UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 245

Development of DNA-based methods for analysis of protein interactions

AXEL KLAESSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018

ISSN 1651-6192
ISBN 978-91-513-0222-5
urn:nbn:se:uu:diva-340078

Dissertation presented at Uppsala University to be publicly examined in A1:107a, BMC, Husargatan 3, Uppsala, Thursday, 15 March 2018 at 13:00 for the degree of Doctor of Philosophy (Faculty of Pharmacy). The examination will be conducted in English. Faculty examiner: Cristina Al-Khalili Szigyarto (Kungliga Tekniska högskolan, Department of Protein Science).

Abstract

Klaesson, A. 2018. Development of DNA-based methods for analysis of protein interactions. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy* 245. 47 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0222-5.

In situ proximity ligation assay (PLA) is a method for detection of protein interactions, post-translational modifications (PTMs) and individual proteins that allows information about their localization in a cell or tissue to be extracted. The method is based on oligonucleotide-conjugated antibodies (proximity probes) that upon binding of two epitopes in close proximity give rise to an amplifiable DNA circle. Rolling circle amplification (RCA) is used to create a DNA bundle of over a thousand repeats to which fluorescently labeled detection oligonucleotides are hybridized. This thesis is focused on improving the existing *in situ* PLA method and on developing new approaches for detection of proteins, protein-protein interactions and PTMs *in situ* in cells and tissues.

In paper I, a new enzyme-independent method capable of *in situ* detection of protein-protein interactions was developed. The method combined the proximity requirement of *in situ* PLA and the amplification of hybridization chain reaction (HCR) creating a proximity-dependent initiation of hybridization chain reaction (proxHCR). Circumventing the need for enzymes resulted in a cost-efficient method that is less sensitive to storing conditions.

Paper II addresses the problem of irregularly formed RCA products that can appear to be split into several fluorescent objects. A compaction oligonucleotide system was designed to crosslink the DNA bundle with itself and thereby reduce the size and increase the brightness of each individual RCA product.

In paper III, the conventional *in situ* PLA was redesigned to increase the detection efficiency of protein interactions and PTMs *in situ*. The new set of proximity probes was designed to have circularization oligonucleotides incorporated that were unfolded through enzymatic digestion. The UnFold *in situ* PLA was able to generate more signals and had a higher sensitivity than the conventional *in situ* PLA.

In paper IV, an oligonucleotide system able to generate signals for individual proteins (A or B) and their interaction (A and B) in a molecular Boolean (MolBoolean) protein analysis was designed. The MolBoolean design was able to generate signals detecting both individual proteins and their interaction *in situ.*

Keywords: In situ proximity ligation assay (PLA), rolling circle amplification (RCA), Hybridization chain reaction (HCR), proxHCR, Oligonucleotide design, Protein-protein interactions, Post-translational modifications.

Axel Klaesson, Department of Pharmaceutical Biosciences, Box 591, Uppsala University, SE-75124 Uppsala, Sweden.

ISSN 1651-6192
ISBN 978-91-513-0222-5
urn:nbn:se:uu:diva-340078 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-340078)

To my family and friends

List of Papers

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

I Koos B, Cane G, Grannas K, Löf L, Arngården L, Heldin J, Clausson CM, **Klaesson A**, Hirvonen MK, Souza de Oliviera FM, Talibov V, Pham N, Auer M, Danielson H, Haybaeck J, Kamali-Moghaddam M, Söderberg O. (2015) Proximity-dependent initiation of hybridization chain reaction. *Nature Communications,* 6, 7294, doi:10.1038/ncomms8294

II Clausson CM, Arngården L, Ishaq O, **Klaesson A**, Kühnemund M, Grannas K, Koos B, Qian X, Ranefall P, Krzywkowski T, Brismar H, Nilsson M, Wählby C, Söderberg O. (2015) Compaction of rolling circle amplification products increases signal integrity and signal-to-noise ratio. *Scientific Reports,* Jul 23;5:12317. doi: 10.1038/srep12317

III **Klaesson A**, Grannas K, Ebai T, Heldin J, Koos B, Leino M, Raykova D, Oelrich J, Arngården L, Söderberg O, Landegren U. Improved efficiency of *in situ* protein analysis by proximity ligation using UnFold probes. *Submitted manuscript*

IV **Klaesson A**, Raykova D, Söderberg O. A Boolean analysis of protein interactions at a molecular level in single cells. *Manuscript*

Related work by the author

Original article

Perland E, Bagchi S, **Klaesson A**, Fredriksson R. (2017) Characteristics of 29 novel atypical solute carriers of major facilitator superfamily type: evolutionary conservation, predicted structure and neuronal co-expression. Open biology, Sep;7(9). pii: 170142. doi: 10.1098/rsob.170142

Book chapters

Koos B, Andersson L, Clausson CM, Grannas K, **Klaesson A**, Cane G, Söderberg O. (2014) Analysis of protein interactions *in situ* by proximity ligation assays. Current topics in microbiology and immunology, 377:111-26. doi: 10.1007/82_2013_334.

Raykova D, Arngården L, **Klaesson A**, Heldin J, Söderberg O. (2017) Selecting the appropriate *in situ* proximity ligation assay protocol. Fluorescence Imaging and Biological Quantification, chapter 8 DOI10.1201/9781315121017-10

Contents

Abbreviations

A	Adenine
AP site	Apurinic/apyrimidinic site or abasic site
BER	Base excision repair
CFP	Cyan fluorescent protein
C	Cytosine
DNA	Deoxyribonucleic acid
DO	Detection oligonucleotide
dsDNA	Double-stranded DNA
EndoIV	Endonuclease IV
ELISA	Enzyme-linked immunosorbent assay
E.coli	*Escherichia coli*
ExoIII	Exonuclease III
FRET	Förster resonance energy transfer
GFP	Green fluorescent protein
G	Guanine
H1	Hairpin 1
H2	Hairpin 2
H&E	Hematoxylin and eosin
HRP	Horseradish peroxidase
HCR	Hybridization chain reaction
IF	Immunofluorescence
Ig	Immunoglobulin
IHC	Immunohistochemistry
ISH	*In situ* hybridization
T_m	Melting temperature
mRNA	messenger ribonucleic acid
dNTP	Nucleoside triphosphate
NA	Numerical aperture
PCR	polymerase chain reaction
PTM	Post-translational modification
PH1	Proximity hairpin 1

PH2	Proximity hairpin 2
PLA	Proximity ligation assay
qPCR	quantitative PCR
RNA	ribonucleic acid
RCA	Rolling circle amplification
SIM	Structured illumination microscopy
T	Thymine
U	Uracil
UNG	Uracil-DNA glycosylase
YFP	Yellow fluorescent protein

Introduction

All cells have a common ancestry and therefore, have similarities in their biomolecules, whether the cell is living as a unicellular organism at the edge of a volcano on the bottom of the ocean floor or if it is a part of a complex organ in a large multicellular organism. Cells have a plasma membrane with receptors and other surface proteins, often with the same evolutionary origin, that allows them to interact with their environment, which provides signals to alter cellular functions accordingly.[1] In multicellular eukaryotic cells, the signal can inform the cells to take a variety of different actions *e.g.* replication, differentiation, secretion, regulation of metabolism and apoptosis and a substantial portion of the signals are emanating from neighboring cells. However, molecules traveling around the organism through its circulatory system also have a prominent effect on the cell. When a receptor is activated by a ligand, it transfers the received information from the outside of the cell into the intracellular space. Signal transduction over the plasma membrane usually involves ligand-induced conformational changes and dimerization of the receptor, which often result in post-translational modifications (PTMs) on the intracellular receptor part. This starts a signaling cascade of protein interactions and protein modifications, like phosphorylations, which is amplifying the original signal. The signaling network of activated proteins often results in regulation of gene expression by binding of transcription factors to promotor regions of the genomic deoxyribonucleic acid (DNA). Genes are transcribed to messenger ribonucleic acid (mRNA) via pre-mRNA and thereafter translated into proteins by the ribosomes. Proteins perform the vast majority of cellular functions *i.e.* structural support, transport, a huge array of different enzymatic functions and the aforementioned signaling.[2]

These inter- and intracellular communications are strictly regulated in order for cells and organisms to function normally and to keep homeostasis. Errors in the cellular communication that sidestep the regulatory functions are the cause of most non-infectious diseases *e.g.* cancer and diabetes.[3,4] To identify these diseased cells and tissues and to understand how they differ from their healthy counterparts, methods for visualizing the proteins important for the cells' signaling networks are required. Many of these methods have drawn inspiration and borrowed biomolecular tools from nature's own processes. For example, many different methods use antibodies to bind a protein of interest, either to visualize it or to capture it.[5]

Apart from the existence of a protein, also information on interaction partners and PTMs, *e.g.* phosphorylations, are important for comparisons of the status of the cell.[6-8] Furthermore, determining where proteins are expressed and whether they are interacting with each other, provides information on functional status and allows us to differentiate between diseased and healthy cells within a specimen.[9,10]

During my PhD work, I have been focused on improving existing methods and developing new approaches for detection of proteins, protein-protein interactions and PTMs *in situ* in cells and tissues.

Visualization of the cell

For visualizing proteins in a cellular context, magnification of the cell is required. The first known microscope used for visualizing biology was developed in the 17th century[11] and was a compound microscope. In the beginning, it was mainly larger histological structures that were explored. However, as the microscope technology improved and different chemical dyes for staining were developed, further histological details were gained. These different staining protocols are still used today and are a vital tool for diagnosis by pathologists all over the world. Originally, specimens were mainly stained with different chromophores that are visible under a normal light microscope. One example of this is the widely used hematoxylin and eosin (H&E) staining. Hematoxylin is a base that mainly stains negatively charged structures and molecules, like nucleic acid, with a dark blue stain. Eosin is an acid that has a preference to stain positively charged molecules, like some proteins. Due to the high protein content, it stains cytoplasm and extracellular matrix pink/red. The H&E stains complement each other and give contrast thereby enhancing morphological structures.[12] An alternative to using the, by eye visible, chromophores is to stain the tissue using fluorophores instead and visualizing it with an epi-fluorescence microscope.[13] Fluorophores are molecules that are excited by light in a certain spectrum (*i.e* a range of wavelengths) and thereafter emit light in another spectrum. Both normal light and epi-fluorescence microscopes are based on the same type of compound microscope used in the 17th century with an eyepiece and objective lens setup. Most microscopes have the ability to change between objectives with different magnification depending on how detailed a certain structure should be studied. The resolution of a microscope is the distance between two objects that still can be resolved. The magnification and other characteristics of the objective define its numerical aperture (NA) and affect the resolution. With the NA, the maximum resolution (diffraction limit) for any given wavelength (λ) can be calculated as follows:

$$Diffraction Limit = \frac{\lambda}{2NA}$$

This, by optical physics limited maximum resolution, will be around 200-300 nm for a normal light or epi-fluorescence microscope, depending on the objective and wavelength imaged.[14] One way to achieve a higher resolution is using electrons instead of photons to radiate the sample. Electrons have a shorter wavelength, which allows for a resolution below 0.1 nm and magnifications above a million times can be achieved. Electron microscopes are, however, technically complex, resulting in high maintenance cost and the need for extensive user training Moreover, they require the sample to be labeled with an electron dense molecule like gold.[15] There are also optical microscopes with a higher resolution that goes beyond the diffraction limit and these are collectively called super-resolution microscopes. These microscopes have optical resolutions ranging from 10-100 nm.[16] One example is structured illumination microscopy (SIM) used in **paper II**. SIM overcomes the optical limitations by acquiring several images under different light patterns creating a Moiré effect. The Moiré pattern of the different images are thereafter reconstructed into one high-resolution image (100 nm resolution).[17] The resolution required determines what type of microscope is necessary and most of the time a normal light or epi-fluorescence microscope is sufficient. The size of cells and histological structures are all on the µm scale and are resolvable without going below the diffraction limit of light. However, some subcellular structures are too small to be resolved, like individual ribosomes on the endoplasmic reticulum. Imaging the cell in an electron microscope allows seeing the folding of the endoplasmic reticulum and the about 20 nm large ribosomes become visible.[15]

Apart from staining different cellular structures *e.g.* the nucleus, it is also common to stain proteins of interest. When staining a cell with different fluorophores, the microscope should only visualize one type of fluorophore at a time. Epi-fluorescence microscopes usually have a wide spectra lamp able to excite all types of fluorophores. A filter cube is placed between the objective and the eyepiece to prevent excitation and visualization of the wrong fluorophores. The light from the lamp is directed into the cubes excitation filter where only wavelengths optimal for the selected fluorophore´s excitation can pass through. Thereafter, the light is reflected towards the objective by a dichromatic mirror and out to the specimen where the fluorophores are excited. The excited fluorophores will then emit light of lower energy and higher wavelength. The emitted light collected by the objective can pass through the dichromatic mirror and the emission filter of the filter cube if it has the right wavelength. The emission filter reduces the amount of autofluorescence able to pass through the filter cube.[14] The ability to dis-

criminate between different fluorophores provides the possibility to simultaneously stain the same specimen with multiple fluorophores. Standard microscopes usually have a few filter sets that can be used simultaneously without extensive overlap of the spectra or any significant bleed-through. Staining different structures and proteins in a specimen each with a different fluorophore enables more information to be extracted, which can be relevant when working with limited patient and animal samples, as well as for single cell analysis. Therefore, it is of interest to be able to maximize the number of colors that can be visualized. At least seven different colors can be added in an image by correcting for bleed-through between the fluorophores computationally[18] or six colors can be created by marking each epitope with two or more fluorophores reading the overlap as an individual color.[19] Both of these methods have their advantages and disadvantages but both are technically/mathematically challenging solutions to the problem. There are also microscopes with condensed filter sets that allow up to six different fluorophores to be visualized with minimal bleed-through.[20] The actual images of the samples are collected by a camera under the microscope and for light microscopes, this is a color camera. However, for epi-fluorescence microscopes, it is common to have a black and white camera resulting in a grayscale image per channel. The images are thereafter given an artificial color by a software and can be exported for image analysis.

Generally, stains are used to highlight morphological structures, mark proteins and PTMs of interest and assess their subcellular localization. This information is of high importance for research and medical purposes, which becomes apparent by the fact that pathologists diagnose and classify cancer from tissue biopsies, usually with the help of stains. Even though tissue staining is routinely used in the clinical setting and is valuable for stratifying patients, the subjectivity of evaluating the stains is a known problem.[21] To help find the correct diagnosis, additional stains/biomarkers can be developed but also finding more standardized ways to interpret the stained specimens can aid in this process. Imaging the samples under a microscope and doing an automated analysis of the images would help pathologists in their interpretation and allow to standardize the procedure. In the long run, this would likely increase both patient safety and consistency in treatment.

Many different software for image analysis exist[22] and one of them is CellProfiler.[23] CellProfiler is a free, open-source software designed to analyze individual cells in images with a high throughput. The software is modular based in its analysis, simplifying the construction of an image analysis pipeline. Applying the same pipeline to all images ensures that information is collected equally from the images and removes the bias of manual evaluation by eye. With the appropriate staining of the cell and nucleus, an analysis pipeline can identify and measure each individual cell in the image. Infor-

mation on the size and shape of the cells can be extracted including measurements of protein amounts and their subcellular localization. Compared to an automated analysis, in a manual analysis, only a subset of the image or tissue features are scored and some differences are too subtle to be noticed by eye. The heterogeneity of the sample is retained by analyzing on a cellular level and possible subpopulations of cells become easier to detect. After a pipeline for analyzing the images has been set up in CellProfiler, it can reduce the time required to analyze a huge data set of images compared to using less high-throughput image analysis strategies. However, image analysis is not always straightforward and it does take time and adjustments to build a pipeline able to extract the necessary information required per experimental setup. Nevertheless, it can give objective data from samples for researchers and possibly together with new assays and biomarkers help pathologists as a tool to complement their own ratings of tissues.

Methods for *in situ* protein detection

Many of the methods and treatments within medical science are based upon already existing compounds found in nature. Nature has, over millions of years, solved problems through natural selection. For example, a drug that has revolutionized medicine, penicillin, was found in fungi, where the molecule protects *Penicillium chrysogenum* from bacteria in its surrounding.[24] By extracting the molecule and improving it, we now have many different antibiotic derivatives deploying that same mechanism first used by the fungus.[25] Apart from the development of nature-based drugs,[26] different biological processes and molecules have also been utilized as an extensive toolbox for the development of molecular tools for research and diagnosis.

To be able to stain a protein of interest in or on a cell, a probe needs to meet two requirements: firstly, that it is targeting a specific structure (epitope) and, secondly, that it is resulting in a detectable signal. These properties are often achieved by borrowing molecules from nature directly or by modifying them for better effect. For example, phalloidin is a toxin from a poisonous mushroom with affinity to bind F-actin, which can be used to visualize actin filaments after it has been equipped with a fluorophore.[27] However, the most commonly used molecules to target proteins are antibodies.

Antibodies or immunoglobulins (Ig) are a part of the adaptive immune system of all jawed vertebrates (*gnathostomes*)[28] and have evolved to be highly variable to bind epitopes of any pathogen the host may come across. Antibodies are produced by B-cells as a response to coming in contact with foreign epitopes. There are five main antibody isotypes each with its own func-

tion, but in research, mainly the IgG isotype is used for detecting proteins. Immunoglobulin G has a Y shaped structure that consists of four polypeptide chains joined by disulfide bonds. Together the molecule is about 10 nm in length and has a weight of 150 kDa. All polypeptide chains have a constant and a variable region. The variable epitope binding regions are created in the B-cells through gene recombination and somatic hypermutation, followed by affinity selection.[29] Antibodies produced for research are either polyclonal or monoclonal, meaning that the IgG is derived either from several different matured B-cell clones or from a single antibody-producing B-cell clone. As a result, polyclonal antibodies bind several different epitopes while a monoclonal antibody recognizes a single epitope. Also, the production of the two antibody types differ and polyclonal antibodies are produced by immunizing the protein of interest or a part of it into a host animal (*e.g.* rabbit, mouse or donkey). Thereafter, the antibodies are extracted from host blood serum, meaning there will be a batch variation of polyclonal antibody quality.[30] Monoclonal antibodies are produced by one hybridoma clone,[31] usually of mouse origin, and have no variation in binding properties between batches.[32,33]

Since the discovery of antibodies in the 1890s,[34] it took about 50 years before Coons *et al.* published a paper describing conjugation of a fluorophore to an anti-pneumococcus III antibody. They found that the fluorescently labeled antibody would specifically mark the type III pneumococci and render them fluorescent under UV light.[35,36] The method described by Coons *et al.* for staining tissues is now called immunohistochemistry (IHC) or immunocytochemistry if performed on *e.g.* cultured cells. Subsequently, antibodies have been a cornerstone for *in situ* detection with their ability to specifically bind most proteins and the antibody staining technique has evolved into many different methods beyond *in situ* detection. Immunohistochemistry with fluorescently labeled antibodies is also commonly called immunofluorescence (IF), which can be done directly with only a primary fluorophore-labeled antibody or indirectly with an unlabeled primary and secondary fluorophore-labeled antibody.[37] Indirect IF requires that first, a primary antibody binds the target protein, followed by a secondary antibody binding specifically to the constant region of the primary antibody. Several secondary antibodies can bind every primary antibody causing an amplification of signal. The constant region of an antibody isotype remains the same when it originates from the same host leading to reusability of secondary antibodies for a variety of different primary antibodies. However, this also causes problems visualizing several proteins simultaneously if the primary antibodies come from the same species. For direct IF the primary antibody is labeled with a fluorophore and is not limited in regards to multiplex-ability by host origin.[37] To increase the signal amplification of both direct and indirect staining, instead of using a fluorophore, the antibody can be labeled with an

enzyme.[38-40] One enzyme that has proven to function exceptionally well for this purpose is horseradish peroxidase (HRP), which, as the name entails, is found in the roots of the horseradish.[41,42] In the presence of hydrogen peroxide (H_2O_2), HRP is able to oxidize and thereby activate a plethora of different substrates. The activated substrate results in a local build-up of either a colored, fluorescent or chemiluminescent product, depending on which substrate was added.[38,39,42] This enzymatic approach is also the basis for the enzyme-linked immunosorbent assay (ELISA), which is a solid-phase method for measurement of a substance in solution.[43,44] The IHC method is today extensively used both within research and for diagnosing patients, due to the large repertoire of commercially available antibodies. The flexibility and variety of probes available enable the user to customize the method according to the lab equipment available and assay requirements. However, the method still has disadvantages including that the probes always have some off-target binding and, therefore, each assay setup requires validation.[45] The cells and tissues stained are usually fixated and if the target is below the plasma membrane, permeabilization is required. This prevents intra-cellular antibody staining of living cells. To visualize protein interactions the resolution required is higher than that of many commonly used microscopes making it difficult to draw conclusions from double-stained cell structures. A cell stained for two highly expressed proteins in the cytoplasm will look similar in the microscope independent of whether the proteins interact with each other.[46]

To visualize intra-cellular proteins in living cells there are other options than to label them with an antibody. With the help of genetic manipulations, we are able to create cells expressing the protein we are interested in linked to *e.g.* a green fluorescent protein (GFP). Hence, no external probe is required to visualize the protein, because the protein becomes fluorescent by itself. After discovering the GFP in the jellyfish *Aequorea Victoria*,[47] it has been developed into an excellent tool for monitoring living cells and the work was awarded the Noble Prize in chemistry 2008.[48,49] Two variants of GFP were created, both with a shifted excitation and emission spectra. The new variants, cyan and yellow-emitting florescent protein (CFP and YFP), allow Förster resonance energy transfer (FRET) between them and they are therefore useful for measuring protein-protein interactions.[50] FRET occurs between two fluorophores that have the same dipole orientation and where the emission spectra of the donor fluorophore overlap with the excitation spectra of the acceptor fluorophore. It is the distance dependency of FRET that makes it suitable for detection of protein-protein interactions, as only proteins within 10 nm of each other can generate a signal.[51,52]

In contrast to detecting proteins with antibodies, using fluorescent fusion proteins is more cumbersome but has the advantage of being able to track

proteins in living cells and to detect if two proteins interact. However, modifying a protein may alter its function. In systems where fusion proteins are overexpressed, it can cause a change in the localization and interaction profile of the proteins. In addition, it is not possible to monitor interactions in tissue sections, as the method requires transfection of fusion constructs. The comparative ease of use of antibodies coupled to a detectable molecule for staining, has established IHC and IF as widespread standards for *in situ* detection of fixated specimens.

Structure and properties of DNA

The genomic DNA is the main blueprint of any living organism and its double-stranded molecular structure makes it a stable carrier of information. In a single eukaryotic cell, the information to create the whole organism is stored. The information is encoded using the four different nucleotides that each consists of a sugar-phosphate and one of the four nucleotide bases, which are cytosine (C), guanine (G), adenine (A) or thymine (T). The sugar-phosphate is a deoxyribose with a phosphate group attached to its 5' carbon and, as the name implies, it has a hydroxyl group on the 3' carbon, but lacks one on the 2' carbon. The nucleotides are linked together into a long polynucleotide chain by phosphodiesters connecting the 5' carbon of the deoxyribose sugar together with the 3' carbon of the next. Already before we knew the exact structure of DNA, it was discovered that extracted DNA from different organisms always contained the same amount of the bases G and C as well as A and T. Therefore, it was predicted that these nucleotides were binding each other and forming pairs.[53] This information together with the X-ray crystallography performed by Franklin and Wilkins allowed Watson and Crick[54] to correctly predict the double helix structure of DNA. In the double helix DNA, two polynucleotide chains are hybridized to each other by hydrogen bonds that can form between the base pairs; with three and two hydrogen bonds between G≡C and A=T, respectively. The strands are antiparallel which means that one strand has a reverse complementary nucleotide sequence to the other and that the 5' and 3' end of the two strands pair up (**figure 1**).[54] The distance between each base pair in the helix is around 3.4 Å and they have an angle of 36° between them. Therefore, the structure bends around its own axis every 10 base pairs at a distance of 34 Å. The innate structure of genomic DNA allows for high fidelity replication of the information carried in the nucleotide sequence due to a precise polymerase and various proofreading enzymes.

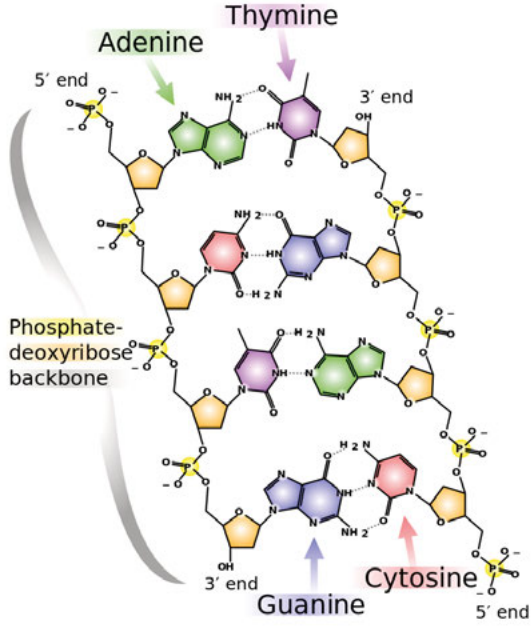

Figure 1. The structure of DNA (Illustration by Madeleine Price Ball)

It has been a common perception that the hydrogen bonds between the bases result in increased stability of double-stranded DNA (dsDNA). However, the stability of dsDNA is an effect of base stacking with π-π interactions between the adjacent base pairs.[55,56] It was shown by Yakovchuk *et al.* in 2006 that the A=T base pairing by itself reduced the stability of the dsDNA while the G≡C base pairing contributed little to no stabilization.[56] Nevertheless, any mismatch in the sequence affects the strength of the hybridization and lowers the average melting temperature (T_m) of an oligonucleotide of 10 to 35 nucleotides by 20°C to 2°C, respectively.[57] The T_m of an oligonucleotide is defined as the temperature at which 50% is hybridized in its ds conformation and 50% is denatured. Even though the base pairing does not stabilize the DNA structure, a mismatch in the sequence is causing structural alterations and therefore destabilizes binding to a variable degree. The severity of the effect depends on the mismatch varying from mild with C-C pairing to severe with G-G pairing.[58] When oligonucleotides hybridize, the binding of the first couple of base pairs is thermodynamically costly (Gibbs free energy, $\Delta G > 0$), before the base stacking is able to affect the stabilization.[59] For predicting the T_m of an oligonucleotide sequence there are three main aspects which have an effect: First, the concentration of cations *i.e.* $[Na^+]$ and $[Mg^{2+}]$ affects the binding strength of the oligonucleotide and a higher salt concentration leads to a higher T_m. Second, the T_m is increased when there is a high concentration of oligonucleotides. Finally, the design of the

oligonucleotide affects the T_m, where increased GC content and length stabilizes the dsDNA resulting in a higher T_m.

The characteristics of DNA together with our ability to synthetically manufacture oligonucleotides have allowed us to develop the DNA nanotechnology field.[60] The predictability of DNA base-pairing and behavior has made it a perfect building block for DNA origami,[61] which creates scaffolds and potential drug delivery systems.[60,61] Moreover, the properties of DNA have also been found suitable for developing dynamic DNA nanotechnology, which is based on strand displacements. One example, created by Yurke *et al.*, is a molecular machine. The machine was opened and closed by the sequential addition of two complementary single strands of fuel DNA, which powered the process and expended its potential energy in forming dsDNA (ΔG = - 1.8 kcal mol^{-1}).[62] In the open version, the "tweezer"-like machine has two single-stranded sequences that, upon hybridization to the added fuel DNA, are forced closed. The fuel DNA hybridized into the closed state of the machine creates a short single-stranded overhang, later called toehold.[63] Thereafter, the addition of the complementary fuel DNA can open the machine again, by hybridizing the toehold and starting a branch migration[64] with a random walk strand displacement.[63] The displacement of the first fuel DNA results in expended dsDNA and a reopened machine.[62] The toehold length affects the strand displacement speed until it is seven nucleotides long.[65,66]

Utilizing the toehold for strand displacement, Dirks and Pierce created an oligonucleotide system able to detect an initiator oligonucleotide, which induces a hybridization chain reaction (HCR). The HCR mechanism is fueled, just like in Yurkes "machinery", by the potential energy of single-stranded DNA,[62,66] which in this case is "hidden" in two metastable hairpins (H1 and H2). Each hairpin has an unpaired toehold and a short single-stranded DNA loop separated by a dsDNA stem. The two hairpins are unable to react with each other, until the addition of an initiator oligonucleotide. The initiator hybridizes the toehold of the first hairpin (H1) and displaces its stem. The opening of H1 enables it to hybridize the complementary strand of the H2 toehold, which in turn is opened and reveals an identical sequence of the initiator. This isothermal reaction quickly forms a nicked dsDNA polymer as more hairpins build on to the complex.[67] HCR can be used as an enzyme-independent method of signal amplification in a variety of different settings.[68] The simplest system is to use HCR for direct detection of a nucleic acid (DNA or RNA) in solution, where the target is the initiator.[69] However, it is not always optimal or possible to use the hairpin as a detector, because the target affects the whole sequence design of the hairpins and thereby their stability. When using HCR to amplify the signal of *in situ* hybridization (ISH), the initiator is integrated into the probe. Normally, ISH uses

DNA probes to hybridize nucleic acid (DNA or RNA) in the cell and the target is visualized by attaching a fluorophore/chromophore or an enzyme, like HRP, to the probe. When combining ISH and HCR, the DNA probe contains a sequence able to hybridize to the target molecule and an initiator sequence. The target is hybridized by the probe and all excess probe is removed. Because only the target is marked with the initiator sequence, the hairpins labeled with fluorophores can start a localized HCR amplifying the signal. Multiplexing the system Pierce and colleagues were able to detect five different mRNAs in zebrafish embryos.[70]

To ensure that the hairpins are not starting the HCR without the initiator, they need to be designed with the right lengths and sequences. Dirks and Pierce have created algorithms to predict oligonucleotide performance,[71-73] which they have integrated and made available on the homepage nupack.org.[74] These allow us to analyze sequences *in silico* and predict how they would perform under different salt concentrations and temperatures. The homepage also has a function to design sequences capable of forming secondary structures and hybridizing each other according to a predefined reaction pathway with minimal crosstalk.[75] However, most available software, including nupack.org, are not built for prediction of non-Watson-Crick base pairing in oligonucleotides.[76] These predictions would be valuable when oligonucleotides form tertiary structures, not following the normal DNA helix, like in aptamers[77] and DNAzymes.[78] Aptamers are synthesized oligonucleotides that have been screened for affinity to a specific target and are one of several molecules that can replace antibodies as a target-binding reagent.[77,79] Common tertiary structures for both aptamers and DNAzymes are G-quadruplexes, often formed in G rich sequences. The G-quadruplexes have the ability to bind molecular structures and in the case of DNAzymes they together with cofactors possess an enzymatic activity.[80]

DNA modifying enzymes

The versatility of DNA has made it an excellent tool for both detection of biomarkers and amplification of signal. Combining it with DNA modifying and producing enzymes increases the possibilities of what can be achieved with DNA as a tool. These enzymes normally function as part of the DNA replication and repair mechanism in different pro- and eukaryotic cells.

One enzyme type that has been crucial for genotyping, cloning and recombinant protein production is the restriction enzymes. These enzymes have naturally evolved in prokaryotes as a defense mechanism against viruses.[81] Restriction enzymes are able to cut dsDNA at enzyme specific DNA sequences.

These restriction sites are around four to eight base pairs long and have a specified DNA sequence, which often is a palindrome. Either the enzyme will create a cut in the middle of the recognition site on both strands, causing a blunt end of the dsDNA, or it can cut the strands creating overhangs on each side, called sticky ends.[82] There also exist natural nicking endonucleases, which only cleave one of the stands creating a nick in the dsDNA. Growing amounts of engineered versions of restriction enzymes are created into nicking endonucleases. One example of this is the Nt.BsmAI enzyme used in **paper IV**, which is cleaving a strand between two undefined nucleotides following the five nucleotide sequence (5'… GTCTCN|N…3') if it is binding dsDNA.[83]

An alternative way of nicking dsDNA used in **papers III** and **IV** is to combine two enzymes, uracil-DNA Glycosylase (UNG) and endonuclease IV (EndoIV), which are both parts of the DNA repair mechanism called base excision repair (BER).[84] One type of DNA damage is deamination of the cysteine base causing it to become a uracil (U) instead. Uracil bases are used in RNA, instead of T, and hence pair up A. If deamination damages are not corrected the cell will upon its next replication get a daughter cell with a point mutation, because U will pair with A instead of G. The UNG is an enzyme with highly conserved function and it can be found in species from *Escherichia coli* (*E.coli*) to humans and even some viruses, like Herpes simplex viruses. In all these, UNG is responsible for detecting any U bases in the DNA and cleaving the N-glycosidic link between the deoxyribose and the base. The remaining abasic or apurinic/apyrimidinic site (AP site) is cytotoxic for the cell.[85] Spontaneous depurination and to a lesser amount depyrimidination are also contributing to the creation of AP sites.[86] Therefore, all cells have AP-endonucleases ready to repair this damage. In *E.coli,* there exist two separately evolved AP-endonucleases EndoIV and exonuclease III (ExoIII), but only the ExoIII family is conserved in mammals.[85] The function of EndoIV is to remove AP sites on dsDNA, by first cleaving the 5' phosphodiester bond, leaving a 3' OH group and a deoxyribose without a base. The abasic deoxyribose is subsequently cleaved off by EndoIV leaving a phosphate group on the 5' end of the gap.[85,87] The following step of the BER pathway is to fill in the missing nucleotide with a DNA polymerase and a DNA ligase.[86] This process is similar to the one we employ for **paper IV** where we need to fill a gap in an oligonucleotide. Polymerases and ligases are both critical enzymes for replication and repair processes.

Today, the golden standard for amplification and quantification of DNA and RNA is polymerase chain reaction (PCR).[88] The method requires a DNA polymerase, two priming oligonucleotides able to hybridize opposite sides of a specific target sequence, called forward and reverse primer, and a DNA target molecule. The target molecule or template has to be DNA or RNA

transformed into DNA through reverse transcription. The method is divided into three parts repeated in cycles, starting with a high temperature to denature the target molecule. Thereafter follows a hybridization (annealing) of the primers and finally an extension of the primer with DNA polymerase. The DNA polymerase incorporates nucleotides complementary to the template in a 5' to 3' direction, by attaching nucleoside triphosphates (dNTPs) to the hydroxyl group of the primers 3' end. Every cycle doubles the amount of target DNA and originally, the DNA polymerase from *E.coli* had to be replenished between each cycle due to the heat sensitivity of the enzyme. However, the discovery of the thermostable Taq polymerase[89] from the thermophile bacteria *Thermus aquaticus* together with the building of a thermal cycler allowed the reaction to become automated. The Taq polymerase also allowed the whole reaction to be performed at a higher temperature, thereby increasing the stringency of primer hybridization, resulting in a higher fidelity of the reaction by reducing unspecific amplification. To quantify the template, the amount of DNA can be measured after each cycle, either by adding SYBR green, which fluoresces when it intercalates dsDNA or by adding Taqman probes. This quantitative PCR (qPCR) detects at which cycle (Ct value) the fluorescence of the sample exceeds a threshold above the background from which the amount of template DNA can be inferred.[90]

The ability to detect and amplify a specific DNA sequence with high specificity has been integrated with protein detection in immune PCR, where a DNA labeled antibody is used as a probe to detect low amounts of proteins on a solid support in an ELISA like approach.[91] The signal is amplified with PCR and detected with agarose gel-electrophoresis or qPCR.[91,92] The DNA labeled antibody probe combines the wide range of epitope binding of an antibody with the flexibility of using DNA as a molecular tool. Attaching an oligonucleotide to an antibody, or any other affinity binder provides a platform for developing DNA based detection systems. One example of this is the proximity ligation assay (PLA), using probes to detect proteins in close proximity to each other. These proximity probes are antibodies or aptamers with single-stranded oligonucleotides, which have predefined sequences. One of the oligonucleotides is connected to the affinity binder at the 3' end while the other is connected at the 5' end. This allows the two free 5' and 3' ends to hybridize to a complementary connector oligonucleotide if the probes are binding to interacting proteins. The connector oligonucleotide positions the 3' hydroxyl group adjacent to the 5' phosphate group and a DNA ligase will join them by creating a phosphodiester link between the oligonucleotides of the proximity probes, thereby creating a DNA template for qPCR. The targets can be detected in a homogeneous solution or on a solid phase support and the qPCR readout serves as quantification of interacting proteins.[93,94] The DNA ligase used comes from the bacteriophage T4[95] and is efficient at joining nicked DNA in the presence of ATP for energy.[96]

However, PCR is not capable of preserving the localization of the amplified signal and, therefore, is not a suitable method to detect proteins, protein-protein interactions and PTMs *in situ.* An oligonucleotide system able to detect genomic DNA *in situ* is the padlock probe.[97] A padlock probe is a single-stranded oligonucleotide with the two end segments complementary to the target, which upon target hybridization form a nicked DNA circle. The circle can be ligated only if the base pairing with the target molecule is correct, with an ability to discriminate between single nucleotide polymorphism.[98] Inspired by rolling circle DNA replication[99] found in *e.g.* bacteriophages, the circular DNA can template the extension by a polymerase creating a concatemer.[100,101] In the presence of a primer and dNTPs, the polymerase will produce a long single-stranded oligonucleotide with multiple repeats of the same sequence in a process called rolling circle amplification (RCA). About 1000 copies of a 100 nucleotides long circle are produced in an hour when using the highly processive phi29 DNA polymerase.[102,103] Each copy is hybridized with a fluorescently labeled single-stranded detection oligonucleotide (DO), resulting in a bright, up to one µm sized object, when visualized in an epi-fluorescence microscope. However, phi29 DNA polymerase also has 3' exonuclease activity and it can digest any single-stranded DNA without protected 3' ends, which prevents the use of unprotected DOs during RCA.[104] This ability also allows it to chew up non-base pairing nucleotides at the 3´end of a DNA molecule binding the ligated padlock probe until it is able to prime the extension of the circle.[98,105]

In situ PLA brings the dual recognition of homogenous PLA into the cell and combines it with the signal amplification of padlock probes.[106] Both oligonucleotides of the proximity probes are required to template the ligation of two additional circularization oligonucleotides into a circular DNA molecule. Each of the circularization oligonucleotides, a long and a short, contain complementary parts to both of the proximity probes in each end, and after hybridizing they form a circle with a nick on each template. The long circularization oligonucleotide has a spacer sequence of about 40 bases in which the DO sequence is included. After ligation, the DNA circle is amplified by phi29 polymerase using one of the proximity probes to prime the extension. The resulting RCA product is hybridized by the DO linked with either a fluorophore or an HRP and can be visualized in a microscope (**figure 2**). For IF, the signal is sometimes difficult to distinguish from the autofluorescence.[107,108] For *in situ* PLA, the signals are easier to detect due to the strong signal amplification making them visible above autofluorescence and using image analysis software, *i.e.* CellProfiler, the RCA products can be quantified. However, when there is a large amount of RCA products, they can start to coalesce preventing differentiation of separate signals and reducing the dynamic range. This problem can be solved by exchanging the long circularization oligonucleotide with three separate ones containing different DO

sequences at a known ratio (1:10:100). The ability to choose which of the three RCA products to quantify at a cellular level increases the dynamic range that is required for *e.g.* heterogeneous tissues with a wide range of expression levels.[109]

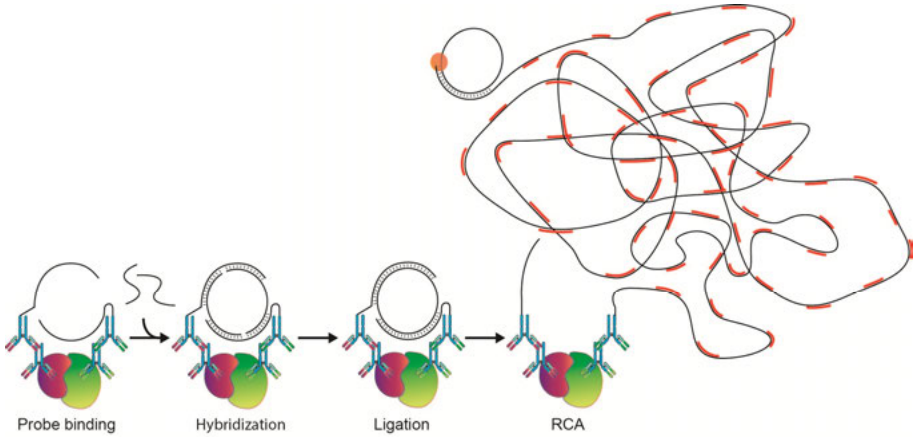

Figure 2. A schematic illustration of *in situ* PLA.

Even though the RCA product is on the scale of 100s of nm, the maximum distance between the epitopes bound by the proximity probes is around 40 nm, when using secondary antibodies and smaller when using directly DNA conjugated primary antibodies. The resolution of an epi-fluorescence microscope is, as stated above, down to around 200 nm, but with *in situ* PLA it is possible to visualize proteins within around 40 nm and, therefore, it is more suitable for detecting interacting proteins than a dual IF staining. Dependent on the antibodies used, *in situ* PLA is able to detect single proteins, their PTMs[110-112] or protein-protein interactions.[106] The method can combine two antibodies raised against the same protein to increase the specificity due to the dual epitope recognition. Although both of the antibodies have undefined off-target binding, they will only generate an amplifiable circle when they bind the same protein and, thereby, the unspecific background signal is decreased.[113] This also means that it is possible to *e.g.* combine a specific receptor antibody with a pan-phosphorylation antibody and still visualize the receptor phosphorylation. Furthermore, *in situ* PLA has been used to detect protein-DNA interactions,[114] mRNA expression[115] and the localization of single-proteins to cellular compartments.[116] Leuchowius *et al.* showed that *in situ* PLA can be multiplexed for detecting different interaction partners of a protein.[117] Besides, PLA can also be combined with other *in situ* detection methods like padlock probes.[118]

Present Investigations

Paper I: Proximity dependent initiation of hybridization chain reaction

Introduction

The proximity-dependent signal generation of *in situ* PLA relies completely on enzymes for it to function. However, enzymes are both relatively expensive reagents and require continuous storage at the right temperature. For example, the phi29 DNA polymerase has a half-life of about 18 minutes without a DNA substrate at 30°C, and at 40°C the enzyme activity is lost almost instantly.[119] It would, therefore, be beneficial to have a method for proximity dependent detection of PTMs and protein-protein interactions without the prerequisite for enzymes. The method would become cheaper and fewer requirements are put on storage capabilities by removing the enzymatic steps. The proximity-dependent initiation of hybridization chain reaction (proxHCR) was developed for this purpose and combines the proximity requirement of *in situ* PLA and the amplification of hybridization chain reaction (HCR).

ProxHCR oligonucleotide design

The proxHCR design has two proximity probes that upon activation can initiate HCR if they are in close proximity. The probes consist of antibodies conjugated to hairpin oligonucleotides. These proximity hairpins (PH1 and PH2) are unable to interact until an activator oligonucleotide is added. The activator oligonucleotide finds a toehold in the loop of PH1 and displaces the stem of the hairpin, thereby opening it up. If the open PH1 probe is bound in proximity to PH2, it starts a strand displacement of that oligonucleotide using the PH2 loop as a toehold. The activator-PH1-PH2 complex results in the presentation of an initiator sequence that is able to start the HCR of two fluorophore-labeled hairpins (H1 and H2, **figure 3**). To prevent the opened PH1 oligonucleotide from initiating HCR and thereby generating a false-positive signal, a two-nucleotide mismatch was introduced. The mismatch was designed in such a way that the toehold of the H1 oligonucleotide has a

perfect hybridization match to initiator sequence of PH2 but has a mismatch to the opened PH1 oligonucleotide. Therefore, HCR can only be started efficiently when the two proximity probes are interacting. All the hairpins (PH1, PH2, H1 and H2) are designed to be metastable and not react with each other when closed. The probes PH1 and PH2 are allowed to bind their targets and any unbound is washed off before the activator oligonucleotide is added to prevent false-positive signal. The HCR will continue until all hairpins are depleted and result in an amplified signal localized where the interaction occurred.

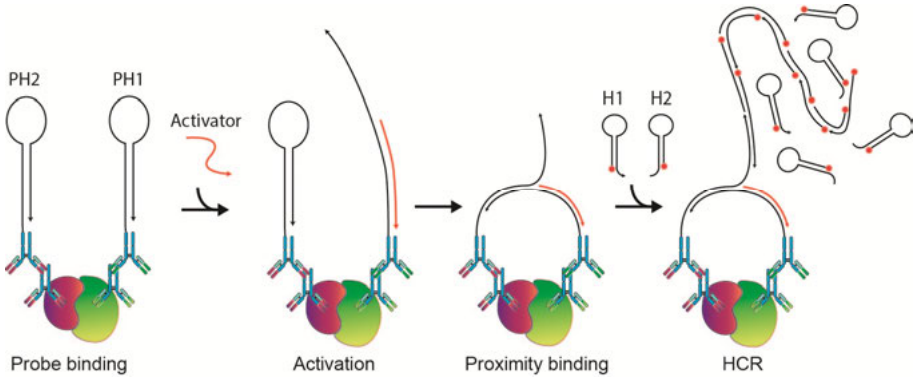

Figure 3. A schematic illustration of proximity-dependent initiation of hybridization chain reaction (proxHCR).

Findings and discussion

The design of proxHCR was optimized by predicting secondary structures and interactions *in silico* using the Nupack.org software. The final design was predicted to have a negative Gibbs free energy (–ΔG) of 40 and 27 kcal/mol for PH1 and PH2 hairpins, respectively, and therefore to be metastable. The kinetic properties of the hybridization reactions were quantified in a surface plasmon resonance instrument. All the different oligonucleotide-binding steps displayed a strong affinity with a quick association and a slow dissociation that was outside of the detection range of the biosensor. Control experiments in which key oligonucleotides were excluded, showed no association of the oligonucleotides suggesting that their secondary hairpin structures are metastable when not activated. This was also confirmed by bead experiments and polyacrylamide gel electrophoresis, where no reaction was detected without an initiator. Attaching an initiator sequence to a magnetic bead allowed for the measurement of how the fluorescent signal increased overtime and showed that after 30 minutes all the HCR hairpins were depleted.

After establishing the stability and robustness of the proxHCR oligonucleotide system, it was evaluated in different *in situ* experiments. The method was tested to detect both protein-protein interactions and PTMs in a variety of different assays and resulted in a significantly increased signal compared to the biological and technical control. The technical controls were performed by excluding both or either of the primary antibodies on cultured cells as well as on tissue and no signal was detected. For the biological controls, the basal level of protein-protein interaction or PTM was detected when cells were not stimulated. The method also produced a similar staining pattern to *in situ* PLA on both cultured cells and on tissue. However, the signal from proxHCR was more diffuse and quantification through enumeration was not possible as with RCA products. Finally, the signal from proxHCR was also clearly detectable on cells passing through a flow cytometer when detecting an extracellular interaction.

Paper II: Compaction of rolling circle amplification products increases signal integrity and signal-to-noise ratio

Introduction

One of the challenges with amplifying the signal by RCA used for *in situ* PLA and other methods is the accurate quantification of RCA products during image analysis. It can be difficult to define the exact edge of a signal and to resolve neighboring signals from each other. As previously described, RCA is performed by a DNA polymerase templated by a circular DNA molecule and results in a linked series of several hundreds of repeated copies of the same DNA sequence. The RCA product collapses into a bundle of DNA and each repeat is hybridized by a fluorescently labeled DO. However, the DNA bundle does not always collapse into a single entity, but sometimes splits into several fluorescent objects. The lack of signal integrity in individual RCA products can cause blurriness around the signal, which is reducing the ability to distinguish signals from each other and from background fluorescence. Therefore, an oligonucleotide system was designed to generate distinct RCA products of a reduced size.

Compaction oligonucleotide design

The RCA product is a long oligonucleotide with a repeated sequence dependent on the template DNA circle, which contains a site for DO hybridiza-

tion for each copy. However, there is a lot of unhybridized single-stranded DNA left in the RCA product and this can be utilized for compacting the DNA bundle. A compaction oligonucleotide with two repeated elements, separated by a few nucleotides long spacer sequence, was designed to hybridize a segment of the RCA product and cross-link the DNA bundle with itself. The compaction elements are reverse-complementary to a segment of the RCA product and hybridize the RCA product at any two repeats of the sequence preventing them from floating apart (**figure 4**). The result is a more uniform and compacted RCA product.

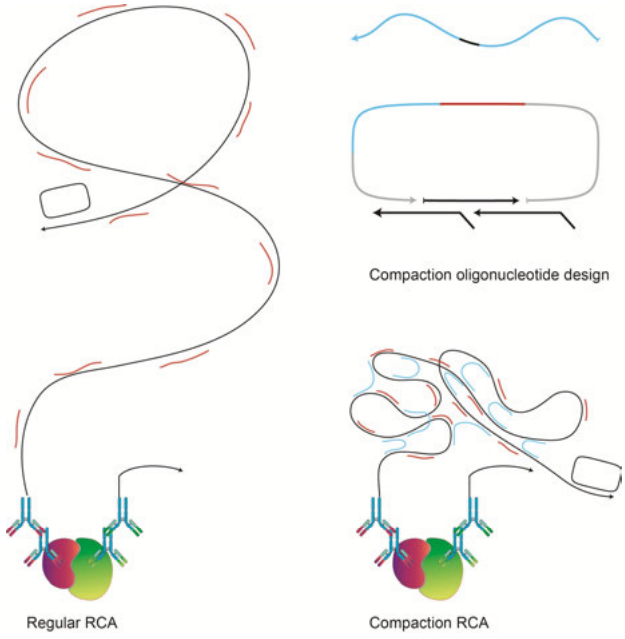

Figure 4. A schematic illustration of the compaction oligonucleotide design and how it can cross-link the RCA product.

Findings and discussion

The compaction oligonucleotide system was validated under different conditions and with different readout methods to confirm its effects on RCA products and to compare it with regular RCA. First, RCA products were produced on slides with synthetic targets and two padlock probes that were identical apart from having different sites for DO hybridization. Compacted and regular RCA products were visualized with two different DOs and compared. The compaction oligonucleotide system resulted in compact and bright RCA products, while the regular RCA products were prone to form small clusters of diffuse fluorescent objects and had a larger object diameter. That compacted products were brighter and had a better signal-to-noise ratio

was confirmed by measuring individual RCA products. To establish whether regular and/or compacted RCA result in single products splitting up in clusters, the observed data for two separately colored RCA products were compared with a Monte Carlo simulation to predict the degree of same color neighbor. For the regular RCA, the simulation predicted a lower amount of expected same colored RCA product neighbors than observed, implying that the RCA products split into clusters. The compacted RCA products had the same amount of same-color neighbors as predicted by the Monte Carlo simulation, thus confirming that the compaction oligonucleotide can prevent clustering. Besides, RCA is also used to amplify signals in solution where the RCA products are quantified in microfluidic systems. For optimal quantification, the RCA product should be as uniform and bright as possible and it is therefore important to evaluate the compaction oligonucleotide system in solution. An RCA reaction in solution is free moving in the liquid and, therefore, there is a risk of the compaction oligonucleotides to cross-link two separate RCA products. Two separately colored RCA products were generated in solution with or without compaction oligonucleotides and they were imaged while flowing through a microfluidic system. In the presence of compaction oligonucleotide, more dual colored RCA products were formed, indicating aggregation of signal at a significantly higher rate. The system was also evaluated on hippocampal tissue, where *in situ* PLA was used to detect progranulin. The compacted RCA products were smaller and sharper compared to the regular RCA signals.

Paper III: Improved efficiency of *in situ* protein analysis by proximity ligation using UnFold probes

Introduction

The conventional *in situ* PLA only detects a fraction of the protein-protein interactions or PTMs. The proportion of detected events is dependent on target-antibody binding and the ability for the oligonucleotide system to generate a detectable signal. Focusing on the signal generation, *in situ* PLA requires two externally added circle oligonucleotides to bind the proximity probes correctly to form a nicked circular DNA molecule, which thereafter have to be ligated at two sites. Improving the signal generation could lead to an increase in both efficiency and sensitivity. This is why, we developed a new design of *in situ* PLA in which one of the proximity probes has an incorporated circularization oligonucleotide that can be unfolded after target binding, thereby increasing the likelihood of signal generation.

UnFold *in situ* PLA oligonucleotide design

UnFold *in situ* PLA was designed to include a circularization oligonucleotide in one of the proximity probes and a template for ligation in the other probe. The circle-forming oligonucleotide of the first probe is shaped like a hairpin loop where the 5' end is conjugated to the antibody and the 3' end has a single-stranded DNA tail after the stem. Three U bases in a row are incorporated into the stem at the 5' end of the loop. The template probe also contains a hairpin oligonucleotide to prevent probe-to-probe hybridization during incubation, by protecting the template oligonucleotide sequence in this dsDNA structure. There are multiple U bases spread out along the stem on one side of the hairpin. After having bound in the sample, the two proximity probes are unfolded by treating them with UNG to remove the U bases and EndoIV to remove the abasic sites generated by UNG. This causes the circle probe to release a padlock-like oligonucleotide still hybridized to the probe, and the other probe to expose the ligation template. A ligase can now close the circularization oligonucleotide with the help of the ligation template. Thereafter, RCA is performed followed by incubation with fluorophore-labeled detection oligonucleotides (**figure 5**). To summarize, one of the UnFold probes carries with it the circularization oligonucleotide and therefore if the two probes are in proximity, only one ligation event is required for signal generation.

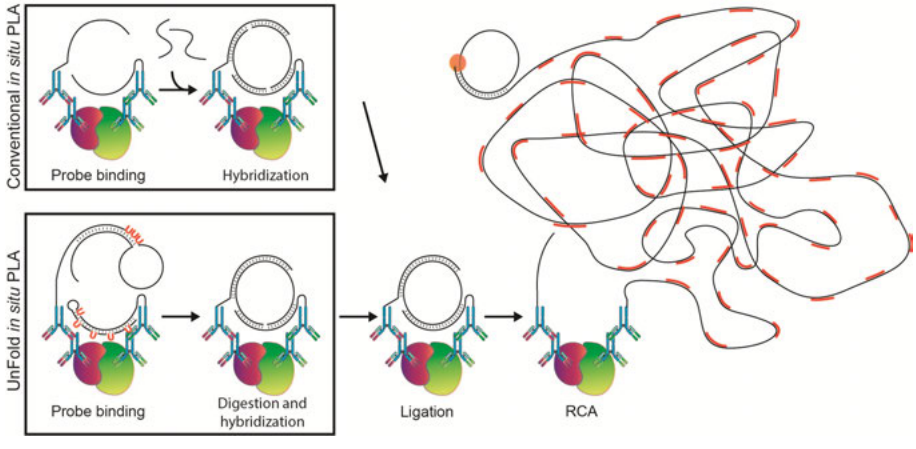

Figure 5. A schematic illustration of conventional and UnFold *in situ* PLA.

Findings and discussion

Comparisons were conducted between conventional and UnFold *in situ* PLA on cultured cells, on tissue and on a solid phase support. First, the signal-to-

noise ratios of the two designs at different concentrations were determined and it was established that the conventional and the UnFold probes performed best at 600 ng/ml and 66 ng/ml, respectively. UnFold *in situ* PLA was able to detect the interacting proteins E-cadherin and β-catenin in both HaCat cells and on skin tissue at a higher efficiency than the conventional design. The technical controls, in which one of the primary antibodies or either of the probes were omitted, showed a low amount of signal, likely resulting from antibody cross-reactivity and/or unspecific binding. The new UnFold design was able to detect more phosphorylation events in BJ hTert cells after stimulation with PDGF-BB, both when investigating the ligands receptor, PDGFRβ, and when detecting the two downstream targets AKT and ERK1/2. The two *in situ* PLA designs were also compared with sandwich ELISA to detect IL-6 immobilized on a microtiter plate. The conventional and UnFold *in situ* PLA probes had a higher sensitivity than the ELISA probes, and the UnFold design had the lowest limit of detection. The UnFold *in situ* PLA design provides increased efficiency for *in situ* analysis and enhanced sensitivity in detecting proteins in a microtiter well format.

Paper IV: A Boolean analysis of protein interactions at a molecular level in single cells

Introduction

Measuring signaling activity in a cell with *in situ* PLA provides information on the abundance of a protein interaction, but not on the abundance of the non-interacting fraction of these proteins. The amount of protein A interacting with protein B is dependent on the free amount of the two proteins and their affinity to bind each other. Developing a method that is able to detect interacting proteins and simultaneously detect the amount non-interacting proteins would reveal whether the fraction of interacting proteins was changed. The aim is to create an oligonucleotide system able to generate signals for individual proteins (A or B) and their interactions (A and B) in a molecular Boolean (MolBoolean) protein analysis.

MolBoolean oligonucleotide design

The MolBoolean design utilizes preformed DNA circles that hybridize to proximity probes and consequently incorporate a tag oligonucleotide into the circle. The circle is thereafter amplified with RCA and a sequence-specific DO visualizes which tags have been incorporated. Each proximity probe is

an oligonucleotide-conjugated antibody and the oligonucleotide consists of a single-stranded sequence with a hairpin loop structure in the middle. The preformed DNA circle has reverse complementary sequences to the single-stranded sequences of both proximity probes. The DNA circle can thus bind to either one of the probes or to both if they are in close proximity. To enable tag incorporation, the circle contains two digestions sites, which can be cut using two different enzyme-dependent strategies. Incorporating a U base in the circle enables it to be nicked by UNG and EndoIV. Alternatively, the digestion can be achieved with Nt.BsmAI that creates a nick in the circle only if it is dsDNA *i.e.* has bound a proximity probe. Both strategies require their own sets of oligonucleotide design with different sequences, but the nicking site is always 3 base pairs away from the hairpin of the probe. After the circle is cut, the tag oligonucleotide invades the hairpin-loop of the proximity probe and opens it up through strand displacement, while the two flanking sequences of the probe remain hybridized to the circle. The circle is reformed by ligating in the tag oligonucleotide and thereby incorporating its sequence into the circle. One of the proximity probes acts as a primer for RCA to generate a concatemer containing hundreds of reverse complementary copies of the tag sequence, which are hybridized by a fluorescently labeled DO. The RCA product is visualized and dependent on the tag(s) incorporated and DOs hybridized, it has either one of two different colors or both (**figure 6**).

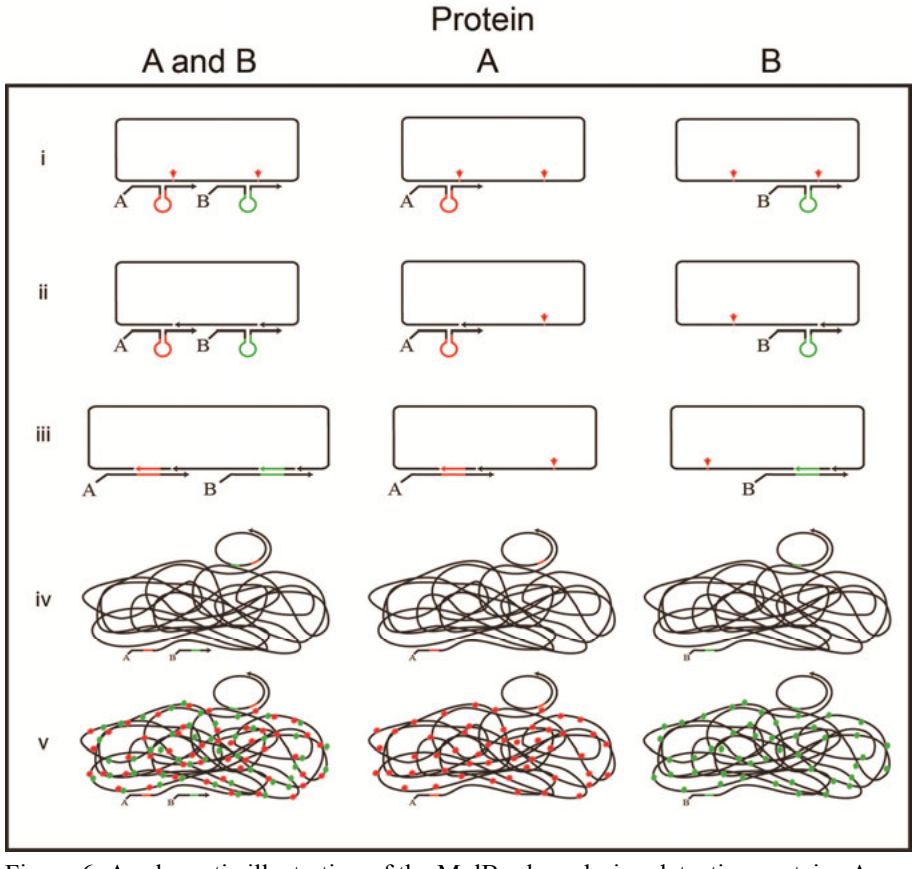

Figure 6. A schematic illustration of the MolBoolean design detecting proteins A or B and their interaction (A and B). i) Hybridization of the preformed DNA circle to the proximity probes. Digestion sites are shown with a red arrowhead. ii) The DNA circle is cut at the digestion sites where it is binding a proximity probe. Note that the proximity probe is unaffected by the digestion. iii) The tag oligonucleotide invades the hairpin-loop of the probe and is integrated into the circle through ligation. iv) Signal amplification by RCA. v) Hybridization of two different detection oligonucleotides (green and red). The resulting RCA product is colored according to which proximity probe gave rise to the signal (A=red and B=green).

Findings and discussion

The two oligonucleotide designs were validated in solution before any *in situ* experiments were conducted. Digestion and ligation efficiencies were evaluated with gel electrophoresis for both digestion strategies with relevant exclusion controls and appeared to produce the expected digestion and ligation products. The conjugated probes were incubated on HaCat cells to detect E-cadherin, β-catenin and their interaction. The functionality of the method was evaluated by excluding either of the primary antibodies or both. The uracil based MolBoolean design successfully detected protein interaction but

was unable to detect the individual proteins. This was likely due to the inability to amplify the signal by RCA when the circle contained an AP site as a result of UNG treatment. To address this issue, two additional steps were added to the protocol, which removed the AP site and filled the resulting gap with the corresponding T nucleotide. With the extended gap-fill protocol we were able to the detect protein interaction as well as E-cadherin and β-catenin individually. The Nt.BsmAI design gave a similar result. These results indicate that our designs are capable of MolBoolean analysis of proteins and have the potential to detect a change in the interaction properties of two proteins.

Concluding remarks

There are estimates mapping all the different protein-protein interactions ranging from 130,000[120] to 650,000[121] in us humans. These interactions are structured into networks and changes in a subset of interactions can be associated with changes in phenotype leading to disease *e.g.* cancer.[120,122] Investigating the level of a known protein interaction using *in situ* PLA in patient samples can be used to predict clinical response to treatment of an inhibitor.[123] The subcellular localization of proteins have also been correlated with disease.[10] Therefore, proximity-based methods for *in situ* protein detection have the potential for clinical applications as well as being used to further enhance our understanding of cellular processes.

The versatility of DNA hybridization and DNA modifying enzymes have enabled us to develop several DNA-based methods for analysis of protein interactions. In this thesis, I present how *in situ* PLA can be improved by compacting the RCA products (**paper II**) or by increasing its efficiency (**paper III**). Furthermore, two new methods were developed for detection of proteins *in situ*, one enzyme-free method (**paper I**) and one that enables detection of a protein-protein interaction and the level of each individual protein (**paper IV**).

In **paper I**, an enzyme-independent method for detection of protein-protein interactions *in situ* was developed. This proximity dependent initiation of hybridization chain reaction (proxHCR)[124] presents an inexpensive and robust alternative to *in situ* PLA. In contrast to RCA products, HCR signals are not quantifiable through enumeration but have to be quantified by integrated fluorescence intensity. The cost efficiency of the method makes proxHCR more suitable for high-throughput screening of proximity detection *in situ* than PLA. The versatility of DNA hybridization allows for multiplexing, which has been performed with conventional HCR measuring mRNA expression.[125] As no enzymes are required, the storage conditions of the reagents are less sensitive and they could be stored in room temperature for longer times. This makes proxHCR a possible candidate for further optimizations into a point-of-care device.

The issue of irregular formed RCA products was addressed in **paper II**. An oligonucleotide system was designed to improve *in situ* PLA and other RCA

based methods by cross-linking and, thereby, compacting each individual RCA product.[126] These compaction oligonucleotides can prevent RCA products from splitting up and improve the signal-to-noise ratio. The compacted and brighter RCA products are easier to quantify during image analysis. The compaction oligonucleotide design can be combined with any RCA based *in situ* method and *e.g.* UnFold *in situ* PLA has been successfully compacted.

A new design of *in situ* PLA was developed in **paper III** to increase the detection efficiency of protein interactions and PTMs *in situ*. For UnFold *in situ* PLA, the proximity probes carry with them the circularization oligonucleotide and template for circularization. This results in a more efficient circle formation, probably due to a higher local concentration of circularization oligonucleotide and the requirement of only one ligation event. The increased efficiency and improved signal-to-noise ratio make the UnFold design suitable for detecting rare events. Even though the protocol includes an additional enzymatic step, it requires fewer proximity probes and has a higher sensitivity. The UnFold *in situ* PLA design is, with slight alterations, suitable for multiplexing due to the high efficiency. By conjugating primary antibody probes and visualization of the RCA product with *e.g. in situ* sequencing[127] or sequential hybridization of DOs[128] it is theoretically possible to reach relatively high levels of multiplexing.

Paper IV is still in the developmental phase but has proven to have the potential for MolBoolean analysis of individual protein levels and their interactions. An increase in interaction is not necessarily due to an altered affinity between the two interacting proteins because of PTMs but can be caused by an increase in expression level of one or both of the interacting proteins thereby shifting the equilibrium of the interaction. The MolBoolean method is designed to increase the insight of why there is a change in interaction.

This thesis illustrates how molecules borrowed from nature in this case DNA, DNA modifying enzymes and antibodies have been combined into molecular biology tools, with a special focus on *in situ* analysis of cell signaling activity.

Acknowledgments

I would like to start by thanking my faculty opponent **Cristina Al-Khalili Szigyarto** and my examination board **Lars Hellman**, **Charlotte Rolny** and **Mathias Hallberg** for taking your time to read my thesis and coming to the defense discussing my work.

I would like to express gratitude to all the people who have inspired and helped me during my PhD time. Beginning with my supervisor **Ola Söderberg**, thank you for allowing me to join your research group and for teaching me how to do method development. Thank you for creating a great work environment with awesome work colleagues.

I would also like to thank my co-supervisor **Mats Nilsson** for the inspiring meetings in both Uppsala and in your lab in Stockholm where our groups were allowed to exchange ideas.

I want to extend a special thanks to all present and former members of Ola Söderberg's group: **Karin** and **Björn** for introducing me to the lab and teaching me how to UnFold probes. **Dorothea** for being a delightful office-mate, also **Linda, Carl-Magnus, Gaelle, Johan H, Mattias** and **Chi-Chih** for all the discussions and help you have offered.

The majority of my time as a PhD student was spent in the MolTools lab at IGP before our group switched faculty and I got to know a lot of great people there. I would like to thank **Ulf Landegren** for creating an environment of discussion and ideas as well as **Masood Kamali-Moghaddam** for organizing the weekly journal club.

I would also like to thank **Felipe** for being an awesome friend and colleague giving me design problems to solve, **Liza** for always having a smile on her face, **Johan B, Marcus, Malte, Agata, Peter, Rasel, Tonge, Sophie, Camilla Reddy, Maria H, Caroline, Hongxing, Lei, Di** and **Phathu** for interesting discussion about scientific and non-scientific topics, **Christina, Johanna, Elin, Erik, Radiosa** and **Johan O** for always being willing to help with all sort of practical problems and thank you **David** for intriguing discussions about both GoT and RCA.

Finally, an infinite thanks to **Sina** for helping to proofread this thesis and for encouraging me throughout the writing process. I'm not sure if it would have been possible without your support.

References

1 Alberts, B. *et al.* in *Molecular Biology of the Cell* Ch. 1, 1-11 (Garland Science, 2008).
2 Alberts, B. *et al.* in *Molecular Biology of the Cell* Ch. 15, 879–903 (Garland Science, 2008).
3 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).
4 Turner, N. C. & Clapham, J. C. Insulin resistance, impaired glucose tolerance and non-insulin-dependent diabetes, pathologic mechanisms and treatment: current status and therapeutic possibilities. *Prog Drug Res* **51**, 33-94 (1998).
5 Uhlen, M. *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. *Molecular & cellular proteomics : MCP* **4**, 1920-1932, doi:10.1074/mcp.M500279-MCP200 (2005).
6 Andreani, J. & Guerois, R. Evolution of protein interactions: from interactomes to interfaces. *Arch Biochem Biophys* **554**, 65-75, doi:10.1016/j.abb.2014.05.010 (2014).
7 Hitosugi, T. & Chen, J. Post-translational modifications and the Warburg effect. *Oncogene* **33**, 4279-4285, doi:10.1038/onc.2013.406 (2014).
8 Gruber, W., Scheidt, T., Aberger, F. & Huber, C. G. Understanding cell signaling in cancer stem cells for targeted therapy - can phosphoproteomics help to reveal the secrets? *Cell Commun Signal* **15**, 12, doi:10.1186/s12964-017-0166-1 (2017).
9 Bauer, N. C., Doetsch, P. W. & Corbett, A. H. Mechanisms Regulating Protein Localization. *Traffic* **16**, 1039-1061, doi:10.1111/tra.12310 (2015).
10 Park, S. *et al.* Protein localization as a principal feature of the etiology and comorbidity of genetic diseases. *Mol Syst Biol* **7**, 494, doi:10.1038/msb.2011.29 (2011).
11 Croft, W. J. *Under the microscope: a brief history of microscopy*. Vol. 5 (World Scientific, 2006).
12 Chan, J. K. The wonderful colors of the hematoxylin-eosin stain in diagnostic surgical pathology. *International journal of surgical pathology* **22**, 12-32, doi:10.1177/1066896913517939 (2014).
13 Lichtman, J. W. & Conchello, J.-A. A. Fluorescence microscopy. *Nature methods* **2**, 910-919, doi:10.1038/nmeth817 (2005).
14 Lichtman, J. W. & Conchello, J. A. Fluorescence microscopy. *Nature methods* **2**, 910-919, doi:10.1038/nmeth817 (2005).
15 Harris, J. R. Transmission electron microscopy in molecular structural biology: A historical survey. *Arch Biochem Biophys* **581**, 3-18, doi:10.1016/j.abb.2014.11.011 (2015).
16 Galbraith, C. G. & Galbraith, J. A. Super-resolution microscopy at a glance. *J Cell Sci* **124**, 1607-1611, doi:10.1242/jcs.080085 (2011).

17 Neil, M. A., Juskaitis, R. & Wilson, T. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Opt Lett* **22**, 1905-1907 (1997).
18 Tsurui, H. *et al.* Seven-color fluorescence imaging of tissue samples based on Fourier spectroscopy and singular value decomposition. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **48**, 653-662 (2000).
19 Ma, B., Winkelbach, S., Lindenmaier, W. & Dittmar, K. E. Six-colour fluorescent imaging of lymphoid tissue based on colour addition theory. *Acta histochemica* **108**, 243-257, doi:10.1016/j.acthis.2006.02.002 (2006).
20 Kijani, S. *et al.* Filter-Dense Multicolor Microscopy. *PLoS One* **10**, e0119499, doi:10.1371/journal.pone.0119499 (2015).
21 Baak, J. P., Langley, F. A., Talerman, A. & Delemarre, J. F. Interpathologist and intrapathologist disagreement in ovarian tumor grading and typing. *Anal Quant Cytol Histol* **8**, 354-357 (1986).
22 Caicedo, J. C. *et al.* Data-analysis strategies for image-based cell profiling. *Nature methods* **14**, 849-863, doi:10.1038/nmeth.4397 (2017).
23 Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100, doi:10.1186/gb-2006-7-10-r100 (2006).
24 Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. *Br J Exp Pathol* **10**, 226-236, doi:10.1590/S0042-96862001000800017 (1929).
25 Raper, K. THE DEVELOPMENT OF IMPROVED PENICILLIN-PRODUCING MOLDS. *Annals of the New York Academy of Sciences* **48**, doi:10.1111/j.1749-6632.1946.tb31753.x (1946).
26 Cragg, G. M. & Newman, D. J. Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta* **1830**, 3670-3695, doi:10.1016/j.bbagen.2013.02.008 (2013).
27 Chazotte, B. Labeling cytoskeletal F-actin with rhodamine phalloidin or fluorescein phalloidin for imaging. *Cold Spring Harb Protoc* **2010**, pdb prot4947, doi:10.1101/pdb.prot4947 (2010).
28 Cooper, M. D. & Alder, M. N. The evolution of adaptive immune systems. *Cell* **124**, 815-822, doi:10.1016/j.cell.2006.02.001 (2006).
29 Borghesi, L. & Milcarek, C. From B cell to plasma cell: regulation of V(D)J recombination and antibody secretion. *Immunologic research* **36**, 27-32, doi:10.1385/IR:36:1:27 (2006).
30 Nakazawa, M., Mukumoto, M. & Miyatake, K. Production and Purification of Polyclonal Antibodies. *Methods in molecular biology* **1474**, 49-59, doi:10.1007/978-1-4939-6352-2_3 (2016).
31 Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-497 (1975).
32 Brorson, K. & Jia, A. Y. Therapeutic monoclonal antibodies and consistent ends: terminal heterogeneity, detection, and impact on quality. *Curr Opin Biotechnol* **30**, 140-146, doi:10.1016/j.copbio.2014.06.012 (2014).
33 Wang, W., Singh, S., Zeng, D. L., King, K. & Nema, S. Antibody structure, instability, and formulation. *Journal of pharmaceutical sciences* **96**, 1-26, doi:10.1002/jps.20727 (2007).
34 Von Behring E, K. S. Ueber das Zustandekommen der Diphtherie-Immunitat and der Tetanus-Immunitat bei Thieren. *Dtsch Med Wochenschr* **16**, 1113-1114 (1890).

35 Coons, A., Creech, H., Jones, R. N. & Berliner, E. Demonstration of pneumoccocal antigen in tissues by use of fluorescent antibody. *J Immunol* **45**, 12 (1942).
36 Coons, A. Imunohistochemistry. *Proceedings of the Society for Experimental Biology and Medicine* **47**, 200-202 (1941).
37 Fritschy, J. M. & Härtig, W. Immunofluorescence. *eLS*, doi:10.1038/npg.els.0001174 (2001).
38 Avrameas, S. & Uriel, J. [Method of antigen and antibody labelling with enzymes and its immunodiffusion application]. *Comptes rendus hebdomadaires des seances de l'Academie des sciences. Serie D: Sciences naturelles* **262**, 2543-2545 (1966).
39 Nakane, P. K. & Pierce, G. B., Jr. Enzyme-labeled antibodies for the light and electron microscopic localization of tissue antigens. *The Journal of cell biology* **33**, 307-318 (1967).
40 Mason, D. Y. & Sammons, R. Rapid preparation of peroxidase: anti-peroxidase complexes for immunocytochemical use. *Journal of immunological methods* **20**, 317-324 (1978).
41 Planche, L. Note sur la sophistication de la résine de jalap et sur les moyens de la reconnaître. *Bull. Pharmacie* **2**, 578–580 (1810).
42 Veitch, N. C. Horseradish peroxidase: a modern view of a classic enzyme. *Phytochemistry* **65**, 249-259 (2004).
43 Engvall, E. & Perlmann, P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry* **8**, 871-874 (1971).
44 Van Weemen, B. K. & Schuurs, A. H. Immunoassay using antigen-enzyme conjugates. *FEBS Lett* **15**, 232-236 (1971).
45 Andersson, S. *et al.* Insufficient antibody validation challenges oestrogen receptor beta research. *Nature communications* **8**, 15840, doi:10.1038/ncomms15840 (2017).
46 Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol* **300**, C723-742, doi:10.1152/ajpcell.00462.2010 (2011).
47 Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* **59**, 223-239 (1962).
48 Morin, J. G. & Hastings, J. W. Energy transfer in a bioluminescent system. *J Cell Physiol* **77**, 313-318, doi:10.1002/jcp.1040770305 (1971).
49 Lippincott-Schwartz, J., Snapp, E. & Kenworthy, A. Studying protein dynamics in living cells. *Nat Rev Mol Cell Biol* **2**, 444-456, doi:10.1038/35073068 (2001).
50 Shimozono, S. & Miyawaki, A. Engineering FRET constructs using CFP and YFP. *Methods Cell Biol* **85**, 381-393, doi:10.1016/S0091-679X(08)85016-9 (2008).
51 Ullman, E. F., Schwarzberg, M. & Rubenstein, K. E. Fluorescent excitation transfer immunoassay. A general method for determination of antigens. *The Journal of biological chemistry* **251**, 4172-4178 (1976).
52 Pietraszewska-Bogiel, A. & Gadella, T. W. FRET microscopy: from principle to routine technology in cell biology. *J Microsc* **241**, 111-118, doi:10.1111/j.1365-2818.2010.03437.x (2011).
53 Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201-209 (1950).
54 Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids. *Nature* **171**, 737-738 (1953).
55 Hunter, C. A. Sequence-dependent DNA structure. The role of base stacking interactions. *J Mol Biol* **230**, 1025-1054, doi:10.1006/jmbi.1993.1217 (1993).

56 Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic acids research* **34**, 564-574, doi:10.1093/nar/gkj454 (2006).
57 Koehler, R. T. & Peyret, N. Thermodynamic properties of DNA sequences: characteristic values for the human genome. *Bioinformatics* **21**, 3333-3339, doi:10.1093/bioinformatics/bti530 (2005).
58 Rossetti, G. *et al.* The structural impact of DNA mismatches. *Nucleic acids research* **43**, 4309-4321, doi:10.1093/nar/gkv254 (2015).
59 SantaLucia, J., Jr. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu Rev Biophys Biomol Struct* **33**, 415-440, doi:10.1146/annurev.biophys.32.110601.141800 (2004).
60 Seeman, N. C. & Sleiman, H. F. DNA nanotechnology. *Nature Reviews Materials* **3**, 17068, doi:10.1038/natrevmats.2017.68 (2017).
61 Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302, doi:10.1038/nature04586 (2006).
62 Yurke, B., Turberfield, A. J., Mills, A. P., Simmel, F. C. & Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605-608, doi:10.1038/35020524 (2000).
63 Green, C. & Tibbetts, C. Reassociation rate limited displacement of DNA strands by branch migration. *Nucleic acids research* **9**, 1905-1918 (1981).
64 Lee, C., Davis, R. W. & Davidson, N. A physical study by electron microscopy of the terminally repetitious, circularly permuted DNA from the coliphage particles of Escherichia coli 15. *Journal of Molecular Biology* **48**, 1-22, doi:10.1016/0022-2836(70)90215-9 (1970).
65 Zhang, D. Y. & of the Society, W.-E. Control of DNA strand displacement kinetics using toehold exchange. *Journal of the American Chemical Society*, doi:10.1021/ja906987s (2009).
66 Yurke, B. & Mills, A. P. Using DNA to power nanostructures. *Genetic Programming and Evolvable Machines* (2003).
67 Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 15275-15278, doi:10.1073/pnas.0407024101 (2004).
68 Bi, S., Yue, S. & Zhang, S. Hybridization chain reaction: a versatile molecular tool for biosensing, bioimaging, and biomedicine. *Chemical Society reviews* **46**, 4281-4298, doi:10.1039/C7CS00055C (2017).
69 Huang, J., Wu, Y., Chen, Y., Zhu, Z. & International …, Y.-X. Pyrene-excimer probes based on the hybridization chain reaction for the detection of nucleic acids in complex biological fluids. *... Chemie International ...*, doi:10.1002/anie.201005375 (2011).
70 Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature biotechnology* **28**, doi:10.1038/nbt.1692 (2010).
71 Dirks, R. M. & Pierce, N. A. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *Journal of computational chemistry* **24**, 1664-1677, doi:10.1002/jcc.10296 (2003).
72 Dirks, R. M. & Pierce, N. A. An algorithm for computing nucleic acid base-pairing probabilities including pseudoknots. *Journal of computational chemistry* **25**, 1295-1304, doi:10.1002/jcc.20057 (2004).
73 Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E. & Pierce, N. A. Thermodynamic analysis of interacting nucleic acid strands. *Siam Review* **49**, 65-88 (2007).

74 Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of computational chemistry* **32**, 170-173, doi:10.1002/jcc.21596 (2011).
75 Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *Journal of the American Chemical Society* **139**, 3134-3144, doi:10.1021/jacs.6b12693 (2017).
76 Idili, A., Ricci, F. & Vallée-Bélisle, A. Determining the folding and binding free energy of DNA-based nanodevices and nanoswitches using urea titration curves. *Nucleic acids research* (2017).
77 Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* **355**, 564-566, doi:10.1038/355564a0 (1992).
78 Breaker, R. R. DNA enzymes. *Nature biotechnology* **15**, 427-431, doi:10.1038/nbt0597-427 (1997).
79 Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822, doi:10.1038/346818a0 (1990).
80 Li, T., Wang, E. & Dong, S. A grafting strategy for the design of improved G-quadruplex aptamers and high-activity DNAzymes. *PloS one* **4**, doi:10.1371/journal.pone.0005126 (2009).
81 Loenen, W. A. & Raleigh, E. A. The other face of restriction: modification-dependent enzymes. *Nucleic acids research* **42**, 56-69, doi:10.1093/nar/gkt747 (2014).
82 Pingoud, A., Wilson, G. G. & Wende, W. Type II restriction endonucleases--a historical perspective and more. *Nucleic acids research* **42**, 7489-7527, doi:10.1093/nar/gku447 (2014).
83 Zhu, Z., Samuelson, J. C., Zhou, J., Dore, A. & Xu, S. Y. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. *J Mol Biol* **337**, 573-583, doi:10.1016/j.jmb.2004.02.003 (2004).
84 Lindahl, T. Instability and decay of the primary structure of DNA. *Nature*, 709–715 (1993).
85 Daley, J. M., Zakaria, C. & in Mutation, R.-D. The endonuclease IV family of apurinic/apyrimidinic endonucleases. *Mutation Research/Reviews in Mutation ...* (2010).
86 Lindahl, T. & Biochemistry, A.-A. Rate of chain breakage at apurinic sites in double-stranded deoxyribonucleic acid. *Biochemistry*, doi:10.1021/bi00769a019 (1972).
87 Hosfield, D. J., Guan, Y., Haas, B. J. & Cunningham, R. P. Structure of the DNA repair enzyme endonuclease IV and its DNA complex: double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell* (1999).
88 Mullis, K. B. *et al.* (Google Patents, 1987).
89 Chien, A., Edgar, D. B. & Trela, J. M. Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus. *Journal of bacteriology* **127**, 1550-1557 (1976).
90 Pabinger, S., Rodiger, S., Kriegner, A., Vierlinger, K. & Weinhausel, A. A survey of tools for the analysis of quantitative PCR (qPCR) data. *Biomol Detect Quantif* **1**, 23-33, doi:10.1016/j.bdq.2014.08.002 (2014).
91 Sano, T., Smith, C. L. & Science, C.-C. R. Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* (1992).
92 Chang, L., Li, J. & Wang, L. Immuno-PCR: An ultrasensitive immunoassay for biomolecular detection. *Analytica chimica acta* **910**, 12-24 (2016).

93 Fredriksson, S. *et al.* Protein detection using proximity-dependent DNA ligation assays. *Nature biotechnology* **20**, 473-477, doi:10.1038/nbt0502-473 (2002).
94 Gullberg, M. *et al.* Cytokine detection by antibody-based proximity ligation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 8420-8424, doi:10.1073/pnas.0400552101 (2004).
95 Armstrong, J., Brown, R. S. & Tsugita, A. Primary structure and genetic organization of phage T4 DNA ligase. *Nucleic acids research*, doi:10.1093/nar/11.20.7145 (1983).
96 Dickson, K. S., Burns, C. M. & Richardson, J. P. Determination of the free-energy change for repair of a DNA phosphodiester bond. *Journal of Biological Chemistry*, doi:10.1074/jbc.M910044199 (2000).
97 Nilsson, M. *et al.* Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science (New York, N.Y.)* **265**, 2085-2088, doi:10.1126/science.7522346 (1994).
98 Larsson, C., Grundberg, I., Söderberg, O. & Nilsson, M. In situ detection and genotyping of individual mRNA molecules. *Nature methods* **7**, 395-397, doi:10.1038/nmeth.1448 (2010).
99 Gilbert, W. & Dressler, D. DNA replication: the rolling circle model. *Cold Spring Harb Symp Quant Biol* **33**, 473-484 (1968).
100 Liu, D., Daubendiek, S. L., Zillman, M. A., Ryan, K. & Kool, E. T. Rolling Circle DNA Synthesis: Small Circular Oligonucleotides as Efficient Templates for DNA Polymerases. *Journal of the American Chemical Society* **118**, 1587-1594, doi:10.1021/ja952786k (1996).
101 Fire, A. & Xu, S. Q. Rolling replication of short DNA circles. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 4641-4645 (1995).
102 Blanco, L. *et al.* Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *The Journal of biological chemistry* **264**, 8935-8940 (1989).
103 Lizardi, P. M. *et al.* Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nature genetics* **19**, 225-232, doi:10.1038/898 (1998).
104 de Vega, M., Lázaro, J. M., Salas, M. & Blanco, L. Mutational analysis of phi29 DNA polymerase residues acting as ssDNA ligands for 3'-5' exonucleolysis. *Journal of molecular biology* **279**, 807-822 (1998).
105 Larsson, C. *et al.* In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nature methods* **1**, 227-232, doi:10.1038/nmeth723 (2004).
106 Soderberg, O. *et al.* Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nature methods* **3**, 995-1000, doi:10.1038/nmeth947 (2006).
107 Abuduwali, N. *et al.* Autofluorescent characteristics of human periodontal ligament cells in vitro. *Ann Anat* **195**, 449-454, doi:10.1016/j.aanat.2013.03.007 (2013).
108 Yang, Y. & Honaramooz, A. Characterization and quenching of autofluorescence in piglet testis tissue and cells. *Anat Res Int* **2012**, 820120, doi:10.1155/2012/820120 (2012).
109 Clausson, C.-M. M. *et al.* Increasing the dynamic range of in situ PLA. *Nature methods* **8**, 892-893, doi:10.1038/nmeth.1743 (2011).

110 Jarvius, M. *et al.* In situ detection of phosphorylated platelet-derived growth factor receptor beta using a generalized proximity ligation method. *Molecular & cellular proteomics : MCP* **6**, 1500-1509, doi:10.1074/mcp.M700166-MCP200 (2007).
111 Koos, B. *et al.* Platelet-derived growth factor receptor expression and activation in choroid plexus tumors. *The American journal of pathology* **175**, 1631-1637, doi:10.2353/ajpath.2009.081022 (2009).
112 Conze, T. *et al.* MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas. *Glycobiology* **20**, 199-206, doi:10.1093/glycob/cwp161 (2010).
113 Liu, Y., Gu, J. & Cellular …, Å. Western blotting via proximity ligation for high performance protein analysis. *Molecular & Cellular ...*, doi:10.1074/mcp.O111.011031 (2011).
114 Weibrecht, I. *et al.* Visualising individual sequence-specific protein-DNA interactions in situ. *New biotechnology* **29**, 589-598, doi:10.1016/j.nbt.2011.08.002 (2012).
115 Frei, A. P. *et al.* Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nature methods* **13**, 269-275, doi:10.1038/nmeth.3742 (2016).
116 Burns, T. J. *et al.* High-throughput precision measurement of subcellular localization in single cells. *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **91**, 180-189, doi:10.1002/cyto.a.23054 (2017).
117 Leuchowius, K.-J. J. *et al.* Parallel visualization of multiple protein complexes in individual cells in tumor tissue. *Molecular & cellular proteomics : MCP* **12**, 1563-1571, doi:10.1074/mcp.O112.023374 (2013).
118 Weibrecht, I., Grundberg, I., Nilsson, M. & Söderberg, O. Simultaneous visualization of both signaling cascade activity and end-point gene expression in single cells. *PloS one* **6**, doi:10.1371/journal.pone.0020148 (2011).
119 Povilaitis, T., Alzbutas, G., Sukackaite, R., Siurkus, J. & Skirgaila, R. In vitro evolution of phi29 DNA polymerase using isothermal compartmentalized self replication technique. *Protein Eng Des Sel* **29**, 617-628, doi:10.1093/protein/gzw052 (2016).
120 Venkatesan, K. *et al.* An empirical framework for binary interactome mapping. *Nature methods* **6**, 83-90, doi:10.1038/nmeth.1280 (2009).
121 Stumpf, M. P. *et al.* Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6959-6964, doi:10.1073/pnas.0708078105 (2008).
122 Vidal, M., Cusick, M. E. & Barabási, A.-L. L. Interactome networks and human disease. *Cell* **144**, 986-998, doi:10.1016/j.cell.2011.02.016 (2011).
123 Smith, M. A. *et al.* Annotation of human cancers with EGFR signaling-associated protein complexes using proximity ligation assays. *Science signaling* **8**, doi:10.1126/scisignal.2005906 (2015).
124 Koos, B. *et al.* Proximity-dependent initiation of hybridization chain reaction. *Nature communications* **6**, 7294, doi:10.1038/ncomms8294 (2015).
125 Choi, H. M. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nature biotechnology* **28**, 1208-1212, doi:10.1038/nbt.1692 (2010).
126 Clausson, C. M. *et al.* Compaction of rolling circle amplification products increases signal integrity and signal-to-noise ratio. *Scientific reports* **5**, 12317, doi:10.1038/srep12317 (2015).
127 Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nature methods* **10**, 857-860, doi:10.1038/nmeth.2563 (2013).

128 Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090, doi:10.1126/science.aaa6090 (2015).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy 245

Editor: The Dean of the Faculty of Pharmacy

A doctoral dissertation from the Faculty of Pharmacy, 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 Pharmacy. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-340078

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2018