 25 copies of this target could be detected using both real time array PCR as well as qPCR. However, as found in figure 5.7, a relatively low efficiency of 58% was found for real time array PCR. Since the hybridization signals are assumed to be proportional to the bulk PCR product concentration, which is confirmed by the mathematical analysis of chapter 6, a reduced overall efficiency can be attributed to a reduced amplification efficiency in the solution.

There are a few possible explanations of the lower amplification efficiency of real time array PCR. A first consideration is that amplicons hybridized to the probes on the microarray surface might not be available for elongation. As a worst case estimate, the maximum number of amplicons that can diffuse towards the microarray surface during the hybridization phase of each PCR cycle was calculated, based on the mathematical analysis presented in chapter 6. The fraction was found to be less than a few % in the vicinity of the microarray surface, and therefore this effect cannot explain reduced the amplification efficiency. This was also confirmed by experiments (data not shown), where a qPCR was performed with and without added capture probes in solution and no difference in results was observed. Furthermore, since the elongation temperature for the 3-step PCR protocol is higher than the calculated melting temperature of the capture probe, it is expected that at least part of the targets will detach and still participate in the elongation process.

A second consideration is loss of polymerase due to either non-specific adsorption to the cartridge material surfaces including needles and connectors, or thermal degradation. This is a well-known phenomenon and could be a source of the reduced PCR efficiency[20]. Bovine Serum Albumin (BSA) is a blocking agent and prevents other proteins from adsorbing onto the surface[20,21]. A strong positive effect of adding BSA was confirmed by experiments: by excluding BSA from the reaction mix, no PCR curves where obtained at all, even at higher input concentrations. An optimum of 0.1 % w/v BSA was found (data not shown), since higher concentrations of BSA are known to inhibit the PCR[20]. Furthermore, enzymes are known to be extremely sensitive to thermal degradation[22] at temperatures exceeding 95 ^{o}C. Thermal degradation could lead to a lack of available enzyme which can become critical particularly at high cycle numbers. As a result C_t-values of amplification reactions using low target input numbers will be affected the most. The heater settings in the thermocycler were offset by 3 ^{o}C to compensate for gradients due to heat transfer to the cartridge. Locally the temperature could reach values higher than 95 ^{o}C during the DNA denaturing phase.

5.3.2 Optimization of mastermix and PCR protocol

5.3.2.1 Optimization of mastermix

The experiments described above were all performed in Geneamp mastermix. The exact composition of this mastermix is not known. Especially the type and amount of the polymerase is not revealed by the supplier. Furthermore, information from the supplier learned that this enzyme could not be purchased separately and therefore not added to study the effect of an increased enzyme concentration. A change to an alternative enzyme and PCR buffer was made to provide more possibilities for investigation and optimization of the PCR composition for real time array PCR. The reagents used for further experiments were Amplitaq Gold enzyme, 10 x PCR buffer II, nucleotides, $MgCl_2$ (all ingredients from

Applied Biosystems, CA, USA). The optimization of this formulation and protocol was done using qPCR.

A concentration of 3 mM of $MgCl_2$ was found to be optimal for amplification of the femA SE (data not shown) and this concentration was used in further experiments. To investigate the effect of the mastermix on the amplification efficiency, a target dilution series using qPCR with the formulation of the home-made mix was compared with the dilution series as obtained with the Geneamp mastermix, and the results are given in figure 5.8. It can be seen that the absolute C_t-values are nearly identical, with the home-made mix having a slightly higher efficiency. Since the home-made mix gave at least as good amplification as the Geneamp mix, no additional optimization of the amplification conditions was needed. Further experiments were done on optimization of the PCR protocol in a real time array PCR setting.

5.3.2.2 Optimization of amplification protocol in real time array PCR

For bulk qPCR, a 2-step PCR protocol resulted in a considerable decrease in the absolute C_t-values when using the home-made mix, which was not found with the Geneamp mastermix (data not shown). The potential of such a 2-step program was therefore investigated for real time array PCR as well, using the home-made mix. Figure 5.9 gives the results of the array PCRs for a complete dilution series. The previously applied 3-step protocol was compared to the 2-step protocol. No significant differences in C_t-values between the two protocols can be observed. Furthermore no amplification curves could be detected for the 100 and 10 input copies regardless of the protocol, while with the Geneamp mastermix it was possible to measure good amplification curves for at least as low as 25 copies as input. This is an indication that with the home-made mix not enough active enzyme is available to have adequate amplification during the higher cycle numbers.

Since the efficiency of the 2-step protocol was slightly higher (68 %) instead of the 3-step protocol (63 %) and resulted in somewhat smaller error bars, the two-step protocol was applied in further experiments.

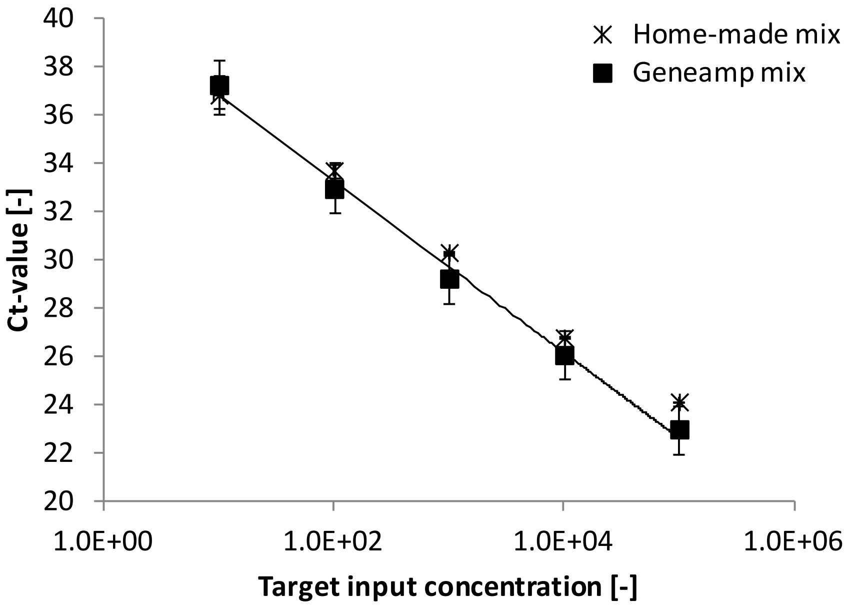


Figure 5.8 C_t-values for different input concentrations using the home-made mastermix (2 Units) and as a reference the ready-to-use Geneamp mastermix, using a 3-step qPCR amplification protocol. The lines represent a semi-log fit between the C_t-value and the input concentration from which the amplification efficiency was calculated (Home-made mix: efficiency 104 %, Geneamp mix: efficiency 92%) (Fit quality R^2 of the home-made mix: 0.998; Geneamp mix: 0.992). Each measurement point consists of a duplicate spot measurement; the error bars represent the maximum and the minimum value of the two measurements.

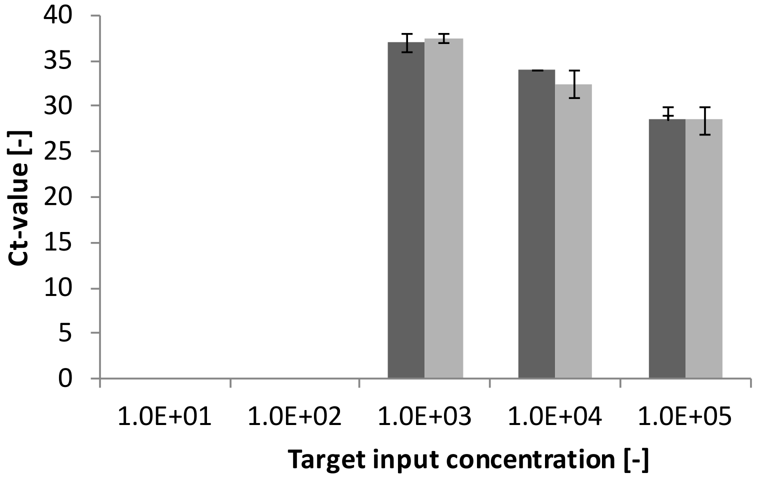


Figure 5.9 C_t-values for different target input concentrations (logarithmic) for a 2-step (dark grey bars) and a 3-step real time array PCR protocol (light grey bars). Each measurement consists of a duplicate measurement on the microarray. The error bars in the graphs represent the maximum and minimum values obtained in this duplicate measurement. For 10 and 100 copies, no C_t-value could be detected.

5.3.2.3 Optimization of enzyme concentration

To investigate if the Taq polymerase becomes limited during the higher cycle numbers, this enzyme concentration was optimized. This was done via two routes: by preventing the loss of enzyme during the amplification process, as well as by the addition of a larger initial amount.

Reduction of non-specific binding

BSA was included in the standard real time array PCR mix to reduce non-specific binding during the amplification process. To fully minimize non-specific binding also during filling of the cartridge, needles and connectors, the effect of a preceding blocking step was investigated by flushing the cartridge twice with a 5 mg/ml BSA solution, prior to filling it with the PCR ingredients and performing the real time array PCR. The C_t-values of a dilution series with and without pre-blocking are given in table 5.3. The protocol was performed using 2 units of Taq polymerase in a 2-step PCR protocol.

Table 5.3 Average C_t-values for a 2-step real time array PCR protocol without and with pre-blocking with BSA. All measurements were done in duplicate. The values in between the brackets indicate the range obtained with the two measurement values.

Input concentration	Without pre-blocking	With pre-blocking
10^5	28.5 (+/-1.5)	28.5 (+/-1.5)
10^4	32.5 (+/-1.5)	32.5 (+/-0.5)
10^3	37.5 (+/-0.5)	35.5 (+/-0.5)
10^2	Not detectable	39.5 (+/-0.5)
10^1	Not detectable	42.0 (+/-1.0)

At higher target input concentrations, C_t-values are comparable, meaning that for the low to medium cycle numbers, sufficient enzyme activity is available. However, at lower target input concentrations, the enzyme activity was reduced to such an extent that no adequate amplification occurred during the higher cycler numbers when no pre-blocking procedure was applied. As a result, higher C_t-values or even no amplification curves at all were found. The pre-blocking protocol ensures that sufficient enzyme activity is maintained throughout all cycles enabling the detection of lower template concentrations as well. Since this step clearly improved the results, further experiments were all performed with the pre-blocking protocol.

Effect of enzyme concentration

In table 5.4 below, the effects of a higher enzyme concentration on PCR performance is given. An extra increase in enzyme activity does not lead to significantly improved results. It can be concluded that there is no need for further improvement in enzyme activity. Amplification efficiencies close to 100 % are achieved.

Table 5.4 Average C_t-values for a 2-step real time array PCR protocol with 2 and 3 units of enzyme. All measurements were done in duplicate. The values in between the brackets indicate the range obtained with the two measurement values.

Input concentration	2 Units	3 Units
10^5	28.5 (+/-1.5)	27.5 (+/-0.5)
10^4	32.5 (+/-0.5)	31.5 (+/-1.5)
10^3	35.5 (+/-0.5)	36.0 (+/-0.5)
10^2	39.5 (+/-0.5)	39.0 (+/-0.5)
10^1	42.0 (+/-1.0)	42.0 (+/-1.0)

The development of the improvement in amplification efficiency during the optimization study is given in figure 5.10. Starting from an efficiency slightly higher than 60 % (though it must be noted that it is difficult to calculate absolute efficiencies based on a such a limited data set due to the fact that no C_t-values could be obtained at the lower input concentrations), an efficiency of 97 % was reached with a 2-step protocol in a pre-blocked cartridge. The largest gain in efficiency is obtained upon inclusion of this pre-blocking step in the amplification protocol. The efficiency is close to the qPCR efficiency, meaning that the amplification itself is nearly optimal and, with these process settings, down to 10 target molecules can indeed be detected.

A more detailed comparison in absolute C_t-values between the initial protocol settings and the final optimized protocol settings is given in figure 5.11. It can be seen that the optimized protocol settings allows for the detection of the lower target concentrations (down to 10 targets).

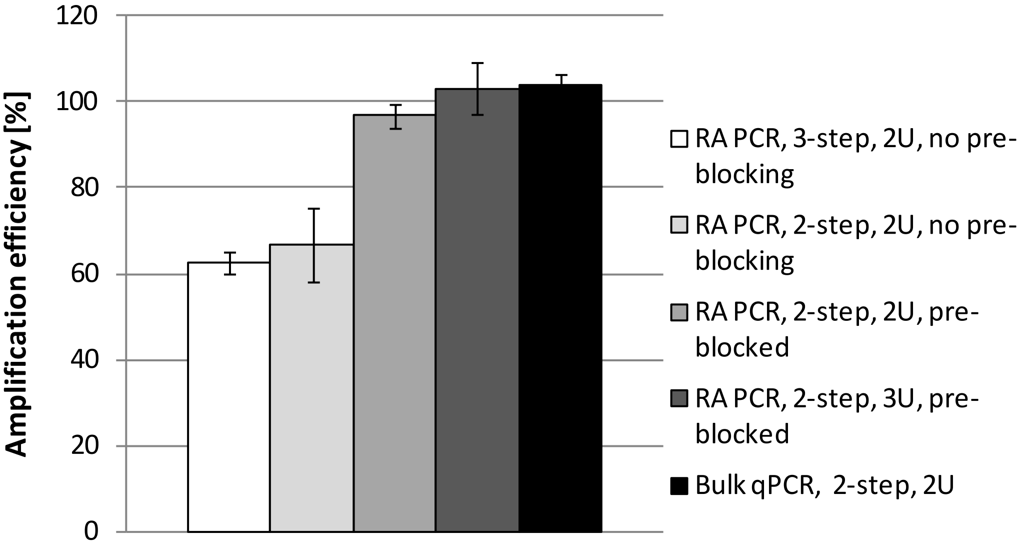


Figure 5.10 Summary of the individual amplification efficiencies as deduced from the C_t-values as presented in the previous graphs and tables for the different protocols tested. Starting from an efficiency of 60 %, several optimization approaches led to an efficiency of 100 %, as high as qPCR. The error bars represent the standard deviations on the calculated amplification efficiency.

It was proven that the enzyme activity in real time array PCR is a crucial factor. This topic will also be addressed in chapter 6 on the mathematical analysis of the real time array PCR.

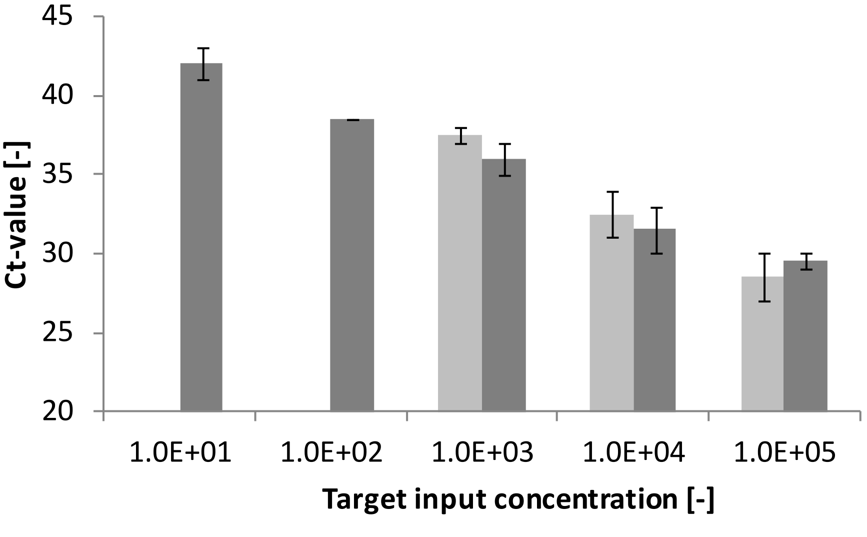


Figure 5.11 C_t-values for the 2 units, 3-step protocol without pre-blocking (light grey), and the 3 units, 2-step protocol with pre-blocking (dark grey). It can be seen that with the latter protocol, it is possible to detect also the lower target input concentrations.

5.3.3 Asymmetric real time array PCR

Asymmetric PCR protocols use unequimolar concentrations of the two primers. This implies that a chosen strand is amplified in excess of the other strand. The advantage of an asymmetric PCR for real time array PCR is the potentially improved hybridization kinetics because surface hybridization is influenced by the competitive reaction of renaturation in the solution[23]. Figure 5.12 gives the symmetric and the asymmetric real time array PCR results for the femA gene of the *MRSA* (femA SA) and the *MSSE* (femA SE).

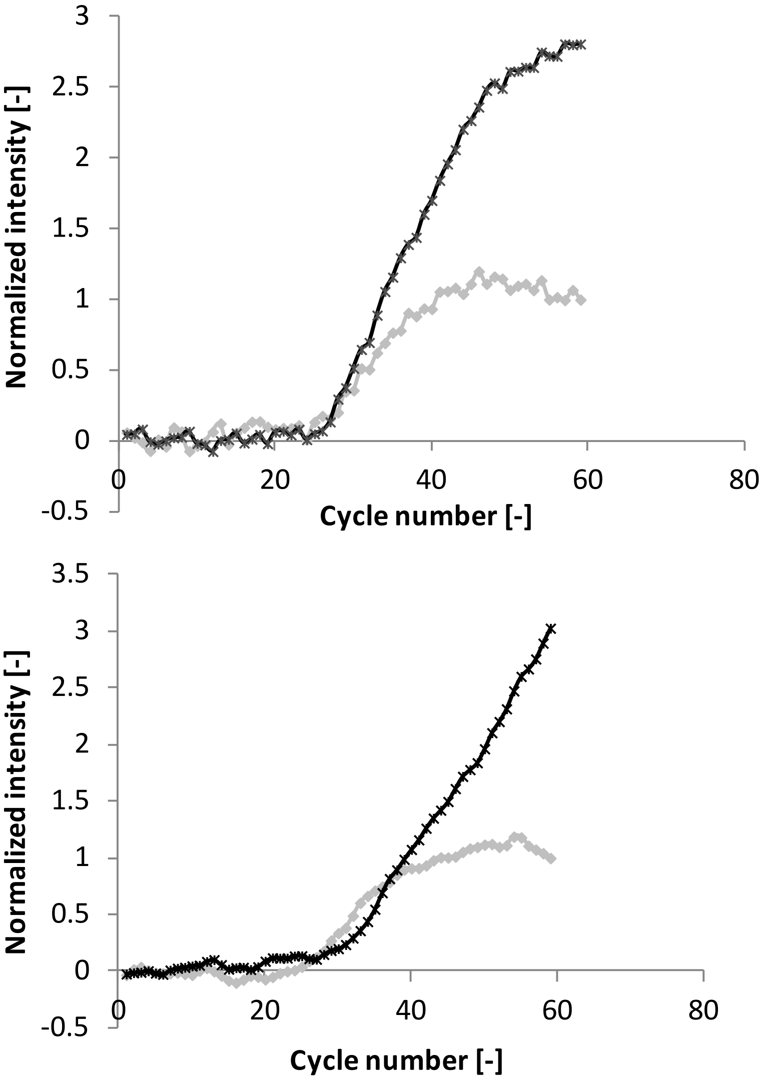


Figure 5.12 Symmetric (300 nM of both primers, grey line with diamonds) and asymmetric (300 nM of the reverse primer and 50 nM of the forward primer, black line with crosses) real time array PCR curves for the amplification of $2.5.10^5$ cps of the *femA* gene of *MRSA* (upper figure) and *MSSE* (lower figure). The results are normalized to the final fluorescence intensity obtained with symmetric PCR. A 3-step temperature protocol was used, no pre-blocking step was applied.

Both graphs clearly show that in the case of asymmetric PCR, signals continue to increase and reach a final fluorescence level up to three times the signals obtained in symmetric PCR. The use of an asymmetric PCR instead of a symmetric PCR can thus positively influence fluorescent signals in a real time array PCR.

5.4 Discussion and conclusions

In this chapter, feasibility of the real time array PCR concept has been shown, i.e. monitoring of the amplification reaction of a DNA target on a microarray using a compact confocal scanner. Dilution series have shown that at least as low as 10 copies of a target can successfully be amplified and detected. Several optimization strategies led to an increase in amplification efficiency from around 60 % to around 100 %, meeting the efficiency of bulk qPCR.

This increase in amplification efficiency was achieved by applying a two-step PCR protocol combined with a blocking step (with BSA) prior to the amplification. The blocking step resulted in a smaller fraction of enzyme activity loss caused by non-specific adsorption to the cartridge materials. The use of such a blocking step is preferred in terms of cost price instead of a higher initial enzyme concentration, since the Taq polymerase is a relatively expensive PCR ingredient. This step might be taken up in the cartridge assembly process flow. After inclusion of a blocking step, an additional increase in enzyme concentration only had a very limited effect, meaning that basically sufficient enzyme activity was available for good amplification throughout all PCR cycles.

A general challenge in post-PCR microarray hybridizations is renaturation of the complementary strands in solution which is competing with the surface hybridization reaction. Renaturation in solution of the complementary amplicons reaches equilibrium quickly which means that hardly any single stranded DNA is left in solution. Surface based reactions do reach equilibrium much slower[24]. Therefore, competition between renaturation in solution and surface hybridization impacts the maximum surface hybridization signals that can potentially be reached. A further improvement identified is the use of an excess of the labeled strand to reduce the effect of renaturation in solution, by asymmetric PCR or a LATE-PCR (Linear-After-The-Exponential[25] which uses primer sequences slightly adapted for the decrease in melting temperature of the rate-limiting primer, due to its very low concentrations in course of the PCR. This method might need less optimization in primer concentrations). Experiments on asymmetric PCR have shown up to three times higher end fluorescent signals measured on the microarray. During the first cycles of the process, the amplification of the asymmetric and symmetric PCR will, to a large extent, be symmetric since both primers are in excess compared to the target concentration. Only after the amount of forward primer has been consumed, amplification becomes asymmetric leading to an excess of the labeled strands. The amplification of this excess of labeled amplicons then follows a linear trend instead of an exponential curve until also the concentration of the labeled primer or other PCR ingredients become limiting. Such a method is also favorable in terms of cost price since, when compared to symmetric PCR, less primer is used. The disadvantage is the need for optimization of primer concentrations for each primer-target combination.

Further improvements can be made in the thermal performance of the instrument. Ramp rates, especially cooling rates, are still relatively slow (1 ^{o}C/s) taking up an essential part of the hybridization time. Increasing the ramp rates will result in more time available for hybridization, and furthermore a reduced exposure time of the enzyme to higher temperatures.

The method as presented in this chapter has shown proof of principle for the quantitative detection of DNA sequences on a microarray. The method itself is not necessarily limited to PCR: in principle alternative methods for nucleic acids amplification can be used as well.

The main advantage of the method discussed is the high multiplexing capability. Multiplexing has not been discussed here. With the current prototype as presented in this chapter, the noise (ideally proportional to the square root of the background signal[26] coming from background fluorescence) is significant, as can be seen in figures 5.5 and 5.6. Though the results of these single-plex assays are promising, for the development of multiple individual primer-target combinations, it will be challenging to continue to achieve significant signals above background without adaptations in the current prototype, method or substrates. This is because the use of multiple primer pairs increases the background proportionally to the concentration of the labels in solution, leading also to an increased noise level. A large part of the background signals is caused by non-specific binding to the microarray substrate (data not shown) and therefore microarray slides and surface chemistries that allow for less non-specific binding will result in less background and thus measurement noise. Additionally, methods can be considered that enable the use of less background fluorescence e.g. intercalating dyes, molecular beacons, or probes that change emission upon target hybridization. Furthermore, the detection format can be further adjusted to aim for smaller detection volumes, either by using different detection methods itself, or by using substrates that allow for evanescent wave detection like a wire-grid based concept[27]. Also an embodiment can be envisioned where the concepts of chapter 4 and 5 are combined, i.e. a flow-through set-up where a porous microarray physically separates the space where denaturation takes place and a detection/annealing/elongation zone. Such a setup reduces diffusion lengths and removes background signals by pumping the labeled molecules through the membrane. Rather than temperature cycling of the complete PCR volume, in this approach the liquid is transported through the different temperature zones.

We conclude that we have shown proof of concept of monitoring PCR amplification on a microarray by confocal scanning using a potentially cost-effective instrument and standard qPCR reagents. This method enables increased multiplexing of qPCR with comparable detection sensitivity as bulk qPCR. Since amplification and detection are carried out in a single, closed cavity thereby minimizing lab contamination and human interaction it is particularly suited for decentralized diagnostic testing. In its current format, the compact confocal scanner equipped with thermal control is especially suitable for applications that needs only a limited number of primer pairs but use multiple capture probes such as genotyping assays, e.g. the HPV assay as discussed in chapter 4 where general primers are used to amplify all types of the HPV virus, and multiple capture probes to discriminate between the multiple genotypes. Furthermore it may also become suitable for the detection of multiple single nucleotide polymorphisms (SNPs).

References

1. Vet JAM, Majithia AR, Marras SAE, Tyagi S, Dube S, Poiesz BJ, Kramer FR. Multiplex detection of four pathogenic retroviruses using molecular beacons. *PNAS* 1999, 96:6394-6399
2. Donhauser SC, Niessner D, Seidel M. Sensitive quantification of Escherichia coli O157:H7, Salmonella enterica, and Campylobacter jejuni by combining stopped polymerase chain reaction with chemiluminescence and flow-through DNA microarray analysis. *Anal Chem* 2011, 83: 3153-3160
3. Hassibi A, Vikalo HM, Riechmann JL, Hassibi B. Real time DNA microarray analysis. *Nucleic Acids Res* 2009, 37:e132
4. Remacle J, Alexandre I, Margain S, Husar D. Real time PCR of targets on a microarray. Eppendorf patent application WO2006053770
5. Trau D, Lee T, Lao A, Lenigk R, Hsing I-M, Ip NY, Carles MC, Sucher NJ. Genotyping on a complementary metal oxide semiconductor silicon Polymerase Chain Reaction chip with integrated DNA microarray. *Anal Chem* 2002, 74:3168-3173
6. Andresen D, von Nickisch-Rosenegk M, Bier FF. Helicase dependent OnChip-amplification and its use in multiplex pathogen detection. *Clin Chem Acta* 2009, 403, 244-248
7. Von Nickisch-Rosenegk M, Marschan X, Andresen D, Abraham A, Heise C, Bier FF. On-chip PCR amplification of very long templates using immobilized primers on glassy surfaces. *Biosensor Bioelectron* 2005, 20:1491-1498
8. Sun Y, Dhumpa R, Bang DD, Høgberg J, Handberg K, Wolff A. A lab-on-a-chip device for rapid identification of avian influenza viral RNA by solid phase PCR. *Lab chip* 2011, 11:457-1463
9. Adessi C, Matton G, Ayala G, Turcatti G, Mermod J-J, Mayer P, Kawashima E. Solid Phase DNA amplification: characterization of primer attachment and amplification mechanisms. *Nucleic Acids Res* 2000, 28(20):e87
10. Khodakov DA, Zakharova NV, Gryadunov DA, Filato FP, Zasedatelev AS, Mikhailovich VM. An oligonucleotide microarray for multiplex real-time PCR identification of HIV-1, HBV, and HCV. *Biotechniques* 2008, 44:241-248
11. Remacle J, Alexandre I, Koehn H, Seippel, M. Lid for PCR vessel comprising probes permitting PCR amplification and detection of the PCR product by hybridisation without opening the PCR vessel, Eppendorf patent application EP1788096
12. Erdogan F, Kirchner R, Mann W, Ropers H-H, Nuber UA. Detection of mitochondrial single nucleotide polymorphisms using primer elongation reaction on oligonucleotide microarrays. *Nucleic Acids Res* 2001, 29(7):e36
13. Pierik A, Zwanenburg C, Moerland E, Broer D, Stapert H, van den Brule AJC. Rapid genotyping of Human Papillomavirus by post-PCR array-based hybridization techniques. *J Clin Microbiol* 2011, 49(4):1395-1402

14. Drobyshev AL, Nasedkina TV, Zakharova NV. The role of DNA diffusion in solid phase polymerase chain reaction with gel-immobilized primers in planar and capillary microarray format. *Biomicrofluidics* 2009, 3:044112
15. Pierik A, Dijksman JF, Raaijmakers ATA. Wismans AJJ, Stapert HR. Quality control of inkjet technology for DNA microarray manufacturing. *Biotech J* 2008, 3:1581-1590
16. Pierik A, Dijksman JF, Lub J, Stapert HR, Broer DJ. Immobilization of homo-oligomer tails onto amine-functionalized solid substrates and the effects on hybridization. Anal Chem 2010, 82:1191-1199
17. Francois P, Pittet D, Bento M, Pepey B, Vaudaux P, Lew D, Schrenzel J. Rapid detection of Methilicin-Resistant Staphylococcus aureus directly from sterile or nonsterile clinical samples by a new molecular assay. *J Clin Microbiol* 2003, 41:254-260
18. Spiess AN, Feig C, Ritz C. Highly accurate sigmoidal filling of real-time PCR data by introducing a parameter for asymmetry. *BMC Informatics* 2008, 9:221
19. Novak L, Neuzil P, Pipper J, Zhang Y, Shinhan L. An integrated fluorescence detection system for lab-on-a-chip applications. *Lab Chip* 2007, 7:27-29
20. Erill I, Campoy S, Erill N, Barbé J, Aguiló J. Biochemical analysis and optimization of inhibition and adsorption phenomena in glass-silicon PCR-chips. *Sensor Actuator-B* 2003, 96:685-692
21. Shin DS, Lee KN, Jang KH, Kim JK, Chung WJ, Kim YK, Lee YS. Protein patterning by maskless photolithography on hydrophilic polymer-grafted surface. *Biosens Bioelectron* 2003 19:485-494
22. Sambrook J, Russell DW. Molecular Cloning, 2001. Cold Spring Harbor Laboratory Press, 3rd edition
23. Gyllensten UB, Erlich H. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. *Proc Natl Acad Sci* 1988, 85:7652-7656
24. Gao Y, Wolf LK, Georgiadis RM. Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res* 2006, 34(11):3370-3377
25. Sanchez JA, Pierce KE, Rice JE, Wangh LJ. Linear-after-the-exponential-PCR: an advanced method of asymmetric PCR and its uses in quantitative real time analysis. *PNAS* 2004, 101(7):1933-1938
26. Sandison DR, Webb WW. Background rejection and signal-to-noise optimization in confocal and alternative fluorescence microscopes. *Appl optics* 1994, 33(4):603-615
27. Klunder DJW, van Herpen MJW, Stapert HR. Wiregrid monitor device. Philips Electronics Patent Application US 2010/0108908.

Chapter 6

Mathematical analysis of the real time array PCR process

This chapter is based on:

J.F. Dijksman, A. Pierik, *under revision*

Abstract

The real time array PCR method[1], as described in chapter 5, is a novel biochemical tool to monitor DNA amplification on a microarray. In this chapter, a mathematical model is presented that describes the integral biochemical processes that are taking place: in particular the interplay between the amplification reaction happening in the bulk and the surface reaction that occurs on the microarray substrate. This model has been used to investigate the effect of different parameters on the reaction efficiencies. Increasing for instance the annealing time enhances the amplification and the hybridization signals. From a clinical application point of view, turnaround times should preferably be minimized. There is thus a trade-off between signal quality and sensitivity on the one hand and turnaround time on the other hand. One of the most critical process parameters appeared to be the potential polymerase enzyme loss during the reaction due to thermal degradation or surface adsorption. This effect is likely to occur in this novel PCR method since the materials of choice have not yet been optimized for minimal enzyme adsorption. Another important parameter is the degradation of the capture probes that may cause loss of signal quality and sensitivity.

6.1 Introduction

Real time array PCR is a novel biochemical method to monitor real time the amplification process of a high number of different DNA targets sequences that can possibly be present in a PCR sample on a microarray[1,2]. It combines the real time capabilities to retrieve data during the amplification process as done in real time PCR and thereby obtaining quantitative information on the initial number of target molecules present in the sample (qPCR), with the high multiplex capability of a microarray based assay. The concept is summarized in figure 6.1 below. The concept and principle of real time array PCR is explained in detail in chapter 5.

The real time array PCR process consists of, in the mathematical analysis as performed in this chapter, a 3-step temperature profile. Each PCR cycle consists of a short denaturation step, an annealing/hybridization step and an elongation step, as indicated in figure 6.1.

Considering the real time array PCR process, compared to qPCR, there are a number of differences. First of all, qPCR is a bulk process, which is usually symmetric meaning that commonly equal amounts of sense and antisense primers are elongated everywhere in the bulk. Amplicons are detected by measuring fluorescence in the bulk volume. In real time array PCR, amplification takes place in the bulk as well. It is, however, not only a bulk process because at the microarray surface, specific molecules are captured and taken away from the solution, causing a concentration gradient in a small boundary layer above the microarray. This results in mass transport towards the microarray surface. Because one specific species is captured by the probes on the microarray surface, locally the PCR process is not symmetric anymore even during one PCR cycle.

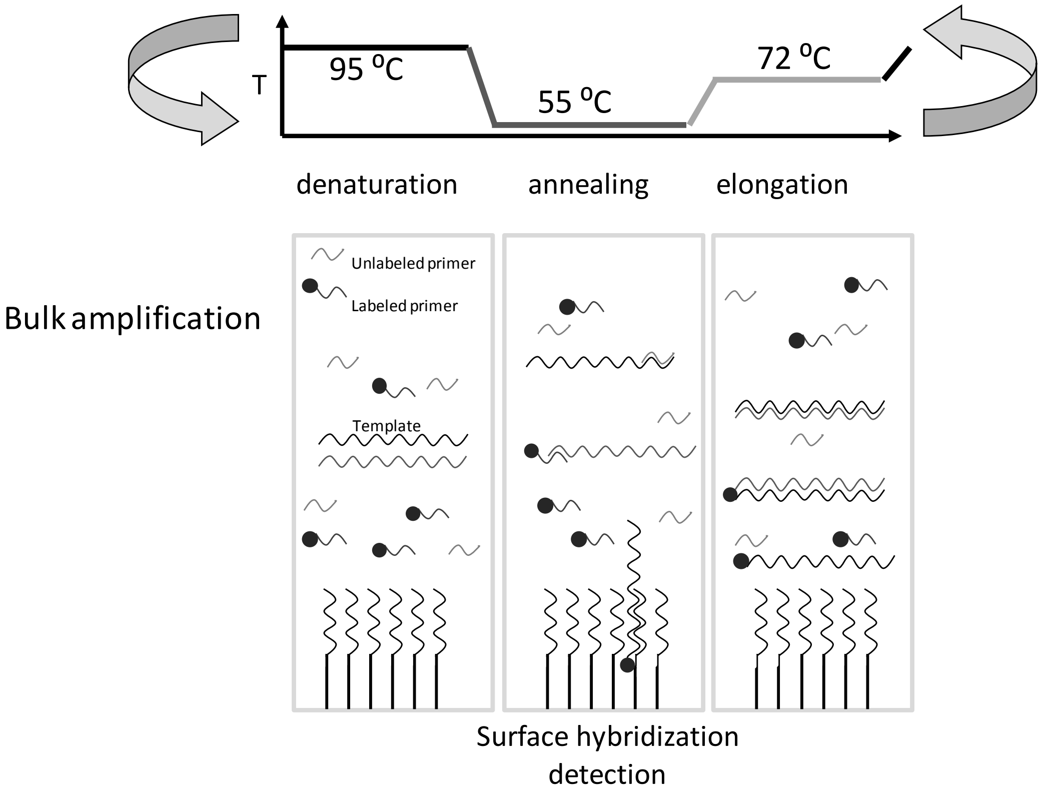

T
95 ºC
55 ºC
72 ºC
denaturation
annealing
elongation
Bulk amplification
Unlabeled primer
Labeled primer
Template
Surface hybridization
detection

Figure 6.1 Schematic view of the real time array PCR. The upper part of the figure shows the temperature profiles of a three-step PCR reaction, whereas the middle and lower part depict the bulk and the surface effects taking place at the respective temperatures, respectively.

Finally, because of the surface reaction a different time constant is playing a role, namely the time it takes before the surface concentration has reached equilibrium.

Per cycle the concentration of amplicons increases and more and more targets are captured by specific probes in the spots of the microarray. After each annealing and hybridization period the hybridization is measured and in this way the dynamics of the amplification process are followed in time.

Mathematical models of the real time PCR process are known[3,4] as well as models of hybridization on a microarray surface[5]. In this chapter, a model that describes the real time array PCR process is presented, thereby linking the real time PCR process to the microarray hybridization process.

6.2 Mathematical modeling

6.2.1 Approach

The mathematical modeling of the real time array PCR process can be split in a number of different building blocks:

- The reactions taking place during the different temperature steps of the PCR process including the surface reaction.
- The various rate equations describing the time evolution of the reaction equations.
- The solutions of the equations describing the rate of change of the different species in the bulk.
- The bulk concentrations will be used as input to solve the kinetic equation describing the hybridization process at the microarray surface.
- The end concentrations of the different species of one cycle will be the initial conditions for the next cycle.

The main result of the analysis will be the evolution of the target sequence concentration in the bulk, as well as the fraction of occupied capture probes on a spot of the microarray at the end of the annealing step, determined cycle after cycle.

6.2.2 Model assumptions

In order to develop the analytical PCR model we assumed a number of conditions:

- The PCR protocol investigated is a three-step temperature profile.
- The formation of primer-dimers is out of scope.
- Unlike in reference 3, renaturation of the amplicons is included.
- The capture probes are not elongated during the process. It has experimentally been verified that this is indeed not or very limited occurring (data not shown).
- Tertiary complex formation between complexes and enzymes and elongation take place during the time window indicated as elongation phase.
- Elongation of the primers is complete at the end of the elongation phase: no intermediate products are formed.
- The parameters describing the surface hybridization rate are assumed to be constant in course of the process. In reference 4, it is mentioned that the reaction rates might change in course of time, but since in the real time array PCR the hybridization time scale is short, a constant reaction rate is assumed.
- The reaction rates for both primer annealing as well as ds DNA renaturation are assumed to be different. In this sense our analysis can be considered as an extension of the paper by reference 6. It is assumed that transport of amplicons to the capture probe spots is only through diffusion[7]. The temperature profile is assumed to be homogeneous

inside the real time array PCR volume, and therefore does not lead to thermally induced convection.

- All reverse primers are labeled.
- The reaction rate of the surface hybridization is independent on the surface density of the capture probes. This is only a point of attention where the number of available capture probes decreases due to degradation.
- During denaturation, all double stranded DNA and complexes of single stranded DNA with primers will be denatured and all captured targets will detach from the array surface.
- Initially the template concentration is very low. The template molecules are usually long DNA fragments of which during amplification only a specific sequence that is much shorter is amplified. After only a few cycles the concentration of original templates is negligible with respect to the concentration of the much shorter stranded amplicons.
- Only hybridization of free single stranded DNA is considered, as the complexes are most probably not able to hybridize due to steric hindrance. This despite the fact that by design the primer sites and the hybridization site on the amplicon are not overlapping.

6.2.3 Basic reaction equations

During the denaturation step double stranded DNA (ds DNA) splits into equal amounts of single stranded complementary DNA (sense and antisense). In the following, sense single stranded DNA (ss DNA) is referred to as species A and likewise antisense ss DNA as species B:

$$dsDNA \rightarrow A + B \tag{1}$$

After denaturation, the temperature is lowered and a number of processes start in parallel. In the bulk the single stranded DNA reacts with primers to form binary complexes, moreover species A and species B react with each other again (renaturation) to form double stranded DNA[6]. During amplification more and more species A will be formed during each cycle, the primer annealed to species B will be elongated to form labeled species A. Thus, the concentration of labeled reverse primer decreases in time, while at the same time the concentration of labeled species A increases. Likewise for species B. This can be summarized to the following scheme (the * below indicates the presence of a fluorophore):

$$A* + B \rightarrow dsDNA* \tag{2}$$

$$A* + primer_{forward} \rightarrow complex_1 * \tag{3}$$

$$B + primer_{reverse} * \rightarrow complex_2 * \tag{4}$$

During initial cycling, the complex formation is the preferred reaction because of the vast excess of primers compared to the initial load of target molecules. At the end of the real time PCR process the situation is reverse: the preferred reaction is the renaturation since the system is running short of primers. More extended descriptions of the reactions involved in real time PCR can be found in reference 3 and 4.

On the spot located on the microarray surface, the hybridization takes place as explained above:

$$A * + capture\ probe_{surface} \rightarrow complex_{surface\ without\ primer} * \qquad (5)$$

Only species A participates in the hybridization reaction. All other reactions involving capture probes with binary or tertiary complexes are unlikely because of the steric hindrance effect of the primers and enzymes on the hybridization reaction, especially for the short strands considered (molecular weight binary complex (100 bases amplicon and 30 bases primer) roughly 60-70 kDa, Taq enzyme 94 kDa).

For elongation, the temperature is raised to the elongation temperature. The binary complexes form tertiary complexes with the polymerase:

$$complex_i + polymerase \rightleftarrows tertiary\ complex_i\ ,\ i = 1{,}2 \qquad (6)$$

This tertiary complex yields the reaction product ds DNA and the enzyme according to:

$$tertiary\ complex_i + dNTPs \longrightarrow dsDNA + polymerase\ ,\quad i = 1{,}2 \qquad (7)$$

Finally the system will be heated to the denaturation temperature and all double stranded DNA molecules (both the ds DNA molecules consisting of the elongated primers as well as the annealed primers that have not elongated) in the bulk denature and the hybridized targets detach from the spot surface. Also complexes in the bulk that did not elongate will be split into single stranded components again. The temperature is lowered to the annealing temperature and the whole procedure is repeated again until the last cycle is performed. After each anneal/hybridization step the hybridization on the surface is calculated in terms of the fraction of capture probes occupied.

Furthermore, during temperature cycling part of the polymerase will lose its activity and will not be active anymore during the elongation step[8]. Also part of the enzymes may be caught by the walls of the reaction chamber and will not participate any longer[9].

During thermocycling, the capture probes might be damaged or can detach from the surface, leading to a lower number of available capturing sites on the microarray surface. Another mechanism can be envisioned as well: fouling of the microarray surface. This is not specifically taken into account in this analysis.

6.2.4 Rate equations

The concentrations of species A and B in the bulk (any interference with the microarray is neglected here) decrease in course of time according [6, 10]:

$$\frac{\partial c_A}{\partial t} = -kc_A c_{pF} - k_{dsDNA} c_A c_B, \quad \frac{\partial c_B}{\partial t} = -kc_B c_{pR} - k_{dsDNA} c_A c_B \tag{8}$$

Here k is the rate constant for the formation of complexes while k_{dsDNA} describes the renaturation of complementary single stranded DNA to double stranded DNA. All the time the following conditions are valid:

$$c_A + c_{pR} = c_{p0} + c_{template} \approx c_{p0}, \quad c_B + c_{pF} = c_{p0} + c_{template} \approx c_{p0}, \quad c_A = c_B \tag{9}$$

These conditions convert the second order reaction equations into:

$$\frac{\partial c_A}{\partial t} = -kc_{p0}c_A + (k - k_{dsDNA})c_A^2, \quad \frac{\partial c_B}{\partial t} = -kc_{p0}c_B + (k - k_{dsDNA})c_B^2 \tag{10}$$

From the beginning of thermal cycling up to almost the very end, the concentration of both species A and B remains small with respect to the initial primer concentration. This means that the second terms in the right hand sides of the two equations (10) are small and the first terms dominate the time evolution of the concentrations. Therefore the following approximate equations are used:

$$\frac{\partial c_A}{\partial t} = -kc_{p0}c_A, \quad \frac{\partial c_B}{\partial t} = -kc_{p0}c_B \tag{11}$$

For c_{A0} and c_{B0} the initial target input concentrations are used.

The concentration of double stranded DNA follows from:

$$\frac{\partial c_{dsDNA}}{\partial t} = k_{dsDNA} c_A c_B \tag{12}$$

Because of conservation of species A and B, the concentrations of $complex_1$ and $complex_2$ are given by:

$$c_{complex_1} = c_{A0} - c_A(t) - c_{dsDNA}(t); \qquad c_{complex_2} = c_{B0} - c_B(t) - c_{dsDNA}(t) \tag{13}$$

During the elongation step, the rate of consumption of either $complex_1$ or $complex_2$ is decribed by the so-called Michaelis-Menten kinetics (no distinction is made between the complexes)[11]:

$$\frac{dc_{complex}}{dt} = -\frac{R_{max}\, c_{complex}}{K_{mDNA} + c_{complex}} \tag{14}$$

Where R_{max} is the maximum reaction rate and K_{mDNA} the Michaelis constant for the enzyme controlled process of the formation of tertiary complexes defined by:

$$R_{max} = k_2 c_{TAQ}, \quad K_{mDNA} = \frac{k_{-1}+k_2}{k_1} \tag{15}$$

The rate constants k_1 and k_{-1} describe the formation and dissociation of tertiary complexes (complexes with enzymes) and k_2 is defined as the rate constant of the elongation reaction. The basic assumption of the Michaelis-Menten kinetics is that the formation of elongated ds DNA is much faster than the rate of formation of the tertiary complexes. The formation of the tertiary complexes is thus the rate limiting process.

The value of R_{max} depends on the concentration of nucleotides (dNTPs). Usually the nucleotides are present in excess, so the reaction between the tertiary complexes and nucleotides runs at a constant rate. Only when the concentration dNTP's becomes so low that it becomes equal or smaller than K_{mdNTP} the value R_{max} decreases. Therefore we propose that the dynamics of the Taq polymerase driven elongation process can be described by:

$$\frac{dc_{complex}}{dt} = -\frac{R_{max}\frac{c_{dNTP}}{K_{mdNTP}+c_{dNTP}}}{K_{mDNA}+c_{complex}} c_{complex} \tag{16}$$

In the neighborhood of the spots of the microarray the diffusion is essentially three dimensional. This can be simplified by assuming the concentration profile to be axi-symmetric. In case of a thought experiment the flat spot is replaced by a spherical cap (dome) with the same surface area, the only co-ordinate that remains is the distance r from the centre of the spherical cap (spherical symmetry, see figure 6.2). In that situation the diffusion processes can be modeled as a one-dimensional transient problem with respect to a spherical co-ordinate system. As the surface density of the probes stays the same the radius of the semi-spherical dome a and the radius of the flat spot b are related by:

$$2\pi a^2 = \pi b^2 \rightarrow a = \frac{b}{\sqrt{2}} = \frac{1}{2} b\sqrt{2} \tag{17}$$

The components of the non-stationary diffusion equation (being Fick's second law) are defined with respect to a spherical co-ordinate system: r, θ, φ [12]. The only component that remains after carrying out the steps outlined above is the *r*-component.

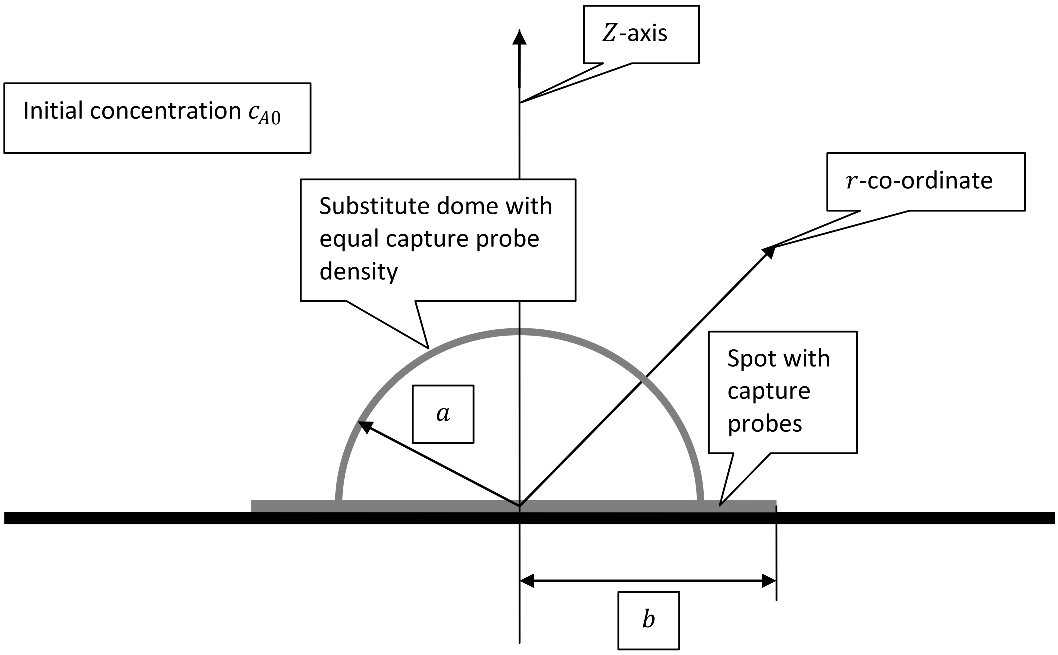

Z-axis
Initial concentration c_{A0}
Substitute dome with equal capture probe density
r-co-ordinate
a
Spot with capture probes
b

Figure 6.2 Definition of substitute semi-spherical dome carrying the capture probes. The surface areas of the spherical cap and the original spot surface are equal, thereby keeping the capture probe density equal.

In order to understand the hybridization kinetics at the spot surface we use the Langmuir kinetic model[14-16] (θ is the fraction of hybridized probes at the spot surface with respect to the total number of capture probes, k_{ass} and k_{diss} are the association and dissociation rate constants and $c_{r=a}$ is the concentration of targets in the bulk at the surface of the spot):

$$\frac{d\theta}{dt} = k_{ass}c_{r=a}(1-\theta) - k_{diss}\theta \tag{18}$$

During the first number of cycles, the fraction of hybridized targets is very small (no competition) and the Langmuir kinetic equation can be approximated by:

$$\frac{d\theta}{dt} = k_{ass}c_{r=a} \tag{19}$$

This equation can be rewritten to:

$$\frac{d\theta n_{probes}/n_{avogadro}}{\pi b^2 dt} = -N_{r=a} = k_{ass}c_{r=a}\frac{n_{probes}}{n_{avogadro}}\frac{1}{2\pi a^2} \tag{20}$$

In the bulk the molar flux of the targets is given by:

$$N_r = -\mathcal{D}\frac{\partial c}{\partial r} \tag{21}$$

The flux boundary condition at $r = 0$ (being referred to as a radiation condition, that holds true for any t) is given by:

$$\frac{\partial c}{\partial r}|_{r=a} = \frac{k_{ass}}{\mathcal{D}}\frac{n_{probes}}{n_{avogadro}}\frac{1}{2\pi a^2}c_{r=a} = hc_{r=a} \tag{22}$$

In this equation, h is a short hand writing for the constant in the radiation condition. For the case that free species A can hybridize, which is the default case, c should be replaced by c_A.

6.3 Results and discussion

6.3.1 Geometrical, process and material data

a:	radius semi-spherical dome [m]
b:	radius spot [m]
c^*:	dimensionless concentration [-]
c_A:	molar concentration species A (labeled sense ss DNA) [mol/m^3]
c_B:	molar concentration species B (unlabeled antisense ss DNA) [mol/m^3]
$c_{complex_1}$:	molar concentration $complex_1$ [mol/m^3]
$c_{complex_2}$:	molar concentration $complex_2$ [mol/m^3]
c_{dsDNA}:	molar concentration double stranded DNA (ds DNA) [mol/m^3]
c_{dNTP0}:	initial concentration dNTP's [mol/m^3] (for each dNTP's : dATP, dCTP, dGTP and dTTP, total concentration of triphosphates 4 c_{dNTP0}), typically our protocol uses 0.2 mol/m^3 per dNTP
c_{dNTP}:	concentration dNTP's during process [mol/m^3] (for each dNTP's : assuming equal consumption of dATP, dCTP, dGTP and dTTP, total concentration of triphosphates 4x c_{dNTP}
c_0:	initial concentration of species A or B at the beginning of a cycle [mol/m^3]
c_{p0}:	initial concentration of forward and reverse primers [mol/m^3] (total initial concentration primers 2c_{p0})
c_{pF}:	concentration forward primers [mol/m^3]
c_{pR}:	concentration reverse primers [mol/m^3], typically $c_{pF} = c_{pR}$
c_{TAQ}:	concentration TAQ polymerase [mol/m^3]
$\mathcal{D}$:	diffusion coefficient [m^2/s]
f_{probes}:	reduction of active probes on the microarray per cycle [-]
f_{Taq}:	reduction of active enzyme per cycle [-]
E :	PCR efficiency [-]
E_{M-M}:	elongation factor based on Michaelis-Menten enzyme kinetics [-]
k_{ass}:	association rate constant Langmuir surface reaction [m^3/s/mol]
k_{diss}:	dissociation rate constant Langmuir surface reaction [1/s]
k:	annealing rate constant binary complexes [m^3/s/mol]
k_{dsDNA}:	renaturation rate constant [m^3/s/mol]
k_1:	association rate constant tertiary complexes [m^3/s/mol]
k_{-1}:	dissociation rate constant tertiary complexes [1/s]
k_2:	rate constant enzyme reaction [1/s]
K_{mDNA}:	Michaelis-Menten constant for tertiary complex formation enzyme DNA [mol/m^3]

K_{mdNTP}:	Michaelis-Menten constant for tertiary complex formation enzyme dNTP [mol/m^3]
n:	cycle number [-]
N_{base}:	number of base pairs to be elongated [-]
N_r:	molar flux species in r-direction [mol/m^2/s]
n_{cycle}:	number of templates captured on spot surface [-]
n_{probes}:	number of probes in spot [-]
R:	radius semi-spherical control volume [m]
r:	radial co-ordinate measuring the distance from the centre of the impermeable spot surface [m]
R_{max}:	elongation rate constant [mol/m^3/s]
ρ:	density [kg/m^3]
$n_{avogadro}$:	Avogadro's number [6.022*10^{23} /mol]
θ:	fraction of hybridized probes [-]
θ_{eq}:	equilibrium fraction of hybridized probes [-]

The model developed as described in this chapter allows for a parameter study to investigate the influence of the different parameters involved around a standard protocol. The solutions and the graphs as presented in this chapter all make use of the following starting points, unless stated otherwise:

- Initial amount of templates (targets): 100.000 copies
- Initial load enzyme: 1 Unit
- Decay rate enzyme activity: 2 % per cycle
- c_{p0} : $0.3*10^{-3}$ mol/m^3
- c_{dNTP0}: 0.2 mol/m^3
- k: 60 m^3/mol/s
- k_{dsDNA}: 30 m^3/mol/s
- R_{max} $1.59*10^{-6}$ mol/m^3/s
- K_{mDNA} $3.68*10^{-5}$ mol/m^3
- K_{mdNTP} $15*10^{-3}$ mol/m^3
- $\mathcal{D}$ $8.54*10^{-11}$ m^2/s
- Annealing / hybridization time: 120 seconds
- Elongation time: 40 seconds
- Denaturation time: 20 seconds
- Total number of cycles: 60

The rationale behind the values of these constants and some additional constants is given below.

The annealing is described by a second order reaction. For k the Wetmur/Davidson relation[10] will be used to obtain a first estimation of the annealing constant

$$k = 300\frac{\sqrt{L}}{N}\ [\frac{\mathrm{m}^3}{\mathrm{mol\ s}}] \qquad (23)$$

This relation holds true under strict conditions (salt concentration 1 M [Na$^+$] and 25 °C below the melting temperature).The value of k strongly depends on the salt concentration. The number of bases is denoted by L. The value of N is referred to as the complexity, the number of base pairs in non-repeating sequences, for short strands: $N = L$. The number of base pairs in the amplicons in our case is about 100. In case of strands of different lengths anneal, the data for the shortest should be entered, which is the length of the primers (20-40 bases) leading to the conclusion that the rate constant for annealing is about 50-70 m^3/mol/s. Though the experimental salt conditions deviate from the conditions described by the Wetmur/Davidson relation, the values obtained are nicely in line with the literature[16-18])). In our calculation a value of $k = 60$ m^3/mol/s is used and for the renaturation reaction with strands of 100 bases, $k_{dsDNA} = 30$ m^3/mol/s.

This annealing reaction rate can be used to make a first estimation of the annealing time constants: the solution of equation 10, which gives the concentration of species A (the same reasoning holds true for species B) as a function of time, is given in the appendix in equation A1. This solution: $c_A = c_{A0}e^{-kc_{p0}t}$ can be re-written into $c_A = c_{A0}e^{-t/\tau}$, with $\tau = 1/(kc_{p0})$, and τ being a time constant. With the reaction rate constants and primer concentrations used as mentioned above, after a typical annealing time of 120 seconds, almost 90 % of the primer annealing is complete. In the experiments of chapter 5, a PCR efficiency of almost 100 % is achieved, meaning that this annealing rate constant might be on the lower side. A further analysis of the impact of changing the annealing constants is given in the concluding section which lists the potential improvements of the model.

In reference 16, data are reported that suggest that the surface hybridization reaction is 20-30 times slower than the bulk reaction due to steric hindrance, and therefore the following is used:

$$k_{ass} = k/20 = 3\ \mathrm{m}^3/\mathrm{mol/s} \qquad (24)$$

One enzyme unit (U) is defined as the amount of enzyme which will convert 10 nmoles of dNTP‘s to an acid-insoluble form in 30 minutes at 72°C under the assay conditions. From this definition we can derive the rate constant for the elongation process, driven by 1 U polymerase (for 70 base pairs):

$$R_{max} = \frac{10*10^{-9}}{30*60}\frac{1}{V_{PCR\,chamber}}\frac{1}{N_{base}} = 3.17 * 10^{-6}\ [\mathrm{mol/m^3/s}] \qquad (25)$$

As there are two complexes in the mix, in order to calculate the appropriate constants of the Michaelis-Menten equation, per complex half of the number of units must be entered.

The value of the rate constant determining the formation of tertiary complexes between the complex DNA and the Taq polymerase cannot be easily derived since hardly any literature can be found, nor can it be easily deduced from experimental data. Therefore, an optimization of the model and comparison with the experiments has been made. Using the following rate constants based on the Wetmur/Davidson relation[10], a reasonable match with experiments was obtained and these were therefore used in the model.

$$k_1 = 325\left[\frac{m3}{mol*s}\right];\ k_{-1} = 0[\tfrac{1}{s}] \tag{26}$$

To calculate k_2 the number of enzymes per unit must be entered. In reference 8, a value of 4.10^{12} is mentioned. Other numbers up to a factor of 100 different can be found[19, 20]. In order to have a consistent set of data that comply with the experiments, the value given by reference 8 has been used here. So for 1 Unit the following relation is valid:

$$k_2 = \frac{R_{max}}{c_{Taq}} = \frac{3.17*10^{-6}}{2*\,4*10^{12}}\, n_{avogadro}\, V_{PCR\ chamber} = 1.19*10^{-2}[\tfrac{1}{\mathrm{s}}] \tag{27}$$

$$K_{mDNA} = \frac{k_{-1}+k_2}{k_1} = 3.68*10^{-5}\ [\tfrac{\mathrm{mol}}{\mathrm{m}^3}] \tag{28}$$

For K_{mdNTP}, for each of the dNTPs a value of $10\text{-}15*10^{-3}$ mol/m^3 is found in literature[21].The diffusion coefficient for ss DNA of 100 nucleotides in length at 55 ^{o}C is taken equal to[22]:

$$\mathcal{D} = 8.54*10^{-11}\ [\tfrac{\mathrm{m}^2}{\mathrm{s}}] \tag{29}$$

During thermal cycling the polymerase degrades and per cycle some amount is lost. To take this effect into account an assumption is made that per cycle 2 % of the catalytic properties of the enzyme content is gone. (This 2 % is larger than would be expected in real time PCR, since the temperatures are not yet as strictly defined as in real time PCR, and a larger amount of enzyme is likely to be lost due to adsorption to the wall.) This effect only affects the value of R_{max} according to (n cycle number):

$$R_{max}(n) = 3.17*10^{-6}\,(1-f_{Taq})^n\ [\mathrm{mol/m^3/s}] \tag{30}$$

In case a certain percentage of the probes deactivates during thermal cycling, the radiation condition will be affected since the number of probes is present in this equation. Since the assumption is made that the value of k_{ass} is independent on the surface density of the capture probes, only the radiation condition is influenced.

Given the reduction of active probes by f_{probes} , the radiation condition becomes:

$$\frac{\partial c}{\partial r}\Big|_{r=a} = h(1-f_{probes})^{n} c_{r=a} \tag{31}$$

The decay of the activity of the capture probes will initially be set to zero. In the results and discussion section a series of results with a decay rate for the capture probe activity is included.

6.3.2 Typical results: bulk results

In Appendix A, the solutions for real time PCR and real time array PCR are derived. The presented solutions allow monitoring the concentrations of the different species during the PCR process. A typical result is shown in figure 6.3.

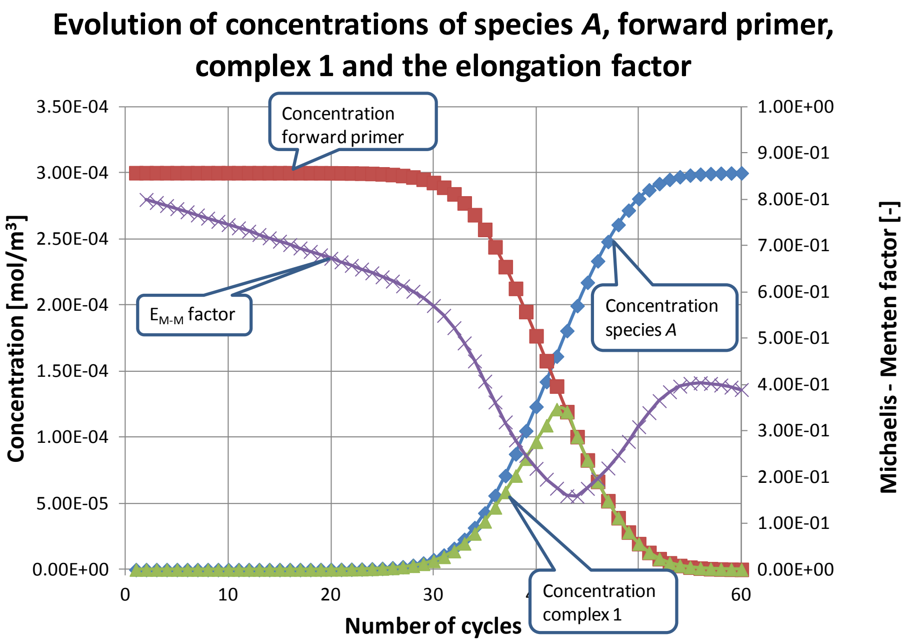


Figure 6.3 Evolution of concentrations of species A, forward primer and $complex_1$ during thermal cycling. Also displayed is the Michaelis-Menten factor E_{M-M} (see Appendix A).

From the results as displayed in figure 6.3 it can be seen that the calculated Michealis-Menten factor at the start of the thermal cycling is about 0.8, which is in line with the value given by reference 8. Initially the concentration of the substrate ($complex_1$) is low ($c_{0,complex} \ll K_m$), and the $E_{M\text{-}M}$ factor does not depend on the concentration of $complex_1$. The gradual decrease of the $E_{M\text{-}M}$ factor is due to the decrease of activity of the enzyme (2 % per cycle). After a number of cycles, the concentration of complexes increases to such an extent that the activity of the enzyme slows down because of the high concentration of complexes. At the end of the PCR process because of the lack of primers the complex concentration is that low that the activity of the enzyme increases again. The elongation rate, however, remains low since only a few primers are still elongated.

6.3.3 Typical results: microarray surface hybridization

The solution for the reaction equation for the microarray hybridization is given in the Appendix. As explained earlier, only free species A can bind to the capture probes. A typical example of the concentrations of species A develops over time and space above the spot is shown in figure 6.4.

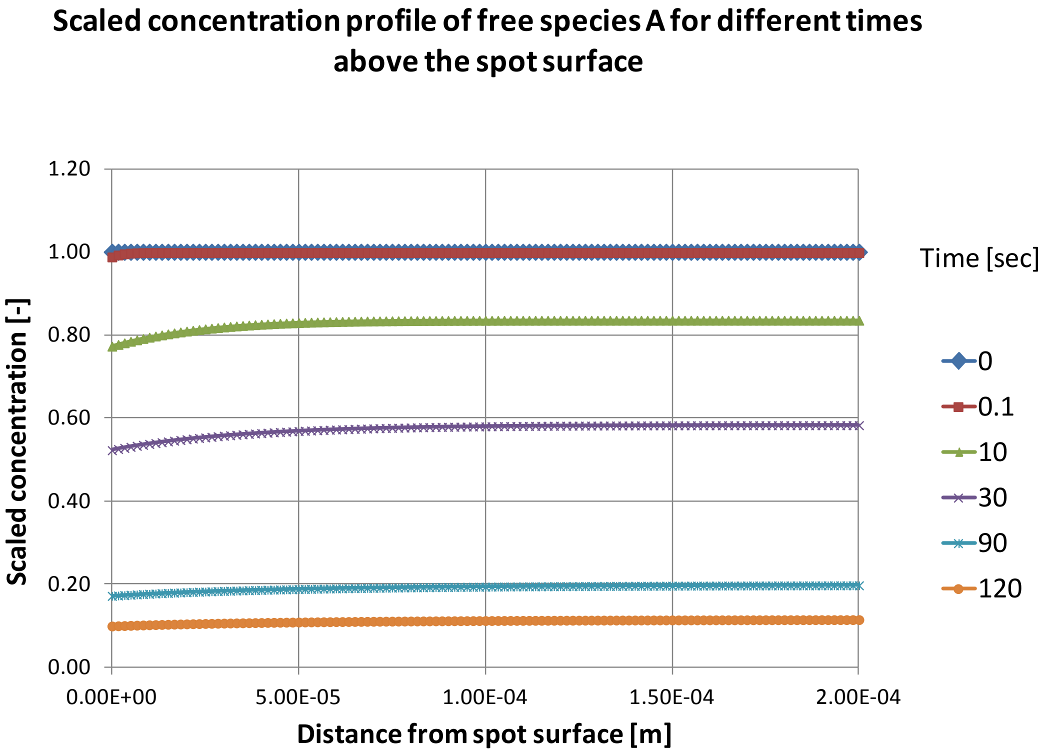


Figure 6.4 Normalized (with respect to the initial concentration of species A in the bulk c_{A0}) concentration profiles above spot surface (spherical cap). In the legend the times related to the different concentration profiles are indicated.

In this figure, it can be seen that the concentration of free species A decreases in course of the hybridization time, due to the renaturation reaction. Furthermore, since the reaction rate at the surface of the microarray is relatively low, the concentrations close to the microarray surface decrease only to a limited extent.

Figure 6.5 shows the fraction of capture probes hybridized in course of time with targets diffused from the bulk to the array surface at the end of the process (which means that the concentration of target molecules is equal to the initial concentration of primers). Clearly is visible that the surface concentration has not yet reached the maximal value, so within 120 seconds, no steady state is reached.

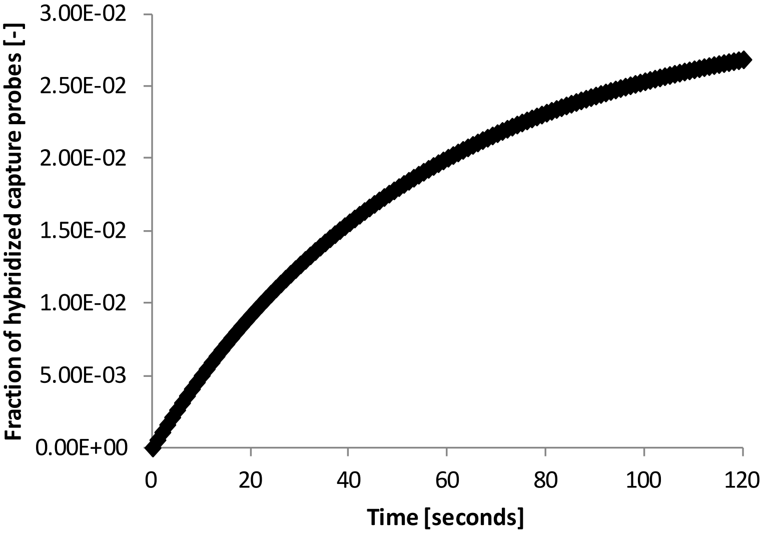


Figure 6.5 Time evolution of the fraction occupied capture probes on the microarray surface at the end of the last PCR cycle.

6.3.4 Optimization of experimental parameters

The model described allows for a parameter study to investigate the influence of the different parameters involved around a standard protocol. The parameters used in the standard protocol have been mentioned in paragraph 6.3.1. Around the standard protocol the sensitivity of the amplification curve on the following parameters is investigated:

- The annealing/hybridization time has been varied between 0 and 180 seconds
- The initial load of enzymes (with a decay rate in enzyme activity of 2 %)
- The decay rate of the enzyme load (for 1 unit Taq polymerase)
- The decay rate of the capture probes
- Initial load of primers (decay rate activity enzymes 2 % per cycle and initial number of templates 1.10^5)
- Initial load of dNTP's (decay rate activity enzymes 2 % per cycle and initial number of templates 1.10^5)

All results are shown either for the concentration of species *A* in the bulk (for bulk results) at the beginning of the annealing phase, or as the fraction of capture probes occupied on the microarray surface at the end of the annealing phase (for hybridization results).

In the subsequent paragraph, the outcomes of the model are compared with the experimental data as given in chapter 5. The C_t-values for a dilution series will be investigated for the qPCR and real time array PCR.

6.3.4.1 Optimization of annealing/hybridization time

The amplification efficiency is defined as the ratio between the concentration of species *A* at the beginning of the annealing phase of cycle *n* with the concentration of species *A* at the beginning of the previous annealing phase during cycle *n-1*:

$$Eff_{amplification} = 100\ \% * \frac{c_{a,n}}{c_{A,n-1}} \tag{32}$$

The amplification efficiency increases with longer annealing times, visualized in figure 6.6.

In figure 6.7, the concentration of species *A* in the bulk at the beginning of each annealing step is given for different hybridization/annealing times. It can be seen that for the case of the shorter times, amplification is not optimal due to the lower annealing efficiency and therefore, the C_t-values are higher. Increasing the annealing/hybridization times results in an improved amplification.

Figure 6.6 and 6.7 present the bulk results; in figure 6.8, the results of the hybridization onto the capture spots are given. Compared to the results in figure 6.6, it can be seen that there is a two-fold positive effect of increasing the annealing/hybridization time: next to a better annealing process, the hybridization process improves since more time is available which leads to a higher fraction of occupied capture probes. To increase the annealing/ hybridization time thus has a double advantage since both annealing and hybridization are improved.

However, there is a trade-off between the optimal efficiency and the overall time to results, since an increased annealing time leads to a longer overall turn-around time, which is undesirable from a clincal perspective. From the results as given in these two figures, a minimum time of around 100 seconds is preferred in order to reach sufficient amplification without too much giving in on reaction times. For the experiments as described in chapter 5, and also for the remaining calculations, an annealing time of 120 seconds is used.

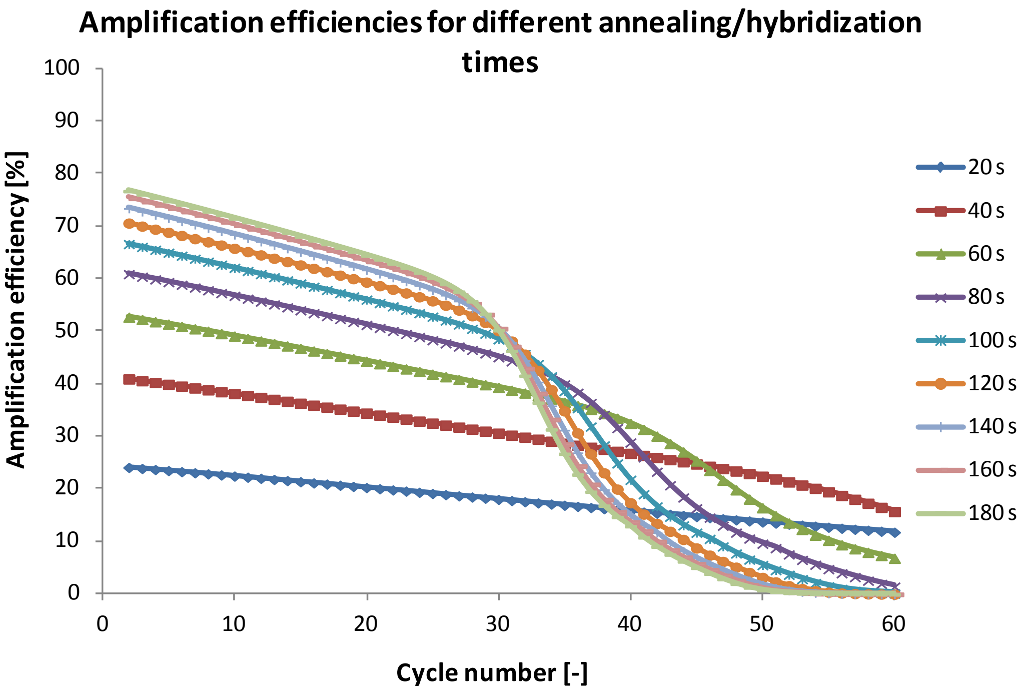


Figure 6.6 Amplification efficiencies as defined in equation 32 for different annealing/hybridization times.

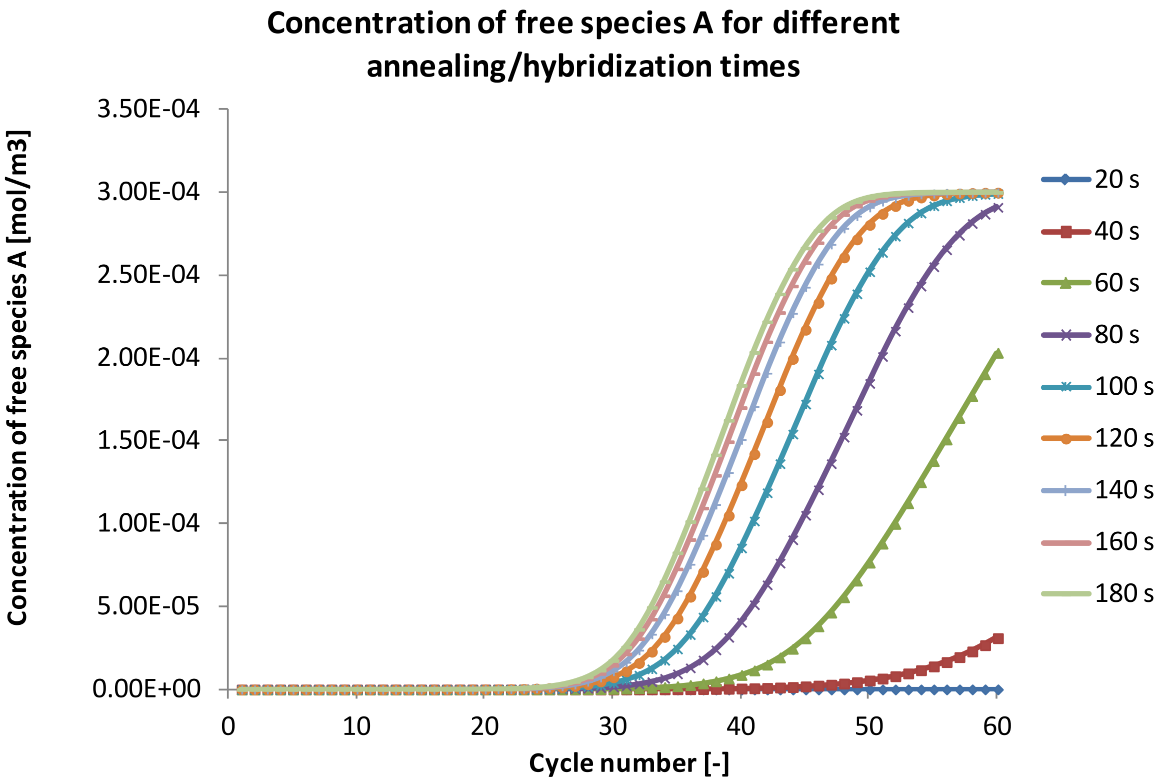


Figure 6.7 Concentration of species *A* at the end of each denaturation phase for different annealing/hybridization times.

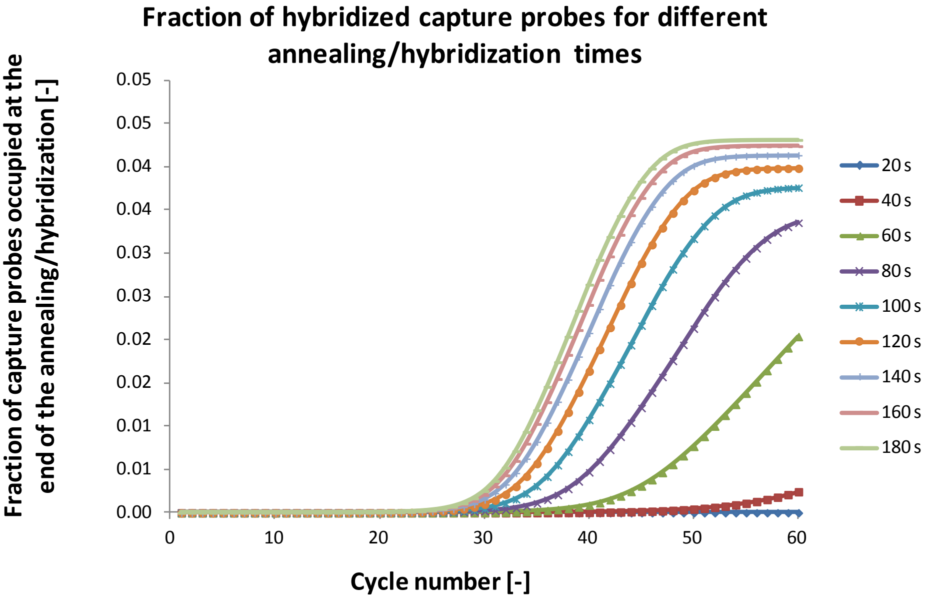


Figure 6.8 Fraction of hybridized capture probes at the end of the annealing phase for different annealing/hybridization times.

The elongation and denaturation times of the amplification process have not been considered in this analysis. The denaturation is considered to be 100 %; the model does not include the denaturation kinetics.

6.3.4.2 Influence of concentration of Taq enzyme

Figure 6.9 gives the amplification curves in terms of free species *A* at the end of each denaturation step as a function of the initial concentration of Taq enzymes.

Since the PCR enzyme is the most expensive components of the assay, the concentration should preferably be minimized. Judging the results as presented below, an enzyme load of 1 Unit per reaction should result in reasonable results when not too much enzyme is lost during amplification.

6.3.4.3 Influence of decay rate of the enzyme

As explained in the introduction, enzyme activity can be lost by thermal degradation or by enzyme adsorption to the glass surface. Enzymes are extremely sensitive for high temperatures. Following reference 8, the half time of Taq enzyme at 95 °C is 40 minutes (meaning for 15 seconds denaturation time, around 0.5 % of the enzyme is lost). At 97.5 °C, however, this is reduced to only 5-6 minutes (if the whole PCR chamber would be at this temperature, a loss of 3 % per cycle would be obtained). So a small increase from the set melting temperature will drastically change the amplification curve. For the experimental setup as described in chapter 5, temperatures higher than 95 °C cannot be fully excluded during experiments since the heater has a small offset of a few degrees Celsius higher compared to the real time array PCR chamber. Though the effect of this offset has been characterized (not shown), the actual temperature profile everywhere inside the PCR chamber is not fully known and therefore thermal degradation higher than conventional real time PCR cannot be excluded.

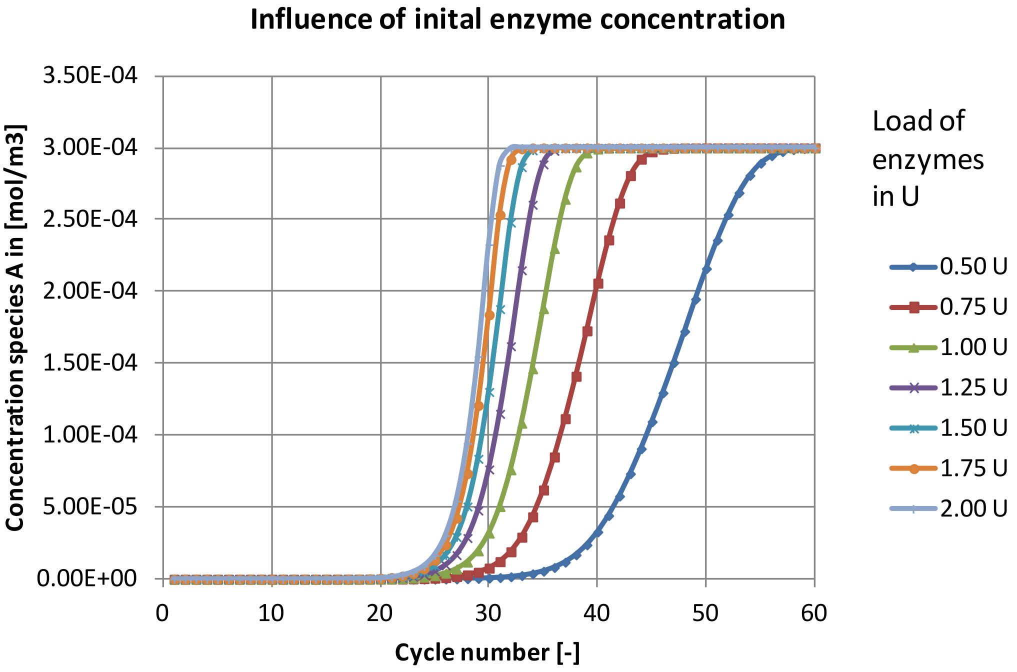


Figure 6.9 Concentration of species *A* for different initial concentration of Taq polymerase. In this case, no decay of the enzyme has been taken into account.

Concerning the adsorption of enzymes: for real time array PCR, this effect is not unlikely since glass is used for the microarray surface and proteins and enzymes generally adsorb onto glass surfaces[9]. Compared to real time PCR, the effect of enzyme adsorption to the wall might be higher: for real time PCR generally a low DNA binding PCR plate is used which is optimized in its surface to volume ratio. The flat PCR chamber as explained in chapter 5 is in that respect less favorable. Though BSA is added to the reaction mix to reduce surface adsorption, it may not be prevented fully that part of the reaction components and especially the enzyme adsorb onto the glass. In order to investigate the effect of protein adsorption and thermal degradation onto the amplification curves, this effect is visualized in figure 6.10.

Figure 6.10 gives the net effect if the enzyme activity decreases with a certain percentage and shows that effects that are likely to occur can have a large effect on the PCR efficiency especially when more than 1 % of the enzyme is degraded per PCR cycle. Minimizing the decay in enzyme activity can thus significantly improve the amplification and thus the hybridization results.

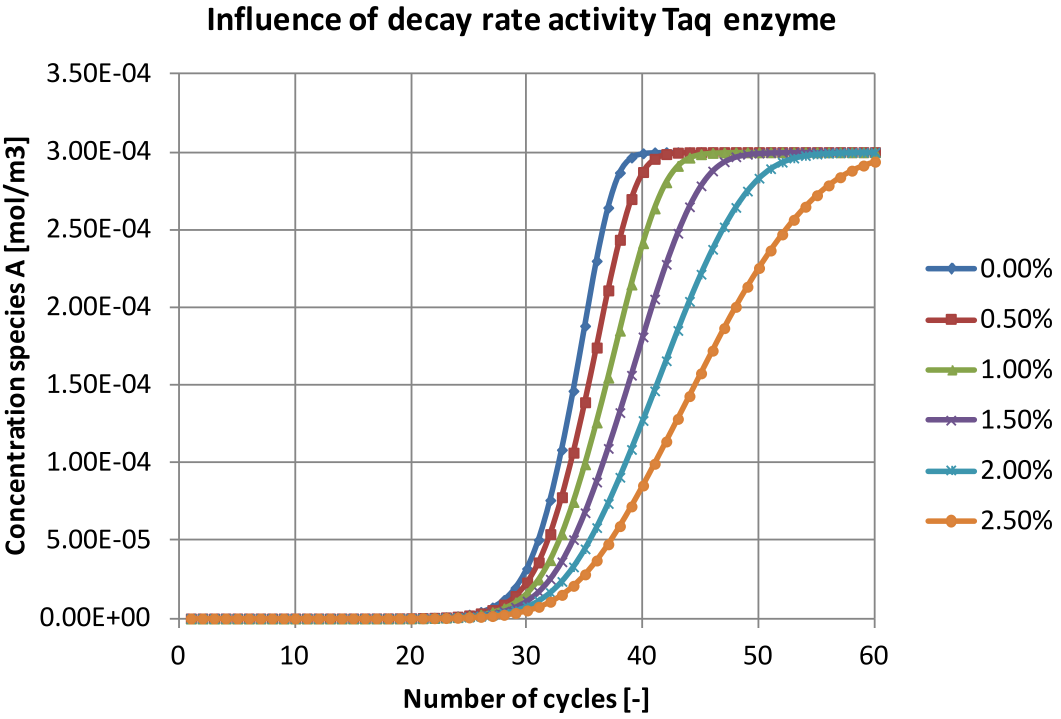


Figure 6.10 Influence of the decay rate of enzyme on the amplification curve. The points indicate the concentration at the start of the annealing/hybridization.

6.3.4.4 Influence of decay of capture probes

Not only enzymes may become inactive during thermal cycling, the same may be true for the capture probes. The effect on the amplification curves can be considerable as is displayed in figure 6.12. There is only a very small influence on the C_t-value, the most pronounced effect is the decrease in fluorescence in the plateau phase. This has indeed been observed in experiments (data not shown) and a decrease in amount of available intact capture probes might be the explanation for this.

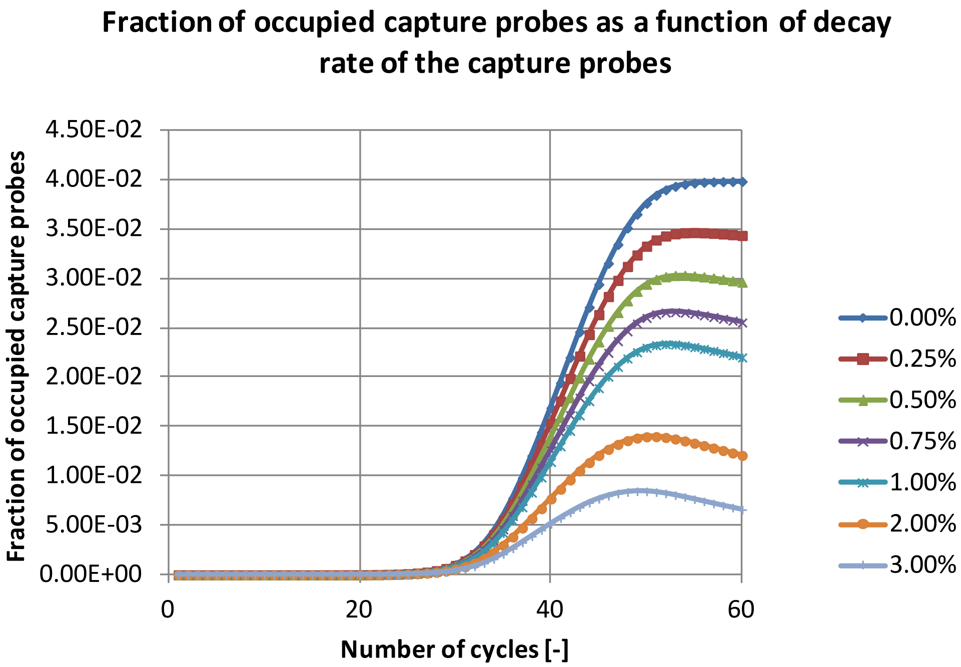


Figure 6.11 Fraction of capture probes occupied for different decay rates of the capture probes (the fraction is related to the initial number of probes).

6.4 Comparison with experiments

In this section, the model as explained in the previous paragraphs is compared with the experiments. This is done to investigate if the parameters included in the model can explain the performance differences between the qPCR and the real time array PCR as explained in chapter 5 (figure 5.7). For the qPCR experiments, a situation is modeled where no enzyme degradation takes place and no initial enzyme loss due to non-specific adsorption to the wall. This no-degradation case may be considered as representative for a standard qPCR process, because of the tight temperature control, highly temperature stable Taq and the use of plastic vials. For the array PCR, a 1 % decay of activity of Taq per cycle is included as a first starting point to correct for degradation of the enzyme. The 1 % decay case may be close to the real time array PCR process because of the potentially somewhat higher temperature in the bulk and the use of glass as material for the cartridge containing the reaction chamber. In order to estimate the C_t-values based on the amplification curves, a concentration of 1.10^{-5} mol/m^3 (10 nM) has been assumed as a cut-off value.

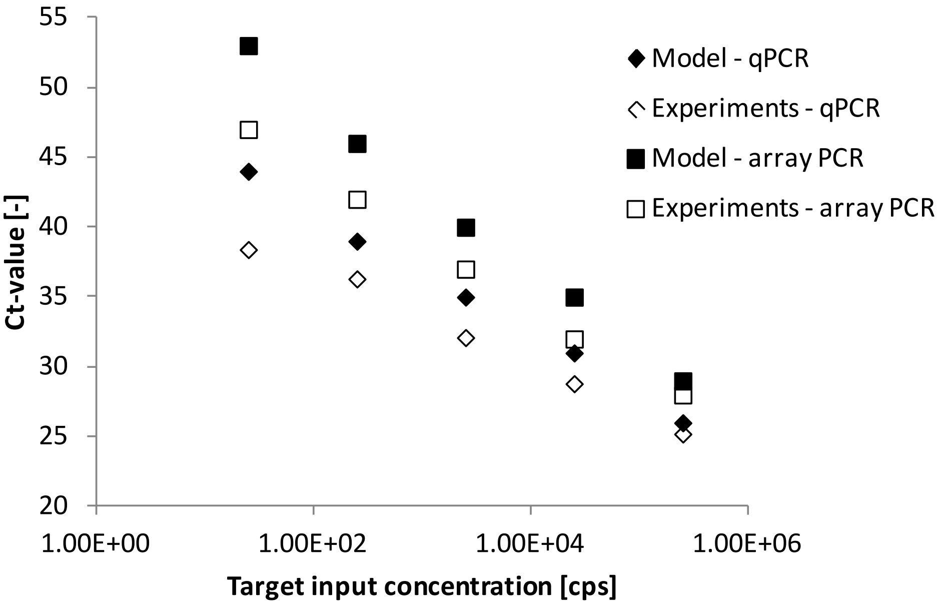


Figure 6.12a C_t-values as a function of target input concentrations for the qPCR model (modeled as no enzyme degradation) and experimental data, and for the array PCR model (modeled as 1 % enzyme degradation per cycle) and experimental data.

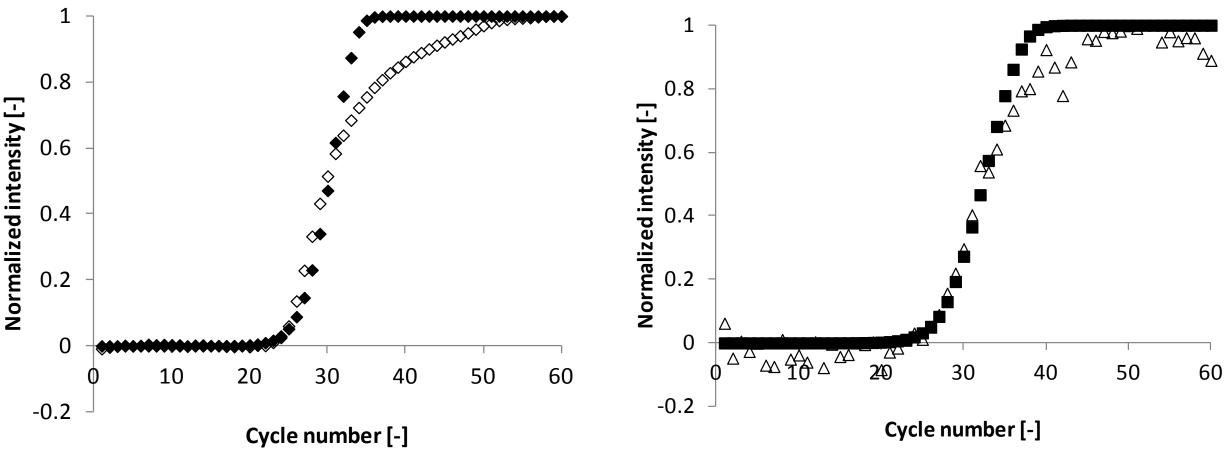


Figure 12b Comparison between the experimental data (open markers) and the model (closed markers) of qPCR (left) and the real time array PCR (right). In this case 2.10^5 copies were used as target input concentration.

In figure 6.12a, it can be seen that the differences in C_t-value between the qPCR and the array PCR can also be predicted by the model when including an enzyme activity decrease of 1 % per cycle.

The absolute C_t-values as obtained in the experiments are lower than the values predicted by the model. This can be largely attributed to an incomplete annealing when using an annealing constant of 60 m^3/mol/s. in the section below, a parameter sensitivity analysis has been done in order to identify the most critical parameters. The value chosen as the default value of the model (60 m^3/mol/s) is most probably too low since annealing at the end of the annealing time is incomplete, while in the experiments full annealing is observed. At a reaction rate constant of around 100 m^3/mol/s, full annealing has been accomplished. When including this value in the model and keeping all other parameters identical, the experiments can be described still rather well by the model for both the qPCR as the array PCR situation, see figure 6.13. At the lowest input concentrations, the difference between the experiments and the model is the largest.

Figure 6.12b gives the comparison between the experiments and the model. It can be seen that the model predicts the experimental data reasonable well. In the qPCR curve, it can be seen that the normalized intensity starts to deviate at the higher cycle numbers. The amplification efficiency is lower in the experiments, which might be due to some enzyme degradation taking place, while the model does not take any enzyme degradation at all into account. The real time array PCR curve for the experiment and the model show a good comparison.

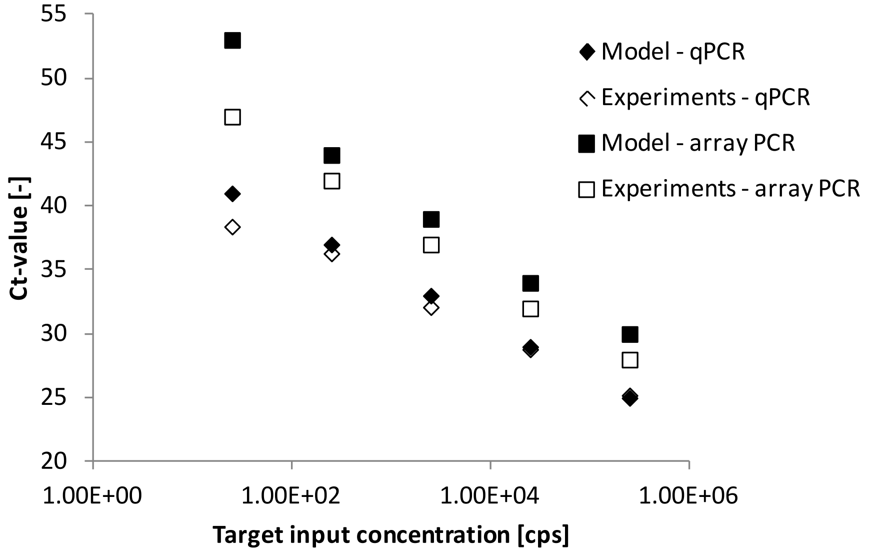


Figure 6.13 C_t-values as a function of target input concentrations for the qPCR model (modeled as no enzyme degradation) and experimental data, and for the array PCR model (modeled as 1 % enzyme degradation per cycle) and experimental data. In this case, an annealing constant of 100 m^3/mol/s has been chosen.

There is an off-set between the C_t-values as predicted by the model and the experimental data. This can e.g. be explained by the dependency on the cut-off value, used to determine the C_t-value for the bulk results.

6.5 Parameter sensitivity analysis

In this section, a short analysis of the sensitivity of parameters included on the model is given. This has been executed to identify the most critical parameters of the model focused on the surface kinetics to verify that the loss in and degradation of the polymerase enzyme are the most likely factor to explain the differences in performance between the qPCR and array PCR as observed in the first part of chapter 5.

Different parameters and their effect on the PCR results were investigated.

- Diffusion coefficient at 53 ^{0}C: in the model, a value of $8.54.10^{-11}$ m^2/s was used. Within a window of 1.10^{-12} and 1.10^{-9}, no differences in C_t-value were found.
- The annealing rate constant has been varied between 30 and 100 m^3/mol/s. The results are given in figure 6.14. Though the annealing constant has a profound effect on the amplification efficiency, since it mainly influences the bulk kinetics, it cannot explain the differences in surface kinetics.
- The association constant was varied between 1 and 10 m^3/mol/s. The association constant is an important factor in the hybridization rate. However, a change in the association rate does not explain the differences between qPCR and real time array PCR, since only the absolute C_t-values are influenced by the association rate constant, while the PCR efficiency turned out to be hardly dependent thereon.
- Since the whole Michaelis-Menten kinetics influences the bulk reaction kinetics, it cannot explain differences between qPCR and real time array PCR. Therefore, variations in the Michaelis-Menten kinetics are not taken into account in this chapter.

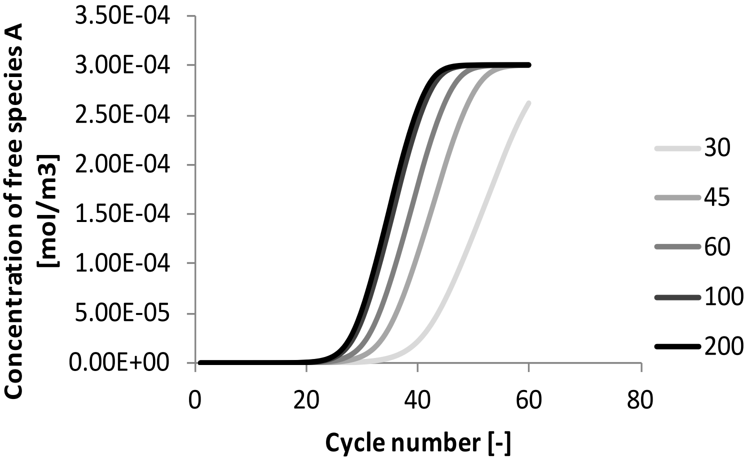


Figure 6.14 Concentration of free species A for different annealing constants. At around 100 m^3/mol/s, 100 % annealing is accomplished.

6.6 Discussion and conclusion

In this chapter, a model has been presented that describes the biochemical processes taking place during real time array PCR. This model combines two processes that have been described individually and combines these into one model: the processes taking place in the bulk as in real time PCR, and microarray hybridization. Thereby, it is possible to perform a parameter study and investigate the most important parameters taking place. Furthermore, the influence of bulk properties that change throughout every PCR cycle on the microarray hybridizations can be studied.

The annealing time has a positive effect on the concentration of hybridizable strands (optimized annealing and thus more targets available) as well as on the hybridization signals since more time is available for hybridization. Therefore there is a two-fold gain in signal when this time is extended. Increasing the elongation time also leads to higher target concentrations. There is, however, a trade-off between increasing the reaction times for optimized reaction efficiencies and the time needed to do a full analysis, which is from a clinical aspect an important factor to minimize.

The most important experimental parameters to take into account are the concentration of Taq polymerase enzyme and potential degradation of this enzyme. Either thermal degradation or enzyme adsorption to the cartridge surfaces are likely to occur in real time array PCR.

The model as presented here is not yet complete and does not fully describe the experimental conditions of chapter 5. Specific potential improvements are identified below:

- The model is not yet depending on the temperature, but has been built around a standard 3-step PCR protocol using temperatures of 95, 55, and 72 ^{o}C for the denaturation, annealing and elongation, respectively using material and reaction rate data at these temperatures. These temperatures have been optimized during the experimental phase and thus were used as input for the model. E.g. to make the model temperature-dependent would lead to a temperature-dependency of the diffusion constant, using the Einstein-Stokes relation[14].
- The elongation factor $E_{M\text{-}M}$ and the overall amplification efficiencies found in the model (using 1 U and no thermal degradation) are lower than the overall efficiency achieved in the experiments (close to 100 %). To create the model, data were needed on the reaction rates and number of molecules of enzyme per unit activity. In general, the concentration of enzymes per unit is cannot be found readily for different polymerases, and for the enzymes used in the experiments described in chapter 5 no data are available on the reaction rates. Assumptions had to be made based on data present in literature. There is a large variety in quality of the polymerases available which impact the reaction rate constants. Recent developments in enzyme design and formulation have produced improved thermostable polymerases that can form more efficiently

tertiary complexes (leading to a higher k_1). The reaction rate constants used in the Michaelis-Menten kinetics in the model might be for lower quality enzymes than actually used in the experiments. Another effect to consider is the annealing rate constant for the formation of complexes between single stranded target molecule and primers, which is not 100 %. In the experiments, 100 % efficiency was reached. When these experimentally found results are included in the model by increasing the reaction rate, the efficiency also approaches 100 %.

- It has been assumed that in the PCR chamber that the temperature is constant and equal everywhere. The PCR chamber, however, is heated from one side. This means that in practice there is always a temperature gradient across the PCR chamber. This temperature difference can power thermal convection

To conclude: in general, the data set used for material properties and reaction constants is consistent in the sense that it is possible to describe the experimental findings listed in chapter 5. The actual numbers used were based on literature data. Since there are different sources with varying content for these numbers, the actual numbers may be different from the ones used in this chapter. For this analysis, however, the trends were studied and not too much focus has been put on absolute values. The trends could, with the data set as applied, be investigated rather well.

References

1. Remacle J, Alexandre I, Margaine S, Husar D. 2006, Real-time quantification of multiple targets on a micro-array. Eppendorf Array Technologies patent EP1659183
2. Remacle J, Alexandre I, Koehn H, Seippel M, 2007. Lid for PCR vessel comprising probes permitting PCR amplification and detection of the PCR product by hybridisation without opening the PCR vessel, Eppendorf Array Technologies patent EP1788095
3. Booth CS, Pienaar E, Termaat JR, Whitney SE, Louw TM, Viljoen HJ. Efficiency of the polymerase chain reaction. *Chem Eng Sci* 2010, 65: 4996-5006
4. Gevertz JL, Dunn SM, Roth CM. Mathematical Model of Real-Time PCR Kinetics. *Biotechnol Bioeng* 2005, 92(3):346-355
5. Pappaert K, Van Hummelen P, Vanderhoeven J, Baron GV, Desmet G. Diffusion-reaction modelling of DNA hybridization kinetics on biochips, *Chem Eng Sci* 2003, 58:4921-4930
6. Whitney SE, Sudhir A, Nelson RM, Viljoen HJ. Principles of rapid polymerase chain reactions: mathematical modeling and experimental verification. *Comp Biol Chem* 2004, 28:195-209
7. Gadgil C, Yeckel A, Derby JJ, Hu W-S. A diffusion-reaction model for DNA microarray assays. *J Biotech* 2004, 114:31-45
8. Sambrook J, Russell DW. Molecular Cloning, 2001. Cold Spring Harbor Laboratory Press, 3rd edition
9. Erill I, Campoy S, Erill N, Barbe J, Aguilo J. Biochemical analysis and optimization of inhibition and adsorption phenomena in glass-silicon PCR chips. *Sensor Actuat B-chem* 2003, 96, 685-692
10. Wetmur JG, Davidson N. Kinetics of renaturation of DNA*. J Mol Biol* 1968, 31, 349-370
11. Truskey GA, Yuan F, Katz DF. Transport Phenomena in Biological Systems, 2010. Pearson Prentice Hall Bioengineering, 2nd edition
12. Bird RB, Stewart WE, Lightfoot EN. Transport Phenomena, 2002. John Wiley & Sons, 2nd edition
13. Fiche JB, Buhot A, Calemczuk R, Livache T. Temperature Effects on DNA Chip Experiments from Surface Plasmon Resonance Imaging: Isotherms and Melting Curves. *Biophys J* 2007, 92, 935-946
14. Mocanu D, Kolesnychenko A, Aarts S, De Jong AT, Pierik A, Coene W, Vossenaar E, Stapert H. Quantitative analysis of DNA hybridization in a flow-through microarray for molecular testing. *Anal Biochem* 2008, 380(1):84-90
15. Mocanu D, Kolesnychenko A, Aarts S, De Jong AT, Pierik A, Coene W, Vossenaar E, Stapert H. Mass transfer effects on DNA hybridization in a flow through microarray. *J Biotechnol* 2009, 139(2):179-185

16. Gao Y, Wolf LK, Georgiadis RM. Secondary structure effects on DNA hybridization kinetics: a solution versus surface hybridization comparison. *Nucleic Acids Res* 2006, 34(11):3370-3377
17. Wang J-Y, Drlica K. Modeling hybridization kinetics. *Math Biosci* 2003, 183:37–47
18. Serak MMA, Bloch W, St John PM. Comparative study of sequence-dependent hybridization kinetics in solution and on microspheres. *Nucleic Acids Res* 2005, 33(1):366-375
19. Mamedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R. A. Subramanian, Viljoen HJ, A fundamental study of PCR amplification of GC-rich DNA templates, *Comp Biol Chem* 2008 32:452-457
20. Sprangler R, Goddard NL, Thaler DS, Optimizing Taq Polymerase Concentration for Improved Signal-to-Noise in the Broad Range Detection of Low Abundance Bacteria, *PLoS one* 2009, 4(9): e7010, 1-9
21. Landgraf A, Wolfs H, Taq Polymerase (E.C. 2.7.7.7) with Particular Emphasis on Its Use in PCR Protocols, *Methods Mol Biol* 1993, 16:31-58
22. Reinick P, Wienken CJ, Braun D. Thermophoresis of single stranded DNA. *Electrophoresis* 2010, 31:279-286
23. Zhang Y, Hammer DA, Graves DJ. Competitive hybridization kinetics reveals unexpected behavior patterns. *Biophys J* 2005, 89:2950-2959
24. Carslaw HS, Jaeger JC, Conduction of Heat in Solids, 1959. Oxford at the Clarendon Press, 2nd edition
25. Squires TM, Messinger RJ, Manalis SR. Making it stick: convection, reaction and diffusion in surface-based biosensors. *Nature Biotech* 2008, 26(4):417-426

Appendix A: Solutions of the equations

Appendix A1: Solution of real time PCR

The equations describing the annealing of amplicons to primers and renaturation are given by equation 13:

$$\frac{\partial c_A}{\partial t} = -kc_{p0}c_A + (k - k_{dsDNA})c_A^2 \approx -kc_{p0}c_A,$$

$$\frac{\partial c_B}{\partial t} = -kc_{p0}c_B + (k - k_{dsDNA})c_B^2 \approx -kc_{p0}c_B$$

When starting the amplification process the concentration of primers is much larger than the concentrations of both species A and B. That means that the first terms of the right hand sides of the two rate equations above are strongly dominant. So the rate equations reduce to first order equations. This will be true most of the time, only at the end of the amplification process, renaturation will be dominant over the formation of complexes.

During annealing the solutions of the approximate rate equations for species A and B are given by (with c_{A0} and c_{B0} the initial concentrations of species A and B):

$$c_A = c_{A0}e^{-kc_{p0}t},\ c_B = c_{B0}e^{-kc_{p0}t} \qquad \text{(A1)}$$

The concentration of free species A and B goes to zero for sufficiently long time time t ($t \gg 1/(kc_{p0})$). Because of the fact that the PCR reaction is symmetric in species A and species B, at the start of each hybridization/anneal step it holds:

$$c_{A0} = c_{B0} = c_0 \qquad \text{(A2)}$$

The concentration renaturated ds DNA follows from (using the approximate expressions for the concentration of species A and B):

$$\frac{\partial c_{dsDNA}}{\partial t} = k_{dsDNA}c_Ac_B = k_{dsDNA}c_0^2e^{-2kc_{p0}t} \qquad \text{(A3)}$$

With the following solution:

$$c_{dsDNA} = \frac{k_{dsDNA}}{k}\frac{c_0^2}{2c_{p0}}(1 - e^{-2kc_{p0}t}) \qquad \text{(A4)}$$

The concentration of ds DNA goes to a constant value $c_0^2k_{dsDNA}/(2kc_{p0})$ for sufficiently long time t ($t \gg 1/(2kc_{p0})$). Note that the concentration of ds DNA reaches much faster its equilibrium value than that the concentration of ss DNA species goes to zero.

Because of conservation of species A the following equation can be used:

$$c_{complex_1} = c_0 - c_A - c_{dsDNA} = c_0\left(1 - e^{-kc_{p0}t}\right) - \frac{k_{dsDNA}}{k}\frac{c_0^2}{2c_{p0}}(1 - e^{-2kc_{p0}t}) \qquad \text{(A5)}$$

A similar expression holds true for the concentration of $complex_2$.

At the end of the annealing/hybridization phase ($t = t_{hyb}$) , the temperature is raised to the elongation temperature when all anneal reactions stop and the elongation phase starts. Part of the complexes will be elongated, controlled by the factor E_{M-M} the so-called elongation factor (Michaelis-Menten factor). After elongation the temperature is raised again to the denaturation temperature during which all complexes are denatured. At the end of the denaturation, species A and B can given as follows: the unreacted species (no primer annealing and elongation), the species onto which the primers have annealed but no elongation took place, and the species which are fully elongated:

$$c_A = c_A(t_{hyb}) + c_{dsDNA}(t_{hyb}) + 2E_{M-M}c_{complex_1}(t_{hyb}) + (1 - E_{M-M})c_{complex_1}(t_{hyb}) \quad \text{(A6)}$$

Note that factor 2 in front of the elongation factor comes from the fact that new amounts of species A (and likewise for B) are produced from both elongated $complex_1$ and $complex_2$. A similar expression holds true for species B.

The amplification equation is subjected to two limiting effects, namely the depletion of concentration of primers available for making complexes and the Michaelis-Menten kinetics of the enzymes used for elongation. Moreover the PCR process stops when all dNTP's have been consumed during elongation. For all cases discussed it is assumed that the PCR process stops because of depletion of primers and that the other PCR ingredients like nucleotides are present in excess. The above description holds true as long as:

$$c_{complex_1}(t_{hyb}) < c_{p0} - c_0 \quad \text{(A7)}$$

When not enough primers are available to anneal to all strands available the following expression holds true:

$$c_A = c_B = 2c_0 - c_{p0} + 2E_{M-M}[c_{p0} - c_0] + (1 - E_{M-M})[c_{p0} - c_0] \quad \text{(A8)}$$

The depletion of primers also influences the rate at which ds DNA is formed. This effect is small and will not be discussed here.

The rate of consumption of either $complex_1$ or $complex_2$ is given by Michaelis-Menten equation (18) (no distinction is made between the complexes). This relation describes that for low concentrations of the complex the catalytic action of an enzyme scales with the concentration of the complex involved. For high concentrations, however, the rate of conversion becomes a constant. The amount found at the end of the elongation phase follows from (equation 18) :

$$\frac{dc_{complex}}{dt} = -\frac{R_{max}\frac{c_{dNTP}}{K_{mdNTP}+c_{dNTP}}c_{complex}}{K_m+c_{complex}} \quad \text{(A9)}$$

With solution:

$$K_m \ln \frac{c_{complex}(t)}{c_{0,complex}} + c_{complex}(t) - c_{0,complex} = -tR_{max} \frac{c_{dNTP}}{K_{mdNTP}+c_{dNTP}} \tag{A10}$$

This equation is transcendental and cannot be solved in closed form. The Michaelis-Menten factor is defined by:

$$E_{M-M} = 1 - \frac{c_{complex}(t_{elongation})}{c_{0,complex}} \tag{A11}$$

It should be mentioned that for the calculation of the elongation factor E_{M-M} the reduction of the elongation rate because of depletion of dNTP's has been considered a constant. This is reasonable because the concentration of dNTP's is high and therefore a slowly decreasing function of the cycle number. Per cycle, however, the value of the concentration of dNTP must be adapted in order to calculate the correct elongation rate.

During thermal cycling the polymerase degrades and per cycle some amount is lost. To take this effect into account an assumption should be made on the amount of degradation per cycle. This effect only affects the value of R_{max} .

Appendix A2: Solutions for Real Time Array PCR

The case that only species A can hybridize with the capture probes on the microarray substrate is discussed. The concentration of the different species during the annealing phase as function of time are given in paragraph 6.3.2.

At the very moment targets are captured, the concentration just above the spot surface (the surface $r = a$) becomes lower and diffusion of targets towards the surface starts, governed by Fick's second law[25]:

$$\frac{\partial c_A}{\partial t} = \mathcal{D} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial c_A}{\partial r} - kc_{p0}c_A + (k - k_{dsDNA})c_A^2 \tag{A12}$$

The rate constants k and k_{dsDNA} are of equal magnitude and during most of the thermal cycles $c_A \ll c_{p0}$, so the last term of the right hand side of equation (33) can be neglected:

$$\frac{\partial c_A}{\partial t} = \mathcal{D} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial c_A}{\partial r} - kc_A c_{p0} \tag{A13}$$

We can solve for the reaction part:

$$c_A = c_{A0} e^{-kc_{p0}t} \tag{A14}$$

As solution for the diffusion equation including the bulk chemical reaction, the following equation is used:

$$c_A = f(r,t)c_{A0}e^{-kc_{p0}t} \tag{A15}$$

Where $f(r,t)$ follows from:

$$\frac{\partial f(r,t)}{\partial t} = \mathcal{D}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial f(r,t)}{\partial r} \tag{A16}$$

With initial condition $t = 0, r > a, f(r,t) = 1$, and the flux condition for $r = a$:

$$\frac{\partial f(a,t)}{\partial z} = \frac{k_{ass}}{\mathcal{D}}\frac{n_{probes}}{n_{avogadro}}\frac{1}{2\pi a^2}f(a,t) = hf(a,t) \tag{A17}$$

The solution of the differential equation (A17) subjected to initial and boundary conditions (A18) can be found in[24]. The concentration as a function of time now becomes (h' being a short hand writing defined below):

$$\frac{c_A}{c_{A0}} = e^{-kc_{p0}t}\left[1 - \frac{ha^2}{r(ah+1)}\left\{\mathrm{erfc}\frac{r-a}{\sqrt{4Dt}} - e^{h'(r-a)+h'^2Dt}\,\mathrm{erfc}[\frac{r-a}{\sqrt{4Dt}} + h'\sqrt{Dt}]\right\}\right] \tag{A18}$$

With $h' = h + \frac{1}{a}$

The concentration as a function of time at the spot surface $r = a$ is given by:

$$\frac{c_{A,r=a}}{c_{A0}} = e^{-kc_pt}\left[1 - \frac{h}{h'}\left\{1 - e^{h'^2Dt}\,\mathrm{erfc}(h'\sqrt{Dt})\right\}\right] \tag{A19}$$

The total number of captured targets can be found by integration by parts of (A19):

$$\begin{aligned} n_{cycle} = \theta n_{probes} &= \int_0^{t_{hyb}} \frac{d\theta n_{probes}}{dt}dt \\ &= k_{ass}c_{A0}n_{probes}\int_0^{t_{hyb}} e^{-kc_{p0}t}[1 - \frac{h}{h'} + \frac{h}{h'}e^{h'^2Dt}\,\mathrm{erfc}(h'\sqrt{Dt})]\ dt \end{aligned} \tag{A20}$$

The final result reads (with $\xi_{hyb} = h'\sqrt{Dt_{hyb}}$):

$$
\begin{aligned}
n_{cycle} = k_{ass} C_{A0} n_{probes} \Bigg[& \frac{1}{k c_{p0}} \left(1 - \frac{h}{h'}\right) \left(1 - e^{-k c_{p0} t_{hyb}}\right) \\
& + \frac{h}{h'} \frac{1}{\left(-k c_{p0} + h'^2 D\right)} \Bigg\{ e^{\frac{(-k c_{p0} + h'^2 D)}{h'^2 D} \xi_{hyb}^2} \operatorname{erfc} \xi_{hyb} \\
& - 1 \frac{1}{\sqrt{\frac{k c_p}{h'^2 D}}} \operatorname{erf}(\xi_{hyb} \sqrt{\frac{k c_{p0}}{h'^2 D}}) \Bigg\} \Bigg]
\end{aligned}
\tag{A21}
$$

Chapter 7

Technology assessment

Technology assessment

Though many microarray based research platforms exist, the translation of these tests from research into the clinic has been slower and more complicated than anticipated. This can be attributed to reproducibility issues and lack of robust analytical performance[1]. In this thesis, a number of technical aspects regarding the use of inkjet printed microarrays in assays for clinical diagnostics have been evaluated. Throughout the thesis, the focus is on investigation of potential solutions for addressing performance limitations in the use of low density microarrays for molecular diagnostics. The following paragraphs summarize the most important conclusions and put these into perspective.

7.1 Microarray printing

Inkjet printing is a common method used for low density microarray manufacturing. A stable and reproducible inkjet process is key for producing high quality microarrays. However, no detailed information is available on inkjet reliability and failure mechanisms. In chapter 2, a system is presented that was developed to study failure mechanisms and failure rates during the inkjet process. Moreover, this system was used for quality control of the inkjet process by imaging and analyzing each jetted droplet, and intervention of the manufacturing process upon detection of a jetting failure. A non-destructive quality control check is crucial to include in a manufacturing process for microarrays that are going to be used in clinical diagnostics. In order to optimize the yield of the production process, it is required to minimize the time in between jetting failure and the detection of this event. Therefore, a system as presented in chapter 2 that detects real time jetting failure can be a valuable tool for producing high quality microarrays with minimal yield losses.

Apart from detecting jetting failure, it was found that in about half of the cases, the droplet detection system (e.g. droplet detection algorithm built in a standard computer driven by a standard operating system) itself failed rather than droplet jetting, which resulted in unnecessary process interruptions. Therefore, before such a system can be implemented in a production environment, the system itself should be further optimized for minimization of these unnecessary process interruptions where droplet recognition fails rather than droplet emission. Furthermore, the failure mode analysis can be automated, thereby reducing operator intervention time and handling. In a number of cases it was found that jetting failure could be predicted based on a decreasing droplet velocity during the preceding droplets. Repairing missing spots on microarrays by using an additional print stroke is technically possible but can be time consuming. An additional potential improvement of such a system is measuring the droplet velocity automatically and performing a process interrupt (i.e. emptying the print head, cleaning and aspiration of fresh liquid) when the droplet velocity starts to decrease.

Inkjet reliability is determined by a combination of electronics, mechanics, software and process failures, which makes it a complex matter. Monitoring jetting failure is needed for manufacturing high quality microarrays for use in clinical diagnostics. In chapter 2, a camera system has been presented that visualizes every droplet emitted. Alternative technologies like acoustical measurements or capacitive measurements are in development and can also be a good choice for monitoring droplet jetting[2-3]. Using such a monitoring system can lead to more insights into failure mechanisms and statistics and ways to improve the manufacturing process. Additional work needs to be done to implement this in a high yield production line.

7.2 Surface chemistry

Different surface chemistries are used for linking the deposited oligonucleotides to the microarray substrate[4-7]. The choice of a surface chemistry is important since it is a factor contributing to the overall microarray performance and reproducibility. A surface chemistry commonly used is 254 nm UV-induced cross-linking of DNA to amine-functional surfaces. This is an advantage above surface chemistries where active or pre-activated groups are on the surface since for these surface chemistries, degradation of the active groups starts upon exposure of the microarray substrate to the manufacturing environment. This means that differences in reactivity of the grafting sites can exist for the first and the last solution deposited on the array. For the amine-chemistry, activation of the microarray surface is done for all spots in a single run, and therefore it can be scaled up for manufacturing relatively easy. Furthermore it is also a relatively fast and cheap process. However, the UV-induced immobilization process is poorly understood[4]. Since DNA is susceptible to degradation by 254 nm UV light[8], DNA degradation might be a potential cause for reduced microarray performance. In chapter 3, a method is presented which was used to independently studying the grafting of model oligonucleotides onto amine-containing nylon membranes and the hybridization efficiency. The active role of thymine in the immobilization process was confirmed[9] whereas also to a limited extent guanine and furthermore the DNA backbone plays a role in the UV-induced immobilization process. Within a certain UV-dose range, no significant DNA degradation caused by UV-exposure was found that would result in a reduced hybridization efficiency.

7.3 Microarray-based tests for clinical diagnostics

In chapters 4 and 5, two microarray-based analytical methods are presented which integrate a number of consecutive assay steps in a single cartridge and device. In this way, the workflow of a diagnostic analysis and ease-of-use can be increased, hands-on time can be reduced and as a result of this, variations caused by operator handling can be minimized. Additionally for the method presented in chapter 5, a large degree of multiplexing with quantitative DNA detection is possible.

In chapter 4, an evaluation of a microarray flow-through post-PCR measurement method is described using the Human Papilloma Virus (HPV) as a model system and including the Reverse Line Blot as a reference. Good agreement between the two methods was found.

Compared to the reference, the presented system inherently has additional technical advantages: the availability of kinetic binding curves enables a more reliable evaluation as compared to conventional end-point measurement which provides only a single measurement point per spot. Kinetic measurements also provide the possibility of detecting cross-hybridizations which makes the method especially useful for detection of homologous sequences like single nucleotide polymorphism (SNP) detection or HPV genotyping. Both the binding curves but especially the melting curves can be used for the detection of cross-hybridizations since perfect match hybrids have a higher melting temperature than mismatch hybrids. For full implementation in the clinic, the device needs to be further improved in terms of throughput: the Reverse Line Blot enables the parallel detection of up to 42 different samples as compared to a single sample per test for the current real time array PCR device.

Chapter 5 describes an experimental method that integrates hybridization and detection with DNA amplification into a single cartridge (real time array PCR). The cartridge is placed in a unit that performs all the heating and cooling steps and the read-out. This method allows for quantitative detection like real time PCR, but with a higher multiplex capability since the different targets can be discriminated spatially on a microarray rather than with discriminating differently colored fluorophores as in qPCR. Feasibility of quantitative detection over at least 4 orders of magnitude is shown using a singleplex model system. Compared to other systems currently in development, this method makes use of an additional degree of specificity since probes are immobilized on the microarray surface rather than primers which are then elongated[10]. Moreover use has been made of standard detection components based on optical recording technology which would allow a rapid development for commercial use. Additionally, standard reagents and disposables were used as well.

For high multiplex PCR detection on the microarray, the detection itself should preferably be further optimized by further confining the detection volume to increase the fluorescence background suppression. Implementing convection of the flow-through device as used in chapter 4 during the real time array PCR process, significant enhancement of hybridization signals might be achieved. Additionally, instead of temperature cycling of the complete real time array PCR chamber, also the use of different temperature zones for denaturation, annealing/hybridization and elongation can be considered[11]. The use of a separate hybridization zone will give rise to cumulative and thus higher hybridization signals, and result in less capture probe degradation since exposure to the higher temperatures needed for denaturation is prevented.

The mathematical analysis of real time array PCR as presented in chapter 6 provides insights into the biochemical processes that are taking place. It is especially suited for investigation of trends to study the influence of specific parameters on the performance of the real time array PCR process. The model and the experiments were compared for a limited data set and upon optimization of a number of parameters, reasonably good accordance was achieved. This kind of molecular biological processes are difficult to analyze quantitatively since there are still a number of biological processes taking place that are either not yet understood or very difficult to model in itself. For example, the use of different PCR enhancers in the PCR reaction mixture to increase PCR efficiency or fluorescence[12] (e.g. polyethylene glycol, BSA, glycerol, non-ionic detergents, DMSO, or betaine) is usually empirically optimized for each primer/probe/template combination. Current knowledge does not make it possible to predict the effect of these enhancers in real time PCR[13] and it is therefore very difficult to incorporate these effects into a quantitative model. The trends, however, are less depending on the detailed compositions and can therefore be used to optimize the real time array PCR process.

In conclusion, a number of aspects in the use of microarrays have been investigated in this research project, and potential improvements have been analyzed. Further directions for a larger application area for microarrays, e.g. real time array PCR to target quantitative high multiplex DNA detection in a closed chamber, have been addressed. These improvements have all been directed to, in a broader sense, stimulating the implementation of microarray-based diagnostics into the clinic. The powerful properties of microarrays especially regarding their speed and parallelism in detection, are now at a point where they are outweighing the issues associated with the use of microarrays.

References

1. Li X, Quigg RJ, Gu W, Rao N, Reed E. Clinical utility of microarrays: current status, existing challenges and future outlook. *Curr Genomics* 2008, 9:466-474
2. De Jong J, de Bruijn G, Reinten H, van den Berg M, Wijshoff H, Versluis M, Lohse D. Air entrapment in piezo-driven inkjet printheads. *J Acoust Soc Am* 2006, 120(3):1257-1265
3. Wei J, Yue C, van der Velden M, Chen ZL, Liu ZW, Makinwa KAA, Sarro PM. Design, fabrication and characterization of a femto-farad capacitive sensor for pico-liter liquid monitoring. *Sensor Actuator A* 2010, 162:406-417
4. Taylor S, Smith S, Windle B, Guiseppi-Elie A. Impact of surface chemistry and blocking strategies on DNA microarrays. *Nucleic Acids Res* 2003, 31(16):e87
5. Wu P, Castner DG, Grainger DW. Diagnostic devices as biomaterials: a review of nucleic acid and protein microarray surface performance issues. *J Biomat Sci-Polym E* 2008, 19(6):725-753
6. Wang H-Y, Malek RL, Kwitek AE, Greene AS, Luu TV, Behbahani B, Frank B, Quackenbush J, Lee NH. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biology* 2003, 4:R5
7. Nierzwickibauer SA, Gebhardt JS, Linkkila L, Walsh K. A comparison of UV cross-linking and vacuum baking for nucleic acid immobilization and retention. *Biotechniques* 1990, 9:472-478
8. Ravanat J-L, Douki T, Cadet J. Direct and indirect effects of UV radiation on DNA and its components. *J Photoch photobio B* 2001, 63:88-102
9. Brown TJ, Anthony RMJ. The addition of low numbers of 3'thymine can be used to improve the hybridization signal of oligonucleotides for use within arrays on nylon supports. *J Microbiol Meth* 2000, 42:203-207
10. Khodakov DA, Zakharova NV, Gryadunov DA, Filato FP, Zasedatelev AS Mikhailovich VM. An oligonucleotide microarray for multiplex real-time PCR identification of HIV-1, HBV, and HCV. *Biotechniques* 2008, 44:241-248
11. Klunder DJW, Pierik A, Kolesnychenko A, Boamfa MI, Schroeders RJM. Amplification of nucleic acids using temperature zones. Philips Electronics Patent Application WO2009156895.
12. Horáková H, Polakovičová, Shaik GM, Eitler J, Bugajev V, Dráberová L, Dráber P. 1,2-propanediol-trehlaose mixture as a potent quantitative real-time PCR enhancer. *BMC Biotechn* 2011, 11:41
13. Shaik GM, Dráberová L, Dráber P, Boubelík M, Dráber P. Tetraalkylammonium derivatives as real-time PCR enhances and stabilizers of the qPCR mixtures containing SYBR Green I. *Nucleic Acids Res* 2008, 36(15):e93.

Summary

Summary

DNA microarrays become increasingly important in the field of clinical diagnostics. These microarrays, also called DNA chips, are small solid substrates, typically having a maximum surface area of a few cm^2, onto which many spots are arrayed in a pre-determined pattern. Each of these spots contains multiple copies of one oligonucleotide, called the capture probe, which has a sequence complementary to one single specific target molecule. A DNA microarray assay is based on the preferential binding between a DNA target molecule with its complementary probe present on the microarray surface. Provided appropriate conditions, these two single stranded DNA molecules with complementary DNA sequences form a double stranded molecule during a process called hybridization. When the target molecules are labeled with e.g. a fluorophore, the fluorescent microarray pattern achieved upon incubation of the sample with the microarray provides information on the presence of the target sequences in the sample. The key advantage of microarrays is their large degree of parallelism per sample due to the simultaneous reaction between multiple capture probes and target sequences. This thesis focuses on low density DNA arrays that are limited to a few hundred spots for diagnostic use, for the rapid detection of a constrained and specific set of target sequences. One major application of low density microarrays lies in the detection and identification of pathogens in the field of infectious diseases. So, high density arrays containing as high as millions of different spots that allows for the screening and detection of a much higher number of targets simultaneously, usually applied for studying gene expression studies, are beyond the scope of this thesis.

The transfer of microarray-based assays from the research lab into the clinic has taken more time than anticipated and it faces a number of technical and regulatory challenges. These are especially related to quality control and data reproducibility caused by a lack of fundamental understanding of the biochemical processes involved. In this thesis, a number of these challenges are addressed and investigated, and potential routes for improvement are identified. This work focuses on the improvement of microarray technology in order to enable a faster implementation of microarray-based assays into clinical practice.

Inkjet printing is a suitable technology for manufacturing low density microarrays. It is, however, an open-loop process and does therefore not provide any feedback on the quality of manufactured microarrays. E.g. missing spots are unacceptable since this could lead to potentially false negative results. Moreover, microarrays are generally being considered as high complexity products by the Clinical Laboratory Improvement Amendments (CLIA), or as high risk applications by the Food and Drug Administration (FDA). Therefore specific quality control measures on microarrays for clinical diagnostics are needed. In chapter 2, a closed-loop inkjet printing system equipped with an optical droplet detection system to investigate failure mechanisms of the printing process is described. It was found that of all microarrays analyzed, in 1.6 % jetting failed at some point in time. In 14 % of these cases, these failures could have been detected in advance by detection of changing droplet velocities during

jetting. Real time analysis of droplet characteristics and closed-loop control can enable a higher yield of the microarray manufacturing process. The improved production process as explained in chapter 2 is the basis for the microarray experiments carried out for the results in experimental chapters 3-5.

Many different surface chemistries exist that are applied to graft the DNA probes onto the microarray surface. A lack of understanding on the characteristics of DNA immobilization and hybridization as separate processes hampers the improvement in reproducibility of the use of microarrays. Immobilization and hybridization rates of oligonucleotides equipped with specific tails on amine-functionalized surfaces by using 254 nm UV-light are studied. This surface chemistry is attractive because of its ease-of-use, robustness and low cost. A method was developed that enables a systematic study of the immobilization and subsequent hybridization processes independently. It was found that immobilization efficiencies are greatly influenced both by the UV dose applied and by the composition, as well as the length of the specific tails attached to the oligonucleotides. Hybridization is mainly influenced by the length of the oligonucleotides, and much less by the composition of the tail as observed for immobilization. This means that within the UV-window as applied in this study, no significant UV-damage was found.

The standard microarray procedure runs as follows: microarrays are incubated with the sample, followed by a washing step to remove non-specific binding, then dried and scanned. This process flow implies that only a single measurement (end) point is used to evaluate target and capture molecule interactions. Since these interactions depend on many factors, including the preceding steps and the composition of the mixture itself, a complete hybridization measurement would provide additional and more reliable data. This means, however, that measurements should be done in high background signals requiring the deployment of methods that enable background suppression. Chapter 4 describes a method that measures real time hybridization and melting curves based on the repeated flow of the solution containing the targets through a porous microarray substrate. During each hybridization cycle, background was suppressed by pumping the liquid through the substrate. This method, using the Human Papilloma Virus (HPV) assay as a model assay, was evaluated with clinical samples and benchmarked against a method commonly used in the clinic (Reverse Line Blot). Both methods show comparable results, while the flow-through method also enabled the identification of cross-hybridizations by the analysis of the binding and melting kinetics. This method can contribute to identification of false positive signals.

The Polymerase Chain Reaction (PCR) is a method to amplify the concentration of specific DNA target sequences and is often performed prior to a microarray hybridization. Sample preparation variations including this amplification step contribute to a lower reproducibility of microarrays. In the real time array PCR concept the steps of real time amplification (qPCR) and detection are integrated, thereby combining the advantages of both methods: the high multiplexing capabilities of microarray based assays and the quantitative characteristics of real time PCR. In this procedure the hybridization of the formed amplicons on the microarray surface is monitored, during each annealing step of the amplification process. Consequently an amplification curve on the microarray surface is obtained from which quantitative information on the target input concentration can be derived for a much higher number of targets compared to qPCR. After the amplification, an additional hybridization and melting curve step can be performed to further assess the specificity. The detection of low signals within high background levels in relatively short times is the greatest technical challenge of this concept. This was achieved by the use of a confocal fluorescence scanner that significantly reduced background levels. A proof of concept study was carried out with a prototype instrument that was designed and built from standard components known from the optical storage technology. In order to minimize costs of goods, standard products (reagents and disposables) were also used for the biochemical processes of amplification and hybridization as much as possible. Amplification and detection were performed in a single and closed chamber, minimizing risks on cross-contamination and reducing manual steps, making this method especially suited for diagnostic testing. As low as 10 copies of a target sequence could be detected using this method.

Chapter 6 contains a mathematical analysis of the real time array PCR process which was carried out to provide insights in the biochemical processes taking place. Models of qPCR and Langmuir adsorption were combined into a new model describing the different steps in the real time array protocol. Reasonably good agreement between the model and the experiments was found. Using this result for the experimental array-based assay, a number of measures have been identified and tested to increase the overall amplification efficiency which resulted in a similar performance compared to bulk qPCR.

Nederlandse samenvatting

Samenvatting

In desoxyribonucleïnezuur (DNA) is de erfelijke code van vrijwel alle organismen opgeslagen. Een DNA molecuul bestaat uit twee lange strengen die aan elkaar complementair zijn zodat ze kunnen binden. Ze zijn beide opgebouwd uit een veelvoud van 4 verschillende bouwstenen, de DNA basen. De erfelijke informatie van een organisme wordt bepaald door de volgorde van deze basen, de DNA-sequentie. Sinds de ontdekking van de DNA-structuur en evenals de polymerase ketting reactie (PCR) waarmee kleine hoeveelheden DNA kunnen worden vermenigvuldigd, heeft de DNA-technologie een vlucht genomen. Een specifiek voorbeeld van DNA-technologie is het gebruik van DNA-microarrays voor de analyse van patiëntenmonsters voor diagnostische toepassingen.

DNA-microarrays spelen een steeds grotere rol in de moleculaire diagnostiek. Op een microarray zijn vele korte stukjes DNA geïmmobiliseerd in spots op een drageroppervlak van enkele vierkante centimeter oppervlakte. Iedere spot bevat specifieke korte stukjes synthetisch DNA, die ook wel ‘capture probes’ genoemd worden, met een gegeven bekende DNA-sequentie. Iedere capture probe sequentie is exact complementair aan een sequentie waarvan men de aanwezigheid in een monster wil aantonen (de ‘targetsequentie’). Een microarray-analyse is gebaseerd op de unieke interactie tussen het targetmolecuul met een complementaire capture probe op het oppervlak. Deze interactie vindt plaats tijdens een incubatiestap van de microarray met het (voorbewerkte) monster met daarin al dan niet de aanwezige targetsequenties. Onder de juiste condities kunnen de targetmoleculen binden aan de geïmmobiliseerde capture probes in een proces dat hybridisatie genoemd wordt. Door het meten van de gebonden targetsequenties, bijvoorbeeld door middel van fluorescentiebepalingen nadat alle targetsequenties in een voorgaande processtap fluorescent zijn gemarkeerd, kan aangetoond worden of de targetsequenties zich in het monster bevinden. Doordat veel spots, die elk een verschillende capture probe sequentie bevatten, op het oppervlak zijn geïmmobiliseerd, kan er een groot aantal aan parallelle bindingen plaatsvinden op het oppervlak. De kracht van een op een microarray gebaseerde test is dan ook dat met een enkel experiment het monster kan worden onderzocht op de aanwezigheid van zeer veel verschillende targetmoleculen.

Hoge-dichtheid microarrays bevatten een paar miljoen verschillende spots. Dit type microarrays wordt voornamelijk gebruikt voor genexpressie analyses. De focus van deze thesis is echter gericht op lage-dichtheid microarrays met een beperkte specifieke set aan verschillende spots (maximaal een paar honderd) en die daardoor geschikter zijn voor klinische diagnostiek. Een van de meest gebruikte toepassingen voor dit type microarrays is de detectie van ziekteverwekkers voor het vaststellen van infecties.

Er worden nog relatief weinig microarray-gebaseerde moleculaire testen gebruikt in de klinische diagnostiek. Ondanks de mogelijkheden van het gebruik van microarrays, is er ook een aantal technologische problemen waarvan het oplossen meer tijd kost dan oorspronkelijk gedacht. De technologische uitdagingen liggen voornamelijk op het gebied van kwaliteitscontrole en reproduceerbaarheid, mede veroorzaakt door een gebrek aan fundamentele begripsvorming van de onderliggende biochemische processen. In dit proefschrift is een aantal van deze problemen geanalyseerd en zijn mogelijke oplossingsrichtingen aangedragen. De doelstelling van dit proefschrift is om daarmee een bijdrage te leveren aan de daadwerkelijke verbetering van de technologie om een snellere acceptatie van microarray–gebaseerde testen in een ziekenhuisomgeving te bereiken.

Inkjetprinten is een geschikte technologie om lage-dichtheid microarrays te produceren. In dit proces wordt echter normaliter geen terugkoppeling verkregen over het verloop ervan en daarmee ontbreekt informatie over de kwaliteit van de geproduceerde microarrays. Voor dit type medische producten dat door de Amerikaanse Clinical Laboratory Improvement Amendments (CLIA) wordt ingedeeld als een hoog complex product en door de Food and Drugs Administration (FDA) als een hoog risico toepassing, is een hoge mate van kwaliteitsborging essentieel. Ontbrekende spots veroorzaakt door een niet goed werkend productieproces zijn niet acceptabel en specifieke kwaliteitscontroles op het fabricageproces zijn daarom noodzakelijk. In hoofdstuk 2 is een inkjetprinter beschreven die uitgerust is met twee optische systemen. Met deze camera’s kan iedere verspoten druppel worden gefotografeerd en de aan- of afwezigheid van druppels kan daarmee gemakkelijk worden vastgesteld. Op basis van deze foto’s kunnen de verschillende mechanismes die leiden tot defecten in het productieproces gecategoriseerd worden. Het real-time analyseren van de geprinte druppels is de basis voor het terugkoppelingsmechanisme dat in het proces is ingebouwd waardoor het automatisch wordt onderbroken wanneer het niet meer aan de specificaties voldoet. In 1.6 % van alle geanalyseerde microarrays had het printen op enig moment in de tijd niet goed gefunctioneerd. In 14 % van deze gevallen had op basis van het meten van druppelsnelheden dit gedrag voorspeld kunnen worden. Het terugkoppelen en het voorspellen van uitval kan leiden tot een efficiënter productieproces met hogere opbrengst. Het verbeterde productieproces van hoofdstuk 2 heeft als basis gediend voor de microarray-experimenten beschreven in de hoofdstukken 3 tot en met 5.

Er zijn veel verschillende typen oppervlaktechemie beschikbaar om capture probes vast te zetten op het drageroppervlak. Binding van DNA aan amine-gemodificeerde oppervlaktes door middel van 254 nm UV-belichting is een aantrekkelijk proces gezien de hoge mate van gebruiksgemak en de lage kosten die hiermee gepaard gaan. Er is echter slechts beperkte kennis beschikbaar betreffende het exacte mechanisme van dit proces, de invloed van procesparameters op het rendement van de reactie en het effect van deze processtap op de daaropvolgende hybridisatiereactie. In hoofdstuk 3 is dit onderzocht via het bepalen van de afzonderlijke rendementen van de immobilisatie- en hybridisatiestap. Dit is gedaan voor verschillende 5’-uiteinde gemodificeerde oligonucleotiden mee te nemen om zo ook informatie te krijgen over de invloed van de samenstelling van het DNA op deze processtappen. De UV-dosis, de samenstelling en de lengte van de modificatie aan het 5’-uiteinde van de capture probe hebben een grote invloed op het rendement van de bindingsreactie. De hybridisatiereactie wordt voornamelijk beïnvloed door de lengte van de modificatie aan het 5’-uiteinde. Binnen het onderzochte bereik van UV-blootstelling werd dus geen invloed gevonden op de hybridisatie-efficiency, hetgeen impliceert dat er geen meetbare degradatie van de capture probes door blootstelling aan het UV-licht plaatsvond.

Voordat de targetsequenties in een monster aangetoond kunnen worden, is het gebruikelijk om het monster eerst voor te bewerken om de concentratie aan targetsequenties te verhogen en om de moleculen te markeren zodat ze gedetecteerd kunnen worden. Na deze voorbewerking worden de targetmoleculen geïncubeerd op de microarray. Vervolgens wordt het microarray-substraat gewassen en gedroogd om niet-specifieke binding te reduceren. Pas op het einde van deze processen wordt de intensiteit van de spots op het oppervlak gemeten via een eindpuntmeting. Deze intensiteit, en de reproduceerbaarheid hiervan, wordt beïnvloed door de voorafgaande processtappen. Door real-time hybridisatie te meten op het oppervlak in plaats van een enkele eindpuntmeting uit te voeren kan de kwaliteit van een dergelijke bepaling verbeterd worden doordat enerzijds er meer meetpunten verkregen worden en anderzijds de variatie veroorzaakt door de nabehandelingsstappen geen invloed meer heeft. In hoofdstuk 4 is een analyse beschreven van een methode waarbij een poreuze microarray is gebruikt waar de targetmoleculen doorheen gepompt worden. Doordat de fluorescerende vloeistof door het oppervlak heen wordt gepompt zijn de fluorescente moleculen die normaliter voor een significante achtergrond zorgen niet meer op de microarray aanwezig en daardoor kan na iedere cyclus de intensiteit van alle spots gemeten worden. Een aantal klinische monsters is met deze methode geanalyseerd en vergeleken met de veel gebruikte Reverse Line Blot als referentie. Als model organisme is het Humaan Papillomavirus (HPV) gebruikt. Beide methodes geven vergelijkbare resultaten waarbij de real-time methode de mogelijkheid biedt om de specificiteit van de reactie nader te bestuderen via een smeltcurve, wat kan bijdragen aan een reductie van het aantal foutpositieve diagnoses.

Voor het vermenigvuldigen van de targetmoleculen wordt vaak de Polymerase Ketting Reaction (PCR) gebruikt. Hierbij worden targetsequenties vermenigvuldigd totdat de reactie is verzadigd om zo een voldoende hoge concentratie te verkrijgen. Dit betekent echter wel dat er geen kwantitatieve bepaling gedaan kan worden met betrekking tot het aantal targetmoleculen dat voor de amplificatie aanwezig was in het monster, zoals dat bij kwantitatieve PCR (qPCR) wel kan worden gedaan omdat hierbij de reactie in de tijd gevolgd wordt. Het nadeel van qPCR is dat het aantal targetsequenties dat parallel geanalyseerd kan worden beperkt wordt tot circa 6 tot 8. In het laatste deel van dit proefschrift is een methode uitgewerkt waarbij deze amplificatiestap wordt geïntegreerd met de microarray-hybridisatie en –detectie (real-time array PCR). Het doel hiervan is om een kwantitatieve bepaling uit te voeren op een hoger aantal targetsequenties dan met qPCR mogelijk is. Doordat deze processen worden geïntegreerd, wordt daarnaast het aantal manuele handelingen en daarmee het risico op contaminatie verminderd. De grootste technische uitdaging van dit concept is het meten van zeer lage signalen in een hoge achtergrond. Dit is bereikt door middel van confocale detectie met behulp van dvd-gebaseerde technologie.

Door gebruik te maken van standaard componenten uit de wereld van de optische recording lijkt het mogelijk om eenvoudige apparatuur te ontwikkelen en zo te komen tot een potentieel een lage kostprijs. Het meten van een enkelvoudige amplificatie op een microarray is succesvol aangetoond en tot 10 kopieën van een target sequentie konden worden gedetecteerd, enkel gebruik makend van standaard laboratoriummaterialen en reagentia.

Een mathematische analyse van het real-time array PCR proces is uitgevoerd en beschreven in hoofdstuk 6 om de fundamentele biochemische processen die optreden in deze methode te onderzoeken. Een nieuw model is verkregen door modellen van real-time PCR en Langmuir adsorptie te combineren. Hiermee kunnen de experimenten goed beschreven worden en de meest belangrijke parameters zijn geïdentificeerd.

Dankwoord

Toen ik een jaar of vier geleden begon met het schrijven van publicaties die uiteindelijk hebben geleid tot dit proefschrift, had ik eigenlijk nooit verwacht dat daadwerkelijk het moment zou komen van het schrijven van het dankwoord. Het feit dat dit toch gelukt is, komt door de bijdrage van een groot aantal mensen en hen wil ik hier graag bedanken.

Allereerst wil ik graag Dick Broer noemen: bedankt dat je me hebt willen begeleiden met mijn promotie-traject. Je hebt me op een prettige manier begeleid en veel vrijheid gegeven om mijn onderzoek uit te voeren en daardoor heb ik het werken en promoveren goed kunnen combineren. Het was even schrikken toen je begin dit jaar vond dat het dit jaar maar eens afgerond moest worden! Maar dat bleek uiteindelijk de benodigde energie op te leveren om er daadwerkelijk een punt achter te kunnen zetten.

Gedurende mijn werk bij Philips heb ik altijd veel samengewerkt met Frits Dijksman. Frits, je hebt me kunnen overtuigen om de eerste, en daarmee de belangrijkste, stap te zetten met promoveren. Je hebt me in dit proces altijd geholpen waar mogelijk, al dan niet met uitvoerige discussies en zeer nauwgezette correcties, alsook met de 'soms-eist-het-leven-een-daad'-methodiek die heeft bijgedragen om mijn submitknopdrukvrees (enigszins) te overwinnen. Bedankt voor de prettige en leerzame samenwerking.

Henk Stapert, je hebt mijn beslissing om te gaan promoveren altijd vol enthousiasme gesteund, initieel als projectleider van Pasteur, maar daarna ook bij Biocartis. Ondanks alle drukte en soms ook enige chaos in het project heb je waar kon geholpen en tijd vrij gemaakt. In een omgeving met veelal biologen en fysici, waren onze wat meer chemisch technologische discussies voor mij erg waardevol. Daarnaast bedankt dat je wilt deelnemen in mijn promotiecommissie!

Ook wil ik de overige leden van mijn promotie-commissie bedanken voor hun deelname in de commissie en hun feedback op mijn proefschrift: Menno Prins, Luc Brunsveld, Jan van Hest en Cees Bastiaansen.

Vanuit Philips Research hebben mijn groepsleiders Paul van de Wiel en Christopher Busch mij de gelegenheid gegeven om deze promotie respectievelijk af te maken en te beginnen. Hun steun hierin was voor mij zeer waardevol.

Het voordeel van het combineren van werken en promoveren binnen het bedrijfsleven is dat ik heb kunnen profiteren van de infrastructuur die binnen projecten is opgezet voor het reguliere onderzoek. Daardoor hebben direct of indirect veel mensen een bijdrage hebben geleverd aan de uiteindelijke totstandkoming van dit proefschrift. Het werk dat hierin beschreven is, is afkomstig vanuit een verschillend aantal projecten uitgevoerd binnen het Philips Research programma. Allereerst wil ik de mensen vanuit het Pasteur project bedanken voor de ontzettend goede en gezellige samenwerking. Hierbij mag natuurlijk het team met wie de MDx-printers zijn gebouwd, ingeregeld en gebruikt zijn niet vergeten worden. Wim Schram en Faustin Usabuwera die een ontiegelijke hoeveelheid membraantjes in het printlab hebben gemaakt en gemeten wil ik hier nog even extra noemen, want hun werk in het maken van goede kwaliteit microarrays was noodzakelijk voor de experimenten beschreven in dit proefschrift.

Daarnaast Marius, Daniëlle, Martijn (SMS: 'jaaaaaa we hebben spotjes'), Dion en Richard vanwege de goede samenwerking in het Artemis-project. De eerste drie wil ik speciaal bedanken voor het feit dat ze gewillig slachtoffer waren bij mijn eerste stapjes als projectleider: 'het waren drie fantastische maanden…'. Ook in de tijd daarna hebben we nog toch een behoorlijke output weten te bewerkstelligen die bevestigd wordt door de wet van Marius: $\lim_{FTE \to 0} \frac{Output}{FTE} = \infty$

En ik zou graag ook de Apollo-projectteamleden willen bedanken voor de gezellige tijd en de prettige samenwerking: Henk, Jessica, Mona, Margot, Martijn, Paul, Pieter, Rene, Sonja en Tibor! Ook de andere Apollo-projectleiders, met name op vrijdagmiddag: Bart, Marloes, Rudi, Tim en Wim: "dat hebben jullie goed gedaan!!".

De bijdrage van een aantal studenten was belangrijk voor de totstandkoming van dit proefschrift: Martijn van Zelst, Chris Zwanenburg en Hanneke Peeters. Speciaal wil ik Martijn hier noemen: ook na je studie tijdens je werk bij Philips heb je een zeer waardevolle bijdrage geleverd aan wat uiteindelijk heeft geresulteerd in hoofdstuk 5. Je zeer goed ontwikkelde capaciteit in het multitasken op het lab heeft ertoe geleid dat je een onmogelijk aantal experimenten uit kunt voeren op één dag, en dat heeft er mede voor gezorgd dat hoofdstuk 5 er uiteindelijk toch gekomen is! Bedankt voor al je 'combined results'.

De groepsleden van de afdeling Molecular Diagnostics wil ik bedanken voor de gezellige tijd en de fijne samenwerking en ook voor hun enthousiasme. In het bijzonder wil ik Harma bedanken voor haar bijdrage aan de voorkant van deze thesis, en Reinhold, Daniëlle (en ook je echtgenoot bedankt!), Ron, Ellen (ex-MDx, maar toch…) en Joke voor hun bijdrage aan de binnenkant van deze thesis of voor of even bijkomen op de fatboy.

Bernadet, bedankt voor je praktische hulp en ook je mooie relativeringen... Johan, bedankt voor het doorlezen van m'n proefschrift en de feedback daarop! Verder zijn er nog veel meer collega's die ik hier niet allemaal persoonlijk kan noemen maar die de afgelopen jaren wel tot een heel aangename tijd hebben gemaakt.

Verder heeft nog een aantal mensen die ik nog niet genoemd heb een essentiële bijdrage geleverd aan een van de hoofdstukken van deze thesis en hun bijdrage met de nodige discussies was hierin zeer waardevol: Ton Wismans, Adrie Raaijmakers, Johan Lub, Elna Moerland en Adriaan van den Brule.

Roy en Rogier: het feit dat jullie mijn paranimfen zijn, gaat mij niet in de koude kleren zitten. Mijn beeld van paranimfen zal nooit meer hetzelfde zijn…. Roy, het delen van jouw ervaringen van het schrijven van publicaties en het proefschrift hebben mij geholpen. Ik wil jullie beiden bedanken voor de gezellige tijd (moeten we vaker doen), en voor af en toe een mooi momentje van relativering!

Ondanks dat ik me aan het begin had voorgenomen om het promoveren niet significant te laten interfereren met sociale activiteiten, is dat de laatste tijd helaas niet helemaal gelukt. Mijn vrienden wil ik bedanken voor hun interesse en geduld de afgelopen tijd! Ik zal de vanaf nu hopelijk wat gezelliger en socialer zijn! En Angela, super gepland zo!

Ook mijn schoonfamilie, en dan in het bijzonder Jan en Ineke wil ik hier noemen. Na de eerste salontafel-performance, heb ik me altijd thuis gevoeld bij jullie! Bedankt voor al jullie hulp.

Daarnaast mijn zus Paulien: hoe origineel ook, vooral omdat jullie nu keigezellig weer terugkomen naar Brabant (hopelijk zien we Mark en jou de komende tijd weer wat vaker!), ik heb jouw Brabantse samenvatting toch maar niet opgenomen in dit proefschrift… En Rosalieke, kom je logeren??

Papa en mama: jullie hebben me vanaf het begin af aan gesteund om te gaan promoveren en daarnaast ook geholpen waar mogelijk. Bedankt dat jullie er altijd zijn wanneer het nodig is. Laat ik alles maar samenvatten met: ‘ram te woepie’.

Als allerlaatste Milan: we kennen elkaar tot nu toe twee maanden korter dan de tijd die nodig was voor de totstandkoming van mijn proefschrift en vanaf het begin bestond voor jou de optie om te stoppen niet! Zonder dit begrip alsmede al je hulp (thx Mr Ed) en de vele drankjes naar de 1^e verdieping was het nooit gelukt. Kortom, bedankt voor alles!

Curriculum Vitae

Anke Pierik was born on February 20th 1979 in Hoensbroek, the Netherlands. After finishing her secondary education (VWO) at the Bisschoppelijk College Broekhin in Roermond (1997), she started to study Chemical Engineering and Chemistry at the Eindhoven University of Technology. As part of her study, she did a four-month internship at SINTEF Petroleum Research in Trondheim, Norway. In 2003 she graduated in the Chemical Reactor Engineering group of prof.dr.ir. J.C. Schouten.

In 2004, she started at the Philips Research Laboratories Eindhoven in the ‘Healthcare devices and Instrumentation’ department on the subject of inkjet technology for OLED displays. Subsequently she worked on inkjet technology and surface chemistries for a flow-through microarray detection platform. In 2009, she made a change to the department of ‘Molecular Diagnostics’. She now works as a senior scientist working on the development of an integrated nucleic acid detection platform.

From 2008 onwards, she started to publish the results of the research performed within the Philips Research program, leading to the publication of this thesis under supervision of prof.dr. D.J. Broer.