DNA Tools and Microfluidic Systems for Molecular Analysis

JONAS JARVIUS
Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, C11, Uppsala, Thursday, September 21, 2006 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

Improved methods are needed to interrogate the genome and the proteome. Methods with high selectivity, wide dynamic range, and excellent precision, capable of simultaneously analyzing many biomolecules are required to decipher cellular function. This thesis describes a molecular and microfluidic toolbox designed with those criteria in mind. It also presents a tool for graphical representation of nucleic acid sequences.

Proximity ligation is a novel protein detection method that requires dual and proximate binding of two oligonucleotide-tagged affinity reagents to a protein or protein complex in order to elicit a signal. The responses from such recognition reactions are the formation of specific nucleic acid reporter molecules that are subsequently amplified and quantitatively detected.

A scalable microfluidic platform suitable for fluorescence detection, cell culture, and actuation is also described. The platform uses rapid injection molding to produce microstructures in thermoplastic materials. By applying a thin layer of silica to the structures, a lid made of silicone rubber coated onto a thermoplastic support can be covalently bonded to generate enclosed channels.

A method is presented for precise biomolecule counting, termed “amplified single-molecule detection”. The method preserves the discrete nature of biomolecules, converting specific molecular recognition events to fluorescence-labeled micrometer-sized objects that are enumerated in microfluidic channels.

I also present a novel microarray-based detection method. To attain high selectivity and a wide dynamic range, the method is based on dual recognition with enzymatic discrimination and amplification. Upon target recognition in solution, DNA probes are subjected to thousand-fold amplification in solution, followed by selective detection on arrays and another hundred-fold amplification of reporter molecule created from the first amplification reaction.

Lastly, I describe a novel graphical representation of nucleic acid sequences using TrueType fonts that can be of value for visual inspection of DNA sequences and for teaching purposes.

Keywords: Proximity ligation, Microfluidics, Single molecule detection, Microarray, Bonding

Jonas Jarvius, Department of Genetics and Pathology, Rudbecklaboratoriet, Uppsala University, SE-75185 Uppsala, Sweden

© Jonas Jarvius 2006

ISSN 1651-6206
ISBN 91-554-6616-8
urn:nbn:se:uu:diva-7079 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7079)
If these tools don’t turn you on
You don’t have a switch
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals:


In paper III and IV the first two authors contributed equally to the work.
## Contents

Introduction...................................................................................................11

Biomolecular detection methods...............................................................................................................................13
  Enzymes, tools provided by nature .......................................................13
  Protein detection methods ........................................................................14
    Probe-based methods .................................................................14
      Engineered antibodies ................................................................15
      Other types of affinity reagents ..................................................16
      Homogenous assays .....................................................................16
      Heterogenous assays ..............................................................17
  Nucleic acid detection methods ........................................................18
  Detecting the crowd or the individual ....................................................19
    Averaging measurements ..............................................................20
    Single-molecule detection .............................................................20
      Amplified single-molecule detection “ASMD” ..............................22
  Microfluidics, a short background .....................................................24
    Production methods .....................................................................26
      Traditional silicon manufacturing .............................................27
      Rapid prototyping .................................................................28
  Applications of microfluidics in biology ............................................29

Present investigation .....................................................................................32
  Paper I. Protein detection using proximity-dependent DNA ligation assays ..............................................................32
    Incubation of proximity probes and sample .....................................33
    Connector, enzymes and buffer addition ........................................33
    Read-out using qPCR ...................................................................34
    Perspectives on paper I ..................................................................35

  Paper II, Thermoplastic microfluidic platform for single-molecule detection, cell culture, and actuation. ............................37
    Perspectives on paper II ..............................................................38
    Is there a need for high-throughput production methods? ................39

  Paper III. Digital quantification using amplified single molecule detection .................................................................41
    Perspectives on paper III ..............................................................42
Paper IV, Enhanced molecular analyses by rolling-circle amplification on paired-tag arrays.................................................................44
Perspectives on paper IV.................................................................45

Paper V, DNA Skyline: fonts to facilitate visual inspection of nucleic acid sequences .................................................................................................................................48
Perspectives on paper V.................................................................48
The history of TrueType.................................................................49

Discussion.......................................................................................50

Acknowledgements.......................................................................52

References.....................................................................................56
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>ASMD</td>
<td>Amplified single-molecule detection</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>C2CA</td>
<td>Circle-to-circle amplification</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td>ECM</td>
<td>Extra cellular matrix</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme linked immunosorbent assay</td>
</tr>
<tr>
<td>ESCA</td>
<td>Electron Spectroscopy for Chemical Analysis</td>
</tr>
<tr>
<td>LOCI</td>
<td>Luminescent oxygen channeling assay</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDMS</td>
<td>Poly(dimethylsiloxane)</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative PCR</td>
</tr>
<tr>
<td>RIA</td>
<td>Radioimmunoassay</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SELEX</td>
<td>Systematic evolution of ligands by exponential enrichment</td>
</tr>
<tr>
<td>S/N</td>
<td>Signal to noise</td>
</tr>
<tr>
<td>SMD</td>
<td>Single-molecule detection</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>TAS</td>
<td>Total analysis system</td>
</tr>
<tr>
<td>UV-light</td>
<td>Ultraviolet, wavelengths &lt; 380 nm</td>
</tr>
</tbody>
</table>
Introduction

Novel analysis platforms should fulfill a number of important criteria such as: high selectivity among targets and on-target molecules, low limit of detection (LOD) ideally extending to single molecules, a wide dynamic range, and excellent quantitative precision. Furthermore a perfect assay should be: simple to perform, reliable, offer multiplex capabilities, enable rapid answers, and be cost efficient\(^1\). Although difficult criteria to meet in all their aspects, they are all important and prioritizing among them depending on the application is crucial for a successful analysis method.

A long-term goal in biological research is to achieve complete knowledge of the human body and mind (and those of other species). We are however a great distance from such a state of knowledge. To approach this goal we need a wide variety of tools in our toolbox. Physiology deals with processes and functions of organisms, but due to the complexity cell biologists and molecular biologists prefer to interrogate systems of reduced complexity, such as cells and their molecular building blocks. With the completion of the human genome sequence a few years ago\(^2,3\) research on the function of genes and their transcripts has been greatly stimulated. The advent of the recording of the complete human genome sequence with the complexity of 3 billion base-pairs requires tools selective enough to interrogate differences between even the smallest sequence variants between individuals, but also to measure expression levels of transcripts.

Another important area is the study of proteins - their interactions, modifications, and concentrations. The proteome is defined as “the protein complement expressed by the genome of an organism or cell type\(^4\). Analyses of the proteome face even greater challenges of selectivity than those that pertain to the genome. A single cell can express in the order of 100 million protein molecules, each occurring in numbers that can vary from a few copies to millions per cell. Furthermore, the presence of splice variants in around half of the genes together with numerous different possible post-translational modifications (PTMs) increases the complexity of the proteome. Therefore, exceedingly strict criteria must be placed upon methods for protein analyses.

This thesis presents a novel protein detection method termed proximity ligation. The method uses dual and proximate binding of pairs of affinity reagents each equipped with nucleic acid strands. Upon binding to the protein or protein complex, the free ends of the nucleic acids can be enzymatically joined by a DNA ligase and detected.
I also present two novel analysis platforms. Firstly, a new method to construct microfluidic channels is reported. The method uses rapid injection molding techniques adopted from the compact disc (CD) industry to produce microstructured CDs. A thin layer of silica is applied to the CD surface, followed by covalent bonding of a lid to create channel structures. These microchannels were combined with a set of molecular tools to create a platform for digital enumeration of biomolecules. Secondly, a microarray-based platform suitable for highly resolving multiplexed analyses is described.

As an introduction I will give some background to techniques used for either protein or nucleic acids detection, and I also provide some background to microstructures.
Biomolecular detection methods

Enzymes, tools provided by nature

Before I start describing some of the man-made reagents for biomolecular detection, I will discuss some of the building blocks provided by nature. The use of enzymes has enabled many of the techniques described later in the introduction, but it is equally true that enzyme characteristics have been unraveled using man-made techniques. I will take some important enzyme-based methods as examples, but with no intention to provide a comprehensive account. In the first paper we used quantitative PCR (qPCR) for amplification and detection. PCR was first invented by Kary Mullis and co-workers in the mid 80s\(^5\). Initially they used a thermolabile enzyme (Klenow fragment of *E. coli* DNA polymerase I) requiring addition of enzyme for every cycle of PCR. In 1988 the method’s original inventor improved the technique by using a thermostable polymerase\(^6\) from the *Thermus aquaticus* bacterium (abbreviated *Taq*), originally isolated by Chien and co-workers in 1976\(^7\). The 5’-3’ exonucleolytic property of *Taq* polymerase was later used by Holland *et al\(^8* to simplify readout using electrophoresis. Later followed the use of cleavable fluorescent probes by Ken Livak and co-workers to develop the real-time quantitative PCR method\(^9\). The pioneers of PCR also realized that is important to avoid non-specific amplification. The enzyme should remain inactive until all samples are ready for amplification. This has been achieved by various means to obtain a “hot start”. Two common methods use either a thermostable but temperature-sensitive mutant version of the enzyme, which needs high temperatures to become active, or an antibody blocking the active site\(^10\). In my work I have also used an alternative amplification mechanism that proceeds in a linear rather than exponential fashion, as distinct for PCR. In this alternative amplification procedure, a circular DNA molecule is continuously replicated by a polymerase that must be highly processive. The phi 29 (ϕ29) polymerase was originally isolated by Luis Blanco and Margarita Salas in the early 1980s\(^11\). ϕ29 is characterized by high strand-displacement activity and it is also highly processive\(^12\). ϕ29 furthermore has an efficient 3’-5’ exonucleolytic activity\(^13\). These important features have been put to use in isothermal amplification strategies that provide alternatives to PCR\(^14\-16\).

We also used DNA ligases in paper I, III and IV to join two juxtaposed oligonucleotides. This allowed for improved detection since only ligation prod-
ucts are detected. Ligases catalyze the formation of a phosphodiester bond between the 3´ hydroxyl group and the 5´ phosphate group on juxtaposed nucleic acids\(^1\). The most commonly used ligases are the T4 DNA ligase and the *Tth* DNA ligase, originating from the T4 bacteriophage and the *Thermus thermophilus* bacterium, respectively. *Tth* DNA ligase is thermostable in contrast to the T4 enzyme. The prominent sequence selectivity of DNA joining by ligases, in combination with the requirement for dual hybridization has enabled highly specific genetic analyses with a resolving power of single basepair changes in total genomic DNA\(^1\),\(^2\). Recently, Shendure et al demonstrated the use of sequential ligation reactions for high-throughput sequencing\(^3\). In conclusion, without the availability of enzymes for copying nucleic acid sequences, cleaving reporter molecules at specific sequences, and selectively uniting ends of oligonucleotides this thesis would not have been possible.

**Protein detection methods**

Protein detection methods can be categorized in two major groups: separation-based, and probe-based methods. Separation-based methods are excellent for highly multiplexed protein analysis without the need for analyte-specific reagents. A common mean to obtain an overview of protein complement is to use two-dimensional gel electrophoresis (2D electrophoresis). Separation in the first dimension is based on charge and uses isoelectric focusing. In the second dimension molecules are separated according to size using sodium dodecyl sulfate (SDS) polyacrylamid gel electrophores. After electrophoresis, separated proteins are visualized by different staining techniques or analyzed using mass spectrometry (MS). Using 2D-MS about thousand of proteins can be resolved\(^4\). The combination with mass spectrometry enables detection of PTMs, something that is much more difficult using probe-based assays. 2D-MS have downsides such as poor performance of hydrophobic proteins. The method is also very time consuming and typically not quantitative\(^5\).

**Probe-based methods**

The most used assays for protein detection both in research and diagnostics are probe-based methods. They all require prior knowledge about the protein to be detected, and in particular affinity reagents capable of binding to that protein must be available. The methods can be classified in two major groups, homogenous (solution-phase) and heterogenous (solid-phase) assays, although the distinction is not absolute. The read-out typically represents an average measurement of protein levels, but lately several single-molecule
techniques have emerged where individual proteins are detected and enumerated as discrete events. The latter methods will be discussed separately (page 19).

Antibodies are the most commonly used probes or affinity reagents in protein assays. Traditionally antibodies are made by immunizing a host animal with the protein or molecule you wish to detect, followed by isolation of the antibodies. The immune response towards an antigen by the B-cells of the immune system first results in production of IgM antibodies, and after a second challenge to the same antigen IgG antibodies are produced. The affinities of the IgG antibodies are usually higher and can normally reach a $K_d$ in the nanomolar range. Dissociation constant “$K_d$” is defined as the concentration of free binder at half maximal binding to an antigen under conditions where [binder]>>[antigen]. The $K_d$ is measured in the unit $M^{-1}$. Serum from immunized animals contains polyclonal antibodies and no separate batches from different or even the same animal will be identical. A solution to reduce variability can be to generate monoclonal antibodies as first described by Georges Köhler and César Milstein. The principle is based on immunization followed by retrieval of B-cells from the spleen and subsequent fusion of B-cells and an immortalized cell line (i.e. malignant cells) to generate hybridoma cells. Selection of the hybridoma cells expressing the desired antibody results in immortalized cell clones, each producing a unique antibody. Using the physiological machinery for antibody production is fairly straightforward, but this also compromises the possibility to generate binders towards epitopes that are highly conserved among species, and also towards toxic molecules. The $K_d$ of native antibodies is typically never below 0.1 nM.

To improve affinity or stability of native antibodies a number of techniques have been developed to either engineer antibodies or produce affinity reagents that are not based on antibodies. Many of these binders can be selected for affinity towards “any” molecule and they can also often be expressed in vitro making production straightforward.

**Engineered antibodies**

To increase affinity of an antibody based affinity reagent, extensive mutagenesis of a specific sequence element is commonly used, followed by in vitro selection. Chen et al used mutagenesis on an α-VEGF antibody fragment expressed using phage display and they reported a 120-140-fold improved affinity compared to native α-VEGF antibodies. Boder et al used yeast surface-display to increase the affinity of an α-fluorescein antibody and reported a $K_d$ of 270 fM in PBS buffer and 48 fM in a low salt buffer, about thousand-fold higher affinity than native antibodies. Alternative expression systems to phages and yeast include methods that do not require any cells for binder production. Hanes et al used ribosome display, a fully in
vitro based system without a need for expression in cells to improve the affinity of single chain antibody fragments with about 40-fold, reaching a $K_d$ of 82 pM for binding to insulin$^{27}$.

**Other types of affinity reagents**

During the last couple of decades several alternatives to antibodies as affinity reagents have been proposed involving other protein scaffolds or using nucleic acid. Aptamers selected using the SELEX (systematic evolution of ligands by exponential enrichment) procedure have proven to be an interesting alternatives to antibodies. The SELEX procedure typically uses a chemically synthesized randomized sequence of 20-40 bases flanked by two fixed PCR primer sequences as a starting library containing up to $10^{15}$ molecules. The library is used for 8-15 rounds of *in vitro* selection and enrichment by PCR. After the last selection, the pool of specific binders is cloned and sequenced to identify successful aptamers. The first papers describing the isolation of aptamers in 1990 employed RNA sequences$^{28, 29}$ but for stability purposes DNA aptamers$^{30}$ seem to be a more viable alternative as affinity reagent. The affinity of aptamers is typically higher compared to antibodies, and $K_d$ values below 100 pM have been achieved$^{31}$.

Protein scaffolds other than antibodies have also been extensively explored. Many of the binders in this class are much more robust compared to antibodies, making them suitable for use in affinity columns and potentially in microarray applications that risk denaturing antibodies. Examples of scaffolds that are being explored include the Z-domain of *Staphylococcus aureus* (Affibodies) lipocalins, and fibronectin III (reviewed in$^{32, 33}$).

**Homogenous assays**

In homogenous assays both target recognition and signal detection is performed in solution without any washes or separations. One advantage with homogenous assays is that they are less labor intense compared to heterogeneous assays and they are suitable for automation. Performance of the assay typically depends strongly on the affinity of binders and type of reporter system. To generate a strong signal compared to background many commonly used homogenous assays require dual recognition to elicit a signal. The field of homogenous immunoassay was pioneered by Edwin Ullman and co workers who have developed some of the most commonly used techniques to date. The first assay presented in 1972 was a competitive enzyme-coupled assay for detection of morphine, and LOD in the nanomolar range was reported$^{34}$. Later developments requiring proximal binding to elicit a signal such as in the LOCI (luminescent oxygen channeling assay) assay$^{35, 36}$ have enabled low concentration of proteins to be detected. The assay format
often used in clinical laboratories today, involving chemiluminescent detection, was first described in 1976 by Schroeder et al\textsuperscript{37}. In recent years fluorescence resonance energy transfer (FRET) have become popular. In early work by Ullman and co-workers, they used a fluorescein-labeled antigen and a rhodamin-labeled antibody in a competition based assay where rhodamin acted as a quencher\textsuperscript{38}.

**Heterogenous assays**

By introducing a solid phase, larger sample volumes can be analyzed, and higher concentrations of binders may be used compared to homogenous assays. Most solid phase assays also use extensive washing to remove both excess reagents and potential inhibitors for subsequent steps. Because of the higher concentrations of probes, the dynamic range is often larger compared to homogenous methods.

On the downside, a solid support causes binding reactions to be restricted by mass transport kinetics compared to the free diffusion in homogenous assays.

The most common form of heterogenous assay used today is the sandwich immunoassay, and specifically the enzyme-linked immunoassay (ELISA)\textsuperscript{39}, which evolved from the RIA assay developed by Berson and Yalow in 1960\textsuperscript{40} to the dual specific assay described by Leif Wide and co-workers in 1967\textsuperscript{41}. The ELISA technique uses enzymes, usually horseradish peroxidase (HRP) or alkaline phosphatase (AP) linked to a secondary binder (often a species-specific antibody) to convert a substrate to a detectable signal for either colorimetric or chemiluminescent readout. Recently, use of powerful nucleic acid amplification strategies have enabled some of the most sensitive protein detection reactions to date, some examples of procedures include the immuno-PCR assay\textsuperscript{42} and the bio-bar code technique \textsuperscript{43, 44}. In immuno-PCR the secondary antibodies are equipped with an amplifiable nucleic acid sequence used for detection. In the bio-bar code assay, primary antibodies coupled to beads capture the antigen. This step is followed by incubation of smaller detection beads coupled with other antibodies toward the same antigen and also with nucleic acid bar-code sequences. After washing the barcodes are released and detected, directly or after PCR amplification. A problem with using powerful amplification techniques coupled to a directly amplifiable tag is that any nonspecifically bound tags will contribute to background, thereby requiring extensive washing to lower background.

The so far described assays are usually employed to measure one protein at a time. By using differently labeled fluorescent microparticles, multiplex analyses of up to 15 protein have been performed in one tube using flow cytometry\textsuperscript{45, 46} with similar LOD but increased dynamic range compared to the conventional ELISA format.
The use of a solid support enables spatial separation of numerous antibodies bound to a surface, thereby increasing multiplexing capabilities over solution phase bead-based arrays. Schweitzer et al performed simultaneous detection of 75 cytokines on microarrays using rolling-circle amplification (RCA) to increase sensitivity.47

Nucleic acid detection methods

Similar considerations as those for protein detection also apply for nucleic acid detection; you either know what you are looking for or you don’t. The most common way to de novo decipher nucleic acid sequences still involves variants of Sanger sequencing using chain terminators as originally described 1977.48

Due to the nature of papers described in this thesis I will focus on probe-based methods for nucleic acid interrogation and specifically microarray-based techniques. The field of DNA microarrays has made a tremendous difference in how molecular analyses can be performed, enabling many thousand parameters to be analyzed in a sample in parallel. Microarrays have been extensively used for a number of investigations, such as SNP analysis, expression profiling, resequencing, and copy number measurements, to mention some examples as reviewed in.49 The field has evolved along several trajectories. In particular, the nature of the probes and production methods for arrays have developed rapidly. In this context I will refer the oligonucleotide of known sequence spotted on the array as the probe, while the target is the unknown sequence being interrogated.

Early reports used cDNA products about 1 kb in length spotted on activated glass slides for expression analysis. In a landmark paper published by Patrick Brown’s group in 1995, expression of 45 Arabidopsis genes were presented as a model system for parallel expression analysis.50 The probes used were PCR products, physically deposited on the array surface followed by cross-linking. An alternative to deposition is direct synthesis on the array surface. The most commonly used method for site-directed synthesis has evolved from integrated circuit fabrication techniques using photolithography.51 The use of light-sensitive phosphoramidite deoxynucleosides enabled stepwise light directed synthesis using a series of standard photolithographic masks. Using this technology feature sizes of 8x8 μm have been used to produce arrays containing about 2.5 million unique sequences with overall dimensions of 1.25 cm squared, as commercialized by Affymetrix. Matsuzaki et al used Affymetrix microarrays to simultaneously genotype over 100 000 SNPs.52 An alternative to chromium masks is using digital micromirrors for addressing light-directed synthesis, and feature sizes of 16x16 μm has been produced.53 One problem with light-directed synthesis is that
the stepwise yield has been limited to around 95% (Pirrung et al reported a range of 80-97% \textsuperscript{54}) compared to >99% for conventional chemistry. This makes synthesis limited to shorter oligonucleotides and a length of 25 bases is typically used.

Other approaches using standard chemistry combined with inkjet printer technology enables longer oligonucleotides to be array-synthesized, but typically generates larger features (about 100 \textmu m)\textsuperscript{55}. Another interesting alternative is the combination of standard chemistry combined with microfluidics for DNA array production, theoretically enabling higher coupling efficiency without increasing feature size\textsuperscript{56}.

Most often, microarrays are tailored for specific applications. Accordingly, different arrays are needed for every new application. By contrast, the introduction of bar-code or zip-code arrays in the mid 90s enabled separation of the biomolecular assay from the microarray hybridization\textsuperscript{57}. By using this strategy detection reactions can be performed in solution, followed by subsequent decoding using universal microarrays. Zip codes are oligonucleotides of some 20 nucleotides in length that are \textit{in silico} designed so as not to cross hybridize to any known genomic sequence and to have similar thermodynamic properties. Zip code arrays have also been used for genotyping\textsuperscript{58, 59} to analyze splice variants\textsuperscript{60} and to measure point mutations\textsuperscript{61}. We have further developed this approach in Paper IV.

Detecting the crowd or the individual

Most of the commonly used biomolecule detection methods, both homogenous and heterogenous, are based on average measurements across many molecules. This enables use of relatively low cost instrumentation for readout, or as in point-of-care tests, the human eye. A typical question to be answered is “what is the concentration of a specific biomolecule in a particular sample?”. In many cases measured ensemble averages suffice to answer such questions. However, another possibility is to enumerate molecules in discrete steps. If it had been possible to study individual molecules with the same ease as performing averaging analysis this would probably be the method of choice since in theory it provides the ultimate level of quantification and sensitivity. Single-molecule analysis has mainly been held back by the considerable technical demands raised by such analyses. The possibility to study single molecules - not just their numbers but also their localization and changes over time - would provide valuable insights in how molecular functions are coupled in larger complexes, ultimately in whole cells or organisms. Single-molecule analyses will also reveal if an elevated expression of a protein or a transcript is present in the whole cell population or only in a
minor sub population having a proportionately larger increase. Clearly, insights in cellular heterogeneity will be critical for understanding diseases such as cancer. Furthermore, single-molecule analyses can reveal differences between apparently identical biomolecules. A fundamental consideration for single-molecule detection is the challenge to ensure that it is the correct one.

Averaging measurements
The LOD in averaging measurements is usually set by unspecific binding of the reporter molecule to the solid phase or to irrelevant target molecules, and also by the limited affinity of the binder used for detection. The LOD also depends on how sensitively the reporter molecule can be detected *per se*.

In order to detect the recognition event between the binder and the target, a measurable response needs to be created, and this response must relate to the amount of binding for use in quantitative assays.

Early techniques used radioactivity, *e.g.* detected as ionizing radiation acting on photographic films where the degree of darkening of the film was used for quantification. In scintillation techniques, the radioactive decay is converted to light that is subsequently detected. Radioactive detection can be very useful for competition-based assays since the labels used are small and therefore do not change the properties of the molecules to be investigated. Drawbacks of these labels are the fact that they are hazardous to work with and have limited half-lives. To enable more convenient means of detection, although not necessarily more sensitive, a number of different properties are currently used for averaging measurements. The least demanding instrumentation-wise are probably agglutination and colorimetric techniques that usually are detected based on absorption of specific wavelengths of light or by the human eye. Using the properties of fluorescent dyes to exchange excited state energy more selective responses can be achieved in various FRET-based techniques. qPCR is also a common method for averaging measurements of nucleic acids. In qPCR the accumulating amplification product is detected due to an increase in fluorescence. During amplification the polymerase degrades probes equipped with a fluorescent and quenching moiety hybridized to the target sequence, thereby releasing the fluorophore from the quencher and resulting in an increase in fluorescence. The round of amplification at which the fluorescence reaches above a threshold value is inversely proportional to the logarithm of the amount of target.

Single-molecule detection
The field of single-molecule detection (SMD) has undergone a dramatic development during the last couple of decades and only some aspects will be discussed here. The study of single fluorescent molecules for investigation of various physio-chemical properties is itself a large field but this section will
focus on SMD for biomolecule analysis. Although fluorescence is among the most common properties in use for detection, other techniques have also been used such as patch-clamping to monitor current fluxes of individual ion channels. Detection methods for single-molecule analysis may be classified according to if they are aimed at studying single-molecule properties such as motion, turn-over, localization, conformational changes, and so forth; or if they are using the discrete nature of single molecules for improved biomolecule analysis. The first group can potentially unravel new biological findings impossible to reveal using averaging techniques; the second group mainly focuses on increasing precision and sensitivity, but also to study subpopulations of biomolecules in “snap shot” situations.

The first SMD analysis was performed by Thomas Hirschfeld in the mid-70s. He detected IgG antibodies labeled with a polyethyleneimine chain to which 80-100 fluoresceine molecules had been conjugated. In 1990, improved instruments were able to detect single rhodamin molecules in hydrodynamically focused flow using a confocal setup. After 1990 there has been a tremendous increase in SMD studies of biomolecules accompanied by improved instrumentation. Single motor proteins have been successfully characterized in several studies, based on energy consumption, motility, and mechanical properties. Tracking of receptor movements in living cells is another interesting application for SMD.

One difficulty when measuring proteins is to attach fluorescent labels without inducing conformal changes and thereby altering or disrupting function. This is much more trivial when working with nucleic acids, where labels can be incorporated by enzymes or synthesized in vitro. Alternatively, studies of DNA tethered to solid phases can reveal novel findings of enzyme activity and mechanical properties of native DNA without prior labeling of the nucleic acid itself. In a study performed by Strick et al they attached a double stranded DNA of 11 000 base pairs in length to a solid support and a magnetic bead of 0.5 μm in diameter in each end respectively. By applying a rotating magnetic field it was possible to induce supercoiling of the DNA, which subsequently could be relaxed using a topoisomerase. Experimental data suggest that topoisomeras II unwinds two supercoils for every cycle and that it is dependent on ATP concentration. Other studies of enzymes active during genome replication are studies of single DNA polymerases. Wuite et al analyzed replication rate of T7 DNA polymerase on a single stranded target tethered between two beads. Different stretching forces on the ssDNA using optical tweezers were applied during replication. The amount of stretching affected polymerization rates, and forces of about 6 pN yielded the highest speed of ~200 bases s⁻¹. Interestingly, when forces above 40 pN were applied the enzyme shifted from polymerization to exonuclease. The authors also noted that different polymerization rates were recorded for individual enzyme molecules, suggesting differences among the T7 DNA polymerase population.
The interest in polymerases extends beyond characterization of molecular properties, and a particularly important emerging field for single-molecule analysis is single-molecule sequencing applications. As previously mentioned the sequencing of the human genome was performed using Sanger sequencing with capillary electrophoresis for detection. This is a costly method and techniques capable of producing sequence data more rapidly and less expensive are in high demand. Many variants have been proposed, and some are more likely to succeed compared to others. Early approaches were based upon complete fluorescence labeling of sequences of interest with distinct spectra for each base, followed by trapping individual molecules in a flow system. Thereafter a nuclease would degrade the sequence stepwise and the released fluorophores detected in real time. The principle was originally described in 1989 by Jett et al and later explored by Rudolf Rigler’s group and others as reviewed in\cite{71}. More recently, another single-molecule approach based on sequencing by synthesis has been proposed by Stephen Quake and co-workers\cite{72, 73}. This technique may outperform the degradation-based approach. In this method DNA molecules are bound to a solid support and fluorescent sequencing primers are added followed by registration of positions on a planar surface. Thereafter, fluorescence labeled nucleotides are added in cycles, coupled to detection after each addition. Before every new addition the incorporated fluorophores are bleached to enable subsequent base information to be collected. In a follow-up paper the process was automated\cite{73}. So far, read length is limited to a few bases and improvements in labeling efficiency, enzyme incorporation, and bleaching is likely to be needed to enable longer reads.

**Amplified single-molecule detection “ASMD”**

A key element in SMD analysis is to improve signal over background and this is usually performed by reducing the investigated volume by various means. If the volume is small (< 1fl), background can be reduced to a minimum since it scales linearly with volume. The signal however does not change with decreased volume and thus signal over background is increased. Furthermore, high quantum efficiency detectors with low dark noise and high quality optics are needed. A common way to reduce background is the use of confocal systems, where a pinhole efficiently reduces analysis volume by excluding light generated outside the conjugate plane, but other approaches using levitated microdroplets, evanescent illumination techniques or physical confinement using microstructures have also been used\cite{74}.

Although proven possible, single molecule detection on the fluorophore level is still demanding. An alternative to circumvent the difficulties of SMD but still enable the benefits is not to identify single molecules but detect their amplified product. Interestingly, this strategy only advanced a decade ago, almost two decades after Hirschfelds early experiments. And perhaps the focus on achieving the “ultimate” level of detection held back development
of the amplified techniques. The golden standard of nucleic acid analysis is still qPCR, the power to amplify even single copies of target and allow quantification over a dynamic range spanning seven orders of magnitude is quite powerful. However, there are some downsides of qPCR. The technique suffers from poor quantitative precision and multiplexing capabilities.

Many ASMD approaches use dilution or compartmentalization of targets prior amplification using PCR, thereby combining benefits from single molecule analysis with high amplification power to facilitate digital detection.

An early report used dilution of genomic DNA templates to the single molecule level for determination of the haplotype of that individual (where haplotype being “A set of associated SNP alleles in a chromosome region”). The diluted fragments were amplified using PCR followed by digestion using an endonuclease that specifically cleaved a polymorphic site. The products were subsequently analyzed using gel electrophoresis. Although the authors did not use enumeration of the products, they showed the potential of ASMD. Other reports have diluted samples in 96-well plates followed by amplification and detection.

Chetverina et al. diluted viral DNA or RNA molecules in polyacrylamide gel “caging” individual nucleic acid sequences followed by amplification using PCR and subsequent detection using autoradiography. Individual PCR colonies originating from one viral sequence was enumerated. In a recent development from the same group they used real time detection of growing colonies. Mitra et al combined colony (as they named polony) technique with single basepair extension for high precision genotyping measurements. To increase throughput, micro-emulsions have been used to confine single target molecules bound to micrometer sized magnetic particles followed by emulsion PCR. The amplification products remains bound to the bead, and subsequent rupture of the micro-emulsion followed by magnetic purification of the beads enables fluorescence labeling and detection using a standard flow cytometer. The approach has been named BEAMing based on the principal components used in the assay (beads, emulsion, amplification, and magnetics). The authors used the technique for high precision genotyping and to study uncommon variations. A further addition to the toolbox was recently published by Shendure et al. They combined RCA of circularized genomic fragments followed by emulsion PCR on magnetic beads. An enrichments step separating un-amplified from amplified beads was performed prior loading in a microfluidic channel. And in the final stage sequencing by ligation was performed. In a demonstration experiment they successfully sequenced an Escherichia coli strain with an error rate of less than one in a million bases.

In my work I have used RCA as an alternative to the confinement methods for ASMD described so far. In RCA a circular target is continuously copied, resulting in a several kilobases long concatenated amplification
product (RCP) as originally described by Fire et al. The product collapses into an amorphous blob of DNA and therefore creates its own compartment. If a specific sequence element is encoded in the circular target molecule, it enables subsequent labeling of the RCP with fluorescent oligonucleotides. Lizardi et al used RCA coupled to a solid support for amplification of individual allele specific ligation products. Paper III in this thesis describes an alternative method for ASMD detection using RCA coupled to padlock and proximity ligation probes.

Microfluidics, a short background

In our daily lives we are all surrounded by microfabricated structures without most of us thinking about it. I am referring to microelectronic devices such as the processors of computers, cell phones, the controllers that make our car moving and so on. The prediction made by Gordon Moore in 1965 that the development of transistor density would double every 18 months has nearly been realized by the rapid development in the field of microelectronics. Although, the doubling rate has been about 24 months it is still impressive. However, the mode of transistor operation has not changed alongside with the downsizing. Fluid behavior however change drastically when forced to operate in micro-dimensional environments compared to macro-dimensional, and this is one effect explored in the field of microfluidics. The common denominator of the two fields is that manufacturing techniques and materials used in the microelectronic community was the foundation of which microfluidics initially was built. Microfluidics typically handles liquids in channels of sizes ranging from 10-100 μm in width. There are many benefits of scaling down reaction volumes, except the obvious such as lower reagent consumption, increased kinetics thereby shortening analysis time, increased heat dissipation due to increased surface to volume ration, and possibility to perform many reactions in parallel on a small area. Perhaps less obvious benefits is the behavior of fluid when scaled down in microchannels. One of these effects is that viscous forces dominate over inertial forces in microchannels. This is described in the dimensionless Reynolds number, Re = ρvL/μ, where ρ is the fluid density (kg/m³), v is the fluid velocity (m/s), L is the cross-sectional dimension (m) and μ is the fluid viscosity (kg/m·s). When Re >2300 fluid flow becomes turbulent (as in most every day experience), however in microchannels Re is normally close to 1 meaning laminar flow. In laminar flow a particles velocity moving in the fluid is not a random function of time, and the effect if two fluid streams flows parallel to each other mixing between them only occurs via diffusion (as observed in figure 1). This and many other effects of fluid physics in the micro-scale is described in. The effect of laminar flow has enabled a multitude of novel biological studies inaccessible on the macro-scale. Microflu-
idics has been around for some decades now and the initial prediction of rapid commercial successes has not been realized. Although, there are some exceptions, one being the capillary electrophoresis chips commercialized by Agilent and Bio-Rad (both empowered using Caliper microstructures) and the DNA sequencing instruments used in sequencing the human genome, to mention some. The field has so far also been limited to a relatively few research group and has not been readily adopted by other research disciplines such as the cell-biological community (who may benefit the most). The future will tell if microfluidics moves even further than from the most zealous researchers. I am convinced that this will happen, although a bit later than originally expected.

Figure 1. Effect of laminar flow in microchannels. Hydrodynamic flow of two different fluids in a microchannel. In one of the fluids a blue dye is added, and diffusion of the dye into the transparent fluid stream is apparent. With increasingly higher flow rate the boundary between the fluids becomes sharper as a result of shorter diffusion time. Experiment performed by Jonas Melin.
Production methods

As mentioned above the realization of a broad community using microfluidics has not been realized, and this may depend on that many techniques used for production of microstructures are limited to a few specially equipped laboratories. A number of alternative production methods have been proposed that simplifies production and making it available to non-specialists. This section will describe the two extremes, the conventional microelectronic approach and the rapid-prototyping approach (described in more detail in 86 and 87 respectively). There is considerable overlap between the methods, and the following section will give a brief overview and some pros and cons of the two approaches. Figure 2 gives a schematic illustration of the different fabrication methods.

*Figure 2. Different methods used for production of microfluidic structures. Traditional silicon lithography is illustrated in “a” and “b”, where in “a” the fabricated substrate is used for microfluidic experiments, and in “b” the etched silicon substrate serves as a mold for subsequent electroplating to produce a master in nickel. The nickel master is then used for molding in PDMS or for rapid injection molding replication strategies. In “c” the rapid prototyping method is illustrated. The colored boxes in the lower right corner illustrate the materials or components used in every step.*
Traditional silicon manufacturing

“Traditional” is referring to methods originally developed in the microelectronics industry for production of semiconductor circuits. But, they can also be used for production of microfluidic systems. Using traditional manufacturing it is possible to create structures of dimensions below 100 nm. The materials used are mainly glass or silicon and there are basically two formats for channel production.

The first step in both methods is to design the structures using a computer program, this two dimensional drawing is then transferred to a mask made of glass coated with a thin layer of chromium. The resulting mask is transparent in the regions of the channels and opaque outside the structures. In parallel a glass or silicon substrate is coated with a thin layer of photoresist (photoresist being a light sensitive material that if “positive” becomes soluble and “negative” becomes insoluble to the developing solution after exposure to UV-light). The substrate is aligned with the mask and subjected to UV radiation. For this type of construction usually a positive photoresist is used. The soluble part of the resist is removed and underlying substrate etched using “wet” or “dry” techniques to create trenches. From hereon the method diverges:

In the first format, the channels are created in either glass or silicon and the resulting structures are also used for the experiments (figure 2a). To form enclosed channels a lid is bonded to the substrate. Glass is usually used, but PDMS (Poly(dimethylsiloxane)) or plastic materials have also been used. This strategy was used in some of the first microfluidic applications, where structures were created in glass and later used for electrophoresis of small fluorescent molecules or peptides. This strategy is beneficial when materials such as glass and silicon are required due to their compatibility to several organic solvents often needed for organic chemistry applications. Furthermore, in manufacturing of channels less than 100 nm in width (often referred to as nanofluidics) glass and silicon are well suited. This method is also suitable for structures containing geometries that are difficult to replicate, such as structures with large differences in aspect-ratio, that is depth-to-width of structures. It is also possible to include active elements, such as heating devices with relative ease or chemical sensors, something impossible using rapid prototyping. The drawback of using glass and silicon is that production time is long, the cost for large scale production scales almost linearly. Furthermore, for production of complex systems it is necessary to integrate active elements such as pumps and valves, which is difficult in rigid materials.

The second way of using glass and silicon is to produce a master from which subsequent structures are replicated (figure 2b). To do this, the silica substrate of negative polarity (i.e. channels are represented as valleys) is converted to a mold of positive polarity using electroplating. In electroplat-
ing a negatively charged substrate is submerged into a solution of the salt of the metal to be deposited, permitting the metal ions to be reduced to the metallic form on the negatively charged surface. This master can thereafter be replicated using molding of PDMS or more suitably be used for rapid injection molding of thermoplastics. This technique is well suited for cost-effective large-scale production of microchips, but so far it has only been used to a limited extent by the research community. A number of methods for producing enclosed channels in thermoplastics have been presented, such as variants of the thermal annealing process, using adhesive tapes or solvent bonding using sacrificial structures. However, many of these techniques are not suitable for optical detection or for incorporation of active elements. Paper II in this thesis describes an alternative production method that solves some of these problems.

Rapid prototyping

The invention of the rapid prototyping strategy for use in microfluidics was made by George Whitesides group in the mid 90s, and the technique was then further developed in Stephen Quakes group creating multilayered soft lithography. The technique enables production of microstructures without costly equipment and has by far been the most frequently used in the research community. The structures are designed in the same way as described before using a computer program. The drawing is then printed using high resolution laser printers on transparency films to generate the mask. The substrate (usually silicon) is subsequently coated with a negative photoresist followed by mask alignment and irradiation using UV light. Uncured photoresist is removed, leaving a mold with channels protruding from the plane (figure 2c). In rapid prototyping the material of choice is PDMS due to excellent material properties. PDMS is optically transparent down to 300nm, biocompatible and easy to seal to itself or glass using oxidation as reviewed in. The flexibility of PDMS enables integration of active elements, such as pumps and valves. The benefits of this technology are short time from design to functional chip, relatively inexpensive for prototype construction, easy to interface and the before mentioned material properties of PDMS. On the downside, the method is not suitable for large-scale production due to occupation of the master for 30-90 minutes (as 10-20 seconds in injection molding), making it difficult to commercialize systems developed using rapid prototyping. The flexibility of PDMS is crucial for production of active elements, but it may not be optimal for production of routine products. Furthermore rapid prototyping is usually not suitable for features smaller than 50 µm, although feature sizes down to 10 µm have been reported using high resolution photoplotters instead of laserprinters.
Applications of microfluidics in biology

To cover the field of microfluidics within this introduction is impossible, but I will give some examples where microfluidics have improved or enabled biological analyses, and of some more technical oriented achievements. A number of recent reviews can be found here106-112.

As mentioned before early uses of microfluidics were mainly for separation based analysis. Later on in the development of the field, numerous macroscopic analyses have been scaled down to the micro-scale, and in some cases drastically improving their performance. One technical demanding task in microfluidics has been to include a multitude of functionalities in the same chip to get a fully functional liquid processor unit, termed total analysis system or TAS113. Some recent examples have been produced both in silicon and in PDMS. Pal et al used a silicon based platform with integrated modules for fluid handling such as valves, electrophoresis but also thermal control and sensing. The chip was used to identify viral DNA using PCR amplification followed by RFLP analysis using electrophoresis114. Balagaddé et al used soft plumbing in PDMS to monitor and automatically control growth of bacteria. The chip housed six mini-reactors of 16 nl, each separated from the others enabling different growth conditions to be evaluated115. This type of device may be used to monitor effects of antibiotics in screening applications but also enable selection of sub-populations for following molecular analysis. The possibility to use μTAS for chemical synthesis is also an interesting option. Miniaturized systems have proven useful for synthesis of CdSe nanocrystals (quantum dots)116, but also for production of short-lived isotope compounds used in PET (positron emission tomography) studies117. In these systems fluid handling is of major importance and means to move and enclose fluid is important. For systems in soft materials pumping and valving is relatively straight forward as originally described using two layer systems103. For hard materials this is a more demanding task, the use of hydrophobic barriers for valving118 is an interesting technique. Hydrophobic barriers have been used in “lab on a CD” applications coupled with centrifugal forces to propel fluids oriented in a radial geometry119. Recently light has been proposed as an alternative method to precisely control fluid movement in microchannels without need for complex microfabrication. The method uses conversion of optical energy to hydrodynamic force via nanocrescent particles120.

Many enzymatic assays are product inhibited and the possibility to use microfluidics has enabled major improvements in performance compared to reactions performed in test tubes. In pyrosequencing, stepwise incorporation of unmodified nucleotides are converted to flashes of light used for sequence assembly, in the first paper a solid phase was used121, and in a later version the assay was performed in solution122. Although an elegant way for obtaining sequence data, sequencing of long stretches has proved difficult due to
build-up of intermediate byproducts and enzyme exhaustion. Early instruments were capable of performing 96 simultaneous reactions with a read length of 20-40 bases. In a recent development, the assay was performed in a microfluidic chamber containing about 1.6 million reaction wells harboring a volume of 75 pl each. The wells were filled with beads to which the DNA to be sequenced is attached. The optically transparent bottom of the wells was fibre-coupled to a CCD-detector for image capture. Using this format, improvements in read-length and capacity have been dramatic, with an average read of 110 bases and the possibility to perform 1.6 million sequence-reactions simultaneously. However, this systems also suffers from much of the same problems as the original format, and perhaps perfusion through the bottom of the wells may improve read-length. This type of flow-through instead of flow-by setup has proven to work efficiently in a similar microfluidic setup\cite{123}.

The possibility to interrogate cellular functions using microfluidics offers very interesting manipulations inaccessible on the macro scale. Some examples are the possibility to expose a cell or parts of a cell to different stimulus using laminar flow, but also to precisely define gradients of either extracellular matrix (ECM) components or molecules present in the growth media. Microfluidics also enables handling of individual cells or a multitude of cells and exposing them for different growth conditions. Microstructures within such systems can also be used to interrogate how adherent cells are affected by physical contact.

Early work in cell culture was typically using generically designed channels, later reports have made dedicated vessels for culturing applications. Hung et al\cite{124} recently developed a microfluidic cell chip capable of performing 100 simultaneous culturing reactions. Viable and dividing cells were recorded for up to ten days of culturing. The design also enabled finely tuned gradients to be applied column-wise, and possibility to harvest cell after culture. Other reports have enabled 3-dimensional cell culture to interrogate cell-spreading in all dimensions\cite{125}.

The use of laminar flow to expose parts of cells with chemicals has also been extensively used to study intra-cellular transport mechanisms. Much of this work has been done in George Whitesides group. Some examples are introduction of mitochondrial selective dyes to parts of cells followed by tracking mitochondrial movement within the cell, but also for region specific disruption of the cytoskeleton\cite{126}. They also studied movement of neutrophils towards chemoattractant-gradients gradients of chemoattractants over growing and monitored directed movement\cite{127}.

Adherent cells are normally surrounded by a microenvironment of extra cellular matter, and studies of cellular interactions with different types of biomolecules are of importance. Théry et al performed micro contact printing using a PDMS stamp to control ECM deposition followed by culturing of cells onto the printed surface. They linked geometrical distribution of
focal adhesion points to the ECM with orientation of the mitotic spindle of HeLa-B cells. Microfluidic systems can also be used to create controlled 3-dimensional orientation of ECM components within microchannels that could be used to study cellular responses. Using microstructures within channels it has also been possible to measure individual muscle cells contractive force using micro-pillars.
Present investigation

The Present investigation was conducted in order to develop enhanced detection methods for proteins and nucleic acids as well as improved read-out formats; a tool for visual inspection of nucleic acid sequences is also described. In paper I we present a protein detection method that uses multiple binding of affinity reagents to convert the presence of a protein to a nucleic acid reporter molecule by means of enzymatic ligation. In paper II we describe a scalable platform for microfluidic analyses; this platform was used in paper III for high precision biomolecular analysis using amplified single molecule enumeration. In paper IV we describe a microarray-based platform that enables cDNA expression analysis over a wide dynamic range with high selectivity. In paper V we describe a set of true type fonts to graphically represent nucleic acid sequences that can be used for manual inspection of sequence motifs and for teaching purposes.

Paper I. Protein detection using proximity-dependent DNA ligation assays

In this paper a novel protein detection method is presented. The method can be thought of as reverse translation, where a detected protein gives rise to a specific nucleic acid sequence that can be detected. Two so-called proximity probes, each equipped with a protein affinity moiety and a specific nucleic acid sequence, bind to the target protein. Upon binding, the nucleic acid sequences are brought in close proximity, and a third complementary oligonucleotide termed “connector” allows the nucleic acids to be enzymatically joined. The newly formed sequence can thereafter be detected using qPCR\textsuperscript{9,132}.

The homogenous format of the assay can be broken down in three consecutive steps (figure 3). 1) Incubation of proximity probes and protein, 2) addition of connector oligonucleotide and all required reagents for ligation and subsequent PCR amplification, 3) Detection using qPCR. A more in-depth description of these three steps will serve to explain the basis of the proximity ligation assay.
Figure 3. Illustration of the proximity ligation work-flow. First proximity probes are incubated with the sample (1), this is followed by addition of all components required for ligation and qPCR (2) and last in (3) amplification and detection of reporter molecules using qPCR.

Incubation of proximity probes and sample
In paper I the protein binding moieties used were DNA aptamers, nucleic acid sequences that have been selected using the SELEX technique to specifically bind either the homodimeric plateled derived growth factor BB (PDGF-BB) or the monomeric alpha-thrombin protein. The proximity probe concentration and the incubation volume are key factors in homogenous proximity ligation, and they were optimized for high signal to noise in figure 2b, paper I. A concentration of 20 pM proximity probes and an incubation volume of 5 microliters were selected and subsequently used in all experiments. A suitably low proximity probe concentration reduces background ligations to a minimum, but does not adversely affect protein derived ligation levels. At higher probe concentrations signal to noise (S/N) decreases mainly due to increased target independent background ligations. The incubation volume of five microliters was chosen since the proximity probe concentration determines the probability of detecting the target protein, and larger volumes increase background ligations.

Connector, enzymes and buffer addition
Following incubation all components required for enzymatic ligation and detection by qPCR was added. The joining of two proximity probes is orchestrated by the connector oligonucleotide. The optimization of the connector concentration and ligation time is illustrated in figure 2a, paper I. At high connector concentration all proximity probe arms are occupied by a connec-
tor oligonucleotide. However, the local concentration of two proximity probes bound to the same protein is about a million fold higher compared to freely diffusing probes when probe concentration is 20 pM. This effect enables rapid ligation of juxtaposed proximity probes and simultaneously renders unbound probes unavailable for subsequent ligation. At high connector concentrations signal and background ligation change little over time resulting in a robust system. A connector concentration of 400 nM was determined to be suitable for the assay.

Read-out using qPCR
In paper I we used qPCR for detection. Each proximity probe was equipped with a primer sequence, allowing the ligation product to be selectively amplified. To avoid unspecific PCR products that could arise by being templated by the connector oligonucleotide, we optimized the length of the hybridizing part to the proximity probes. We found that ten hybridizing bases to each proximity probe rendered ligation independent signals to a minimum. Furthermore, three additional bases without complementarity to the proximity probes were added at each end of the connector oligonucleotide, also introduced for the same purpose.

Detection of PDGF-BB using the homogenous format of proximity ligation performed equally well in fetal calf serum, cerebrospinal fluid, EMEM cell culture medium as in buffer. LOD was 24 000 protein molecules, 1000 times lower than a commercially available ELISA for the same protein. The affinity of the binders is of great importance for the performance of the assay as reflected in the alpha-thrombin assay, where the lower affinity of the binders resulted in a LOD about 15-times higher. The requirement for dual recognition events in order to generate a signal has proved essential to ensure selectivity for nucleic acid detection techniques. This property was further investigated in figure 6, paper I. where, we compared a solid phase proximity ligation format to Immuno-PCR. First, a polyclonal antibody was adsorbed to the bottom of a reaction vessel, followed by addition of sample and proximity probes. The same aptamers were used in both cases with the difference that Immuno-PCR uses a single binder equipped with a directly amplifiable tag with no requirement for pair-wise binding and ligation. In immuno-PCR all unspecifically adsorbed immuno-PCR probes contribute to background resulting in higher LOD compared to dual recognition using proximity ligation. Solid phase proximity ligation also adds additional selectivity by requiring three specific recognition events.
Perspectives on paper I

The stimulus for the presented work was the need for improved protein detection methods to monitor the presence of proteins or protein complexes. As protein binders we used SELEX aptamers, thereby simplifying construction of proximity probes since they can be synthesized using conventional oligonucleotide synthesis chemistry. Furthermore, aptamers are stable reagents that survive long-time storage and are easily shared within the research community either as DNA probes or as sequence information. When this paper was published about 100 aptamers had been raised for different proteins\textsuperscript{31}. To enable large scale production of aptamers, several groups have automated the process\textsuperscript{133-135}, but the main focus on aptamer production has shifted from generating protein binders to producing inhibitors for protein binding as drug candidates\textsuperscript{136}. To increase the generality of proximity ligation we suggested that it should be possible to use other classes of binders, such as antibodies. This was later demonstrated by Gullberg et al who used both polyclonal and monoclonal antibodies for cytokine detection\textsuperscript{137}. To generate a proximity probe pair, a polyclonal antibody batch was divided in two aliquots and coupled to either the 3’ or 5’ proximity probe arm using a biotin-Streptavidin linkage. In the same paper antibodies were also covalently coupled to oligonucleotides, thereby increasing the stability of the complex. This was performed using a heterobifunctional linker for coupling oligonucleotides to matched monoclonal antibodies raised against different epitopes of insulin. To detect homodimeric proteins as in paper I, a single binder is sufficient to generate a proximity pair, although the probability of ligation is reduced by half. Gustafsdottir et al, used the same strategy for viral detection with a LOD of only a few copies of porcine parvovirus (PPV). This low LOD was enabled by multiple bindings to the repeated capsid protein of the virus. To achieve the same LOD when only one copy of each epitope is present, the likelihood of binding between proximity probes and protein must be increased. Most straightforward would be to increase probe concentration, but as illustrated in figure 2b, paper I, this is not possible due to increased background. Schallmeiner et al (submitted), describes an elegant solution where three simultaneous binding events are required in order to elicit a signal, thereby drastically lowering background and LOD of less than hundred targets protein molecules or complexes can be achieved.

The low sample consumption of one microliter in the proximity ligation method is beneficial when sample is limited, for instance in biobank samples and studies of small experimental animal.

By enabling protein analyses using standard nucleic acid detection strategies, Proximity ligation takes a step towards unifying platforms of both protein and nucleic acid analysis. The coefficient of variation (CV) for PDGF-BB analysis achieved in Paper I was 35 %, very close to the 32 % for 88 PCR amplified replicates of diluted amplicons using the same instrument. To
resolve small changes in concentration, a more highly resolving analysis platform compared to qPCR will be needed. In paper III, we described a first step towards a detection platform that enables high precision analysis, unifying both DNA and protein detection. The use of nucleic acid tag sequences also opens the possibilities to microarray based readout. Here, the multiplexing capabilities are not limited by the microarray platform\textsuperscript{52} but rather by production of proximity reagents. Making relative measurements using proximity ligation is in theory trivial, simply by equipping otherwise identical probe pairs with different tag sequences each designed to hybridize to a specific fluorophore labeled detection oligonucleotide. One limiting factor for protein analysis is that traditional microarray platforms have a relatively high LOD. For nucleic acid analyses this is compensated for by sample amplification, but there is no traditional way of amplifying proteins. The “reverse translation” using proximity ligation enables this type of signal amplification.

Paper IV in this thesis presents a microarray based technology that can be used for decoding proximity probe reactions. In conclusion, Proximity ligation enables protein detection using small sample volumes and the homogeneous format enables easy sample handling using only serial additions of reagents enabling automation. Furthermore, by giving rise to a nucleic acid signal in response to the presence of a specific protein or protein complex many techniques for nucleic acid analysis can be used for efficient detection.
Paper II, Thermoplastic microfluidic platform for single-molecule detection, cell culture, and actuation.

This paper presents a scalable platform for microfluidic analyses and we used it for fluorescence detection, cell culture, and to construct peristaltic pumps and valves. This is made possible due to a novel production method for layered microfluidic structures. The process uses rapid injection molding of thermoplastics for generation of structured substrates similar to techniques used for production of music CDs (compact disc). This is followed by functionalizing the surface by applying a thin layer of silica. To form enclosed channels, an unstructured lid composed of PDMS coated to a thermoplastic film is covalently bonded to the silica surface using oxidization. The injection molding used in paper II was performed by Åmic AB, an Uppsala-based company specialized in CD replication technology.

A key element in the presented technique is the invention of using silica deposited onto thermoplastic materials for bonding purposes, and for this purpose we used electron beam evaporation.

By applying silica to the thermoplastic materials, we could use standard techniques for PDMS-glass bonding. To characterize silica deposition we used atomic force microscopy (AFM) and electron spectroscopy for chemical analysis (ESCA), to study surface topology and composition respectively. Silica deposition increases surface roughness slightly as shown in figure 2, paper II. Perhaps the most striking result of the AFM measurements is the extremely smooth surface of the injection molded surface prior to deposition. We have not seen any negative effects of the slightly rough silica surface in any experiments performed to date. To further characterize silica deposition we performed ESCA analysis, figure 3, paper II. As expected the surface composition changes from mainly carbon to a silicon rich surface after deposition. Electron beam evaporation can be used for deposition on nearly all thermoplastics, as long as the temperature in the chamber is maintained below the transition temperature "T_g" for the plastic to be coated.

To investigate the strength of the silica-plastic interface we bonded a thick slab of PDMS to the silica surface and performed a tensile test. The results indicated that rupture occurred in the PDMS bulk before detachment of the silica-plastic interface. A tensile strength of 30 N/cm² was recorded,
and in conclusion the silica-plastic interface is stronger than the PDMS polymer used. Tensile tests on PC (polycarbonate), PMMA (polymethyl methacrylate) and Zeonor (a cyclic olefin copolymer) all gave the same results regarding bonding strength. For further experiments we exclusively used Zeonor as the material of choice, due to the favorable optical properties.

The lid is typically produced using spin-coating of uncured PDMS on top of a Zeonor support film, although we have also used spray-coating. Figure 4, paper II illustrates the possibility to vary thickness of the lid in the range of 20 to 500 micrometers. This makes it easy to produce the optimal lid thickness for a given application. The Zeonor support film was left in place for fluorescence detection (figure 6, paper II) and removed for actuation and cell culture experiments (figure 7 and 8 paper II). In the actuation experiments a flow rate of about 9 nl/s was achieved in spite of less than optimal channel geometry created by the anisotropic etching technique used for master production.

Both the PDMS part of the lid and the silica surface of the thermoplastic chip are oxidized prior bonding. We used a Corona surface instrument for oxidization at atmospheric pressure. It has been proposed that silanol groups (Si-OH) is formed due to oxidation of methyl groups of the surface followed by dehydration. After activation, the lid is simply placed on top of the thermoplastic chip. Since no external pressure is applied low aspect ratio structures can be sealed without structure deformation.

Perspectives on paper II

In this paper a novel production method for microfluidic structures is presented. The method described reduces production time compared to traditional PDMS micro fabrication techniques, but offers the same versatility.

To enable large-scale production of microfluidic chips all steps in the production process need to be fast or possible to do in batch processes. Table 1, paper II compares our platform with traditional PDMS molding and the industry standard thermal annealing process. PDMS molding is probably the most widely used technique for microchip production in the research community. The technique is based on curing PDMS on top of preformed structures to form microchannels, thereby occupying the master for the entire curing time (typically 30-90 minutes). Our technique, which also relies on PDMS as a component for channel production differs in one important aspect for production time. We only use unstructured PDMS components, thereby enabling batch-wise production. In our case the master is only occupied for a few seconds during each injection molding cycle of the thermoplastic compact disc, greatly decreasing production time and costs when many identical structures are produced.
After the paper was published we have shifted from electron beam evaporation to sputtering for silica deposition. The transition to sputtering has enabled faster deposition times, mainly due to a more suitable instrument equipped with a load-lock chamber. We are also currently investigating possibilities to use sputters fully adopted for the CD format, enabling cycle times of a few seconds.

Integration of active elements in microstructures fabricated using multilayer PDMS technology has been pioneered by Stephen Quakes group. They have produced structures ranging from simple systems\textsuperscript{103} to perhaps the most complicated active chips created to date\textsuperscript{141,142}. In the industry standard thermal annealing technique, it is inherently difficult to integrate pumps and valves due to the rigid nature of the lid. The performance of the peristaltic pumps described in the paper could probably be improved if valve closure is complete. In newly designed channels we have used a combination of isotropic and anisotropic etching techniques, to generate channel geometries with an arched cross section in the valve area and rectangular geometries in the rest of the channels. The use of a thin flexible lid and rigid walls should allow higher packing densities of valves compared to traditional PDMS based chips, since no deformation of the fluid network occurs during actuation.

We have also started to work in collaboration with cell biologists to study chemotaxis of endothelial cell lines and primary cell cultures. Söderberg \textit{et al} developed an \textit{in situ} version of the proximity ligation technique for single-molecule analysis in cells and tissue-sections. We are currently applying microfluidics to perform such analysis faster and with lower reagent consumption (Johansson, Jarvius, M in progress). The possibility to use microstructures to generate precise gradients is also an interesting opportunity for optimization of reaction conditions that we are currently exploring.

\textbf{Is there a need for high-throughput production methods?}

To build functional microfluidic chips for bioassays, it is necessary to integrate functional elements such as pumps and valves as well as allowing optical inspection and providing bio-compatibility. The material properties of the PDMS technology and ease of integration of active elements has enabled production of a number of devices for biological analyses.

Fu \textit{et al} built a microfabricated fluorescence activated cell sorter (FACS)\textsuperscript{143} and they used it for sorting fluorescence labeled \textit{E. coli} from unlabeled \textit{E. coli}. Takayama \textit{et al} studied sub-cellular compartments using laminar flow of two mitochondrial specific dyes across individual cells\textsuperscript{144}, something extremely challenging using traditional techniques. By using controlled gradients Jeon \textit{et al}. studied chemotaxis of neutrophils using a PDMS fabricated device\textsuperscript{127}. Some of these devices can be reused, but they are preferably used once and then disposed to eliminate contaminations.
Despite the extensive research on PDMS based devices few PDMS-based microfabricated structures have been made commercially available. Although, PDMS remains the material of choice for rapid development of new microfluidic channel geometries. We believe, however that our thermoplastic-PDMS platform will enable high-throughput production of microfluidic devices, taking the advantage of the attractive properties of PDMS for actuation and gas permeability. This could help to introduce microfabricated structures on a larger scale in areas such as cell- or molecular-biology, rather than limiting the powerful technology to a small number of research groups who often do not have these areas as their main focus.
Paper III. Digital quantification using amplified single molecule detection

In paper III we describe a novel read-out technique for molecular analyses named amplified single molecule detection. The method use microfluidic structures described in paper II and a set of molecular tools for biomolecular analysis.

Figure 1, paper III, gives a schematic illustration of the method, a key element is to convert detected proteins or nucleic acid sequences to circular reporter molecules using highly specific molecular probing reactions. For nucleic acid detection we used padlock probes$^{10}$, and protein detection was performed using a modification of the proximity ligation technique described in paper I. Circular DNA-strands that form as a consequence of specific probing reactions are amplified by RCA. The $\delta 29$ polymerase we use for RCA was characterized by Margerita Salas group$^{11}$ and later used for various RCA applications by Lizardi et al.$^{22}$ and is in our opinion the most suitable enzyme due to its high processivity.

Following amplification, the RCA product was labeled by hybridizing fluorescent oligonucleotides, corresponding to a sequence tag originally encoded in the padlock or proximity probe backbone. After one hour of amplification the RCA product has a size of several tens of kb that forms an amorphous “blob” of DNA roughly one micrometer in diameter (figure 6, paper II, illustrated blob size as a function of amplification time).

RCA products flowing through the microfluidic channels were recorded using a confocal microscope (although standard epi-fluorescence microscopes can also be used) operated in line-scanning mode perpendicular to the microchannel (w x h, 200$\mu$m x 40 $\mu$m). Due to an approximately 100 times enrichment of fluorescence in each blob, there was no need to remove remaining un-hybridized fluorescence oligonucleotides prior detection and enumeration.

In figure 2, paper III several analysis parameters were characterized with regard to dynamic range and precision. Figure 2b, paper III illustrates a comparison of theoretically predicted and observed quantitative precision. The observed counts nicely follow the Poisson predicted values up to about 1000 counts, where statistical constrains are no longer limiting. Due to the high precision, small differences in concentration can be analyzed (figure 2c, paper III). The linear range in figure 2a, paper III spanned four orders of
magnitude, sufficient for many applications. We also demonstrated the possibility to shift the linear range by amplifying the number of blobs using circle-to-circle amplification “C2CA”\textsuperscript{16} to extend detection from a few molecules to seven orders of magnitude higher concentrations as shown in supplementary figure 1, paper III. We also used the method for detection of the globally significant pathogen \textit{Vibrio cholerae} as shown in figure 3, paper III. It was possible to detect concentrations much lower than the oral infectious dose and both linearity and precision was similar to the idealized situation shown in figure 2, paper III. In homogenous detection methods, inhibitors in the sample may drastically reduce efficiency. To evaluate performance in difficult samples, we performed the analysis in increasing concentration of whole blood and pig feces (supplementary figure 2, paper III). Assay performance was almost unaffected in 12.5% blood and in 0.05% pig feces. Simultaneous analysis of four different DNA circles was demonstrated in figure 3, paper III. To achieve higher multiplex capabilities than the number of available discernible colors we used dual labeling for detection of one class of blobs, thereby four classes of blobs could be detected using three fluorophores.

Perspectives on paper III

We have developed a novel analysis platform for single molecule enumeration with unified detection of both nucleic acids and proteins. The method preserves the discrete nature of biomolecules, enabling high precision digital analyses and making data interpretation straightforward. In paper I we used qPCR for detection and the CV of 35% allowed 2-fold difference in protein concentration to be distinguished. Using the amplified single molecule method we typically obtain a CV of 3 percent, drastically increasing precision. Currently data acquisition is relatively slow depending on the instrument used. This is reflected in a LOD in the femtomolar range for an analysis time of 30 seconds. We used C2CA to shift the detection window to a more suitable atto- or zeptomolar range depending on number of C2CA cycles used. An alternative would be to extend the microfluidic analysis time but this rapidly becomes unpractical due to the linear relationship between sampling time and LOD. We are currently exploring the possibility to construct a more suitable instrument with increased sampling rate based on line illumination and recording techniques. A faster system would also simplify the assay by avoiding the need to perform C2CA.

In this paper we performed a low level of multiplexing, but this level could be increased to at least ten analytes, comparable to numbers in use in bead-based immunoassays\textsuperscript{46}. Since we cannot use properties such as size, an
important aspect for multiplexed analysis using our technique is the quality of the fluorophore-modified oligonucleotides used for labeling. Many suppliers provide fluorophores spanning the visible spectrum in ever smaller increments. An example is the Alexa dyes from Molecular probes that presently include 19 distinguishable entities in the range from 400 to 800 nm. This in combination with microscopes equipped with spectral detectors should enable still higher multiplexing capabilities using our method.

A common problem with traditional fluorophores in multiplexed assays is the need for multiple wavelength sources for excitation. Development in the field of nanocrystals or quantum dots for use in molecular biology was pioneered by Moungi Bawendi and Paul Alivisatos in the 90’s and may provide an elegant solution to this problem. Quantum dots are nanometer-sized clusters of atoms (usually mixes of cadmium with selenium or tellurium), often subsequently coated with Zinc sulfide to further improve optical properties. Quantum dots behave basically like fluorophores, but with the difference that they can all be excited in the UV range and depending on their size will emit light of different wavelengths. It would be interesting to use quantum dots for multiplexed experiments in the future as they hold many interesting properties.

In conclusion the amplified single molecule detection method can be a read-out mode, unifying both nucleic acid and protein analysis in the same platform. Key features are the high precision and the large linear range. The method creates clusters of identical sequence-compartments without the need to use micro emulsion or solid phase techniques. The homogenous format of the assay, requiring only serial additions simplifies automation. Further, samples follow a “random access” pathway making the technique suitable for diagnostic assays. Some future projects are now in the starting phase (Göransson et al) where the high precision of the platform is used to detect gene copy numbers and to monitor antibiotic effects on sub population cultures of bacteria. We are also developing assays to detect potential agents of biowarfare in collaboration with the Swedish armed forces. The possibility to combine both protein and nucleic acid based recognition will increase selectivity. The proximity ligation has been shown to permit spore detection, which may prove simpler or valuable as a complement to nucleic acid techniques used today.
Paper IV, Enhanced molecular analyses by rolling-circle amplification on paired-tag arrays

This paper describes a novel technique for improved microarray-based analysis. As illustrated in figure 1, paper IV the technique is initiated by specific target recognition using padlock probes, followed by amplification of reacted probes using RCA. To enable selective on-array ligation, the concatenated RCA product is cleaved to present two specific 3’ and 5’ tag elements. Upon hybridization next to each other on an oligonucleotide immobilized in an array feature, they can be enzymatically joined forming a new circular product. The circle is then amplified on the solid phase, with the product covalently linked on the array. A general sequence encoded in-between the specific tag sequences is used for detection by hybridization of a fluorescence labeled oligonucleotide.

The technique can be viewed as a set of modules allowing detection of a wide range of biomolecules to be read-out on microarrays with excellent performance as illustrated in figure 4.

![Diagram of the paired-tag array technique](image)

*Figure 4. Illustration of the modules in the paired-tag array technique. From left to right, target recognition using padlock probes, homogenous rolling circle amplification (RCA) of reacted padlock probes, cleavage of the RCA product using MlyI to release an in silico designed reporter sequence, ligation of the reporter molecule to the microarray and last RCA amplification of circularized reporter molecule on the array surface. The two RCA reactions result in a total of 100 000-fold amplification.*

The problem of cleaving many RCA products having different pairs of tag sequences was solved using the type IIs restriction enzyme MlyI (reference 146 is a review of type IIs class restriction enzymes). First an oligonucleotide was hybridized to a general sequence motif present in all RCA products. Then MlyI was used to cleave at the junction of the general and specific
sequence (in the region exactly outside the double stranded sequence), thereby creating single stranded reporter molecules with distinct end sequences. The digested reporter molecules consist only of in silico designed sequences, for optimal performance without interference of target-specific sequence elements. The ability to freely choose sequences allows general purpose tag microarrays to be deployed\textsuperscript{57, 60}. The combination of RCA and ligase mediated detection, first in solution and then again on the array, provides for a low LOD and high selectivity. In figure 2a, paper IV, the level of amplification achieved either using solely on-array RCA or RCA first in solution and then on-array is compared to standard hybridization. RCA on the solid phase generates roughly 100 fold amplification compared to hybridization while two RCA reactions resulted in a total of 100 000-fold amplification. The amplification power of the method is illustrated in figure 2b, paper IV where a dilution series of cDNA was detected using our technique or qPCR. Both methods detected the same $\beta$-actin sequence and the LOD was for our methods about 30 copies of the transcript. The LOD corresponds to about 1.6 pg of total RNA, close to the 0.1-1 pg amounts of a single eukaryotic cell. This bypasses the need to use RNA amplification techniques often required in traditional expression analyses, as reviewed by Ginsberg\textsuperscript{147}. Figure 2b, paper IV, also shows that the dynamic range extends almost five orders of magnitude. This range is similar to that of the instrument used (16 bits of gray levels). The resolving power of the assay as illustrated in figure 2c, paper IV showed that about 1.5-fold copy number changes could be detected. We also demonstrated that the approach permits single molecule detection on the array as shown in figure 3b, paper IV. Both the analysis of small copy number differences and the single molecule analysis were performed in two colors, and were used to detect a cDNA sequence and a single nucleotide polymorphism respectively. We therefore used two distinct probes for each sample or genotype. In conclusion, the presented paired-tag technique can be used for cDNA expression analyses with a very low LOD and with read-out by measuring total fluorescence or digital recording of amplification products.

Perspectives on paper IV

In this paper we presented a novel microarray-based platform for expression analyses, using padlock probes for cDNA expression analyses. Padlock probes have previously been used for SNP detection in genomic DNA\textsuperscript{58, 59}, cDNA expression analysis\textsuperscript{148}, and also for in situ detection of mitochondrial DNA variants\textsuperscript{149}. While requirements for a wide dynamic range is limited in SNP analyses, it is very important for mRNA expression studies where number of transcripts can vary from a few copies to millions per cell. The
dynamic range in paper IV is currently limited by hardware, and can probably be increased with improved instruments. The presented methods converts any target recognized to a unique linear molecule equipped with two tag sequences at both ends. This allows for use of standard tag microarrays as previously described in a similar format\cite{57, 60}. The specificity of the assay is ensured by dual hybridization combined with enzymatic discrimination using ligases. Although it has been shown that hybridization can offer a very high degree of selectivity\cite{150-152} alone, we believe the extra selectivity will be necessarily when the molecules to be analyzed are presented at wide differences in concentrations. The use of enzymatic based discrimination adds additional selectivity and often enables unique functions not possible to achieve by mere hybridization.

The low LOD is ensured by amplification of the reacted products. We use RCA for amplification both in the homogenous and solid phase steps of the analysis, although PCR could probably be used in the first amplification. RCA as described earlier in this thesis generates an ideal substrate for subsequent hybridization steps. The reaction is not product inhibited a vital property for attaining a wide dynamic range in multiplexed analyses. Earlier reports have used RCA on the solid phase for both nucleic acid and protein detection\cite{82, 153-155}, but they have exclusively used preformed circles for amplification, an approach that is prone to generate nonspecific background since the exonucleolytic properties of the \( \phi 29 \) polymerase\cite{13} can remove mismatched nucleotides and initiate promiscuous amplification if the circle is partially hybridized to the surface probe. We found that excessive amplification on the solid phase can result in poor reproducibility, probably due to loss and delocalization of rolling circle products, so we limited the time for array RCA to 45 minutes. Figure 5 illustrates RCA performed on the solid phase for about 120 minutes; although beautiful it is clearly shown that the amplification products extend outside the spot, making analysis difficult.
Figure 5. Effect of excessive solid phase amplification. To the left microarray feature a fluorescence labeled complementary oligonucleotide has been hybridized and to the right feature a pre-formed circular target molecule. The circular target was subsequently amplified using RCA for about 120 minutes and the product labeled with a fluorescence oligonucleotide.

Although, the presented method can be used for improved cDNA expression analysis it may prove more important in more demanding analysis. Our future plans are to use the paired-tag strategy to measure protein levels and to decode protein interactions. By using a multitude of specific proximity reagents with attached oligonucleotides it should be possible to identify any protein interaction among a set of proteins. It will also be feasible to monitor disruption of these interactions using different drugs (Ericsson et al, ongoing projects). It will furthermore be interesting to use the paired-tag arrays to decode mRNA splice patterns using a variant of this technique.

The proximity ligation technique described in paper I and the modification to generate circular reporter molecules in paper III combined with the low LOD in paper IV, will enable large-scale protein detection using the same microarrays as for the other analyses described herein.
Paper V, DNA Skyline: fonts to facilitate visual inspection of nucleic acid sequences

This paper describes a set of TrueType fonts to enable visual inspection of nucleic acid sequences. The font design is similar to musical scores. Where the G nucleotides occupy the top level followed by As, Ts and lastly Cs. The extreme top and bottom levels is represented by Gs and Cs basepairing via three hydrogen bonds, and the middle two positions are occupied by As and Ts interacting via two hydrogen bonds. The design enables a precise fit when two complementary sequences are placed anti-parallel to each other. At the top-most level is an arrow indicating the 5’ to 3’ direction. To display degeneracy each character has a window placed in the right half of each character, and the level of blocks indicates possible bases. The letter “B” has blocks at all levels except the “A” level, corresponding to “any nucleotide but A”. Figure 6 illustrates the complete character set of the “GATC” font. By designing the characters as TrueType fonts truly platform independent operation is possible. Accordingly, the fontfiles can be used by all computer software that can handle TrueType fonts. The fonts can be of value for visual inspection of DNA sequences and for teaching purposes.

Figure 6. The complete character set of the skyline “GATC” font. Characters are illustrated using courier and gate notations.

Perspectives on paper V

In the early days of molecular genetics there were few tools available for nucleic acid sequence inspection. Lately, there has been a tremendous increase of available software suits to rapidly evaluate large amounts of se-
sequence data. This development has been crucially important to complete the sequence of the human genome. However, visual inspection of DNA sequences is nonetheless still frequently used in a laboratory such as ours to inspect sequence motifs and oligonucleotide designs, although dedicated software is used to construct large numbers of probes.\cite{156,157}

The history of TrueType

Since this paper differs in scope from the others presented in this thesis and it has been a fun project to pursue along from molecular research I will take the opportunity to give some background of the TrueType technology used in this paper and which inspired me to use a set of fonts to display the “DNA skyline” originally conceived by Ulf Landegren.

The history of developing scalable fonts, suitable for display in any sizes, started in the mid eighties.\cite{158,159} The intention was a collaboration project between Microsoft and Apple (although it turned out to be Apple who came up with the fonts), but a small number of people, one of them living close to Adobe creek, started it all! Not surprisingly this group was the founders of Adobe systems as we know it today, and they had an idea of producing a unifying language for all printing devices, which they called PostScript and it became an immediate success. This first generation of PostScript came in two versions Type 1 and Type 3, the Type 1 fonts were the better ones, but they also came with a price-tag. Of course Microsoft and Apple were interested in this technology, but the amount of royalty to be paid to Adobe encouraged an in-house development of a set of fonts with similar (or better) performance. Microsoft and apple agreed to cross-license rights of the technology, Microsoft should provide the raster engine (TrueImage) and Apple the fonts. It later became apparent that TrueImage was not needed, so all development of TrueType came from Apple. The person responsible for the developments was Sampo Kaasila and he worked with development between 1987 to 1989. The working name was not TrueType, but “Bass” (to emphasize that these fonts were scalable, as is a fish) it later became “Royal” and finally TrueType. The fonts (Times Roman, Helvetica and Courier) were released on march 1991 to run on Apple computers, while Microsoft user had to wait until 1992 to use Times New Roman, Arial and Courier on windows 3.1. However, the 32-bit architecture developed by Kaasila did not fully work on Windows computers until the release of Windows 95 where Microsoft had abandoned the early 16-bit engine for a 32-bit version. Today most platforms have TrueType engines and display TrueType fonts with ease.
Discussion

Further improvements of analysis methods in genomics and proteomics are needed to decipher the complex interactions that take place deep within the cell or organism. During the last couple of decades there has been a tremendous development of analysis techniques both for nucleic acid and protein-based analysis. Interdisciplinary research efforts have also established a number of “new” research fields that can contribute to enhanced molecular analysis. One such field is microfluidics, but also work in the semiconductor field has been fruitful for the generation of new biological analysis. Interesting effects are observed when traditional assays are miniaturized, often resulting in improved performance. One example is protein crystallization, where it can be difficult to generate high quality crystals using traditional techniques\textsuperscript{160}, and also to crystallize multiprotein complexes available only in small amounts. Hansen et al.\textsuperscript{161} presented a microfluidic protein crystallization device that holds potential for solving some of these problems. Recent work in sensor technology has provided a multitude of methods for protein detection using devices like nanowires\textsuperscript{162}, micro cantilevers\textsuperscript{163}, carbon nanotubes\textsuperscript{164} and different types of nanoparticles used either \textit{in vitro}\textsuperscript{43, 165} or \textit{in vivo}\textsuperscript{166}. The future will tell if label-free methods will achieve the required selectivity to correctly identify proteins when present in complex biological samples.

Most molecular analyses use averaging methods while everything in biology depends on discrete events. Of course, a single event frequently does not suffice for a cellular response, but probably a number of events are required. If multiple proteins are monitored simultaneously using ensemble averaging techniques, small number of events could be overshadowed by other processes. Often these “discrete events” are represented by proteins interacting with other proteins, ions, peptides or nucleic acid sequences. Therefore, better methods to study individual molecules will be important for gaining a deeper understanding of cells, and therefore also of life.

Recently, a number of amplified single molecule detection approaches have been presented, thereby not limit detection to single fluorophores (or other reporters) but rather detect thousand molecular clones arising as a result of a single recognition event making data interpretation easier. Most methods require dilution of molecules to be detected prior amplification followed by compartmentalization (page 22). The combination of technologies from paper I and II in this thesis was further developed to our amplified
SMD approach presented in paper III. The technique creates its own compartment and utilizes the selectivity from padlock and proximity ligation reaction to generate a signal. It will be interesting to further develop the method for multiplexed protein detection, and possibly also to simultaneously assay protein and nucleic acid sequences.
Acknowledgements

The presented work has been performed at the department of Genetics and Pathology in the Rudbeck laboratory. And it would not have happened without my colleagues and supervisors. In the last part of this thesis I would like to express my gratitude to the people that made this possible:

Jag vill tacka följande personer:

Min handledare Ulf Landegren, det har varit otroligt roligt att få jobba tillsammans med dig. Din enorma inspirationsförmåga och dina djupa kunskaper har lärt mig massor under de här åren. Jag är också mycket tacksam för att du låtit mig jobba på projekt som du kanske inte trodde så mycket på initialt men ändå stöttat mig. Din förmåga att samla bra personer till labbet har gett en mycket inspirerande miljö att arbeta i. Tack också för dina boktips som breddat min horisont och slutligen hoppas jag att få fortsätta jobba tillsammans med dig under några år och kanske t.o.m. få uppleva din övergång till den mest framstående av datorplattformar...

Min handledare Mats Nilsson, din återkomst till labbet förändrade min avhandling radikalt. Dina kunskaper både teoretiskt men även praktiskt har hjälpt mig att bli en betydligt bättre forskare och dina djupa mekanistiska insikter har lärt mig tänka i nya banor. Framför allt vill jag tacka dig för ditt intresse av små nystan av DNA, som man kan använda på flera nyttiga sätt. Jag hoppas att vårt framtida samarbete kommer leda till nya upptäckter och att jag förhoppningsvis kan ge lite tillbaka av den inspiration som du gett mig under dessa år. Slutligen vill jag tacka för din insikt om dead-lines och hur dessa kan behandlas...


52

hittar alla coola e-handelsplatser för elektronik, vi borde dessutom jobba tillsammans i större utsträckning, det vore roligt. Till sist vill jag tacka Lena som ser till att labbet fungerar, ibland undrar jag hur det var möjligt före din tid (jag vill också tacka för att dina barn är äldre än Emil, det är ju så praktiskt).

Anders A och Carolina R, för att ni organiserar alla komplexa projekt och hjälper till att komma ihåg dead-lines, utan er skulle jag zwoschat förbi i stort sett alla dylika.


Alla medförfattare till publikationerna beskrivna i denna avhandling.

Alla på Olink, hoppas att vi kommer att ha mycket kul tillsammans i framtiden.

Alla mina studenter under åren, ingen nämnd och ingen glömd. Utan er skulle det ha gått så otroligt mycket långsammare. Ni har även hjälpt mig att utvecklas som handledare och jag hoppas på att bli ännu bättre i framtiden.

Avslutningsvis vill jag påminna dig om att jag ännu inte lyckats dela denna avhandling.

All personal på Åmic. Speciellt vill jag tacka Ove för våra många samtal om microfluidic och för att du alltid ställer upp oavsett om det gäller arbete eller finsnickeri, du har verkligen inspirerat mig! Lasse, Dudas och Hasse tack för all hjälp med replikering, finmekanik och att hitta gamla kassettbandsmaskiner på nätet. Simon, för hjälp i de tidiga dagarna med fräsnings och för att du delat med dig av dina kunskaper i 3D-studio och Maya. Toni och David, för all hjälp med kiselbearbetning och övrigt renrumsjobb.

Ib, för dina kunskaper om yt- och -kemi som du gärna delar med mig av. Samt alla övriga, ni har alla hjält ihåg mig oerhört mycket och bidragit till denna avhandling. Vad man än frågar om, så är den någon som kan det och gärna hjälper till.

Staffan Dryselius, för alla timmar med att försöka få en sil på celler och för allt roligt vi har haft tillsammans med patent, vetenskap, riskkapitalister och alla övriga tekniska diskussioner. Lycka till med att göra vården omänskligare.

Mårten F, för att du delat denna ängest med mig och för allt roligt som vi har haft tillsammans. Såga vad man vill men vi kan detta med dead-lines…
Tack alla vänner som jag haft glädjen att lära känna, aningen genom motorcyklar, studier eller i andra sammanhang. Det är skönt att kunna prata om andra saker än blobbar och Reynolds tal, det finns så mycket annat i livet.

Jag vill också speciellt tacka…………………………………………………………
Fyll i denna rad om du inte blivit tackad, så lovar jag att kontrasignera för äkthetens skull (och i så fall...förlåt)

Till min ingifta familj, Ann Catrin, Björn, Marcus och Malin. Tack för allt kul vi har haft hittills och för glädjen att lära känna er, men kanske speciellt tack för er underbara dotter och syster.

Till mamma Monika, pappa Sven-erik och syster Åsa, för att ni stöttat mig genom livets krokiga väg och för er förmåga att tro på mig och mina projekt va de månde vara. Utan er skulle inte denna avhandling blivit till, ja faktiskt inte jag heller när man tänker på saken.

Sist av allt vill jag tacka min underbara fru (och kollega) Malin för allt. Om du inte hade flyttat upp till mig från Skåne, hade jag nog doktorerat i Lund nu och denna avhandling inte existerat. Det ska bli underbart att följa vår son Emilys väg genom livet med dig och kanske tillsammans granska hans avhandling på ålderns höst.
References

159. Penney, L.
A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-7079