



UPPSALA  
UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations  
from the Faculty of Pharmacy 326*

# Advancing DNA-based proximity methods

MATTIAS LEINO



ACTA  
UNIVERSITATIS  
UPSALIENSIS  
UPPSALA  
2023

ISSN 1651-6192  
ISBN 978-91-513-1731-1  
URN urn:nbn:se:uu:diva-496889

Dissertation presented at Uppsala University to be publicly examined in A1:111a, BMC, Husargatan 3, Uppsala, Thursday, 20 April 2023 at 12:00 for the degree of Doctor of Philosophy (Faculty of Pharmacy). The examination will be conducted in English. Faculty examiner: Professor John Löfblom (Kungliga Tekniska Högskolan, Department of Protein Science).

### Abstract

Leino, M. 2023. Advancing DNA-based proximity methods. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Pharmacy* 326. 79 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-1731-1.

Cellular functions are governed by intricate chains of interactions between proteins. In order to properly understand cellular biology one must look into, not only protein function, but also its interacting network. Furthermore, due to the large heterogeneity between cells within cultures and tissue samples it is important to retain spatial information to enable investigations on a single cell level. In order to achieve this, *in situ* methods play a large importance in further elucidation of these interacting networks.

In order to investigate interactions between macromolecules such as proteins and nucleic acids, many outstanding methods have been developed. Some focusing on larger scale analysis, some on live cell imaging and some on detecting novel interactions. Our own group has focused on *in situ* methods utilizing DNA conjugated antibodies. DNA itself is a great macromolecule to work with, it can be produced synthetically, DNA hybridization is highly predictable and there is a large repertoire of DNA-modifying enzymes. This has been used in the development of methods such as proximity ligation assay (PLA) and Proximity-dependent initiation of hybridization chain reaction (ProxHCR).

Both methods utilize antibodies conjugated with DNA in order to detect proximity events between two proteins. PLA utilizes ligation to confirm proximity, while ProxHCR utilizes a chain of strand displacements to do the same. Both methods work well, but no method is beyond further optimization.

For PLA, a general concern lies in the formation of incorrectly interacting probes, resulting in incorrect ligations that yield linear fragments, incapable of producing visible signal. As a result PLA can produce a substantial amount of false negatives. To address this, we produced a similar method, Unfold, to streamline the probe interactions and ligations to improve efficiency.

For ProxHCR the original method required overly stringent reactions conditions to allow for efficient strand displacements and thus strong signal. Furthermore, signal strength was further compromised by oligonucleotide quality. To improve these issues, the ProxHCR method was completely redesigned and oligonucleotide quality along with signal strength was improved by further purification.

Both optimizations resulted in more efficient and versatile methods suitable for routine lab work and potential diagnostic use.

*Keywords:* in situ, proximity dependent initiation of hybridization chain reaction, hybridization chain reaction, Oligonucleotide design, proximity ligation assay, protein interaction, microscopy, purification

*Mattias Leino, Department of Pharmaceutical Biosciences, Box 591, Uppsala University, SE-75124 Uppsala, Sweden.*

© Mattias Leino 2023

ISSN 1651-6192

ISBN 978-91-513-1731-1

URN urn:nbn:se:uu:diva-496889 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-496889>)

# List of Papers

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

- I. Klaesson, A., Grannas, K., Ebai, T., Heldin, J., **Leino, M.**, Raykova, D., Oelrich, J., Arngården, L., Söderberg, O., Landegren, U. (2018) Improved efficiency of in situ protein analysis by proximity ligation using UnFold probes. *Scientific Reports*, 8(1):5400.
- II. **Leino, M.**, Heldin, J., Rubin Sander, M., Kerpatsou, D., Raykova, D., Koos, B., Söderberg, O. (2019) Optimization of proximity-dependent initiation of hybridization chain reaction for improved performance. *Molecular Systems Design & Engineering*. 4(5), 1058–1065.
- III. **Leino, M.**, Söderberg, O. (2023) Purification of DNA oligonucleotides to improve hybridization chain reaction performance. *In revision*.
- IV. **Leino, M.**, Söderberg, O. (2023) Purification of detection hairpins improves the performance of proximity-dependent initiation of hybridization chain reaction. *Manuscript*.

Reprints were made with permission from the respective publishers.

## Related Work by the author

### Original articles

- I. Francés-Soriano, L., **Leino, M.**, Dos Santos, M., Kovacs, D., Borbas, K., Söderberg, O., Hildebrandt, N., (2021) In Situ Rolling Circle Amplification Förster Resonance Energy. *Analytical Chemistry*. 93(3):1842-1850.
- II. Raykova, D.\*, Kerpatsou, D.\*, Malmqvist, T., Harrison, P.J., Sander, MR., Stiller, C., Heldin, J., **Leino, M.**, Ricardo, S., Klemm, A., David, L., Spjuth, O., Vemuri, K., Dimberg, A., Sundqvist, A., Norlin, M., Klaesson, A., Kampf, C. & Söderberg, O. (2022) A method for Boolean analysis of protein interactions at a molecular level. *Nature Communications*. 13(1):1–17.
- III. Heldin, J., Rubin Sander, M., **Leino, M.**, Lennartsson, J., Söderberg, O. (2019) Dynamamin inhibitors impair platelet-derived growth factor beta-receptor dimerization and signaling. *Experimental Cell Research*. 380(1):69-79.

# Contents

|                                                                                                                            |    |
|----------------------------------------------------------------------------------------------------------------------------|----|
| Introduction.....                                                                                                          | 9  |
| The microscope.....                                                                                                        | 10 |
| Microscopy – through the lens of time.....                                                                                 | 10 |
| Conventional light microscopy.....                                                                                         | 11 |
| Fluorescence microscopy.....                                                                                               | 11 |
| Ways to stain the cell.....                                                                                                | 13 |
| Antibodies.....                                                                                                            | 13 |
| Reporter dyes.....                                                                                                         | 15 |
| Fluorescent proteins.....                                                                                                  | 16 |
| DeoxyriboNucleic Acid (DNA).....                                                                                           | 16 |
| A brief introduction.....                                                                                                  | 16 |
| Physical properties of DNA.....                                                                                            | 18 |
| DNA hybridization.....                                                                                                     | 18 |
| Strand displacement.....                                                                                                   | 21 |
| Hybridization Chain Reaction.....                                                                                          | 22 |
| Synthetic production of oligonucleotides.....                                                                              | 25 |
| Study of protein-protein interactions.....                                                                                 | 30 |
| Conventional protein interaction assays.....                                                                               | 30 |
| DNA-based tools for proximity detection.....                                                                               | 33 |
| Drawbacks with DNA-based proximity methods.....                                                                            | 36 |
| Present investigation.....                                                                                                 | 39 |
| Preface.....                                                                                                               | 39 |
| Paper I – Improved efficiency of in situ protein analysis by proximity<br>ligation using UnFold probes.....                | 39 |
| Background.....                                                                                                            | 39 |
| Method and design.....                                                                                                     | 40 |
| Results and discussion.....                                                                                                | 41 |
| Paper II - Optimization of proximity-dependent initiation of<br>hybridization chain reaction for improved performance..... | 42 |
| Background.....                                                                                                            | 42 |
| Method and design.....                                                                                                     | 43 |
| Results and discussion.....                                                                                                | 43 |
| Paper III - Purification of DNA oligonucleotides to improve<br>hybridization chain reaction performance.....               | 44 |
| Background.....                                                                                                            | 44 |

|                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Design and method .....                                                                                                                          | 45 |
| Results and discussion .....                                                                                                                     | 46 |
| Paper IV - Purification of detection hairpins improves the performance<br>of proximity-dependent initiation of hybridization chain reaction..... | 47 |
| Background.....                                                                                                                                  | 47 |
| Design and method .....                                                                                                                          | 47 |
| Results and Discussion .....                                                                                                                     | 48 |
| Author's perspectives.....                                                                                                                       | 49 |
| On HCR in general .....                                                                                                                          | 49 |
| Designing a ProxHCR system.....                                                                                                                  | 52 |
| Projects that did not come to be .....                                                                                                           | 56 |
| Moving sequences .....                                                                                                                           | 56 |
| Multiplex ProxHCR.....                                                                                                                           | 57 |
| 3-plex ProxHCR .....                                                                                                                             | 58 |
| Mismatch hairpins .....                                                                                                                          | 59 |
| PLA/HCR hybrid.....                                                                                                                              | 60 |
| Locked hairpins .....                                                                                                                            | 61 |
| Non-conventional purification.....                                                                                                               | 61 |
| Conclusions.....                                                                                                                                 | 63 |
| Popular scientific summary.....                                                                                                                  | 64 |
| Populärvetenskaplig sammanfattning .....                                                                                                         | 66 |
| Acknowledgements.....                                                                                                                            | 68 |
| References.....                                                                                                                                  | 69 |

# Abbreviations

|           |                                             |
|-----------|---------------------------------------------|
| DCC       | Dicyclohexylcarbodiimide                    |
| E.coli    | Escherichia coli                            |
| ELISA     | Enzyme-linked immunosorbent assay           |
| IgA       | Immunoglobulin A                            |
| A         | Adenine                                     |
| AD        | Activating domain                           |
| AP        | Apurinic/aprimidinic                        |
| ATP       | Adenosine triphosphate                      |
| BCIP      | 5-Bromo- 4-Chloro-3-Indolyl Phosphate       |
| BD        | Binding domain                              |
| C         | Cytosine C                                  |
| C.Elegans | Caenorhabditis elegans                      |
| DAPI      | 4',6-diamidino-2-phenylindole               |
| DMT       | 4,4'-dimethoxytrityl                        |
| DNA       | Deoxyribonucleic acid                       |
| dsDNA     | Double stranded deoxyribonucleic acid       |
| eGFP      | Enhanced green fluorescent protein          |
| Endo IV   | Endonuclease IV                             |
| FISH      | Fluorescent in situ hybridization           |
| FRET      | Förster resonance energy transfer           |
| G         | Guanine                                     |
| GFP       | Green fluorescent protein                   |
| GRB2      | Growth factor receptor-bound protein 2      |
| H&E       | Hematoxylin and eosin                       |
| HB-HCR    | Hyper-branched Hybridization chain reaction |
| HCR       | Hybridization chain reaction                |
| HPLC      | High-performance liquid chromatography      |
| IgD       | Immunoglobulin D                            |
| IgE       | Immunoglobulin E                            |
| IgG       | Immunoglobulin G                            |
| IgM       | Immunoglobulin M                            |
| IHC       | Immunohistochemistry                        |
| LNA       | Locked nucleic acids                        |
| NA        | Numerical apparatus                         |
| NBT       | Nitroblue tetrazolium                       |

|                |                                                                |
|----------------|----------------------------------------------------------------|
| PAGE           | Polyacrylamide gel electrophoresis                             |
| PCA            | Protein fragment complementation assay                         |
| PCR            | Polymerase chain reaction                                      |
| PDGF-BB        | Platelet derived growth factor BB                              |
| PDGFR- $\beta$ | Platelet derived growth factor receptor $\beta$                |
| PI3K           | Phosphoinositide 3-kinase                                      |
| PLA            | Proximity ligation assay                                       |
| PNA            | Peptide nucleic acid                                           |
| ProxHCR        | Proximity-dependent initiation of hybridization chain reaction |
| PTM            | Post-translational modification                                |
| RCA            | Rolling circle amplification                                   |
| RNA            | Ribonucleic acid                                               |
| SANH           | Succinimidyl 6-hydrazinonicotinate acetone hydrazone           |
| SDS            | Sodium dodecyl sulfate                                         |
| SEC            | Size exclusion chromatography                                  |
| SIM            | Structured illumination microscopy                             |
| ssDNA          | Single stranded deoxyribonucleic acid                          |
| STED           | Stimulated emission depletion                                  |
| STORM          | Stochastic optical reconstruction microscopy                   |
| T              | Thymine                                                        |
| Ts-Cl          | 4-toluenesulfonyl chloride                                     |
| UNG            | Uracil-DNA glycosylase                                         |
| Y2H            | Yeast-two-hybrid                                               |

# Introduction

The main goal in life science is just as the name implies, to systematically study and pursue knowledge about life itself. Perhaps as a curiosity regarding how we, as biological machines work. Or perhaps to apply this knowledge in an effort to cure the many ailments we suffer. In a farfetched extrapolation, one might argue that a state without death and disease is the final application of life science following a complete understanding. But as we are all well aware, disease is still very much present around us. And truly understanding all our biological processes is an exceedingly difficult task as the human body is a remarkably complex machine.

A typical human body contains in the order of  $10^{13}$  cells<sup>1</sup>, around  $2 \times 10^4$  different proteins<sup>2</sup>, many with different splice variants transcribed from mRNA which in turn is transcribed from  $2 \times 10^4$  genes<sup>3</sup> contained in our 3 Giga base pair genome<sup>4,5</sup> packed in 46 chromosomes. In addition to all that, there is also epigenetic gene regulation and many forms of post-translational modifications (PTMs) that further increase the complexity. In the light of this, after years of study, or perhaps because of it, I still find it surprising that our bodies even work. Perhaps it is not too strange that small alterations in this machinery can have dire consequences, such as small viruses infecting and manipulating cells or mutations in the genome causing cancer.

With tremendous advancements in the field over the years, the complete human genome has been mapped<sup>6</sup>, a more or less complete library of our proteome has been recorded<sup>7</sup> and a database tracking the protein expression in various tissues and cell types has been established<sup>2,8,9</sup>. These advancements in genome, transcriptome and proteome analysis also give insights in the human machinery and its many problems. In a sense, we have arrived at a state where we have most of the puzzle pieces, and we can understand when a piece is broken, but we do not completely understand how they fit together.

As the proteins in our cells do not exist as independent entities, but rather as highly intricate networks regulated by non-covalent binding to small molecules, direct modifications or direct binding to other peptides or proteins<sup>10-13</sup>. These regulations will commonly alter the protein conformation, which in turn will modify the function. This leads to further interactions with other proteins and molecules, which propagates signals through other proteins, leading to a cascade event. These signals will eventually trigger very specific responses in cells such as survival, differentiation or even programmed cell

death. It is therefore understandable that alterations in any one of the proteins involved in a cascade can cause drastic changes in a cellular response.

Perhaps what we need now, after thorough investigations of the genome, transcriptome and proteome is a more complete understanding of an “interactome”. A clear map of all the interactions between proteins, and a further understanding of their regulations. There are currently many methods that look at the bindings and interactions between proteins, but any method will have both strengths and weaknesses. In this work, we will take a look at some of the already existing methods and I will present my own work which has been aimed at improving and developing new molecular tools in order to detect protein-protein proximity to infer interaction.

This thesis will mainly discuss two methods to detect and analyze protein proximity; UnFold<sup>14</sup>, and proximity-dependent initiation of hybridization chain reaction (ProxHCR)<sup>15</sup>. Both of these methods work on similar principles, methods that utilize antibodies covalently conjugated with oligonucleotides. These antibodies are then used for staining cells for *in situ* microscopy analysis to detect proximity events. UnFold is an optimized version of the enzyme based proximity ligation assay (PLA)<sup>16,17</sup> while ProxHCR<sup>18</sup> is an enzyme free alternative to PLA that we further improved.

## The microscope

In this work the main approach to analyzing samples has been through microscopy. While both UnFold and ProxHCR could be analyzed using plate readers, this would only yield a population based readout, missing the differences between individual cells. With microscopy one can analyze individual cells along with the minute details within the cell. As such, microscopy has been our method of choice for cell analysis.

### Microscopy – through the lens of time

Mankind has for a long time been able to utilize glass lenses for magnification. Already in the 13<sup>th</sup> century, lenses were used for eyeglasses and simple magnifying glasses, what one could call the early precursor to the modern microscope. It was not until the early 17<sup>th</sup> century that the first telescopes were invented using the combined magnification of several lenses to enhance magnification. Perhaps most notable was the astronomer Galileo who used a telescope to view small objects close up<sup>19,20</sup>. The improvements continued and lead to the first compound microscopes and eventually leading to the well-known high magnification microscope made by Antoine van Leeuwenhoek<sup>21</sup>. The use of microscopes also lead to publications such as Robert Hooks *Micrographia*<sup>22</sup> containing his microscopy observations of plant cells among other things, and Leeuwenhoeks publications of “animalcules”, blood cells,

spermatozoa and more were he, among other things, documented his findings of bacteria and protists<sup>23,24</sup>. These early findings may very well be some of the biggest breakthroughs in biology.

## Conventional light microscopy

Light microscopy essentially revolves around illuminating samples with light that either is reflected or passes through the sample. The light then passes through one or several lenses in order to magnify the view of the sample. Light microscopes are fairly inexpensive and are commonplace in most labs.

Naturally, when using microscopy the most important parameter is resolving power, the degree of which structures can be resolved from one another. Here, the resolution is primarily related to the quality of the objective lens, where superior objectives yields better resolution, depth of focus and contrast.

Another method of improving image quality is through enhancing contrast. The most common, and perhaps most simple technique is bright field microscopy. Here the sample is illuminated from below and viewed from above. The contrast from the sample is produced from the difference in light absorbance by the sample. In opposition, there is also dark field microscopy, here, instead of direct illumination, the sample is illuminated at an angle. When the sample is illuminated, diffracted light that is scattered hits the objective and is visualized while non-diffracted light is not picked up. Also notable is the phase contrast method, similar to dark field, the sample is illuminated at an angle, resulting in both scattered and unscattered light. Unlike dark field microscopy, the phase contrast microscopy utilizes the phase change of the scattered light to enhance contrast<sup>19,25</sup>.

## Fluorescence microscopy

It is however apparent that visual contrasting methods is not enough for analysis of samples, this led to the usage of several dyes to stain the samples for better contrast and to visualize subcellular structures and biomolecules. Especially important were the fluorescent dyes<sup>26-28</sup> that are often used in combination with antibodies<sup>29,30</sup>. Using these dyes you gain an almost unrivaled contrast, where the area of interest is brightly fluorescing while the background is black. To visualize the fluorescent light, the most basic method is an epifluorescence microscope<sup>31</sup>. In this type of microscope, the specimen is illuminated with a wavelength capable of exciting the reporter molecule. The reporter molecule will then, after a short interval, emit light of a longer wavelength. In practice, light from a light source is filtered through an excitation filter that removes all but the excitation wavelength. The light is then channeled through the objective where it illuminates the specimen. The reflected and emitted light is then transmitted through the objective to an emission filter that re-

moves all wavelengths except for the emission wavelength. As there are several fluorophores with varying excitation and emission wavelengths, a sample can contain several fluorophores at the same time. This is assuming that the correct excitation and emission filter settings are used.

Epifluorescence microscopes are what I would argue to be the edge of conventional microscopy, they are a staple in many labs and are quite affordable. Beyond typical epifluorescence microscopy there are several other more advanced techniques available. The most common advanced microscopy method is likely confocal microscopy<sup>32,33</sup>, which aims to improve optical resolution and contrast. In principal it works similarly to an epifluorescence microscope. However, instead of illuminating the entire sample, only a small part is excited. This is done by utilizing a pinhole close to the light source to focus the light to a small area of the specimen. Then a second pinhole is placed in the image plane in order to reduce other scattered light, further improving optical clarity.

Both conventional epifluorescence microscopy and confocal microscopy are excellent methods for visualizing and analyzing samples, but they are still limited by the diffraction limit<sup>31,34-37</sup>. This limit is dependent on the detected wavelength and the numerical aperture (NA) of the objective. The numerical aperture in turn is dependent on both refractive index of the imaging medium and the angle of the light cone. To put it simply, an air objective can at best achieve an NA of 1, while some objectives utilizing immersion oils can reach NAs up towards 1.51. The highest resolutions can be achieved at high NAs and lower wavelengths. In other words, the best resolution possible in typical epifluorescence microscopy with a UV-fluorophore would be around 150nm. The diffraction limit naturally imposes a limit when investigating protein interactions and co-localization within cells.

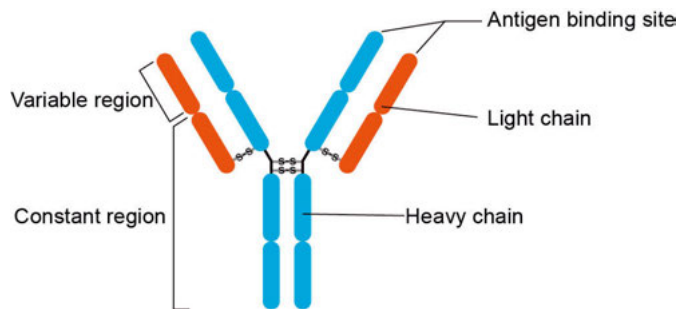
There are of course ways to move past the diffraction limit. Some of the notable methods to me are STED (stimulated emission depletion)<sup>38,39,40</sup>, SIM (structured illumination microscopy)<sup>41</sup> and STORM (Stochastic Optical Reconstruction Microscopy)<sup>42</sup>. They function in different ways but all yield super resolution images. STED, compared to normal confocal microscopy, uses two laser pulses. The two lasers are positioned so that only a tiny point is illuminated with an excitation laser while simultaneously a torus surrounding the point is illuminated with a depleting laser. The depleting laser will force the illuminated fluorophores to emit photons with a slight red-shift. As such, only the center most non-depleted fluorophores are able to emit light. As only a region smaller than the diffraction limit is effectively excited, resolutions down to 20nm can be inferred<sup>43</sup>.

In SIM, the sample is instead imaged several times under different light patterns. This will create a so-called Moiré effect containing spatial information. This effect will thereafter be translated with computational techniques to extract a super-resolution image. STORM assumes that if a low enough amount of fluorophores emit light, the resulting blur from one fluorophore can

be used to estimate its exact spatial origin. This is done by using a Gaussian approximation of the point spread function, resulting in an approximate position of the fluorophore. In practice, this is done by extracting multiple images through cycles of excitation and bleaching resulting in multiple images with only few fluorophores excited simultaneously. While the more advanced methods can result in very high-resolution images, they are often beyond what typical labs can afford unless solely focused on microscopy.

## Ways to stain the cell

Initially microscopes were used simply to provide a magnified image of the viewed sample, however, to further study tissues, cells, organelles and proteins of interest, ways to label the structures of interest were required. Perhaps the most well-known and widely used stain today is the hematoxylin and eosin (H&E) stain<sup>44-46</sup>, initially introduced by Wissowzsky in 1877<sup>47</sup>. In this stain it is not hematoxylin that is used, but rather the oxidized, colored form hematein. When used, hematein is mixed with a mordant to strengthen the positive charge and in doing so increase the binding between hematein and negatively charged complexes. As such, hematein will strongly bind to chromatin complexes and give a strong blue stain of the cell nuclei. Eosin on the other hand has a red/pink color and has an anionic charge that tends to bind to the proteins in the cytoplasm of cells. As the two stains complement each other, the stain enhances morphological structures. The H&E stain is still used by pathologists when analyzing biopsies. Other than the H&E stain there are many other dyes to mark other components of the cell, however, like the H&E stain these are unable to target specific proteins within the cell.



**Figure 1.** Schematic representation of an IgG antibody.

## Antibodies

The lack of specificity towards single protein targets eventually led to the implementation of immunohistochemistry (IHC) in 1941<sup>29</sup>. This procedure instead utilized antibodies that are able to effectively target a single protein of

interest. Antibodies are originally proteins that are a part of the adaptive immune system of jawed vertebrates<sup>48</sup>. In its normal context antibodies are secreted by B-cells as a response to foreign invaders. The secreted antibodies are intended to bind the foreign objects (e.g. bacteria and viruses) and render them less effective, aggregate the objects and eventually signal for phagocytosis or bacteriolysis<sup>49-51</sup>.

Antibodies are large proteins consisting of four separate protein fragments, two heavy- and two light- chains<sup>52</sup> (Figure 1). Depending on the type of heavy chain, antibodies can be divided into five isotypes, IgA, IgD, IgG, IgE and IgM. The heavy- and light- chains are in turn joined by disulfide bonds to a shape most commonly drawn as a Y. It is at the top of this shape where the antibody binds to antigens.

For antibody production, either the entire protein or a peptide sequence of the protein of interest is injected into an animal to induce an immune response<sup>53</sup>. The animal will naturally produce antibodies against the injected foreign body and the antibodies can be extracted from the host blood serum. These antibodies have affinity towards several epitopes of the protein of interest and are classified as polyclonal. Instead of directly extracting the host blood serum, the B-cells from the host animal can be extracted and fused with an immortalized cell, producing a hybridoma clone<sup>54</sup>. The fused cell can in turn produce antibodies indefinitely, these antibodies are capable of targeting a single epitope of the protein of interest, also called a monoclonal antibody.

The antibodies cannot be visualized through a microscope without some form of dye attached to it. With direct visualization, the antibody binding the protein of interest (primary antibody) is tagged with a dye and can be directly visualized. More commonly, dye tagged secondary antibodies targeting the primary antibodies are used instead<sup>55</sup>. These secondary antibodies typically target any antibody produced in a specific host animal and can reduce the cost of having to tag every primary antibody with a dye. Furthermore, the use of secondary antibodies amplifies the signal by allowing several secondary antibodies to bind to a single primary antibody.

As antibodies can be raised against almost any protein of interest along with quite a few PTMs, the method gives an unprecedented detail when analyzing cells. Subsequently, IHC has become a cornerstone in biological sciences and diagnostics.

While antibodies are used extensively in several applications, they are not without their flaws. As with any produced biomolecule, antibodies can suffer from batch-to-batch variations. Additionally, the specificity of antibodies should always be kept in mind<sup>56,57</sup>. When purchased, it is easy to assume that the antibody will be monospecific, however this will largely depend on antibody quality. Furthermore, even good quality antibodies will, at a high enough concentrations, bind unintended low affinity targets.

## Reporter dyes

As mentioned, to visualize antibodies in IHC a reporter dye is required. Usually these fall into one of two categories, enzymatic or fluorophore-based coloration. With enzymatic methods one typically uses either horseradish peroxidase or alkaline phosphatase<sup>58</sup>. Horseradish peroxidase will oxidize nearby substrates in the presence of hydrogen peroxide, usually yielding an insoluble, colored, fluorescent or chemiluminescent product that will accumulate close to the enzyme. Alkaline phosphatase on the other hand will remove phosphate groups from nearby substrates such as nitroblue tetrazolium (NBT) or 5-Bromo-4-Chloro-3-Indolyl Phosphate (BCIP) resulting in a colored, insoluble compound.

While the enzymatic approaches yield a large signal amplification, the most commonly used method for detection utilize fluorescent compounds known as fluorophores<sup>31</sup>. These compounds, when illuminated by photons can sometimes have an electron move to a different orbital and the molecule will enter an 'excited state'. Once in an excited state the molecule will eventually shed the absorbed energy by reverting to its 'ground state'. Typically, the ground state is denoted  $S_0$  and excited states are denoted  $S_1$ ,  $S_2$  etc. When shedding the energy in a fluorophore, the energy lost from jumping from a  $S_1$  vibrational state to a  $S_0$  vibrational state will be lost as a photon. The leaving photon will have a longer wavelength (lower energy) than what was initially absorbed. The shift between the excitation wavelength and emission wavelength is known as a Stokes shift. In short, fluorophores are excited over a span of wavelengths which in turn results in an emission of photons of a span of longer wavelengths. When used in microscopy the sample is illuminated with filtered light matched to the excitation wavelength. The resulting emission is then filtered and visualized in the ocular.

Other than dyes and fluorescently labeled antibodies, other fluorescent molecules that by themselves can target different biomolecules can be used. Perhaps the most common counterstains used to identify individual cells are DNA binding fluorophores such as Hoechst<sup>58</sup> and DAPI<sup>59,60</sup>. Both dyes are excited at UV wavelengths and will normally have poor emission; however when bound to the minor groove of DNA the emission wavelength is slightly shifted to around 460nm and the emission strength is increased. Both of these stains will yield a strong nuclear stain often used to distinguish individual cells.

Another interesting staining molecule used in fluorescence microscopy is phalloidin<sup>61,62</sup>, normally a toxin found in *Amanita phalloides* with strong affinity to F-actin. With this strong and specific affinity, when coupled to a fluorophore it can provide an antibody-free staining methods for the actin cytoskeleton.

## Fluorescent proteins

Finally, there is the protein that resulted in a Nobel Prize in 2008, GFP (green fluorescent protein). This protein is a naturally occurring protein found in jellyfish and acts similarly to a fluorophore. The protein was initially discovered in the 1960s<sup>63</sup> but did not find its use as a molecular tool until 1994 when it was expressed in *E.Coli* and *C.Elegans*<sup>64,65</sup>. Following the initial publication, several improvements were made to increase the spectral characteristics of GFP and eventually yielded the commonly used eGFP<sup>66,67</sup>. Fluorescent proteins have also been extracted from various sources other than jellyfish. With the many sources along with protein engineering these proteins can fluoresce with virtually any color<sup>68</sup>. As this is an entirely protein-based fluorophore it can be expressed in virtually any living organism, including mammalian cells. By coupling fluorescent proteins to a protein of interest one can follow the protein expression in live cells without addition of external probes. This technique has found many uses, including tracking protein expression in the developing embryo, tracking expression of inserted plasmids, detecting protein interactions, photobleaching experiments, promotor tracking etc.

## DeoxyriboNucleic Acid (DNA)

From discussions on microscopes and antibodies to DNA, a sudden jump in most bodies of work but not in this one. All of the proximity methods developed in our lab combines affinity reagents with DNA based molecular tools to infer proximity. But why focus on DNA? This biomolecule has several beneficial properties. It is easy to manipulate with a plethora of enzymes targeting DNA as a substrate. DNA hybridization is easy to understand and predictable, allowing minute manipulation of strand interactions. The oligonucleotide behavior is well studied and can be simulated with readily available resources. With these advantages DNA is a highly suitable and interesting molecule to base molecular tools on.

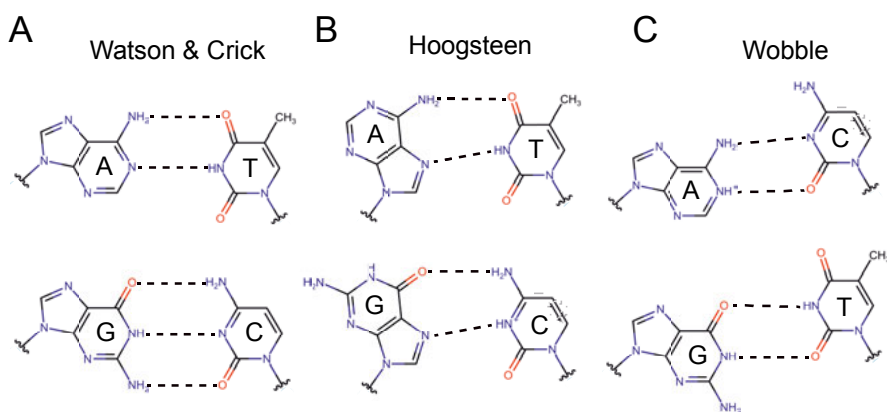
### A brief introduction

DNA, maybe one of the most important molecules known to life. The function of DNA in all organisms can most simply be explained as a way to store the cells blueprint. The information in DNA is stored as a code consisting of four nucleotides, or to be more precise, a phosphorylated deoxyribose coupled to one out of four nucleobases, Adenine (A), Guanine (G), Thymine (T) or Cytosine (C). Where A and G are purines and T and C are pyrimidines. Of course, one nucleotide cannot possibly store any information, so the nucleotides are arranged as a polymer by coupling the 5'- bound phosphate group of one nucleotide to the 3'-hydroxyl group of another. When arranged into a polymer

even a short sequence of 10 nucleotides can contain over a million combinations.

In nature, DNA is typically found as double stranded helices. Here, two separate DNA polymers come together and bind anti-parallel to each other. The two DNA strands typically bind in an orderly fashion by allowing two strands to bind via the Watson and Crick base pairing (Figure 2A)<sup>69</sup>. Nucleotides form a base pair using either two hydrogen bonds between Adenine and Thymine or three hydrogen bonds between Guanine and Cytosine. Due to the antiparallel binding of two complementary DNA strands along with the stacking interaction, the most favorable structure of DNA becomes a helix formation where the DNA twists around itself with about 10 bases per turn. This most naturally occurring double helix is commonly referred to as B-DNA. There are also other forms of DNA<sup>70</sup>, including the similar A-DNA which also contains a right-handed double helix but with a lower rotation per base pair. This form is usually found in solvent free environments. Another more extreme conformation is Z-DNA, which contrary to A- and B-DNA forms a left-handed turn instead, while much more uncommon, it can be induced with certain cations and with certain sequences.

Other than Watson and Crick base pairing, it should be noted that another more rare type of base pairing also exists naturally. This pairing variant was discovered by Karst Hoogsteen and is suitably named Hoogsteen base pairing (Figure 2B)<sup>71,72</sup>. In this variant, like Watson and Crick base pairing A-T and G-C are paired, however here the purines are flipped 180°. This results in two hydrogen bonds between both A-T and G-C. Although the genomic DNA is primarily paired with Watson and Crick pairing, data has suggested that Hoogsteen pairs can exist transiently throughout the genome<sup>73</sup>. Perhaps more interesting is that DNA through Hoogsteen base pairing can form DNA triplexes<sup>74</sup> and G-quadruplexes<sup>75</sup>.



**Figure 2.** Base pairing of DNA with A) Watson and Crick base pairing, B) Hoogsteen base pairing and C) Wobble base pairing.

## Physical properties of DNA

In our molecular methods the physical properties of DNA play an important role. As we utilize DNA hybridization to infer proximity, both length per base and flexibility for ssDNA and dsDNA are important. For dsDNA, length per base and persistence length are commonly quantified as 0.34nm and 50nm respectively. Although it should be noted that the persistence length can differ based on sequence. ssDNA on the other hand has not been as easy to quantify with persistence lengths being reported between 0.7nm and 6nm<sup>7677</sup>. The length per base also varies quite a bit with reports of both 0.475nm<sup>7879</sup> and 0.676nm<sup>80</sup>. In other words, dsDNA of about 150 base pairs or less can be assumed to barely bend at all by itself, while ssDNA is far more flexible.

Another interesting property of DNA is its anionic nature caused by the phosphate backbone. This will naturally cause some amount of repulsion between DNA strands. In solution, this might not matter as much, however this can impact kinetics when interacting near surfaces<sup>81</sup>. Here, a high concentration of DNA situated at the surface can reduce hybridization kinetics.

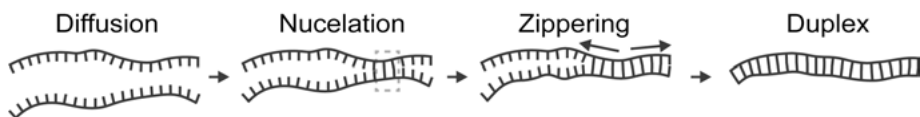
For methods such as PLA, the two probes are 32 nucleotides long and single stranded. In other words, two flexible chains with a length between 15.2nm and 21.6nm. However, when in a hybridized state, 10 bases remain single stranded and 22 become double stranded resulting in a decrease in length to 12.23-14.24nm. The resulting dsDNA also become far more rigid, as the length is far below the persistence length. Presumably, due to the rigid nature of DNA, one could argue that the method not only has a maximum distance, but also a minimum distance allowed.

## DNA hybridization

DNA finds its use in many biological applications, often due to the many DNA-interacting enzymes that allow easy but specific manipulation of the DNA strands. However, even without manipulating DNA with the use of enzymes, several methods like fluorescent *in situ* hybridization (FISH)<sup>82-84</sup>, hybridization chain reaction (HCR)<sup>85</sup> and DNA origami<sup>86</sup> relying solely on the hybridization mechanics exist as well. As previously stated, the DNA hybridizes by allowing A=T and G≡C base pairing. It is often assumed that hydrogen bonds drive the DNA-hybridization, however, it is the stacking interactions between neighboring nucleobases that is the major contributor to DNA-binding<sup>87,88</sup>. In fact, only the G≡C base pair contributes to the double strand stability while A=T pairing slightly destabilize the double strand. Instead, the hydrogen bonds act as a guide for how the DNA should hybridize, as the mismatched base pairs tend to destabilize double strand formation. It should however be noted that the purine-pyrimidine mismatches (A-C and G-T) can form

base pairs where the bases wobble to form two interbase hydrogen bonds (Figure 2C)<sup>89,90</sup>. While these are still destabilizing, these mismatches are not as poor as the homopairs (G-G, A-A, T-T and C-C).

DNA hybridization is usually described as a two-step process<sup>91,92</sup>. Initially, the two strands must diffuse together to form the initial binding between bases, often called nucleation (Figure 3). The initiation of DNA binding is viewed as thermodynamically unfavorable and the duplex formation is only favored after a few bases are bound<sup>93,94</sup>, likely due to the lack of initial stacking interactions. Furthermore, it has been argued that for initial nucleation, already existing intramolecular hybridization must be broken first<sup>91</sup>. After initial nucleation the DNA undergoes what is called ‘zippering’ where the remaining unbound bases bind together and forms the typical helical DNA structure<sup>91,92,95</sup>. This step on the other hand, is thermodynamically favorable, as each added nucleotide continues the stacking interaction within the double strand. As such, smaller double strands will be unstable due to the poor base stacking, and complex stability will increase in proportion to the complex length.



**Figure 3.** Process of DNA hybridization.

Many things govern both hybridization rates and stability of the hybridized DNA. Perhaps the most common method to manipulate DNA hybridization is temperature<sup>96,97</sup>. In fact, one of the ways of measuring double strand stability is by establishing a melting temperature, the temperature where half of the strands exist as double stranded duplexes and half are found as single stranded monomers. Double stranded DNA, like many other complexes will naturally destabilize at high temperatures. Altering temperature is a common procedure in many protocols to either facilitate a stronger binding at low temperatures, reduce non-specific bindings at moderate temperatures (FISH)<sup>83,98</sup> or directly force strand separation (Polymerase Chain Reaction, PCR)<sup>99</sup>. Temperature can also affect the hybridization rate of DNA monomers, as increased temperature will reduce the intramolecular interactions the nucleation rate will increase<sup>91</sup>.

Another factor to consider is monomer concentration, naturally, with a higher concentration of ssDNA an increased amount of interaction events take place. With an increasing amount of interaction events, more ssDNA could potentially hybridize. It is however important to note that it is only the on-rate ( $k_{on}$ ) with dimerization that is concentration dependent<sup>100,101</sup>. While the monomerization of dsDNA, off-rate ( $k_{off}$ ), is concentration independent.

Ions present in the solution can indirectly affect DNA hybridization by altering ionic strength of the solution, thus diminishing the repulsive forces between backbone phosphate groups resulting in a more stable binding<sup>96,102,103</sup>.

As such, the concentration of sodium ions in solutions is a very common way to manipulate DNA and alter its binding properties. Here, an increase in sodium concentration stabilizes the formed double strand. Ions can also directly bind to DNA to alter hybridization. Both commonly used, and found intracellularly is  $Mg^{2+}$  which binds directly to the grooves of dsDNA resulting in increased double strand stability<sup>104,105</sup>. Many other ions also have a stabilizing effect, including  $Mn^{2+}$ ,  $Ca^{2+}$ ,  $Cd^{2+}$  etc<sup>106-108</sup>. Interestingly some of these ions directly binds the nucleobases as well as the phosphate backbone.

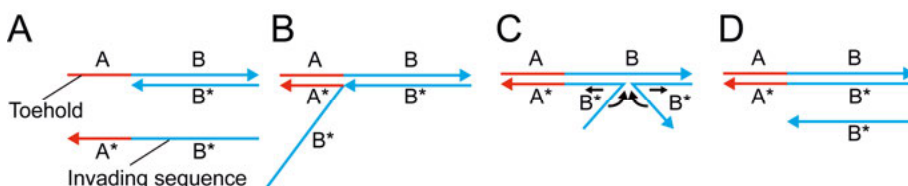
DNA hybridization can be further affected by pH alterations. Typically most DNA hybridization take place at physiological pH levels, however even in a cellular context pH can differ from 8 in the mitochondria to 4.7 in lysosomes. Impacts on DNA hybridization by pH stems from the protonation or deprotonation of the DNA bases. Under alkaline conditions both thymine and guanine can be deprotonated at pH levels above 10<sup>109-111</sup>, resulting in a disruption of both A-T and G-C bonds. Similarly, cytosine can be protonated around pH 4.5, which disrupts G-C bonds<sup>112</sup>. Although in either scenario both highly alkaline and highly acidic conditions can result in DNA degradation.

Other additives can also disrupt DNA hybridization. Perhaps the most commonly used are urea and formamide. Both molecules are capable of acting as hydrogen bond donors and acceptors. This allows both molecules to disrupt interstrand Watson-Crick hydrogen bonding, thus decreasing double strand stability<sup>113,114</sup>. Use of other solvents than water will also impact DNA hybridization and kinetics<sup>115,116</sup>. Organic solvents will typically be poor solvents for DNA, and as a result will cause direct precipitation of DNA. However, for glycerol, ethylene glycol, formamide, methanol, and DMSO, DNA remains soluble even at near anhydrous conditions. Furthermore, these conditions reduce duplex stability by disrupting DNA hydration. Interestingly, some solvents, such as ethanol, can instead improve hybridization kinetics. This is likely due to the aggregating effect that some solvents can have on cation bound DNA strands. This in turn can effectively lower the activation energy of duplex nucleation.

Other alterations have also been made to the DNA molecule itself in order to alter the stability of DNA duplexes. For example peptide nucleic acid (PNA)<sup>117,118</sup>, morpholino<sup>119</sup> and other backbone altering modifications have been tested and used to increase stability of dsDNA. Another approach has come in the form of LNA<sup>120</sup> (locked nucleic acids) where the 2' and 4' carbon on the ribose group is joined via oxygen, the increased rigidity from this modification causes an increased melting temperature. Computational models looking at these modifications along with similar alterations in backbone charge and chain rigidity found that in general an increased strand rigidity results in an increased stability. Similarly, a decreased backbone charge increased stability but also decreased hydrogen bonding time<sup>121</sup>.

## Strand displacement

There are several DNA based nanotechnologies and many of them rely on strand displacement. This is a process in which one strand in a DNA duplex is replaced by another<sup>122-126</sup>. In the simplest scenario, a single DNA strand can replace an identical strand in a DNA duplex. As both the bound and unbound strand contain the same sequence, the binding is thermodynamically equal. While the binding is equal, the exchange is still unlikely to happen, as it requires the binding strand to first partially or completely separate from the duplex. The common technique used to force a strand displacement is via the introduction of a toehold (or foothold) sequence (Figure 4A). This is a single stranded sequence found flanking the double stranded region, which serves as a nucleation site for an invading sequence (Figure 4B). After nucleation, the invading strand can displace the resident strand by pushing off nucleotides through a ‘random-walk’ process. In a double stranded DNA molecule, base pairs found along the edges of the complex, due to the lower stacking interaction, have a chance to unbind spontaneously. This mechanism is typically called fraying and will take place in any double stranded DNA. Typically, the unbound base or bases would rebind and the complete duplex would reform. However, when a nearby competing strand is present, either strand could rebind. As such, the process becomes a repeating process of fraying and binding from both strands. This results in a random process of progression and regression of displacement (Figure 4C). As the resident strand lacks a toehold, and is therefore thermodynamically unfavoured, the invading strand will eventually completely displace it (Figure 4D). One of the most significant factors in strand displacement is the strand length ( $n$ ), and the rate of displacement can be approximated to  $\frac{400}{n^2} \text{ s}^{-1}$ <sup>123</sup>.



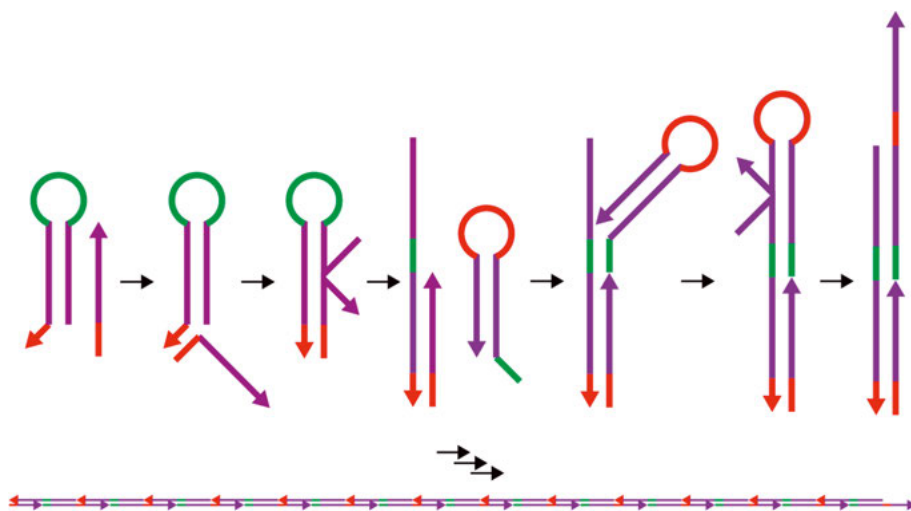
**Figure 4.** Strand displacement of a DNA duplex. A) DNA duplex with flanking toehold sequence. B) Nucleation of invading sequence to toehold. C) Random-walk process with repeated fraying and binding of initial and invading strand. D) Complete displacement of initial strand.

As both a driving force thermodynamically and a nucleation site for the invading DNA strand, the toehold play an important role in the dynamics of strand displacement. Generally speaking, a longer toehold will yield faster kinetics for strand displacement. Furthermore, it has been found that this increase in kinetics plateaus at around seven nucleotides<sup>123,127</sup>. This empirically found plateau will depend on both ion concentrations and temperature. The

toehold needs to allow for proper nucleation and enough binding to keep the invading strand in place for the initial random-walk process. A fringe scenario also emerges when the toehold is found inside a hairpin loop. In this scenario the invading DNA is further hindered due to the poor hybridization between the strand and the bent loop. Here, the kinetics will further depend on the loop size, where larger loops yields faster displacement<sup>127</sup>.

## Hybridization Chain Reaction

Among the different molecular tools that utilize strand displacement, hybridization chain reaction (HCR) has emerged as a novel amplification tool. HCR was initially developed in 2004<sup>85</sup> by Dirks *et.al.* where it initially was coupled to DNA-aptamers to allow for the detection of ATP in solutions. In general, HCR consists of two amplifying hairpins and a trigger sequence (often called initiator). The two hairpins contain three domains; a toehold, a stem and a loop. The trigger sequence is constructed so that it is reverse complementary to the toehold and stem of one of the hairpins. By typical strand displacement the trigger sequence can bind the hairpin and displace the stem (Figure 5). This forces the hairpin to open and expose both the loop and the previously displaced stem. These exposed sequences are, similar to the trigger sequence, reverse complementary to the toehold and stem of the other amplifying hairpin. Again, this will open the other paired hairpin, completing one amplification cycle. The reaction will continue and eventually form a nicked double stranded polymer. The length of the formed polymer typically depends on the amount of present trigger sequence. Lower concentrations of trigger sequence will result in longer polymerized product due to more hairpins being allocated to a single trigger sequence.



**Figure 5.** Schematic representation of HCR. The reaction is started by binding an initiator sequence to the foothold of hairpin<sub>1</sub>. Once bound the initiator will, eventually, displace the stem of hairpin<sub>1</sub>. When displaced, both the loop- and stem segments are exposed and can continue the reaction by acting as a second initiator. The reaction proceeds with the exposed initiator from hairpin<sub>1</sub> binding and opening hairpin<sub>2</sub>. After several cycles the reaction will yield a long nicked double strand. Colours indicate complementary sequences.

Construction of a self-assembling oligonucleotide system is not particularly difficult. The difficulty with HCR is to create two hairpins that do not polymerize until a trigger sequence is added. The design of these hairpins become more difficult and a balance between hybridization kinetics, degree of polymerization and stability is required. Generally, a long stem in the amplifying hairpins will yield greater stability at the cost of slower reactions. At the same time, shorter toeholds and loops also give better stability but provide a worse site for strand nucleation. The loop and toehold are also the driving force in pushing the polymerization, as pairing of these unpaired bases are thermodynamically beneficial. In a paper aimed at defining some definitive design criteria for metastable HCR hairpins a general outline was formed: Stems should be longer than the toeholds and contain a GC content  $\geq 60\%$ , while toeholds should be shorter than 12 nucleotides and contain  $\leq 40\%$  GC bases<sup>128</sup>.

The aforementioned systems yield a strong linear amplification that can result in polymers consisting of hundreds of monomers. To further improve the polymerization potential of HCR, groups have sought to create non-linear amplification systems. To name a few, Xuan *et al.*<sup>129</sup> designed a system where six amplifying oligonucleotides and a trigger form a hyper-branching system. This design did not rely on hairpins for the chain reaction but instead used dsDNA as amplifying oligonucleotides. Here, one of the strands acts as a blocking strand that stops polymerization unless a trigger sequence is added.

Naturally, the successive displacement of blocking strands is what allows the system to proceed with polymerization. As the method relies on dsDNA the method requires thorough purification and removal of non-blocked strands to avoid non-specific triggering. Another hyper-branched system coined HB-HCR (hyper-branched HCR) was also designed<sup>130</sup>. This system also contains six amplifying oligonucleotides. However, four of these oligonucleotides are instead designed as hairpins. Two of these hairpins are designed to contain two triggering sequences, so when the hairpin is opened it can continue the polymerization in two directions. The issue with the non-linear systems lie in their complexity. These systems tend to require both more and larger oligonucleotides. This increases the risk for non-specific triggering resulting in false positive signals. Due to the nature of exponentially growing systems, a single non-specific trigger can cause a large increase in false signal.

HCR as an amplification tool can be used in several applications. Perhaps the most common use for HCR lies in analysis of ribonucleic acid (RNA). Pierce's group, the original inventors of HCR, has done a lot of work on fluorescent in situ hybridization (FISH) where mRNA molecules within cells and tissues are investigated<sup>131-133</sup>. Here, mRNA binding sequences are added to the triggering sequence allowing an HCR polymer to be formed from individual mRNAs. In more simple approaches, the amplifying hairpins are directly triggered by the RNA molecules, allowing measurements of RNA concentrations in solution<sup>134</sup>. Furthermore, HCR can also be used as a tool amplify the signal from northern blots<sup>135</sup>. As an amplification method, HCR also finds its place in detection of proteins and other small molecules. The initial paper used DNA-aptamers that, when bound to ATP, opened and exposed a triggering sequence<sup>85</sup>. The trigger sequence has also been used with *in situ* microscopy, where it has been directly coupled to antibodies to further amplify the signal<sup>136</sup>. Furthermore, our group has used HCR in analyzing protein-protein proximity by introducing a trigger sequence to a proximity probe. This sequence is then only exposed when another proximity probe is bound nearby (detailed explanation in protein-protein interaction section)<sup>18</sup>. Another proximity detection method has also been developed, here instead of hiding the initiating sequence it was instead split in two<sup>137</sup>. The two split parts were then conjugated to separate probes, only when the two probes are close would the split initiators come together and trigger an HCR.

All of the aforementioned methods will result in long nicked double-stranded molecules, however these are not large enough to be detected by themselves. Therefore, most HCR methods further relies on different detection methods to visualize the product. The easiest way to go about detecting the polymerized product, which also is applied in most optimizations of hairpins, is the use of gel electrophoresis<sup>85,128</sup>. *In situ* applications and many of the RNA detection methods instead rely on fluorophore labeled amplification hairpins<sup>131-133,136</sup>. This will yield a large amount of fluorophore labeled oligonucleotides stuck to the trigger sequence, while the monomers are washed

away. Various other means to detect the HCR polymer have also been applied. One of the approaches has been to use Förster Resonance Energy Transfer (FRET) with a donor fluorophore on one hairpin and an acceptor fluorophore on the other<sup>128,138</sup>. This has been utilized in order to receive signal only when both hairpins are within proximity I.E. in a polymerized state. Perhaps one of the more interesting detection methods is the colorimetric use of gold nanoparticles, in these assays the induced polymerization cause a noticeable difference in color<sup>138</sup>. In this particular case, gold nanoparticles absorb onto single stranded DNA and remain non-aggregated. When the hairpins polymerize, the gold nanoparticles are free to aggregate with surrounding salt molecules, resulting in a shift in color from red to purple.

With the several applications and detection methods available for HCR, it may become a staple molecular technology for many other applications. However, more research and exploration is required for the technique to reach its full potential.

## Synthetic production of oligonucleotides

Since the discovery of DNA there has been a clear interest in synthetically producing oligonucleotides. Not only for the interest in producing synthetic genes but also for several other applications. Indeed, the use of DNA is neigh endless, from our own use of DNA-conjugated proximity probes to polymerase chain reaction primers, DNA-aptamers, molecular beacons, DNA-origami, siRNA etcetera.

### Historic improvements

In general, oligonucleotide synthesis is done through a stepwise addition of nucleotides through carefully directed coupling reactions. Over the years much work has been put into improving the synthesis method and several reaction routes have been tested. In order to improve synthesis yield the main point of optimization has been on improving the coupling reaction, typically between a 3'-reactive group and the 5'-hydroxyl group. Otherwise the yield can be improved by reducing unwanted side reactions, this can be achieved through addition of protective groups on phosphate groups, 3'- and 5'-ends, as well as the nucleobases themselves.

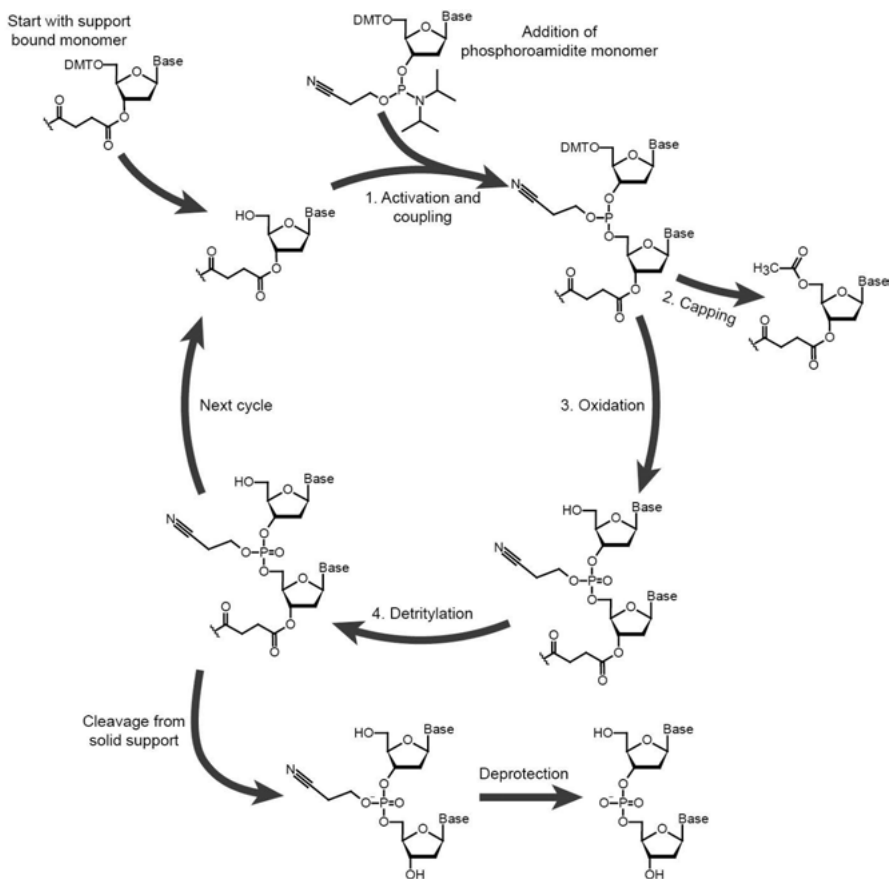
The first chemical synthesis of oligonucleotides was reported in 1955 by Michelson and Todd<sup>139</sup>. The synthesis route used here came to be known as the phosphotriester method due to the interlinking phosphate group formed. Here the coupling reaction is between a 3'-benzyl phosphorochloridate and a 5'-hydroxyl group. No particular blocking was used on the nucleobases, but hydroxyl groups were blocked with acetyl groups and phosphates were protected with benzyl groups.

The next notable synthesis route came a few years later<sup>140-142</sup>, where another group coupled nucleotides with a phosphodiester method instead. Like

the phosphotriester methods name, the diester naming refers to the internucleotide linkage formed during synthesis. The coupling reaction here directly takes place between a 3'-hydroxy group and a 5'-phosphate using a dicyclohexylcarbodiimide (DCC) or 4-toluenesulfonyl chloride (Ts-Cl) catalyst. This directly forms the naturally occurring phosphodiester bond, but the linkage remains unprotected. This synthesis route was eventually dropped, partially due to the unprotected linkage resulting in unwanted side reactions<sup>142</sup>. More notably the same group introduced a new hydroxyl group blocking strategy using 4,4'-dimethoxytrityl (DMT) that is still in use in current synthesis protocols<sup>143</sup>.

At this point synthesis mostly reverted to the phosphotriester method. However, here two noteworthy advancements were made. The first was solid-phase synthesis, initially used in peptide synthesis<sup>144,145</sup>. Use of solid phase synthesis would eventually come to be the staple of oligonucleotide synthesis. With the ability to add excess reactants followed by removal of said reactants allowed higher coupling efficiency and more automated processes. The second advancement was with the testing of several blocking group on the interlinkage phosphotriester bond<sup>142,145-149</sup>. It was here that the 2-cyanoethyl group was discovered, initially found not to be suitable for solution synthesis, the protective group is still in use for solid phase synthesis.

The final major advancement was in the improvement of the coupling reaction, where switching from P(V) to P(III) nucleosides allowed for far more efficient coupling<sup>142</sup>. These coupling reactions form a protected phosphite linkage as an intermediate, which is subsequently oxidized to produce a protected phosphotriester bond. This eventually led to the now well-known phosphoroamidite reactions that are currently in use.



**Figure 6.** The synthesis cycle for solid-phase phosphoroamidite nucleotides.

### Current coupling protocols

The aforementioned advancements have led to the currently used oligonucleotide synthesis protocol. The coupling reaction utilizes nucleosides with 5'-DMT protective groups coupled via 3'-phosphoroamidite groups. Reactions proceed in a stepwise addition of one nucleotide per synthesis cycle where each cycle typically is described as 4 steps<sup>150–153</sup> (Figure 6).

The first nucleoside is typically already attached to the solid support at the 3'-end and needs to be deprotected (detritylation of the DMT protective group) with trichloroacetic acid. This is followed by the first step of the synthesis cycle, where the coupling reaction takes place. Here, phosphoroamidite monomers are added, activated and coupled to the 5'-hydroxy group of the prior bases (step 1).

While the coupling efficiency is expected to be high, some unreacted nucleosides are bound to remain. To stop these unreacted nucleosides from further reacting, a capping reaction is added to avoid base deletions (step 2), here acetic anhydride and *N*-methylimidazole are added. The reaction mixture will

cause a rapid acetylation of the unreacted 5'-hydroxy groups rendering them inert.

At this stage the nucleosides are linked via a phosphite triester which would be unstable in the following detritylation step. Therefore, the phosphite is oxidized to produce a protected phosphotriester instead (step 3).

Finally, to complete the cycle, the protective DMT group from the newly added nucleoside is removed with a detritylation step (step 4). The next cycle can then be initiated by adding and activating the next phosphoramidite monomer. After the final nucleoside has been added, the synthesized oligonucleotides are cleaved from the solid support. Here, the polymers still contain protected phosphotriester bonds as well as protective groups on the nucleobases, both of which are removed in a final deprotection step.

### **Errors and purification**

Like many other biomolecules, DNA is susceptible to various forms of damage, including adduct formation, apurinic/apyrimidinic (AP) sites, strand breaks, and DNA cross-links<sup>154</sup>. Despite the improvements and optimizations in oligonucleotide synthesis, errors and forms of degradation during synthesis remain unavoidable. Coupling reactions are now typically reported with efficiencies above 98.5%<sup>155</sup>, but even at this rate synthesis of a 50-mer would result in around 50% complete oligonucleotides and the remainder with some form of truncation. Although the capping reaction is introduced to avoid deletions, this reaction can also not be expected to be 100% efficient. Other expected errors include non-complete deprotection<sup>153</sup> and depurination during deprotection<sup>156</sup>. As such, for most applications the synthesized oligonucleotides need to be further processed before they can be used.

Perhaps the most simple and somewhat crude method to purify the oligonucleotides is through gel filtration. In this method components are separated according to size<sup>157</sup>. With this method one can expect to separate the larger synthesized oligonucleotide from smaller molecular components and major truncations. However, any near full size truncations are expected to remain.

Perhaps the most commonly used purification strategy is reversed phase HPLC (RP-HPLC). Here the compounds are instead separated in order of hydrophobicity<sup>158</sup>, where more hydrophobic compounds are retarded in the column. Here, one can expect some truncations to be removed due to the differing hydrophobicity caused by oligonucleotide length<sup>159</sup>. Furthermore, any chemical alterations to the oligonucleotide whether they are wanted (E.g. Fluorophores or other modification) or unwanted (E.g. unremoved DMT groups) can alter the hydrophobicity resulting in a different retention.

One can also use anion-exchange HPLC<sup>160,161</sup> to purify synthesized oligonucleotides. Instead of separating the synthesized oligonucleotides according to hydrophobicity they are separated based on charge. Since DNA is essentially a very long polyanion it will be retained in the column. Truncations can

be removed due to a difference in net charge compared to the complete oligonucleotide. Likewise, some chemical modifications can be separated.

While more laborious, polyacrylamide gel electrophoresis (PAGE) can also be used to separate the oligonucleotides based on their mass-to-charge ratio<sup>162</sup>. For a large DNA polyanion this essentially results in a separation based on length. Based on the gel used and oligonucleotides purified this can result in very good separation from similarly sized truncations.

All of the aforementioned methods can remove truncations and modifications to a certain degree, but the extent of purity is also function of the oligonucleotide structure and length. In production, the amount of truncated product will increase exponentially with the oligonucleotide length, therefore, longer oligonucleotides will always carry a larger risk of producing truncations. To add to this problem, potential to remove erroneous oligonucleotides depends on how much the truncated oligonucleotide differs relative to the correct oligonucleotide. E.g. loss of a single base from a 10-mer results in a ~10% mass difference, while the same loss from a 100-mer is just 1%. Another problem arises from oligonucleotides with secondary structure. Regardless of method used, a folded hairpin structure will have different retention compared to linear oligonucleotide with the same amount of nucleotides.

There are different ways to address these problems. Secondary structure can be better resolved by denaturing the oligonucleotide during purification. For RP-HPLC one can increase temperature to reduce inter- and intramolecular interactions<sup>158</sup>. In anion-exchange HPLC, pH can be elevated to the point where guanine and thymine are deprotonated, thus losing the ability to hybridize. And in PAGE purification, one can introduce destabilizing agents such as urea and/or elevate temperature<sup>162</sup>. Issues of oligonucleotide length are naturally harder to improve. One can of course improve the coupling efficiency to reduce errors, but this does not completely resolve the problem. Here groups instead tend to look towards enzymatic approaches to complement the oligonucleotide synthesis. In older methods one would produce several smaller oligonucleotides and ligate them together to improve quality<sup>163-165</sup>. Other approaches involve polymerase chain assembly, where several shorter overlapping oligonucleotides form a larger complex and are elongated in a PCR like reaction<sup>165,166</sup>. Both of these methods essentially relies on a coupling of several shorter DNA segments.

Despite the unavoidable introduction of truncations and subsequent difficulties with purification, when tested, an unmodified 100-mer was examined and a purity over 90% was found for most companies and purification strategies<sup>155</sup>. Furthermore, it should be noted that most of the remaining truncations will be single base deletions which, depending on the method, could be inconsequential. However, one also has to take into account that any increase in complexity in the oligonucleotide could decrease purity, whether it be an increase in length, secondary structure or modifications.

## Study of protein-protein interactions

The affinity reagents mentioned earlier find excellent use in studying the expression and localization of proteins *in situ*. However, the focus of my research and by extension, this thesis, lie in the detection of protein-interactions. These interactions are essential in understanding how cells function. How an outside signal can be transferred from the membrane to the nucleus through a chain of protein-protein interactions, conformational changes and protein modifications.

### Conventional protein interaction assays

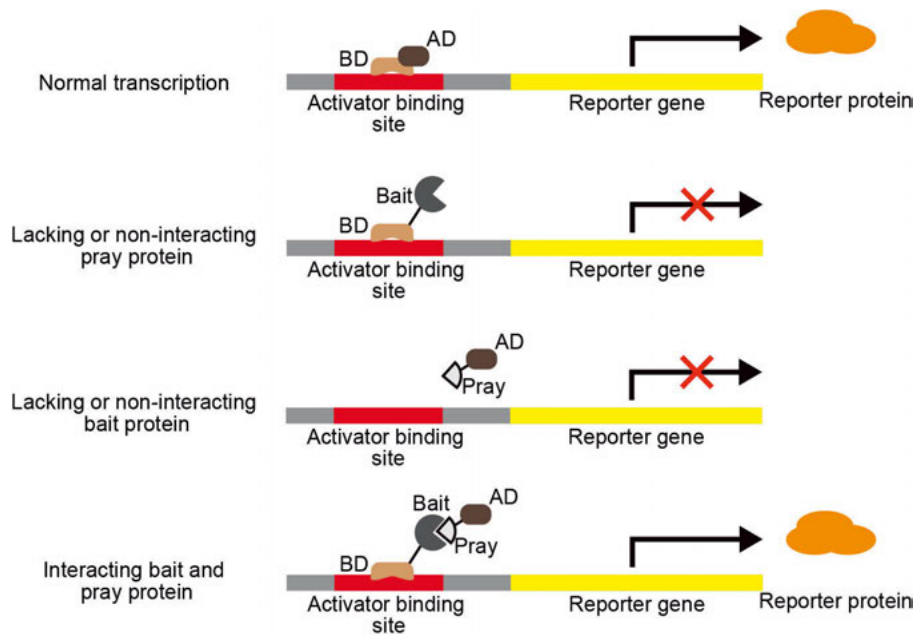
Typical *in vitro* methods for study of protein-interactions have been around for a while. These methods can either extract the proteome from a source of interest: a cultured cell line, a tissue biopsy etc. These methods include gel electrophoresis<sup>167,168</sup> with mass spectrometry<sup>169</sup>, co-immunoprecipitation<sup>170</sup>, pull down assays<sup>171</sup> and functional protein arrays/bead arrays<sup>172</sup>.

In gel electrophoresis the proteins are carefully extracted in order to retain protein complexes. Afterwards, the extract is separated along several gels like native-PAGEs, SDS-PAGEs and isoelectric focusing. The assumption is that proteins with affinity to each other will, with the correct experimental setup, retain their binding through the gel separation. The visualized proteins of interest can then be extracted and identified with mass spectrometry.

In co-immunoprecipitation one of the proteins of interest is targeted and extracted using an antibody. Typically, the antibody is bound to a solid support and mixed together with the lysed proteins. The method relies on both the targeted protein and other proteins with affinity to the targeted protein being extracted together and identified. Similarly, instead of immobilizing an antibody to a solid support, a bait protein can be used to directly bind the interacting proteins in a pull-down assay. Interactions can also be detected using protein array methods. Here, one or several of the target proteins are instead purified and immobilized to a surface. A lysate is then added to the array and any protein with affinity should bind to the immobilized target proteins. All of these methods are quite scalable and can detect multiple interactions in one assay. On the other hand, these methods can easily miss low affinity interactions with fast off-rates while simultaneously presenting abundant proteins non-specifically bound as false positive signals.

Another approach to detecting novel protein-interactions is to express the proteins of interests artificially and then afterwards detect any potential protein-interactions. These methods include the well-known yeast-two-hybrid (Y2H)<sup>173</sup> system (Figure 7) and protein fragment complementation assay (PCA)<sup>174-176</sup>. Of the aforementioned methods, most common is most likely the Y2H system. Initially published in 1989<sup>173</sup>, this method uses a bait and prey

type of process to detect protein-interactions. In the original method the transcriptional activator protein Gal4 which is involved in regulating galactose induced genes was split into two fragments, the activating domain (AD) and the binding domain (BD). These two fragments individually lack the ability to activate transcription, however, when both fragments come together they signal for downstream gene transcription. In the actual system, two fusion proteins, one consisting of AD and bait (one of the proteins of interest) and one consisting of as BD and prey (the other protein of interest) are created. When these are expressed in the yeast cell, if the bait and prey proteins have affinity for each other the AD and BD fragments will come together and a downstream reporter gene will be expressed, allowing for detection.



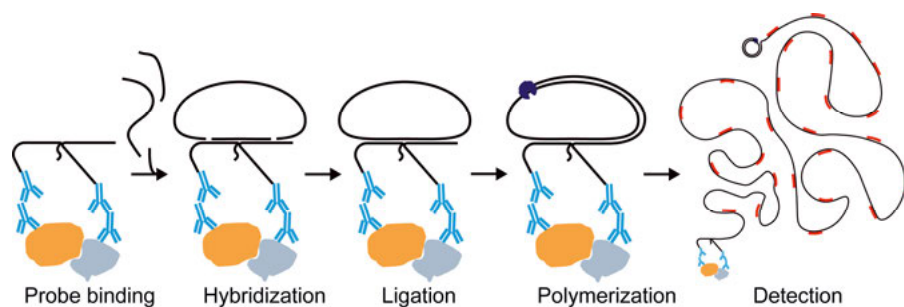
**Figure 7.** Schematic image representing the Yeast-two-hybrid assay. For normal transcription the binding domain (BD) and activator domain (AD) are attached and transcription proceeds as normal. When bait and prey proteins are not interacting the BD and AD will be separated and transcription cannot occur. When bait and prey proteins interact, both BD and AD will be within proximity allowing normal transcription of the reporter gene.

Similar to Y2H, PCA also relies on a similar bait and prey tactic where two separate fusion proteins are created. However, in this method the reporting system is different and can rely on virtually any protein as long as it can be split in two non-functioning halves and come together to form an active protein with a detectable readout. These reporter systems include Focal adhesion kinase<sup>174</sup>, horseradish peroxidase<sup>175</sup>, luciferase<sup>176</sup> etc. The reporter system can also be a split GFP or other fluorescent protein; however, these systems are

often referred to as Bimolecular Fluorescence Complementation Assay (BiFC)<sup>177</sup>. As these methods rely on bait and prey as well as target proteins being expressed in plasmids they can also be introduced in mammalian cells for *in-vitro* study. This allows for the study of protein-protein interaction in live cells. The main advantage of these methods tends to be the scalability and ease of use, as several protein-interactions can be investigated simultaneously. However, since the systems rely on fused protein expressed from plasmids the fusion itself could potentially block certain interactions yielding false negatives<sup>178</sup>. Simultaneously, over-expression from the plasmid can yield high protein concentrations where false positive signals appear.

Closer to what we do in our lab there are methods for studying protein-interactions in fixed samples. This also allows for the detection of interactions in cells and tissues, but instead of relying on genetically altered proteins, the interactions are instead detected through affinity reagents. The simplest approach to detecting interactions is to label both targets of interest with fluorophore-coupled antibodies. If the visualized signals are overlapping and therefore within proximity of each other, one can assume a protein interaction<sup>179,180</sup>. While simple, this method does have some problems: without a super-resolution microscope the exact position cannot be resolved beyond the diffraction limit and it can be difficult to detect low abundant signals.

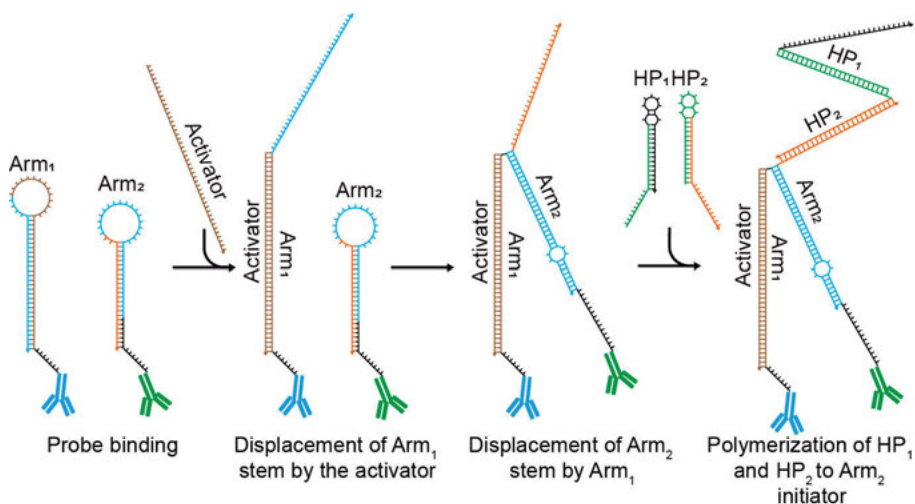
To overcome the first problem scientists have looked towards FRET to improve sensitivity<sup>181</sup>. FRET, relies on the energy transfer from a donor fluorophore to an acceptor fluorophore<sup>182</sup>. By then illuminating the sample with light that excites the donor fluorophore, some of the absorbed energy will be transferred to the acceptor fluorophore, causing it to enter an excited state. The excited acceptor fluorophore can then emit the absorbed energy as a photon when it reverts to its ground state. As such, any detected light emitted from the acceptor fluorophore can be assumed to be within close proximity of a donor fluorophore. The efficiency of the energy transfer decreases exponentially with distance between the fluorophores and therefore usually only occurs between 2-10nm.



**Figure 8.** Schematic representation of PLA. First the probes bind to the proteins of interest. Next, the two circularization oligonucleotides are added and hybridize to the probes. A ligase is added to fill the gaps between the two circularization oligonucleotides, using the probes as a ligation template. Finally, with the probe acting as a primer the circle can be amplified with RCA. To visualize the signal, fluorophore conjugated detection oligonucleotides are added and hybridize to the RCA product.

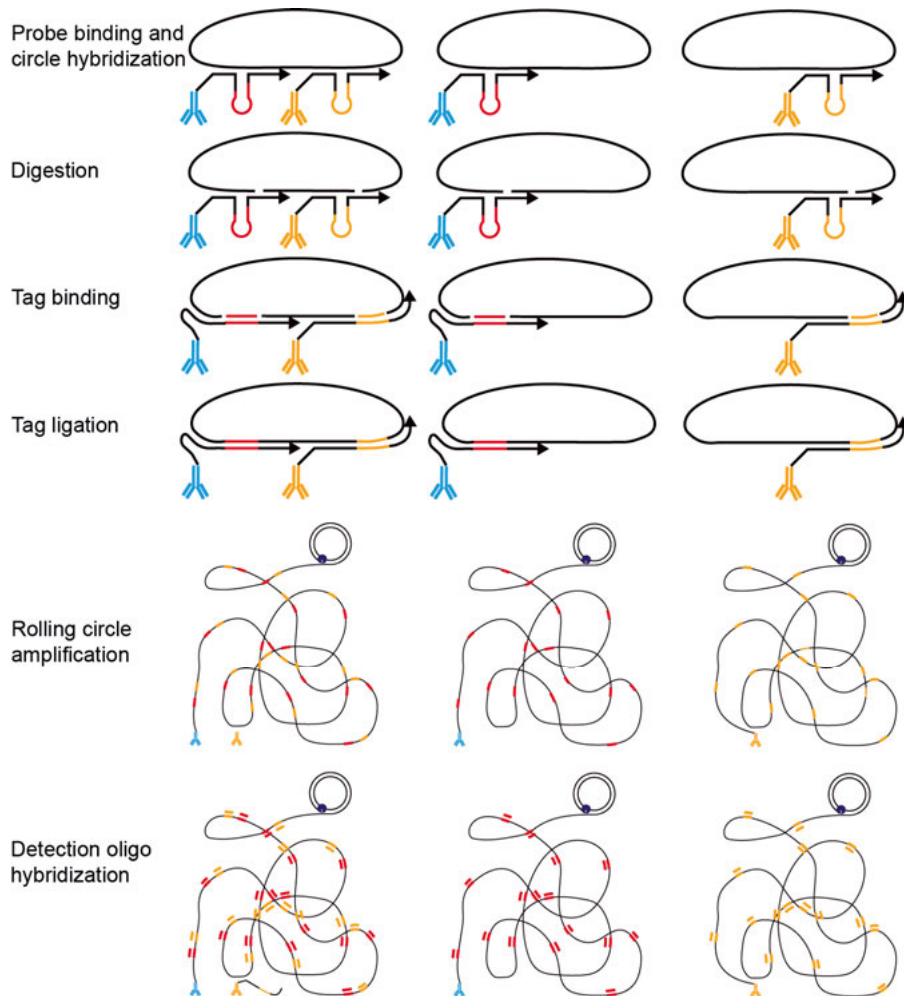
## DNA-based tools for proximity detection

To achieve both better signal and higher flexibility in allowed distances our lab has worked to develop DNA-based methods for proximity detection. Perhaps most notable is the Proximity ligation assay (PLA) (Figure 8)<sup>16,17</sup>. Originally, this method relied on DNA aptamers to bind protein targets in a solution-phase assay<sup>183</sup>. When the 5'-end of one aptamer and 3'-end of the other aptamer are sufficiently close they can, with the help of a third oligonucleotide, be ligated together. The ligated sequence would then be detected by amplifying the ligated sequence with PCR. To instead visualize these probes *in situ*, the method was altered<sup>16</sup>. Instead of aptamers, oligonucleotides were conjugated to antibodies. The conjugated oligonucleotides were also changed so that they would become ligation templates for two additional oligonucleotides. When the two conjugated arms are in proximity of each other, the external oligonucleotides can hybridize to the DNA-arms and be ligated to form a circle, which in turn can be amplified with rolling circle amplification (RCA). The amplified product is then visualized by adding detection oligonucleotides conjugated with fluorophores that can hybridize to the amplified product. As the amplified product is attached to the antibody, it can easily be detected using an epifluorescence microscope.



**Figure 9.** Schematic representation of ProxHCR. Interacting oligonucleotide sequences are color coded. First, ProxHCR probes bind to the target proteins. Next an activator oligonucleotide, when added, will bind and open arm<sub>1</sub>. Once opened, the displaced strand from arm<sub>1</sub> can interact and open arm<sub>2</sub>. This strand displacement exposes an HCR initiator. Once fluorophore conjugated detection hairpins are added, they can bind and polymerize from the exposed initiator sequence.

Another method developed for proximity detection *in situ* was Proximity-dependent initiation of hybridization chain reaction (ProxHCR) (Figure 9)<sup>18</sup>. The aim with this method was to both reduce cost and complexity of the PLA by removing enzymes. Instead of relying on enzymes for detection, the method uses a chain of strand displacements to infer proximity. Here, two antibodies are conjugated with two oligonucleotide hairpins. When an activator sequence complementary to one of the conjugated oligonucleotides is introduced, a strand displacement of said hairpin is initiated. The activator will eventually displace the stem of the first hairpin, instead forcing it to an open conformation. The now opened hairpin can, similarly to the activator, open the second hairpin. This releases a sequence that can initiate HCR (more details in the earlier HCR section). Both of these methods give strong amplified signals that are easy to detect and the distance requirement between targets is adjustable dependent on oligonucleotide length. Typically, the distance allowed between targets is around 40-50nm, it should however be noted that the majority of that length is covered by the antibody length. The distance between attachment points of the oligonucleotides, is however within FRET distance. Hence the distance requirement for these methods are similar to antibody-based FRET.



**Figure 10.** Schematic image of Molboolean assay. Initially, antibody probes bind to the proteins of interest. Next the circular probe is added and will hybridize to one or both of the targets. During the following digestion step the circular oligonucleotide is nicked where it has bound the oligonucleotide probes. If nicked, tags can invade the probe at the nicked site forming a nicked circle with either one or two tags. These are then ligated to fully incorporate them into the circle. Finally, the circle is amplified with RCA and any incorporated tag sequences are amplified as well. To visualize, two fluorophore coupled detection oligonucleotides, complementary to the tag sequences, are added and hybridize to the amplified product.

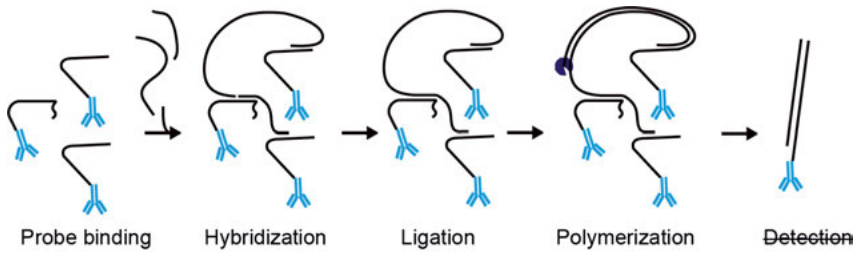
Another method that has recently been published is the MolBoolean method<sup>184</sup> (Figure 10). Like the other two methods, this method also relies on DNA-conjugated secondary antibodies to infer proximity. However, the method differs in that it aims to not only detect proximity events, but also simultaneously visualize single proteins. This allows one to infer information regarding ratio

of interacting to non-interacting proteins as well as translocation or degradation of non-interacting proteins. As a whole, the method is more similar to PLA rather than ProxHCR. Like PLA, the signal is amplified using RCA, where the amplified DNA later is tagged with fluorescent detection oligonucleotides. However, unlike PLA, the circle used in MolBoolean is already preformed before the assay. When the circle is added to the sample, it will bind either a single or two probes simultaneously. When bound, the circle can be opened through a nickase reaction, opening a gap in the circle. After nicking the circle, short tag sequences complementary to the proximity probes are added. If a circle has been bound and nicked, these short tags will invade and hybridize where the circle was nicked. After a following ligation step, the circle will have been barcoded with information revealing which probe it had bound. Finally the circle is amplified using RCA to produce a long DNA strand consisting of complementary repeats of the circle. As the inserted tags also are replicated, these can be used to bind fluorophore coupled detection oligonucleotides. As a result, each amplified product will contain either one of the two fluorophores or both.

## Drawbacks with DNA-based proximity methods

The methods described above are not without their limitations. It is often easy to forget what is actually being investigated in these proximity assays. As the methods above are just that, proximity assays, detecting whether two antibody probes are within close proximity of one another. And with that, comes all the inherent problems with such an assay. Like any other antibody based method, these methods are labile to non-specific antibody interactions. Which in these methods can result in false positives. The problem of antibody specificity is further increased when antibodies are conjugated with oligonucleotides. As we typically have used SANH (succinimidyl 6-hydrazinonicotinate acetone hydrazone) conjugation chemistry targeting primary amines, one can inadvertently target amines close to the antibody binding region resulting in altered affinity<sup>185,186</sup>.

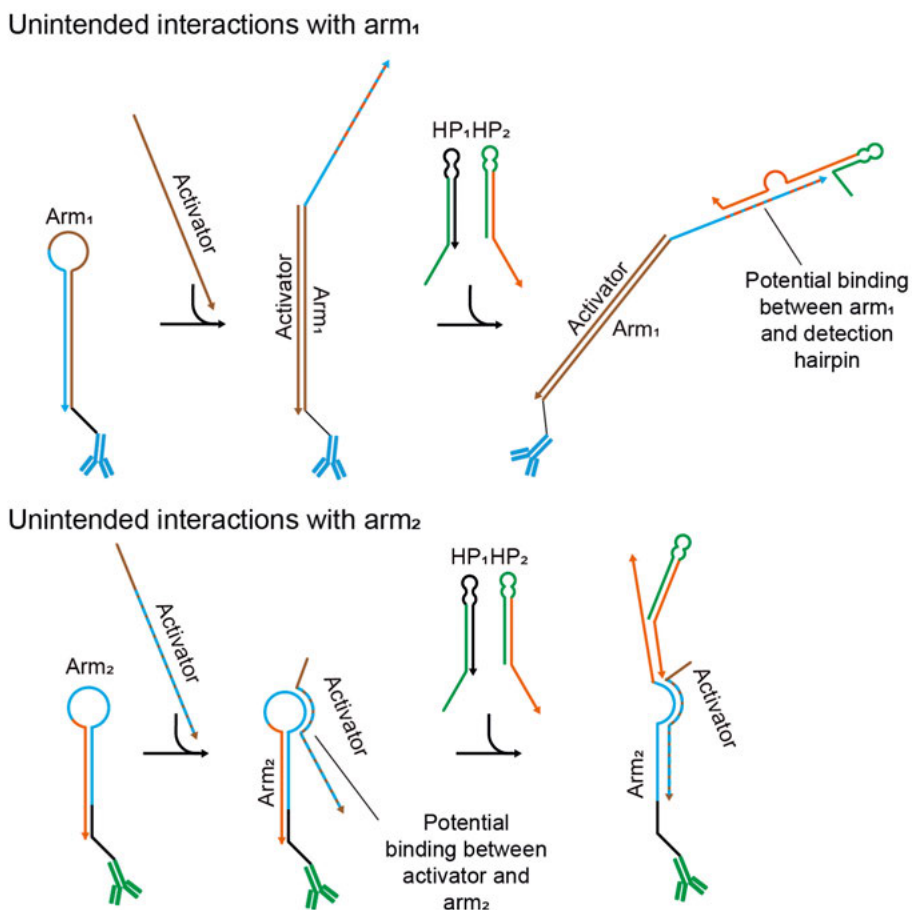
Although not a problem in and of itself, these proximity assays are often used to infer interaction between two proteins. Since one can only detect proximity, one could wrongly attribute coincidental proximity to interaction. This problem is easily understood when considering proximity between two highly expressed, non-interacting proteins present throughout the cell. Conversely, some increased proximity could be missed or underrepresented. As there is a maximal distance of 40-50nm, minor movements, from close proximity to interaction would show a similar amount of signal.



**Figure 11.** Linear fragment formation. Following probe binding, the circularization oligonucleotides binds the same probe<sub>1</sub>, but separate probe<sub>2</sub>. Following ligation, instead of a circular oligonucleotide, a linear oligonucleotide is formed. Once amplified with polymerase only a short double strand incapable of binding detection oligonucleotides is formed.

Any method involving an *in situ* ligation event such as PLA can underrepresent the visualized signals by producing ‘linear fragments’, ligation products that do not form complete circles. As PLA requires the ligation between two circularization oligonucleotides they both need to bind the same pair of probes. If one considers the two circularization oligonucleotides bound to one DNA-arm, both of these oligonucleotides need to bind the same paired probe. The likelihood of this happening becomes directly correlated to the amount of DNA-arms within proximity (Figure 11). Naturally, if the external oligonucleotides fail to form a circle no RCA can occur and no signal is detected. Indeed, this is one of the issues we aimed to reduce in paper I.

ProxHCR also has its own set of related issues. Both proximity arms contain sequences complementary to the detection oligonucleotides (Figure 12). This can give background noise from non-specific amplification from either of the conjugated oligonucleotides, especially with non-metastable detection hairpins. Another issue arises from a slight complementarity between the activator and arm<sub>2</sub>. Furthermore, as ProxHCR relies entirely on a series of strand displacements, visualized signal becomes correlated to any factor related to DNA hybridization, e.g. temperature, salt concentrations etc.



**Figure 12.** Unintended interactions with the ProxHCR method. Due to the nature of the design, arm<sub>1</sub> will contain some complementarity to the detection hairpins (upper panel). Similarly, a potential hybridization between the activator and arm<sub>2</sub> can occur, this could trigger non-specific amplification.

With the original ProxHCR method, experimental conditions were stringent with requirements of high salt concentrations and increased temperatures to achieve good signal, which easily could result in variation between experiments if any parameter was slightly off. These requirements were the target of our optimizations in paper II.

Additionally, as discussed in the oligonucleotide synthesis section, all synthetically produced oligonucleotides will carry some amounts of truncations. While this is also true for PLA, as we show in paper III and IV, this has a large impact on HCR polymerization.

# Present investigation

## Preface

As can be seen in the literature review, both microscopy and staining technologies have been improved over the years. Currently it is one of the most valuable tools for investigating and analyzing cells and tissues. Simultaneously, since the discovery of DNA, its structure and its base pairing, the knowledge of how DNA binds and interacts has only increased. This naturally has led to an increase in DNA computational modeling along with DNA-based technologies. Our lab has taken this opportunity to design molecular tools in the intersection of microscopy, staining techniques and DNA, capable of detecting protein-protein proximity. But as mentioned in the review, our methods still have some limitations that potentially could be improved. This has become the focus of my studies.

## Paper I – Improved efficiency of in situ protein analysis by proximity ligation using UnFold probes

### Background

Conventional PLA is an exceptional tool for investigation of both protein-protein proximity as well as PTMs. However, not every method is without its limitations. The main limitation with PLA lies in the circle ligation of the two external oligonucleotides (backpiece and splint). In this method, a complete circle has to be formed to enable the signal amplification through RCA. When the two external oligonucleotides are added, half of the backpiece and half of splint sequence should bind to one proximity oligonucleotide. The same backpiece and splint should then bend and bind to the second proximity oligonucleotide within proximity. Together the backpiece and splint should then form a nicked circle with the two proximity oligonucleotides as ligation templates. This circle can then be ligated to form a complete circle. However, with the very same protocol one can easily form incomplete circles that do not yield visible signal. As an example, two nearby proximity oligonucleotides occupied by separate backpiece or splint oligonucleotides cannot come together to form a circle. Simultaneously, if the backpiece and splint bound to the first

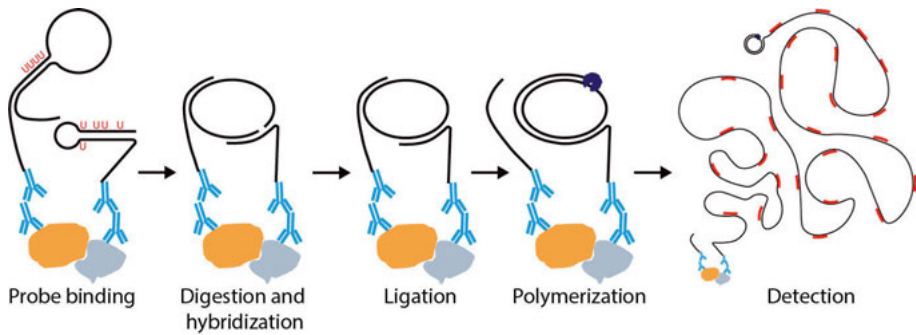
proximity oligonucleotide binds to two separate secondary proximity oligonucleotides it also leads to an incomplete circle. Due to these unwanted hybridizations the method yields less signal. In order to circumvent these problems, we aimed to develop a method with the circularization oligonucleotides already incorporated into the proximity probes. Instead of adding external oligonucleotides, these probes *UnFold* after target binding to form ligatable circularization oligonucleotides.

## Method and design

The UnFold probes (Figure 13), like PLA, consist of two proximity probes conjugated to antibodies. In the UnFold probes, the proximity oligonucleotides are separated into a ligation template probe and a circularization oligonucleotide probe. The circularization probe is constructed as a hairpin with the 5'-end conjugated to the antibody. The circularization sequence, that is to be ligated into a circle for RCA is found between the 5'-end of the loop until the 3'-end of the oligonucleotide.

In order to form a circle, the circularization probe is cut at the start of the circularization sequence, resulting in a free 3'-end and 5'-end that can be ligated together. The template probe carries a complementary sequence for both the 3'-end and 5'-end of the circularization sequence, allowing the binding of both ends to form a nicked circle. However, to prevent probe-to-probe interaction before target binding the template strand is hidden within a hairpin. As such, the template probe also requires the sequence to be cut in order to expose the template strand.

In order to cut these probes, we utilized the enzymes uracil-DNA glycosylase (UNG) and endonuclease IV (Endo IV). UNG can specifically target uracil bases and will cleave the N-glycosidic link between the deoxyribose and the uracil leaving an abasic site. Endo IV on the other hand can specifically target abasic sites and catalyze the hydrolysis of the DNA phosphodiester bond. The combination of these two enzymes will therefore cleave any sequence containing uracil bases. Therefore, in this method the circularization probe is designed with uracil-bases placed before the circularization sequence and the template strand has uracil-bases spread out within the hairpin. When the enzymes are added, the uracil-sites will be cut and allow the circularization sequence to bind to the template strand, the strands can then be ligated using a ligase to form a complete circle. Next, the circle is amplified with RCA using the Phi29 polymerase to form an amplified DNA bundle. Similarly to PLA, the DNA bundle is then visualized by adding complementary fluorophore containing oligonucleotides.



**Figure 13.** Schematic picture of UnFold. First, the antibody probes bind their intended targets. Next, the uracil residues are digested with the help of UNG and EndoIV which releases the circularization oligonucleotide. Once the circularization oligonucleotide unfolds, it can bind to a nearby template strand to form a circle with a single nick. By adding a ligase, the nick is mended and a complete circle is formed. This circle is then amplified using RCA and is later detected using fluorophore coupled detection oligonucleotides.

## Results and discussion

To evaluate the new design, it was directly compared to conventional PLA in cultured cells, tissue slides and solid phase measurements. When investigating the interaction between E-cadherin and  $\beta$ -catenin in both tissue slides and cultured cells the UnFold probes showed a significantly stronger staining with more amplified products per cell. Interestingly, the optimal concentration for the UnFold probes was far lower than that of the conventional PLA. To further test the UnFold design, three phosphorylation events were investigated. BJ-hTERT cells were stimulated with platelet derived growth factor BB (PDGF-BB), which causes the phosphorylation of the PDGF receptor along with the downstream effectors AKT and ERK1/2. These events were analyzed with both conventional PLA and Unfold in both stimulated and unstimulated cells. With all three targets, the new UnFold-design was able to detect more phosphorylation events compared to PLA. As both PLA and Unfold have an extraordinary signal amplification via RCA, both methods were also compared to enzyme-linked immunosorbent assay (ELISA) in detecting IL-6 in a solid support assay. In this assay both PLA and Unfold displayed a higher sensitivity compared to ELISA. Furthermore, Unfold provided a lower limit of detection compared to both PLA and ELISA.

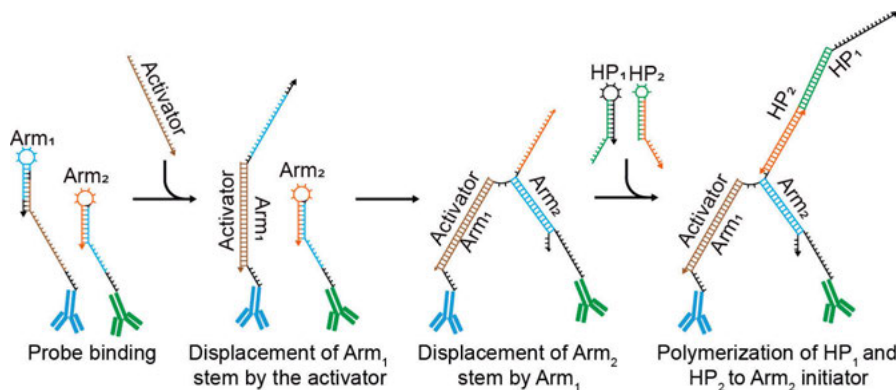
The results clearly indicate that the new UnFold design can detect more proximity events compared to conventional PLA without sacrificing signal strength or background noise. The increased signal is most likely due to the reduced amounts of ligation events required, resulting in fewer faulty ligations. As the circularization oligonucleotide is contained within the proximity probe one gains a high local concentration, resulting in higher likelihood of hybridization between the circularization oligonucleotide and the ligation

template. As the local concentration and the ratio of circularization oligonucleotide to proximity probe is 1:1 there is also less risk of separate circularization oligonucleotides binding the same proximity probe. Put together this resulted in a method that require less probes but yields both higher signal and better signal-to-noise.

## Paper II - Optimization of proximity-dependent initiation of hybridization chain reaction for improved performance

### Background

The original ProxHCR system was initially designed as an enzyme free PLA. Removal of enzymes should allow the method to be more cost efficient, simpler to use and less stringent in terms of experimental conditions. While the method indeed did remove the most costly component of the PLA, enzymes, it still had a similar number of steps involved. Most notably the reaction conditions for ProxHCR were stringent. The method itself requires both high sodium chloride concentrations as well as elevated temperature to yield signal. At the same time, any alterations of these parameters would introduce a variability in the signal. The aim of this paper was to redesign the ProxHCR system to allow for more consistent performance over varying experimental conditions.



**Figure 14.** Schematic picture of the optimized ProxHCR design. Interacting sequences are color coded. Initially the two antibody probes bind to the proteins of interest. To start the strand displacement reaction, an activator complementary to the foothold and stem of arm<sub>1</sub> is added. This activator will open arm<sub>1</sub> and expose the loop and stem. These exposed sequences can in turn bind and open arm<sub>2</sub>, exposing an HCR initiator sequence. This initiator sequence can then start the HCR.

## Method and design

Much like our other proximity-based systems the ProxHCR system relies on two separate probes for proximity detection (Figure 14). These two probes are conjugated with the two proximity hairpins denoted arm<sub>1</sub> and arm<sub>2</sub>. In the original ProxHCR, arm<sub>1</sub> is opened by an externally added activator sequence that binds from the arm<sub>1</sub> hairpin loop and through strand-displacement opens the stem. When the stem from arm<sub>1</sub> is exposed, it can hybridize to the arm<sub>2</sub> hairpin loop and similarly to the activator displace its stem. The exposed region of the arm<sub>2</sub> stem contains the initiator sequence for the HCR. When detection hairpins are added they can bind and polymerize from this initiation sequence to form long nicked double stranded DNA. As each detection hairpin contains a fluorophore, the polymerized product can be visualized.

Based on the initial data from the paper and prior knowledge, we hypothesized that the reduced signal was caused by poor opening of arm<sub>1</sub> and arm<sub>2</sub>. Most likely, the increased salt concentration assisted the nucleation of the strand hybridization and the elevated temperature improved strand displacement. In order to improve the performance of the system we aimed to improve the complete opening of arm<sub>1</sub> and arm<sub>2</sub>. Earlier reports have shown that strand displacement kinetics can be improved both by the placement of the toehold<sup>127</sup> and the amount of displaced nucleotides<sup>123</sup>. As such, we redesigned the ProxHCR system with two main objectives, (1) moving the toehold placements (where the activator binds to arm<sub>1</sub> and where arm<sub>1</sub> binds to arm<sub>2</sub>) from the loop of the hairpin to an external toehold and (2) reducing the number of nucleotides that requires displacement. The resulting system is opened from an external toehold and has 11 bases to displace in arm<sub>1</sub> (down from 30) and 9 bases from arm<sub>2</sub> (down from 24).

## Results and discussion

Initially the parameters required for the previously published ProxHCR system was investigated by detecting the interaction between E-cadherin and  $\beta$ -catenin in formaldehyde fixed HaCat cells. As we had hypothesized, the results indicated notably poor signal when either temperature or salt concentration was lowered during incubation with the activation oligonucleotide. The amplification with the detection hairpins was also slightly worse with low salt reaction conditions. This indicated that the opening of both arm<sub>1</sub> and arm<sub>2</sub> was the main issue.

To test the newly designed system we used an *in solution* approach to analyze how the different oligonucleotides would interact. Here, the oligonucleotides were mixed in a solution and allowed to react for 1h. The resulting hybridizations were then analyzed using gel electrophoresis. The analysis showed no interaction between arm<sub>1</sub> and arm<sub>2</sub> without the addition of the activator, and clear polymerization of the detection hairpins when arm<sub>1</sub>, arm<sub>2</sub>,

activator and detection hairpins were mixed together. Perhaps most importantly, no non-specific polymerization appeared from arm<sub>1</sub> or arm<sub>2</sub> alone.

When comparing the previous ProxHCR system with the new design, the new design showed an improved signal strength in all experimental conditions. While there were some smaller differences in staining between the different conditions, it still produced consistent stains. A potential problem with all proximity-based assays is non-specific background caused by one of the probes alone. Either arm<sub>2</sub> could non-specifically open to initiate HCR or the region in arm<sub>1</sub> that opens arm<sub>2</sub> could potentially bind to the detection hairpins. Investigations of the non-specific signals by only staining with a single primary antibody showed only negligible signal. Phosphorylation of platelet derived growth factor receptor  $\beta$  (PDGFR- $\beta$ ) and Akt was also visualized in PDGF-BB ligand stimulated and unstimulated cells to further show that the method could detect protein-PTM proximities.

Put together, the modified design of ProxHCR gave both more consistent and stronger stain compared to the previous design. As the new system can be used in room temperature and normal salt concentrations it is both easier to use and more likely to find use in other applications. The results also indicate that the previous design may have had a problem with incomplete opening of the proximity oligonucleotides rather than poor polymerization. This would also mean that the new design gives fewer false negatives while not increasing the amount of false positive stains.

## Paper III - Purification of DNA oligonucleotides to improve hybridization chain reaction performance

### Background

While this paper does not contain a single experiment with our ProxHCR method, it is still undoubtedly related. The overall goal with ProxHCR has been to produce a scalable, inexpensive method that potentially could be used in automated staining procedures. The previous optimization was aimed at reducing variability due to minor temperature changes along with dependence on higher salt conditions. However, we were not quite satisfied with the attained signal strength. With the sequences and protocol used we could detect various protein-protein interactions, however signal-to-noise for less abundant reactions was less than optimal, making proper quantification difficult.

In order to improve signal strength, we sought to improve HCR polymerization by improving oligonucleotide purity. As discussed in this thesis, the introduction of truncations during oligonucleotide synthesis is well known. Subsequent purification can reduce the amount of erroneous oligonucleotides to some extent, but some amount of truncated hairpins will remain. As HCR

relies on a linear polymerization, each individual hairpin must be able to bind and displace the following hairpin to propagate the reaction. Introduction of deletions or substitutions in these strand displacement reactions could potentially retard or completely hinder continued propagation. Yet little has been reported with regards to batch-to-batch variations along with the effect of further purifications. Despite this, most papers utilizing HCR seemingly only use a single batch of HPLC purified hairpins.

In this paper, we sought to investigate how continued and improved purification of HCR hairpins impact the degree of polymerization.

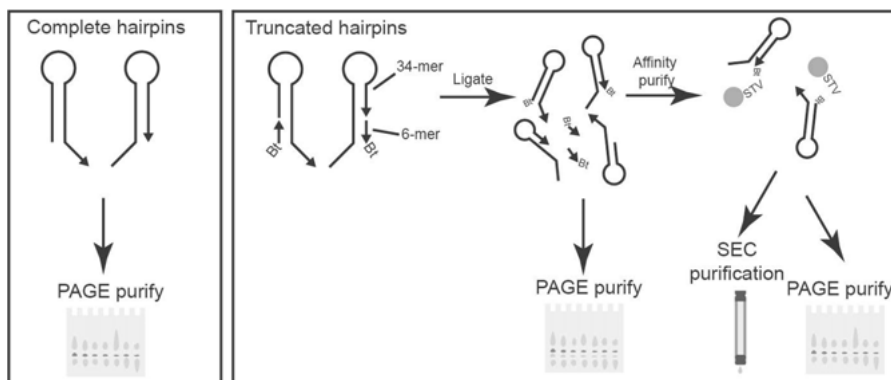
## Design and method

The hairpin design used in this paper had a fairly conservative design comprising of a 6nt toehold and a 14nt stem. While such a design is unlikely to consume all hairpin monomers at equilibrium, the elongated stem should keep it metastable. As such, any non-specific amplifications can more easily be attributed to erroneous oligonucleotides.

For purification we choose to use two approaches (Figure 15), to use and purify complete hairpins as they were or utilize purposefully truncated hairpins that were then ligated and further purified. We hypothesized that introducing a ligation step could improve quality for a few reasons. Firstly, the initially produced oligonucleotides are shorter with less secondary structure, therefore initial production and purification would be easier. Secondly, the ligation requires correct hybridization around the ligation site, therefore, only 'correct' oligonucleotides can ligate properly. Hairpins were then purified mainly using denaturing PAGE, which is one of the better methods for resolving small differences between full length oligonucleotide and truncated product.

For the ligation approach a desthiobiotin was also incorporated onto the smaller ligation fragment. This allowed for a further affinity based purification, where the desthiobiotin fragment along with the ligated hairpin could be extracted. Since this approach also would contain unligated desthiobiotin fragments, oligonucleotides were further purified with PAGE or Size Exclusion Chromatography (SEC).

To assess the effect of the purification strategies we allowed the hairpins to polymerize *in solution*, followed by a PAGE analysis. To evaluate the effect polymerization *in situ* we utilized immunoHCR, where secondary antibodies were conjugated with initiator sequences. Here the signal strength from the stain should be in direct correlation with polymer length.



**Figure 15.** Oligonucleotides were either purchased as complete- or truncated hairpins. Complete hairpins were used as is or further PAGE purified. Truncated hairpins consist of a desthiobiotin 6-mer and a 34-mer that are ligated together. The ligation product is then purified with either PAGE purification or affinity purification (using streptavidin-desthiobiotin affinity). Finally, the affinity purified hairpins were further purified using SEC or PAGE.

## Results and discussion

In this investigation we found that HCR polymerization undoubtedly is correlated to hairpin purity. Both *in situ* and *in solution* analysis showed a large increase in HCR polymerization for purified hairpins. Furthermore, by pre-truncating and ligating the hairpins, amplification could be further improved. Perhaps unsurprisingly, continued purification with affinity purification and PAGE or SEC could further increase amplification. None of these purification strategies can completely remove n-1 truncations, but each subsequent purification would likely reduce the remaining truncations. This also indicates that continued purification could further improve polymerization, however there is a continued trade-off between yield, time committed and purity.

The *in solution* tests also revealed quite a large variation between purchased batches. This is quite concerning, as most published works featuring various HCR designs seemingly only use a single batch of hairpins. Indeed, attempting to assess oligonucleotide quality through analysis of the monomers in a non-denaturing gel show no difference between purifications. It is only when the oligonucleotides were assessed under denaturing conditions that any difference could be seen at all. Extrapolating from this, it is virtually impossible to tell if polymerization ability in other papers is caused by sequence differences or differences in oligonucleotide quality.

While we show that further purification certainly can improve HCR amplification, from a commercial standpoint a sequential purification with RP-HPLC followed by ligation and PAGE would be hard to justify. Both considering time commitment and yield, this approach would not be cost effective.

Therefore, further improvement in more scalable purification strategies for hairpin oligonucleotides are required.

## Paper IV - Purification of detection hairpins improves the performance of proximity-dependent initiation of hybridization chain reaction

### Background

As presented in this thesis, the original ProxHCR design and the optimized design<sup>15,18</sup> could accurately detect proximity events *in situ* without the use of enzymes. Despite the optimization, neither the original system nor the optimization had great signal-to-noise ratio. This resulted in a signal that was difficult to properly quantify. If the method was ever to be used in a diagnostic setting as another dimension of diagnosis, the visualized signal had to be stronger.

As we discovered the correlation between oligonucleotide purity and degree of HCR polymerization in paper III we naturally sought to test how much this would enhance ProxHCR signal. As the hairpins used in paper II were typical RP-HPLC purified detection hairpins, similar to the control hairpins in paper III. We hypothesized that further purification would increase the signal strength for the ProxHCR method, allowing for better signal-to-noise and easier quantification.

### Design and method

As we already had versatile, optimized ProxHCR oligonucleotides available in house, we continued our work using this design. Similarly to paper II, the proximity probes were conjugated to secondary antibodies that could be used to detect proximity events *in situ*. The detection hairpins were either used as is or purified using a denaturing PAGE.

The paper was then separated in to three sections, first a comparison between the purified and non-purified detection hairpins using the interaction between E-cadherin and  $\beta$ -catenin as a model system. The difference was then further clarified by quantifying the proximity signal in the PDGF-BB ligand induced proximity between clathrin and PDGFR- $\beta$ . Next, to establish the viability of the ProxHCR quantifications the method was benchmarked against the PLA. Finally, we investigated whether ProxHCR could accurately track the phosphorylation and subsequent recruitment of downstream proteins to the PDGFR- $\beta$  following PDGF-BB stimulation over time.

## Results and Discussion

The paper has three main findings. First, we show how the addition of an extra denaturing PAGE purification of the detection hairpins substantially increase signal strength without resulting in increased background. This simple improvement would greatly boost signal-to-noise, making the method far more viable in practical use. More specifically, the method could be used with less optimal tissue samples where background autofluorescent signal can be significant.

The second finding was that quantification of ProxHCR signal for proximity between PDGFR- $\beta$  and the early endosome marker EEA-1 as well as PDGFR- $\beta$  and PI3K showed similar induction of proximity compared to a PLA. This would indicate that ProxHCR is a viable method for tracking and quantifying proximity events. Furthermore, the ProxHCR method generally results in far more signal events compared to PLA, while having similar thresholds for detecting proximity events. This allows for a more detailed view of subcellular staining patterns.

Finally, we show that ProxHCR can be used to accurately track well known biological events following PDGF-BB ligand induction in fibroblasts. Here, interaction between PDGF-BB ligand and PDGFR- $\beta$  results in a rapid dimerization followed by autophosphorylation of several tyrosine residues<sup>187,188</sup>. The initial rapid increase of phosphorylation, followed by a decrease in phosphorylation over time could clearly be visualized and quantified with the ProxHCR method. Following autophosphorylation of the PDGFR- $\beta$ , several proteins and complexes are recruited to the receptor to propagate the signal. This could also be tracked using ProxHCR as we could see and quantify a clear increase in proximity signal between PDGFR- $\beta$  and GRB2 as well as PDGFR- $\beta$  and PI3K.

As a conclusion, with extra purification of the detection oligonucleotides signal strength could be vastly increased. This now allows ProxHCR to be a viable enzyme free alternative to both PLA and UnFold without compromising the ability to quantify the results.

## Author's perspectives

I have spent more than five years on working with different proximity methods and DNA-based techniques, tested over 300 DNA sequences and drawn innumerable amounts of hairpins. With this experience, there are many aspects touched upon, but never brought to light. I have had many ideas, undiscussed, several projects, discarded, and many minor details and considerations that will not grace the paper of a journal. I will dedicate this section to informal discussions regarding this thesis and my work.

### On HCR in general

At this point I have worked with HCR for quite some time and have formed a few perspectives. Using DNA as a tool to investigate cellular biology has proven to be very useful. There are multitudes of DNA modifying enzymes, modifications of nucleotides and hybridization interactions are quite easy to manipulate. With the addition of hybridization chain reaction one also has a seemingly perfect way of further amplifying signals from DNA or RNA molecules. Initially, when I first read about HCR, how the polymerization worked and the data that groups published, it seemed great. When eventually testing the hairpins we had in our lab (from the first ProxHCR publication) it seemingly worked great. The hairpins did not polymerize (a lot) without initiator, and with the addition of an initiating sequence they would polymerize. It was essentially a part of the ProxHCR optimization that had been assumed to be “complete”. But over the years and many experiments one comes to realize the issues that exist with the reaction.

First we realized that signal strength for less abundant interactions was poor, something that could not be improved by increasing reaction time. Next, the realization that most of the hairpins remained as unreacted monomers. Then we noticed that hairpins were not as stable when in environments more complex than test tubes with normal buffers. At this point we assumed that our own hairpin designs might be at fault and test other designs from published papers. Here, it turns out that, even with the same sequences, purchased from the same company, with the same purification, published data was hard to reproduce.

At this point, it was obvious that batch to batch quality would differ quite a bit when ordering oligonucleotides. And as explained in the oligonucleotide synthesis and purification section, truncations will happen during synthesis and purification is unlikely to remove all erroneous oligonucleotides. With the data from paper III it is quite easy to conclude that hairpin quality was a major issue. Unfortunately, degree of purification and the correlation to HCR amplification is rarely mentioned in literature, this also makes it hard to conclude which HCR design is optimal. While we show that it is indeed possible for any lab to purify purchased hairpins to achieve better polymerization, our methods will also leave some truncations. For HCR to be a more robust, well-used method, either the purchased oligonucleotides must reach a higher degree of quality or the hairpins need to be redesigned to allow for truncations.

There are quite a few designs published for HCR, the initial design featured a 6-nucleotide stem/loop with an 18-nucleotide stem<sup>85</sup>. This was later altered by the same group to a 12-nucleotide stem/loop with a 24-nucleotide stem<sup>131</sup>. Ang *et. al.* published a few design criteria and suggestions with regards to hairpin design where several smaller hairpins were shown to work<sup>128</sup>. It was also stated that any stem-length longer than the toehold could maintain hairpin stability. Personally I found the paper published by Tsuneoka and Funato<sup>189</sup>, where over 40 different hairpin pairs were tested to be highly interesting. All hairpins were between 36 and 44 nucleotides long and contained 12-nucleotide long stems. Yet, the degree of amplification and metastability varied quite a bit, while few conclusions could be drawn as to what caused leakage and what caused poor polymerization. From an outside perspective it is hard to conclude if the differences were caused by sequence differences or purity differences.

Designing a pair of HCR hairpins is both easy and deceptively difficult. A good HCR design should remain metastable at concentrations up towards 1 $\mu$ M in a typically used buffer and completely polymerize when an initiator is added. Metastability can be enhanced by elongating the stem sequence or reducing the foothold-loop length. Likewise, polymerization efficiency can be improved by increasing foothold/loop length.

Due to the varying quality of hairpins I have been unable to do a proper optimization on hairpin design, but I do have some thoughts on the matter. It is probably easiest to view each addition of an HCR hairpin to the polymer as a single initiator-hairpin interaction and divide this interaction into smaller steps. For the forward reaction, nucleation between the foothold and loop region followed by the displacement of the hairpin stem. And for the backward reaction, initiation of strand displacement of the initiator strand (removing the initiator strand), actual strand displacement of the initiator strand and disengagement of the loop-to-foothold hybridization. Likely, the easiest way to completely push the hairpins to a polymerized state would be to reduce the off-rate between loop and foothold region. If the disengagement never happens, the reaction can only move forward. Indeed, with this line of thought one

will end up with foothold/loop regions with around 12 nucleotides (although it will depend on salt, temperature and sequence), much like the Pierce optimization<sup>131</sup>. I assume that one could also attempt to reduce the reverse strand displacement. Biasing the direction of the random-walk is likely difficult to achieve (although we did attempt it with the mismatch hairpins discussed below), but reducing the chance of initiation is likely possible. This can be done by reducing fraying at the stem edges (by adding G-C bonds) and thus decreasing the chance of a nucleation site. One could also reduce the likelihood of nucleation by increasing distance between the bound initiator strand and the free hairpin strand. In a sense, this would decrease local concentration and thus reduce the nucleation on-rate. This could be accomplished by increasing loop length or potentially increasing the rigidity of the loop<sup>190</sup>. And indeed, the G-C stem edges are within one of the design suggestions proposed by Ang *et. al.*<sup>128</sup>.

Of the above mentioned suggestions to improve polymerization length, increasing foothold and loop length is likely to be the most efficient. This will however affect the metastability of the hairpin pair resulting in leaking hairpins. This problem can of course be alleviated by elongating the stem at cost of reduced kinetics. Otherwise, the G-C content, and thus stability of the stem can be increased. Both of these optimizations will however increase the size of the hairpins, which inadvertently will increase error rate during oligonucleotide synthesis and difficulty in subsequent purification. As a side note on metastability, the HCR leakage will depend on buffer composition, hairpin concentration and sample complexity and should be optimized with that in mind. For *in situ* applications, and other applications where the polymerized product remain bound and remaining unreacted monomers are washed away, the off-rate becomes more important. This is obvious, if the reaction no longer can proceed forward, only the backward reaction will occur and the polymer will shrink in size over time.

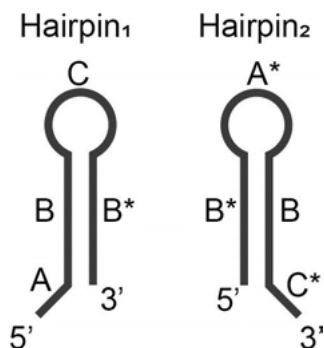
With these considerations in mind, if I were to use or design an HCR-based *in situ* method today, I would likely opt to use hairpins with footholds around 10-14 bases simply to reduce the off-rate. This would of course require usage of longer stems, which, based on experience would require extra purification to function.

Hybridization chain reaction is without a doubt an interesting method for signal amplification. When it works well one can achieve a high degree localized signal amplification and it has a high degree of flexibility. However, I do believe there are some points that needs to be addressed before the method will be more universally adopted. First and foremost, production and purification of hairpins must be improved. While one can purchase and purify hairpins as we did, the PAGE purifications are neither cost- nor time-efficient and will likely be too big of an inconvenience. If sufficient purity can consistently be obtained, sequence design should be reevaluated to yield more robust design criteria.

## Designing a ProxHCR system

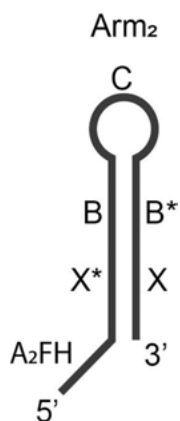
During my studies I have drawn and tested quite a few ProxHCR systems, and this section is here to help other to do the same. These designs are suggestions based on the optimization from paper II and will utilize two probes, both with 5'-footholds and 5'-conjugation. This is not a strict requirement, but 5'-conjugations tend to be cheaper and easier to produce. Furthermore, a conjugation at the stem edge will likely destabilize the stem itself. At the end of this section other conjugations and optimizations will be further discussed.

First grab a pen and paper, doing this by hand is far simpler than doing it on a computer. Start by drawing the detection hairpins, these should be two hairpins, one with a 3'-foothold and one with a 5'-foothold. I would also suggest marking 3'- and/or 5'-ends to keep track of the DNA orientation. Then, we name the different domains, footholds, stem and loops, as in figure 16. Here we need to make sure that the foothold in hairpin<sub>1</sub> is reverse complementary to the loop in hairpin<sub>2</sub> and vice versa.



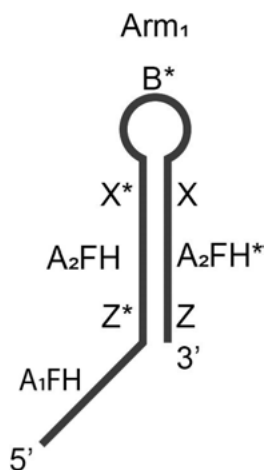
**Figure 16.** Schematic image of hairpins 1 and 2 with added domain names.

Next we will establish the domains in arm<sub>2</sub> (Figure 17), the arm that initiates the HCR reaction. First, draw a hairpin with a 5'-foothold. Arm<sub>2</sub> needs to contain an initiating domain, a sequence reverse complementary to either of the hairpins' foothold and stem (B\*). Here, one technically has a choice to initiate on hairpin<sub>1</sub> (5'-foothold) or hairpin<sub>2</sub> (3'-foothold), both are theoretically possible, but I would suggest initiating on hairpin<sub>2</sub>. This places the reverse complementary sequence to the hairpin<sub>2</sub> foothold (C) in the arm<sub>2</sub> loop. In the arm<sub>2</sub> stem we will place two domains, B and X. For the initiation to work we need to place the B\* domain after (at the 3'-end of) the loop (C). After the initiation sequence we will also add the X domain, a domain that can be added to stabilize the hairpin further to avoid non-specific opening. Also place the complementary stem domains B and X\*. Finally, name the foothold domain of arm<sub>2</sub> A2FH.



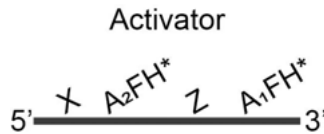
**Figure 17.** Schematic image of arm<sub>2</sub> with added domain names.

To open arm<sub>2</sub> we will also need to draw arm<sub>1</sub> with its complementary domains (Figure 18). Similar to the initiation of HCR by arm<sub>2</sub>, we need to place a strand displacing initiator B\*-X-A2FH\* (5' → 3') in arm<sub>1</sub>. I would suggest to temporarily place B\* in the loop, followed by X and A2FH\*. After these domains, at the 3'-edge of the stem, place another Z domain. Next, place all the complementary sequences in the stem and add the A1FH domain to the foothold.



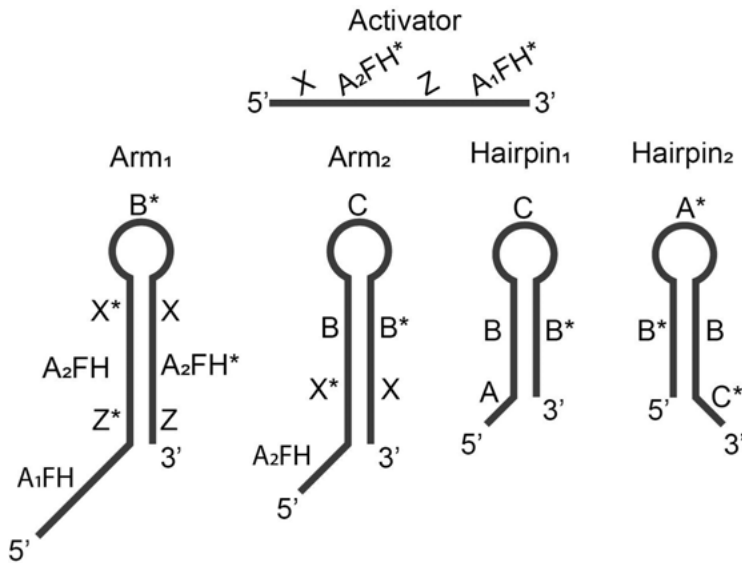
**Figure 18.** Schematic image of arm<sub>1</sub> with added domain names.

Finally we need to have the activator sequence that triggers the chain of strand displacements (Figure 19). This should be a linear fragment only containing the domains X-A2FH\*-Z-A1FH\* (5' → 3').



**Figure 19.** Schematic image of activator with added domain names.

Technically, if we just add random bases with correct complementarity following the domains, we will have a poorly working ProxHCR system (Figure 20). Congratulations. So now we need to address all the minor optimizations and domain lengths that need to take place.



**Figure 20.** Schematic image of a complete ProxHCR system.

First we will discuss the detection hairpins, here I would simply suggest copying a published sequence for HCR hairpins or design one yourself with the Ang *et. al.* design criteria<sup>128</sup>. Regardless, the length of the stem and foothold will impact further optimizations of the ProxHCR system. I will also point out that shorter detection hairpins (like the ones used in paper II and IV) are easier to optimize with this design.

First thing to consider is the B\* domain in arm<sub>1</sub>, the length of this domain will depend on the length of the detection hairpin stem. Therefore, longer detection hairpins will yield a large loop segment which easily can result in an unstable hairpin. This can be balanced out in a few ways, easiest would be to extend the Z domain and aim to make the stem larger than the loop segment, preferably 1.5-times as long. Another method is to remove a few bases from the 5'-end of the B\* domain in arm<sub>1</sub>. This will result in an incomplete opening of arm<sub>2</sub>, but, leaving 2-3 bases unopened will still expose the initiator domain

and trigger HCR. Finally, one can move part of the 3'-end of B\* into the arm<sub>1</sub> stem, thus shortening the loop.

Next we address arm<sub>2</sub>, if the detection hairpin sequence has been decided, domains C and B are already set. With the current design we have complementarity between the activator and arm<sub>2</sub>, if not taken into consideration this can lead to a non-specific stain. Here, the activator can bind and partially displace the stem of arm<sub>2</sub>, however, since the displacement is only partial the activator will unbind as long as A2FH is kept short. During subsequent washing steps the unbound activator will be washed away. Considering that the activator concentration during a wash can be assumed to be zero, the on-rate will be non-existent. The X domain can be introduced to reduce non-specific binding, but could also be skipped entirely. Insertion of this domain, and its length, will depend on the metastability of the detection hairpins. Keep in mind that X domains also are inserted into arm<sub>1</sub> and potentially into the activator, and could result in further optimizations in these oligonucleotides. Another potential optimization would be moving the B\* domain into the loop, this would decrease the stability of arm<sub>2</sub> but reduce complementarity between arm<sub>1</sub> and the detection hairpins.

Finally we need to consider the activator. Technically most of the domains already contain defined sequences from arm<sub>1</sub> and arm<sub>2</sub>. However, how much is incorporated and how long the A1FH should be remains undecided. For the foothold, I would suggest a long domain where one can expect the off-rate to be minimal. For the paper II optimization I decided to use 15-nucleotides. Finally, how much of the A2FH domain along with the X domain that needs to be added will depend on the structure of the arm<sub>1</sub> hairpin. One should aim to open almost the entire stem of arm<sub>1</sub>, leaving at most 2-3 bases. If the activator will contain the full A2FH\*, X-, and potentially a B\* -domain (if it was moved down to the stem in arm<sub>1</sub>) will depend on tweaks done on the other arms. But, it should be kept in mind that the more complementarity between the activator and arm<sub>2</sub>, the more likely arm<sub>2</sub> is to leak due to its binding and will require more washing following the activator addition.

The prior considerations will typically result in an iterative process of making small changes in one domain which will cause alterations somewhere else. Eventually, one will be convinced of a certain design which now only needs actual sequences. I would suggest to use the design tool in Nupack ([www.nupack.org](http://www.nupack.org)) to create sequences. Although one could randomize the sequences or manually add them as well. Next, one should make sure that there are no unwanted interactions between domains. Finally, one just needs to order said sequences, conjugate them to antibodies, test them *in situ*, cry when it doesn't work and do it all again.

There are some final considerations that I would like to add regarding detection hairpin metastability. In the above designed system, arm<sub>1</sub> will contain at least parts of the detection hairpin stem, this domain can bind hairpins non-specifically polymerizing in the solution. Unfortunately, as mentioned in this

thesis, poor quality hairpins can result in compromised metastability. Furthermore, metastability tested in tubes might not translate to an *in situ* context. So either detection hairpins need to be chosen and tested with this in mind and err on the side of extra stability (with longer stem length) or the ProxHCR system must be designed with the assumption that metastability will be compromised. In paper II and IV, we picked the latter option. The detection hairpin used was short, with only 9 nucleotides. Of these 9 nucleotides, only 8 were complementary to arm<sub>1</sub>, as such, any bound hairpins could be washed off.

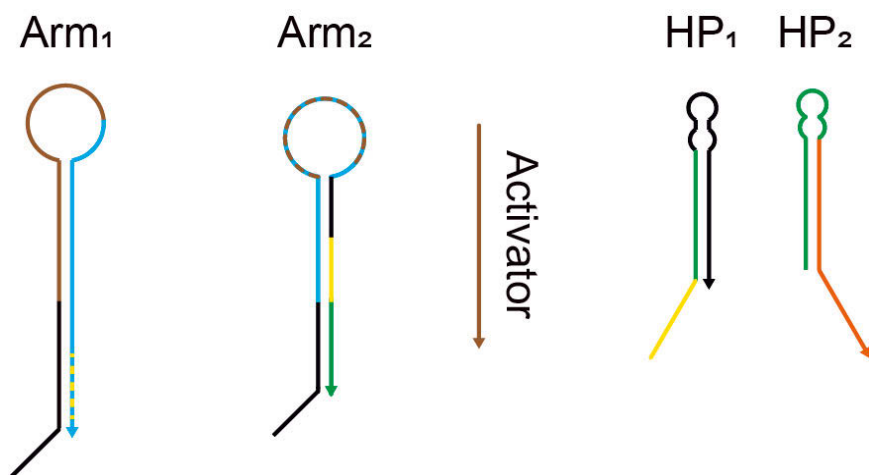
Finally I would like to discuss some of the design regrets I have about the ProxHCR system used in paper II and IV. Unfortunately, no system is truly optimized and I have learned a lot over the years. If I were to redesign the ProxHCR system today I would likely do a few things differently. First, I would try to alter the conjugation of either arm<sub>1</sub> or arm<sub>2</sub> to a 3'-conjugation. This was briefly tested and discarded due to leakage by arm<sub>1</sub>. But in retrospect, this system should be possible. If parts of B\* is moved into the loop of arm<sub>2</sub> and the X domain of arm<sub>2</sub> was extended, then the complementarity between arm<sub>1</sub> and hairpins could potentially be reduced. This variant allows for longer detection hairpins which should polymerize more and yield better signal strength. Furthermore, by reversing one conjugation the proximity distance between antibodies is reduced to almost 0nm. Alas, there was no time and how many ProxHCR optimizations can one actually publish?

## Projects that did not come to be

### Moving sequences

Initially when I started my work with ProxHCR we intended to alter the original design to a new version containing *moving sequences*. One of the problems with the original design and the design in paper II for that matter, was that arm<sub>1</sub> (not supposed to initiate the HCR reaction) contains elements of the initiator. This can be a cause of non-specific signal as this region can bind any spontaneously amplifying detection hairpins. It is quite hard to avoid, as arm<sub>2</sub> must hide the initiation sequence, at least in part, in the stem, which in turn must be complementary to arm<sub>1</sub>. In order to avoid this, we designed a few versions where both arms were initiated from the loop. Here, the rigidity of dsDNA was used to our advantage as binding the loop and parts of the stem was enough to pry the hairpin open. With this design, neither the activator nor arm<sub>1</sub> require complete complementarity to open the next hairpin, thus reducing the sequences complementary to the detection hairpins (Figure 21).

By omitting the sequence that invades the stem region of the detection hairpins from arm<sub>1</sub> it cannot initiate HCR. However, the concept needs to be tested and optimized which wasn't possible within the frame of the current thesis work.



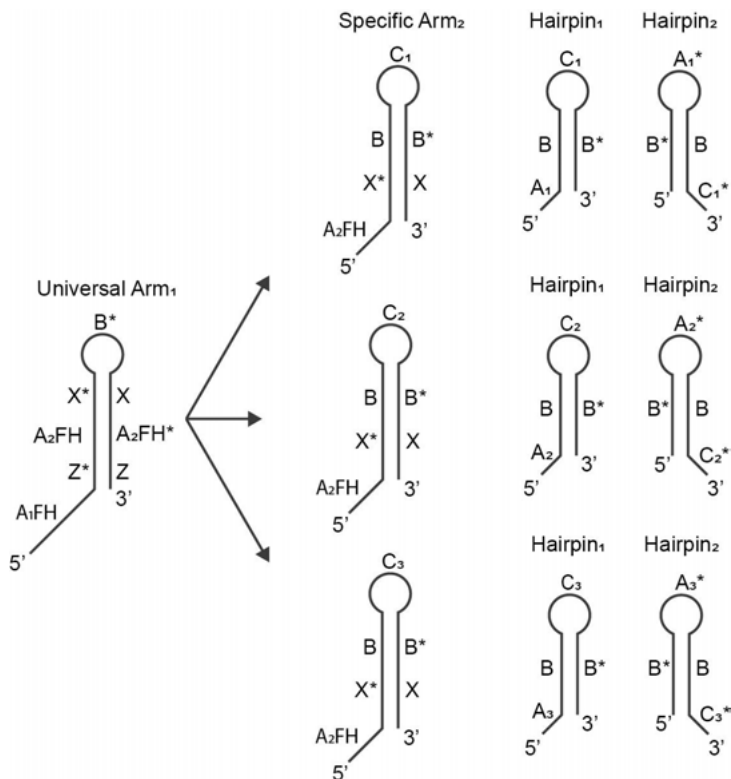
**Figure 21.** Schematic representation of ProxHCR based on moving sequences. Complementary sequences are color-coded. Here, the activator only binds part of arm<sub>1</sub>, which only binds part of arm<sub>2</sub>. Instead, the arms are pried open by the dsDNA rigidity. Compared to the first iteration of ProxHCR there is little complementarity between arm<sub>1</sub> and detection hairpins. Instead, arm<sub>1</sub> is only complementary to the foothold of hairpin<sub>1</sub> (yellow segment in arm<sub>1</sub>). Simultaneously the complementarity between activator and arm<sub>2</sub> is further decreased (brown segment in arm<sub>2</sub>).

## Multiplex ProxHCR

This method is still stuck at a conceptual level. Initially this would have been my continuation project had the signal strength from paper II been sufficient. In theory, it is not too hard to design multiplexed ProxHCR probes. The initial concept would see a single probe with arm<sub>1</sub> capable of opening several variants of arm<sub>2</sub>. This would allow a multiplex analysis between a single protein and several other proteins. In turn the different arm<sub>2</sub> probes would contain loops with differing nucleotide sequences. By producing sets of detection hairpins with the same stem sequence but varying foothold sequences one could have simultaneous detection of several different probes (Figure 22).

There are of course foreseeable problems with the described design, first and foremost, all detection hairpins must remain metastable despite the presence of an exposed stem region. Otherwise one would risk non-specific detection. Typically, one can have around three fluorophores and a nuclear counterstain at once, so the detection would have to be done in a stepwise manner. This would also require stripping of detection hairpins to remove them from the previous stain. At any higher level of multiplexing conjugation of secondary antibodies would no longer be feasible, as one can only reliably use a few species for primary antibodies.

In other words, this method is feasible but would most likely require several design and optimization iterations to function.

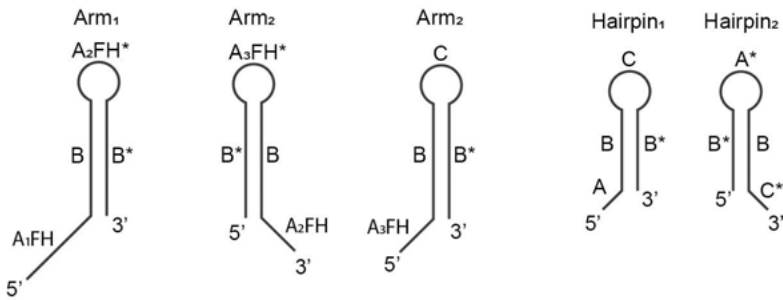


**Figure 22.** Schematic representation of Multiplex ProxHCR. Here a single arm<sub>1</sub> is capable of opening several versions of arm<sub>2</sub>. As each separate arm<sub>2</sub> would contain differing loop sequences, they would open up separate sets of detection hairpins. With varying fluorophores on the detection hairpins, several protein interactions could be visualized at once.

### 3-plex ProxHCR

Like its multiplex brother, this did not come further than the sequence design stages. Here, the ProxHCR method would be utilized for the detection of triple interactions instead of double interactions. Instead of arm<sub>2</sub> initiating the HCR polymerization it would instead open a third arm, which when opened would initiate HCR (Figure 23). Essentially, one would need to move one of the detection hairpins and conjugate it to an antibody instead. Technically, this could be further implemented as 4-plex or 5-plex, although with each addition risk of non-specific leakage increase. One can also assume that there would be some physical hindrances involved in too many chained strand displacements.

The idea was scrapped, or put on hold until stronger signal strength was achieved. Obviously the number of triple interactions would be fewer compared to double interactions and therefore signal strength would appear even weaker.



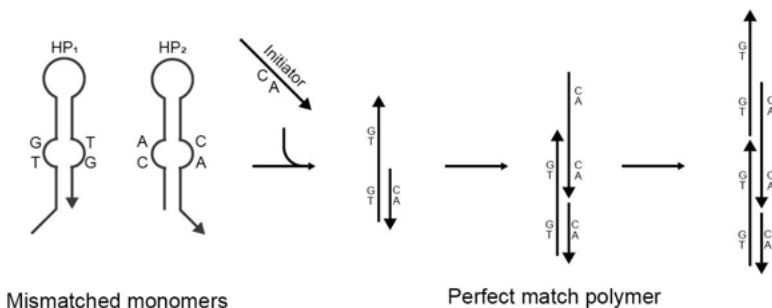
**Figure 23.** Schematic representation of 3-plex ProxHCR. In this design three hairpins would be conjugated to antibodies. Here, arm<sub>1</sub> would open arm<sub>2</sub> which in turn would open arm<sub>3</sub>. Finally, arm<sub>3</sub> would initiate HCR to produce a signal, unless all antibodies are within proximity, no signal would be produced.

## Mismatch hairpins

This was the primary work that took place between paper II and paper III. After the optimization of ProxHCR in paper II we wanted to further improve the signal strength. The main target of optimization was the detection hairpins themselves, assuming that hairpins could be designed better to amplify more *in situ*. Notably, there is a lack of publications detailing the correlation between hairpin sequences and amplification strength for *in situ* methods. We figured that we could capitalize on this lack of data by doing our own optimization and hairpin designs. The idea was to test signal strength and therefore amplification length of already published sequences, our own sequences with varying stem and foothold/loop lengths along with our mismatch hairpins.

The mismatch hairpins had, as the naming suggests, a mismatch incorporated within the stem of the hairpin using Wobble base pairs G-T and C-A (Figure 24). As a monomer the mismatch would remain, but when polymerized the hybridized hairpins would form a correct match instead. Similar to loop/foohold hybridization, the mismatch should result in a larger energy gain per hairpin. Furthermore, by adding a mismatch in the middle of a hairpin stem it would act as a checkpoint, where a random-walk in the reverse direction would be less likely after displacing the mismatched nucleotides.

The idea itself worked somewhat, as we designed mismatch hairpins that remained metastable and polymerized quite well. Unfortunately, due to the difference in quality between the ordered hairpins it was hard to draw consistent conclusions. The design was put on hold until hairpins with better quality could be assured.



**Figure 24.** Schematic representation of mismatch hairpins. Initially, the hairpins remain as metastable monomers with less stable wobble base pairing. Once initiated, the initiator and the complementary hairpin will hybridize, forming correctly paired bases.

## PLA/HCR hybrid

This was a minor side project with potential to work quite well. Here, the design was inspired by PLA, MolBoolean and ProxHCR. The intention was to amplify the signal in a fashion similar to PLA, where two probes would yield a single amplification product using RCA. Simultaneously, the Molboolean method that was being tested in the lab utilized pre-ligated and purified circles for amplification. Initially I had discarded the use of these circles as I had thought they would yield too much background by being primed from other non-specific sequences. But, work on the Molboolean project proved me wrong. To confirm proximity, instead of using ligation as in PLA, one would use strand displacement from ProxHCR, where the initiating sequence from ProxHCR would be changed to a circle priming sequence.

The point of doing this at all was to reduce all linear products from the PLA method. Pre-ligated circles purified with exonucleases should only yield correct circles, and using a good ProxHCR design should have a high efficiency in strand displacement. But unlike the HCR reaction used in ProxHCR, one would have the RCA product instead, which is a far stronger signal. In other words, one would have the number of proximity events from ProxHCR, but the signal strength of PLA.

The initial results looked quite promising, but the probes used to initially test this method was the ProxHCR probes from paper II. These work quite well for the HCR reaction, but binding between the circle and arm<sub>2</sub> was too weak and a lot of signal was lost. To properly get this method working one would need to increase the number of hybridized bases between arm<sub>2</sub> and the circle further. Preferably one would also add some blocking bases to the 3'-end of arm<sub>1</sub> to remove all non-specific signal from this arm. In other words, some small modification to the design would likely yield a functioning method. However, due to lack of time work on paper IV was prioritized instead.

## Locked hairpins

Hybridization chain reaction hairpins have a tendency, when stored, to form dimers, trimers etc. Any hairpin will naturally have a possibility to not only form intrastrand hybridizations but also interstrand hybridizations. These can be somewhat disruptive when polymerizing and typically results in worse metastability. To address this problem, most groups tend to heat the hairpins separately prior to polymerization. This involves a brief heating period to 95°C for about 5min followed by cooling down to room temperature. This will break hybridization, and it is assumed that the intrastrand hybridization will be kinetically favored resulting in predominantly monomers. In a lab setting this is not too inconvenient, but if done anywhere else it is a bit cumbersome. I figured we could produce ‘locked hairpins’ that always would maintain their conformation regardless of storage time.

The idea was to extend the footholds with a non-interacting sequence that would be ligated to the stem, forming several small circles instead. The circles would be more restrained from forming dimers and even if mixed with its complementary hairpin would not hybridize well. Within the non-interacting sequence a few phosphorothioate bonds were added, among several uses they can be cleaved in the presence of iodine. This allowed for an enzyme free way of inducing a strand break.

The project was discarded for various reasons. First and foremost, hydrolysis of the phosphorothioate bonds was not efficient enough. Second, the hydrolyzed hairpins often ended up with overhanging sequences, these are quite disruptive to the polymerization itself and gave much worse signal amplification.

## Non-conventional purification

For paper III we eventually ended up using denaturing PAGE, SEC and affinity based purification. These are quite well known and readily available to most labs, however, we are also well aware that none of these methods will completely remove n-1 truncations. Initially we aimed to instead reduce unwanted sequences through various other methods, mainly focusing on enzymes. The main goal was to show that the hairpins could be purified by virtually any lab.

The first alternative was purification against bound initiator sequences. Both hairpins will have affinity to its complementary initiator sequence, by using bead conjugated initiators one should be able to separate truncated sequences from the correct ones. However, it is quite obvious that n-1 sequences also had good affinity to the same initiators albeit with a slightly lower  $T_m$ . For this purification to actually work one would need to operate with a very fine control over temperature, or other DNA disruptive agents.

Next, we attempted a few variants of the ligation purification used in paper III. Here we also ligate two fragments, with the assumption that only ‘well hybridizing’ fragments will properly ligate. But instead of purifying with PAGE or HPLC, we added nuclease resistant nucleotides onto the fragment edges. As such, properly ligated hairpins would be protected from both ends, while non-ligated fragments would only be protected from one end. By adding exonucleases targeting both 3’- and 5’-ends, one would degrade all non-ligated hairpins. The problem here was that the nuclease resistant nucleotides were worse at hybridizing, drastically decreasing polymerization. This could be overcome by altering the hairpin design, but would be unlikely to be used by other labs as any design would require re-optimization.

Instead we tagged the hairpins with external nuclease protected bases. Simple idea, retain the foothold and stem with unaltered bases and only have a small external sequence added on both edges. Here, the ligation and nuclease treatment worked well, but the external tag also reduced polymerization efficiency.

So we figured that we could just cut off the external tags. Here we utilized the UNG/Endo4 strategy from Unfold. This also worked quite well, however, the enzymatic removal of the tags was not efficient enough to remove every tag. The remainder of the hairpins were enough to disrupt the polymerization, making this strategy unsuitable. Testing the same strategy with FPG enzyme worked better, but efficiency was not high enough.

We also carried out work with the more conventional purifications at the same time, and eventually decided that these were ‘good enough’ and was eventually used in paper III.

## Conclusions

Both ProxHCR and UnFold are systems that have enhanced the original PLA method in various ways. UnFold results in more signal at far lower probe concentrations by increasing efficiency and reducing linear fragment formation. Since the method yields fewer false negatives, the stain better represents the actual cellular state. Similarly, ProxHCR also produces far more proximity events at the cost of a slightly weaker signal amplification. Due to the lack of enzymes, the resulting method is both cheaper and more suitable for automated staining. The optimizations done in paper II, resulted in a more robust system with less variability with regards to temperature and salt concentrations.

However, as seen in the third and fourth paper, HCR hairpins are susceptible to synthesis errors and potential degradation. This had a clear impact on HCR polymerization and therefore ProxHCR signal strength. The problem was alleviated by further purification using methods readily available in labs, showing the importance of HCR hairpin purity.

As it stands, if one uses properly purified oligonucleotides, both UnFold and ProxHCR are viable methods to accurately track proximity events *in situ*. As such they could potentially be adopted for use in other labs or be sold commercially.

With that being said, ProxHCR has potential to be further optimized. As discussed in the previous section, there are many avenues that can be improved, both with regards to detection hairpin quality and probe design. With further improvements in oligonucleotide synthesis and purification being an important factor in improving efficiency and cost effectiveness.

Hopefully this thesis will shed some light on both the potential and challenges with proximity methods. And potentially inspire others to utilize them in their own work.

# Popular scientific summary

The human body consists of over a trillion cells divided into several cell types, and these types must perform very specific functions. In order to execute these functions, these cells are regulated by both intra- and intercellular signaling. The signal transduction is mainly carried out by proteins. The proteins construct a signaling network in order to transmit signals. Each transmission of signals from one protein to another require an interaction between the two proteins.

In our group we have focused on detecting and quantifying these protein interactions within cells using microscopy-based techniques. This will allow us and other scientist to track if, when, and where proteins are interacting. This can play an important role in understanding how these signaling networks work both under healthy conditions, but also under irregular conditions, such as cancer.

Our group has previously produced two notable methods, proximity ligation assay (PLA) and proximity-dependent initiation of hybridization chain reaction (ProxHCR). Both of these methods utilize a pair of antibodies that have been connected to DNA chains with known sequences. The antibodies are intended to bind protein targets of interest, and if they are within a close distance of one another, give rise to an identifiable signal.

For PLA, two external DNA chains are used to bind to the DNA connected to the antibody pair. If a pair of antibodies are within close proximity, these external DNA chains will form a circle with two gaps. By adding an enzyme, a ligase, these gaps can be filled to form a complete DNA circle. If one has formed one of these DNA circles one can perform a ‘rolling circle amplification’, a specific reaction that uses another enzyme, a polymerase. This will result in an amplification of the circle, which will give rise to a large yarn of DNA that can be seen using a microscope.

ProxHCR, on the other hand, attempts to do a similar interaction but without using enzymes. Instead, this method uses what is known as ‘strand displacement’. Here, an antibody pair is connected to another set of DNA chains. These DNA chains have the shape of hairpins, as the DNA has folded and bound itself. This shape hinders the two DNA chains from interacting and binding each other. When the antibodies have bound their intended protein targets, a third DNA chain is added. This DNA chain will act as a key to open one of the antibody bound DNA hairpins. Once opened, this hairpin can bind

the second antibody bound DNA hairpin. This chain of events will expose another 'key' that was hidden within the second DNA hairpin. Once this key is exposed a chain reaction will take place where several sets of 'detection hairpins' can bind to the exposed key. Once many sets have bound, a clear signal can be seen using a microscope. Of course, this chain of events can only take place if the antibody pair has bound close to one another.

Both of these methods work well, but all methods have their own limitations. For PLA, there is a risk of missing some signal events. This is caused by incorrect binding of the external DNA chains. This, in turn, results in a broken DNA circle that cannot be visualized. To improve this method, we produced a similar method called Unfold. Here, the external DNA chains have been incorporated into one of the antibody connected DNA chains. By doing so, the efficiency in creating the circles was improved.

For ProxHCR, there were two main problems. First, the initial version of the method had very rigid conditions that required to be filled to produce a good signal. As a consequence, small variations in experimental conditions would result in large differences in signal, making the data less reliable. The second problem was the quality of the DNA chains used by the method. Due to the hairpin structure of the DNA, the production of said DNA was more difficult and resulted in many incorrect DNA chains which gave little signal. The first problem was fixed by remaking the DNA sequences from scratch to produce a smaller more efficient design. The second problem was tackled by 'purifying' the DNA strands, which resulted in a higher degree of correct DNA chains. Both of these fixes resulted in a more robust method that gave much better signal strength and was less sensitive to minor changes in experimental conditions.

By further improving upon these two methods we hope that they can find their use in other labs when trying to detect protein interactions. In doing so it could help further our understating of how both regular and diseased cells function.

# Populärvetenskaplig sammanfattning

Människokroppen består av över en biljon celler uppdelade i flera olika celltyper, och dessa typer måste utföra mycket specifika funktioner. För att kunna utföra dessa funktioner krävs signalering både i och mellan celler. Denna signaltransduktion utförs framförallt av proteiner. En stor del av denna reglering kräver att signaler överförs från ett protein till ett annat. På så sätt utformar proteinerna ett signalnätverk för att överföra signaler. Varje överföring av signal från ett protein till ett annat kräver en interaktion mellan de två proteinerna.

I vår grupp har vi fokuserat på att detektera och kvantifiera dessa proteininteraktioner inom celler med hjälp av mikroskopibaserade tekniker. Detta tillåter oss och andra forskare att spåra om, när, och var proteiner interagerar. Detta kan spela en viktig roll för att förstå hur dessa signalnätverk fungerar både under friska förhållanden och avvikande förhållanden, som cancer.

Vår grupp har tidigare producerat två anmärkningsvärda metoder, 'proximity ligation assay' (PLA) och 'proximity-dependent initiation of hybridization chain reaction' (ProxHCR). Båda dessa metoder använder ett par av antikroppar som har kopplats till DNA-kedjor med kända sekvenser. Antikropparna är avsedda att binda två proteiner av intresse, och om dessa är inom nära avstånd från varandra, ge upphov till signal.

För PLA används två externa DNA-kedjor för att binda till det DNA som är kopplat till antikroppsparet. Om ett antikroppspar är nära varandra kommer dessa externa DNA-kedjor att bilda en cirkel med två hål. Genom att tillsätta ett enzym (ett ligas) kan dessa hål fyllas för att bilda en komplett DNA-cirkel. Om man har bildat en av dessa DNA-cirklar kan man utföra en "rolling circle amplification", en specifik reaktion som använder ett annat enzym, ett polymeras. Detta kommer att resultera i en amplifiering av cirkeln, vilket kommer att ge upphov till ett stort garn av DNA som kan ses med hjälp av ett mikroskop.

ProxHCR försöker göra en liknande reaktion utan att använda enzymer. Istället använder denna metod vad som kallas "strand displacement". Här är ett antikroppspar kopplat till en annan uppsättning DNA-kedjor. Dessa DNA-kedjor har formen av hårnålar, eftersom DNA:t har veckats och bundit sig själv. Denna form hindrar de två DNA-kedjorna från att interagera och binda varandra. När antikropparna har bundit sina avsedda proteinmål tillsätts en tredje DNA-kedja. Denna DNA-kedja kommer att fungera som en nyckel för att öppna en av de antikroppsbundna DNA-hårnålarna. När den väl öppnats

kan denna hårnål binda den andra antikroppsbundna DNA-hårnålen. Detta handlingsförlopp kommer att exponera en annan "nyckel" som gömms i den andra DNA-hårnålen. När den här nyckeln har exponerats kommer en kedjereaktion att äga rum där flera uppsättningar av "detektionshårnålar" kan binda till det exponerade DNA segmentet. När många uppsättningar av detektionshårnålar har bundit syns en tydlig signal med hjälp av ett mikroskop. Naturligtvis kan detta händelseförlopp endast ske om antikroppsparat har bundit nära varandra.

Dessa metoder fungerar bra, men alla metoder har sina begränsningar. För PLA finns det en risk att man missar en del signaler. Detta orsakas av en felaktig inbindning av de externa DNA-kedjorna. Detta orsakar en felaktig DNA-cirkel som inte kan visualiseras. För att förbättra denna metod tog vi fram en liknande metod kallad Unfold. Här har de externa DNA-kedjorna inkorporerats i en av de antikroppsbundna DNA-kedjorna. Genom denna inkorporering förbättrades effektiviteten av cirkelformationen.

För ProxHCR fanns det två huvudproblem. För det första hade den initiala versionen av metoden mycket snäva förhållanden under vilka signal kunde produceras. Som en konsekvens kunde små variationer i experimentella förhållanden resultera i en stor signalskillnad, vilket gör metoden mindre tillförlitlig. Det andra problemet var kvalitén på de DNA-kedjor som användes av metoden. På grund av DNA:ts hårnålsstruktur var produktionen av nämnda DNA-kedjor besvärligare och resulterade i många felaktiga DNA-kedjor som gav undermålig signal. Det första problemet åtgärdades genom att göra om DNA-sekvenserna från grunden för att producera en mindre, men effektivare design. Det andra problemet åtgärdades genom att rena DNA-strängarna, vilket resulterade i en högre andel korrekta DNA-kedjor. Båda dessa korrigeringar resulterade i en mer robust metod som gav mycket bättre signalstyrka och var mindre känslig för små variationer i experimentella förhållanden.

Genom att ytterligare förbättra dessa två metoder hoppas vi att de kan hitta sin användning i andra forskningslabbs vid undersökning av proteininteraktioner. Förhoppningsvis kommer båda metoderna att fördjupa vår kunskap om hur friska samt sjuka celler fungerar.

# Acknowledgements

As my studies were carried out at the department of pharmaceutical biosciences at Uppsala University I would like to extend my gratitude for allowing me to conduct my studies here.

Next I would like to thank my main supervisor, Ola Söderberg, for giving me the opportunity to work with all these methods. And also for taking his time and listening to all my great (and not so great) ideas.

I would also like to thank my Co-supervisors, Björn Koos, Ola Spjuth and Carolina Wählby.

And finally I would like to thank all my coworkers for all the help I have received over the years and all the interesting discussions we have had.

# References

1. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* **14**, e1002533 (2016).
2. Apweiler, R. The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.* **38**, D142–D148 (2010).
3. Ezkurdia, I. *et al.* Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Hum. Mol. Genet.* **23**, 5866–5878 (2014).
4. Morton, N. E. Parameters of the human genome. *Proc. Natl. Acad. Sci.* **88**, 7474–7476 (1991).
5. Li, W. On parameters of the human genome. *J. Theor. Biol.* **288**, 92–104 (2011).
6. F.S. Collins, E.S. Lander, J. Rogers, *et al.* International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
7. Bateman, A. *et al.* UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169 (2017).
8. Uhlen, M. *et al.* Tissue-based map of the human proteome. *Science (80- )*. **347**, 1260419–1260419 (2015).
9. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science (80- )*. **356**, (2017).
10. Hamm, H. E. The Many Faces of G Protein Signaling. *J. Biol. Chem.* **273**, 669–672 (1998).
11. Trewavas, A. Post-Translational Modification of Proteins by Phosphorylation. *Annu. Rev. Plant Physiol.* **27**, 349–374 (1976).
12. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science (80- )*. **296**, 910–913 (2002).
13. Pickart, C. M. Mechanisms Underlying Ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
14. Klaesson, A. *et al.* Improved efficiency of in situ protein analysis by proximity ligation using UnFold probes. *Sci. Rep.* **8**, 5400 (2018).
15. Leino, M. *et al.* Optimization of proximity-dependent initiation of hybridization chain reaction for improved performance. *Mol. Syst. Des. Eng.* **4**, 1058–1065 (2019).
16. Söderberg, O. *et al.* Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat. Methods* **3**, 995–1000 (2006).
17. Jarvius, M. *et al.* In Situ Detection of Phosphorylated Platelet-derived Growth Factor Receptor  $\beta$  Using a Generalized Proximity Ligation Method. *Mol. Cell. Proteomics* **6**, 1500–1509 (2007).

18. Koos, B. *et al.* Proximity-dependent initiation of hybridization chain reaction. *Nat. Commun.* **6**, 7294 (2015).
19. McNamara, G., Difilippantonio, M., Ried, T. & Bieber, F. R. Microscopy and image analysis. *Curr. Protoc. Hum. Genet.* **2017**, 4.4.1-4.4.89 (2017).
20. Bardell, D. The Invention of the Microscope. *Bios* **75**, 78–84 (2004).
21. van Zuylen, J. The microscopes of Antoni van Leeuwenhoek. *J. Microsc.* **121**, 309–328 (1981).
22. Hooke, R. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses: with observations and inquiries thereupon.* (1667).
23. Lane, N. The unseen World: Reflections on Leeuwenhoek (1677) ‘Concerning little animals’. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).
24. Karamanou, M., Poulakou-Rebelakou, E., Tzetzis, M. & Androutsos, G. Anton van Leeuwenhoek (1632-1723): Father of micromorphology and discoverer of spermatozoa. *Rev. Argent. Microbiol.* **42**, 311–314 (2010).
25. Salmon, E. D. & Canman, J. C. Proper Alignment and Adjustment of the Light Microscope. *Curr. Protoc. Cell Biol.* **00**, 4.1.1-4.1.26 (1998).
26. Sjöback, R., Nygren, J. & Kubista, M. Absorption and fluorescence properties of fluorescein. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **51**, L7–L21 (1995).
27. Benchaib, M., Delorme, R., Pluvinage, M., Bryon, P. A. & Souchier, C. Evaluation of five green fluorescence-emitting streptavidin-conjugated fluorochromes for use in immunofluorescence microscopy. *Histochem. Cell Biol.* **106**, 253–256 (1996).
28. Panchuk-Voloshina, N. *et al.* Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J. Histochem. Cytochem.* **47**, 1179–1188 (1999).
29. Coons, A. H., Creech, H. J. & Jones, R. N. Immunological properties of an antibody containing a fluorescent group. *Proc. Soc. Exp. Biol. Med.* **47**, 200–202 (1941).
30. Coons, A. H. The demonstration of pneumococcal antigen in tissues by the use of fluorescent antibody. *J Immunol* **45**, 159 (1942).
31. Lichtman, J. W. & Conchello, J. A. Fluorescence microscopy. *Nat. Methods* **2**, 910–919 (2005).
32. Minsky, M. Memoir on inventing the confocal scanning microscope. *Scanning* **10**, 128–138 (1988).
33. White, J. G., Amos, W. B. & Fordham, M. An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy. *J. Cell Biol.* **105**, 41–48 (1987).
34. Egner, A., Jakobs, S. & Hell, S. W. Fast 100-nm resolution three-dimensional microscope reveals structural plasticity of mitochondria in live yeast. *Proc. Natl. Acad. Sci.* **99**, 3370–3375 (2002).
35. Chen, F., Tillberg, P. W. & Boyden, E. S. Expansion microscopy. *Science (80-. )*. **347**, 543–548 (2015).
36. Raykova, D. *et al.* Let There Be Light! *Proteomes* **4**, 36 (2016).
37. Schmolze, D. B., Standley, C., Fogarty, K. E. & Fischer, A. H. Advances in microscopy techniques. *Arch. Pathol. Lab. Med.* **135**, 255–263 (2011).

38. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780 (1994).
39. Hein, B., Willig, K. I. & Hell, S. W. Stimulated emission depletion (STED) nanoscopy of a fluorescent protein-labeled organelle inside a living cell. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 14271–14276 (2008).
40. Vicidomini, G., Bianchini, P. & Diaspro, A. STED super-resolved microscopy. *Nat. Methods* **15**, 173–182 (2018).
41. Gustafsson, M. G. L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13081–13086 (2005).
42. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).
43. Göttfert, F. *et al.* Coaligned dual-channel STED nanoscopy and molecular diffusion analysis at 20 nm resolution. *Biophys. J.* **105**, 6–8 (2013).
44. Bancroft, J. D. & Layton, C. The hematoxylin and eosin. *Bancroft's theory Pract. Histol. Tech.* 173–186 (2012).
45. Wittekind, D. Traditional staining for routine diagnostic pathology including the role of tannic acid. 1. Value and limitations of the hematoxylin-eosin stain. *Biotech. Histochem.* **78**, 261–270 (2003).
46. Fischer, A. H., Jacobson, K. A., Rose, J. & Zeller, R. Hematoxylin and eosin staining of tissue and cell sections. *Cold Spring Harb. Protoc.* **3**, 4986–4988 (2008).
47. Wissozky, N. Ueber das Eosin als Reagens auf Hämoglobin und die Bildung von Blutgefäßen und Blutkörperchen bei Säugethier- und Hühnerembryonen. *Arch. für Mikroskopische Anat.* **13**, 479–496 (1877).
48. Cooper, M. D. & Alder, M. N. The evolution of adaptive immune systems. *Cell* **124**, 815–822 (2006).
49. Frank, M. M., Joiner, K. & Hammer, C. The Function of Antibody and Complement in the Lysis of Bacteria. *Rev. Infect. Dis.* **9**, S537–S545 (1987).
50. Burton, D. R. Antibodies, viruses and vaccines. *Nat. Rev. Immunol.* **2**, 706–713 (2002).
51. Hoffman, W., Lakkis, F. G. & Chalasani, G. B cells, antibodies, and more. *Clin. J. Am. Soc. Nephrol.* **11**, 137–154 (2016).
52. Wang, W., Singh, S., Zeng, D. L., King, K. & Nema, S. Antibody Structure, Instability, and Formulation. *J. Pharm. Sci.* **96**, 1–26 (2007).
53. Cooper, H. M. & Paterson, Y. Production of Polyclonal Antisera. *Curr. Protoc. Immunol.* **13**, 1–9 (1995).
54. SHULMAN, M., WILDE, C. D. & KÖHLER, G. A better cell line for making hybridomas secreting specific antibodies. *Nature* **276**, 269–270 (1978).
55. Fritschy, J. & Härtig, W. Immunofluorescence. in *Encyclopedia of Life Sciences* 1–7 (John Wiley & Sons, Ltd, 2001). doi:10.1038/npg.els.0001174
56. Bordeaux, J. *et al.* Antibody validation. *Biotechniques* **48**, 197–209 (2010).
57. Michel, M. C., Wieland, T. & Tsujimoto, G. How reliable are G-protein-coupled receptor antibodies? *Naunyn. Schmiedeberg's Arch. Pharmacol.* **379**, 385–388 (2009).

58. Bratthauer, G. L. Overview of Antigen Detection Through Enzymatic Activity. in *Methods in Molecular Biology* **588**, 231–241 (2010).
59. Kapuscinski, J. DAPI: A DMA-Specific fluorescent probe. *Biotech. Histochem.* **70**, 220–233 (1995).
60. Chazotte, B. Labeling nuclear DNA using DAPI. *Cold Spring Harb. Protoc.* **6**, 80–83 (2011).
61. Lengsfeld, A. M., Low, I., Wieland, T., Dancker, P. & Hasselbach, W. Interaction of Phalloidin with Actin. *Proc. Natl. Acad. Sci.* **71**, 2803–2807 (1974).
62. Adams, A. E. M. & Pringle, J. R. [51] Staining of actin with fluorochrome-conjugated phalloidin. in *Methods in Enzymology* **194**, 729–731 (1991).
63. 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).
64. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. & Cormier, M. J. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* **111**, 229–233 (1992).
65. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. Green fluorescent protein as a marker for gene expression. *Science (80- )*. **263**, 802–805 (1994).
66. Yang, T. T., Cheng, L. & Kain, S. R. Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* **24**, 4592–4593 (1996).
67. Patterson, G. H., Knobel, S. M., Sharif, W. D., Kain, S. R. & Piston, D. W. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophys. J.* **73**, 2782–2790 (1997).
68. Chudakov, D. M., Matz, M. V., Lukyanov, S. & Lukyanov, K. A. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.* **90**, 1103–1163 (2010).
69. WATSON, J. D. & CRICK, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).
70. Dickerson, R. *et al.* The anatomy of A-, B-, and Z-DNA. *Science (80- )*. **216**, 475–485 (1982).
71. Hoogsteen, K. The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **12**, 822–823 (1959).
72. Nikolova, E. N. *et al.* A historical account of Hoogsteen base-pairs in duplex DNA. *Biopolymers* **23**, n/a-n/a (2013).
73. Nikolova, E. N., Goh, G. B., Brooks, C. L. & Al-Hashimi, H. M. Characterizing the protonation state of cytosine in transient G·C Hoogsteen base pairs in duplex DNA. *J. Am. Chem. Soc.* **135**, 6766–6769 (2013).
74. Dingley, A. J. *et al.* Internucleotide scalar couplings across hydrogen bonds in Watson-Crick and Hoogsteen base pairs of a DNA triplex. *J. Am. Chem. Soc.* **121**, 6019–6027 (1999).
75. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).

76. Marin-Gonzalez, A., Vilhena, J. G., Perez, R. & Moreno-Herrero, F. A molecular view of DNA flexibility. *Q. Rev. Biophys.* **54**, e8 (2021).
77. Roth, E., Glick Azaria, A., Girshevitz, O., Bitler, A. & Garini, Y. Measuring the Conformation and Persistence Length of Single-Stranded DNA Using a DNA Origami Structure. *Nano Lett.* **18**, 6703–6709 (2018).
78. Bockelmann, U., Essevaz-Roulet, B. & Heslot, F. DNA strand separation studied by single molecule force measurements. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* **58**, 2386–2394 (1998).
79. Jain, S. & Larson, R. Validity of the bead-spring model for describing the linear viscoelastic properties of single-strand DNA under strongly denaturing conditions. *Rheol. Acta* **46**, 1153–1160 (2007).
80. Chi, Q., Wang, G. & Jiang, J. The persistence length and length per base of single-stranded DNA obtained from fluorescence correlation spectroscopy measurements using mean field theory. *Phys. A Stat. Mech. its Appl.* **392**, 1072–1079 (2013).
81. Rao, A. N. & Grainger, D. W. Biophysical properties of nucleic acids at surfaces relevant to microarray performance. *Biomater. Sci.* **2**, 436–471 (2014).
82. DeLong, E., Wickham, G. & Pace, N. Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science (80-. )*. **243**, 1360–1363 (1989).
83. Amann, R. I., Krumholz, L. & Stahl, D. A. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* **172**, 762–770 (1990).
84. Hirsch, F. R. *et al.* Increased EGFR Gene Copy Number Detected by Fluorescent In Situ Hybridization Predicts Outcome in Non-Small-Cell Lung Cancer Patients Treated With Cetuximab and Chemotherapy. *J. Clin. Oncol.* **26**, 3351–3357 (2008).
85. Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci.* **101**, 15275–15278 (2004).
86. Rothmund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
87. Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564–574 (2006).
88. Petersheim, M. & Turner, D. H. Base-Stacking and Base-Pairing Contributions to Helix Stability: Thermodynamics of Double-Helix Formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry* **22**, 256–263 (1983).
89. Leonard, G. A., Hunter, W. N. & Brown, T. Studies of the structure and stability of base pair mismatches, base pairs involving modified bases, and DNA-drug complexes. *J. Chem. Crystallogr.* **24**, 5–15 (1994).
90. Rossetti, G. *et al.* The structural impact of DNA mismatches. *Nucleic Acids Res.* **43**, 4309–4321 (2015).
91. Chen, C., Wang, W., Wang, Z., Wei, F. & Zhao, X. S. Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization. *Nucleic Acids Res.* **35**, 2875–2884 (2007).

92. Yin, Y. & Zhao, X. S. Kinetics and dynamics of DNA hybridization. *Acc. Chem. Res.* **44**, 1172–1181 (2011).
93. Sugimoto, N., Nakano, S. I., Yoneyama, M. & Honda, K. I. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.* **24**, 4501–4505 (1996).
94. SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 415–440 (2004).
95. Kuznetsov, S. V. & Ansari, A. A kinetic zipper model with intrachain interactions applied to nucleic acid hairpin folding kinetics. *Biophys. J.* **102**, 101–111 (2012).
96. Schildkraut, C. & Lifson, S. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* **3**, 195–208 (1965).
97. Owczarzy, R. *et al.* Predicting sequence-dependent melting stability of short duplex DNA oligomers. *Biopolymers* **44**, 217–239 (1997).
98. Fuchs, B. M., Syutsubo, K., Ludwig, W. & Amann, R. In Situ Accessibility of Escherichia coli 23S rRNA to Fluorescently Labeled Oligonucleotide Probes In Situ Accessibility of Escherichia coli 23S rRNA to Fluorescently Labeled Oligonucleotide Probes Downloaded from <http://aem.asm.org/> on December 11, 2013. *Appl. Environ. Microbiol.* **67**, 961–968 (2001).
99. Mullis, K. *et al.* Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51**, 263–273 (1986).
100. Bielec, K. *et al.* Kinetics and equilibrium constants of oligonucleotides at low concentrations. Hybridization and melting study. *Phys. Chem. Chem. Phys.* **21**, 10798–10807 (2019).
101. Xu, S. *et al.* Real-time reliable determination of binding kinetics of DNA hybridization using a multi-channel graphene biosensor. *Nat. Commun.* **8**, 1–10 (2017).
102. Owczarzy, R. *et al.* Effects of Sodium Ions on DNA Duplex Oligomers: Improved Predictions of Melting Temperatures. *Biochemistry* **43**, 3537–3554 (2004).
103. Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A. & Wälder, J. A. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* **47**, 5336–5353 (2008).
104. Chiu, T. K. & Dickerson, R. E. 1 Å crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium. *J. Mol. Biol.* **301**, 915–945 (2000).
105. Guéroult, M., Boittin, O., Mauffret, O., Etchebest, C. & Hartmann, B. Mg<sup>2+</sup> in the Major Groove Modulates B-DNA Structure and Dynamics. *PLoS One* **7**, e41704 (2012).
106. Duguid, J. G., Bloomfield, V. A., Benevides, J. M. & Thomas, G. J. Raman spectroscopy of DNA-metal complexes. II. The thermal denaturation of DNA in the presence of Sr<sup>2+</sup>, Ba<sup>2+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, Mn<sup>2+</sup>, Co<sup>2+</sup>, Ni<sup>2+</sup>, and Cd<sup>2+</sup>. *Biophys. J.* **69**, 2623–2641 (1995).
107. Hackl, E. V. *et al.* Study of Ca<sup>2+</sup>, Mn<sup>2+</sup> and Cu<sup>2+</sup> binding to DNA in solution by means of IR spectroscopy. *J. Mol. Struct.* **408–409**, 229–232 (1997).
108. Oliva, R. & Cavallo, L. Frequency and effect of the binding of Mg<sup>2+</sup>, Mn<sup>2+</sup>,

- and Co<sup>2+</sup> ions on the guanine base in watson-crick and reverse watson-crick base pairs. *J. Phys. Chem. B* **113**, 15670–15678 (2009).
109. Bivehed, E., Söderberg, O. & Hellman, B. Flash-comet: Significantly improved speed and sensitivity of the comet assay through the introduction of lithium-based solutions and a more gentle lysis. *Mutat. Res. - Genet. Toxicol. Environ. Mutagen.* **858–860**, (2020).
  110. Stewart, R. & Harris, M. G. Amino group acidity in nucleotide bases. *Can. J. Chem.* **55**, 3807–3814 (1977).
  111. Verdolino, V., Cammi, R., Munk, B. H. & Schlegel, H. B. Calculation of p K<sub>a</sub> Values of Nucleobases and the Guanine Oxidation Products Guanidinohydantoin and Spiroiminodihydantoin using Density Functional Theory and a Polarizable Continuum Model. *J. Phys. Chem. B* **112**, 16860–16873 (2008).
  112. Ashwood, B., Sanstead, P. J., Dai, Q., He, C. & Tokmakoff, A. 5-Carboxylcytosine and Cytosine Protonation Distinctly Alter the Stability and Dehybridization Dynamics of the DNA Duplex. *J. Phys. Chem. B* **124**, 627–640 (2020).
  113. Blake, R. & Delcourt, S. Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.* **24**, 2095–2103 (1996).
  114. Raghunathan, S., Jaganade, T. & Priyakumar, U. D. Urea-aromatic interactions in biology. *Biophys. Rev.* **12**, 65–84 (2020).
  115. Wong, K. L. & Liu, J. Factors and methods to modulate DNA hybridization kinetics. *Biotechnol. J.* **16**, 2000338 (2021).
  116. Nakano, S. ichi & Sugimoto, N. The structural stability and catalytic activity of DNA and RNA oligonucleotides in the presence of organic solvents. *Biophys. Rev.* **8**, 11–23 (2016).
  117. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science (80-. )*. **254**, 1497–1500 (1991).
  118. Park, H. *et al.* Effect of ionic strength on PNA-DNA hybridization on surfaces and in solution. *Biointerphases* **2**, 80–88 (2007).
  119. SUMMERTON, J. & WELLER, D. Morpholino Antisense Oligomers: Design, Preparation, and Properties. *Antisense Nucleic Acid Drug Dev.* **7**, 187–195 (1997).
  120. Koshkin, A. A. *et al.* LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607–3630 (1998).
  121. Ghobadi, A. F. & Jayaraman, A. Effect of backbone chemistry on hybridization thermodynamics of oligonucleic acids: A coarse-grained molecular dynamics simulation study. *Soft Matter* **12**, 2276–2287 (2016).
  122. Mills, A. P., Simmel, F. C., Turberfield, A. J., Neumann, J. L. & Yurke, B. A DNA-fuelled molecular machine made of DNA. *Nature* **406**, 605–608 (2002).
  123. Zhang, D. Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **131**, 17303–17314 (2009).

124. Biala, K., Sedova, A. & Flechsig, G. U. Sequence and temperature influence on kinetics of DNA strand displacement at gold electrode surfaces. *ACS Appl. Mater. Interfaces* **7**, 19948–19959 (2015).
125. Machinek, R. R. F., Ouldridge, T. E., Haley, N. E. C., Bath, J. & Turberfield, A. J. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* **5**, 1–9 (2014).
126. Duose, D. Y. *et al.* Configuring robust DNA strand displacement reactions for in situ molecular analyses. *Nucleic Acids Res.* **40**, 3289–3298 (2012).
127. Yurke, B. & Mills, A. P. Using DNA to power nanostructures. *Genet. Program. Evolvable Mach.* **4**, 111–122 (2003).
128. Ang, Y. S. & Yung, L. Y. L. Rational design of hybridization chain reaction monomers for robust signal amplification. *Chem. Commun.* **52**, 4219–4222 (2016).
129. Xuan, F. & Hsing, I. M. Triggering hairpin-free chain-branching growth of fluorescent DNA dendrimers for nonlinear hybridization chain reaction. *J. Am. Chem. Soc.* **136**, 9810–9813 (2014).
130. Bi, S., Chen, M., Jia, X., Dong, Y. & Wang, Z. Hyperbranched Hybridization Chain Reaction for Triggered Signal Amplification and Concatenated Logic Circuits. *Angew. Chemie - Int. Ed.* **54**, 8144–8148 (2015).
131. Choi, H. M. T. *et al.* Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Dev.* **145**, 1–10 (2018).
132. Choi, H. M. T. *et al.* Mapping a multiplexed zoo of mRNA expression. *Dev.* **143**, 3632–3637 (2016).
133. Choi, H. M. T., Beck, V. A. & Pierce, N. A. Next-Generation in Situ Hybridization Chain Reaction: Higher Gain, Lower Cost, Greater Durability. *ACS Nano* **8**, 4284–4294 (2014).
134. Guo, J., Mingo, C., Qiu, X. & Hildebrandt, N. Simple, Amplified, and Multiplexed Detection of MicroRNAs Using Time-Gated FRET and Hybridization Chain Reaction. *Anal. Chem.* **91**, 3101–3109 (2019).
135. Schwarzkopf, M. & Pierce, N. A. Multiplexed miRNA northern blots via hybridization chain reaction. *Nucleic Acids Res.* **44**, e129 (2016).
136. Li, P. *et al.* A hybridization-chain-reaction-based method for amplifying immunosignals. *Nat. Methods* **15**, 275–278 (2018).
137. Ang, Y. S., Tong, R. & Yung, L. Y. L. Engineering a robust DNA split proximity circuit with minimized circuit leakage. *Nucleic Acids Res.* **44**, 0–8 (2016).
138. Huang, J. *et al.* Fluorescence resonance energy transfer-based hybridization chain reaction for: In situ visualization of tumor-related mRNA. *Chem. Sci.* **7**, 3829–3835 (2016).
139. Michelson, A. M. & Todd, A. R. Nucleotides part XXXII. Synthesis of a dithymidine dinucleotide containing a 3': 5'-internucleotidic linkage. *J. Chem. Soc.* **22**, 2632–2638 (1955).
140. Khorana, H. G., Razzell, W. E., Gilham, P. T., Tener, G. M. & Pol, E. H. SYNTHESIS OF DIDEOXYRIBONUCLEOTIDES. *J. Am. Chem. Soc.* **79**, 1002–1003 (1957).

141. Gilham, P. T. & Khorana, H. G. Studies on Polynucleotides. I. A New and General Method for the Chemical Synthesis of the C 5''-C 3'' Internucleotidic Linkage. Syntheses of Deoxyribo-dinucleotides 1. *J. Am. Chem. Soc.* **80**, 6212–6222 (1958).
142. Reese, C. B. Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Org. Biomol. Chem.* **3**, 3851–3868 (2005).
143. Schaller, H., Weimann, G., Lerch, B. & Khorana, H. G. Studies on Polynucleotides. XXIV. 1 The Stepwise Synthesis of Specific Deoxyribopolynucleotides (4). 2 Protected Derivatives of Deoxyribonucleosides and New Syntheses of Deoxyribonucleoside-3'' Phosphates 3. *J. Am. Chem. Soc.* **85**, 3821–3827 (1963).
144. Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
145. Letsinger, R. L. & Mahadevan, V. Oligonucleotide Synthesis on a Polymer Support 1,2. *J. Am. Chem. Soc.* **87**, 3526–3527 (1965).
146. Letsinger, R. L. & Ogilvie, K. K. Convenient method for stepwise synthesis of oligothymidylate derivatives in large-scale quantities. *J. Am. Chem. Soc.* **89**, 4801–4803 (1967).
147. Letsinger, R. L., Ogilvie, K. K. & Miller, P. S. Nucleotide chemistry. XV. Developments in syntheses of oligodeoxyribonucleotides and their organic derivatives. *J. Am. Chem. Soc.* **91**, 3360–3365 (1969).
148. Eckstein, F. & Rizk, I. Synthesis of Oligonucleotides by Use of Phosphoric Triesters. *Angew. Chemie Int. Ed. English* **6**, 695–696 (1967).
149. Eckstein, F. & Rizk, I. Oligonucleotide Syntheses by means of 2,2,2-Trichloroethyl Phosphorodichoridate. *Angew. Chemie Int. Ed. English* **6**, 949–949 (1967).
150. Caruthers, M. H. The Chemical Synthesis of DNA/RNA: Our Gift to Science. *J. Biol. Chem.* **288**, 1420–1427 (2013).
151. Beaucage, S. L. & Caruthers, M. H. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
152. Caruthers, M. H. A brief review of DNA and RNA chemical synthesis. *Biochem. Soc. Trans.* **39**, 575–580 (2011).
153. Hughes, R. A. & Ellington, A. D. Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb. Perspect. Biol.* **9**, a023812 (2017).
154. Martin, L. J. DNA damage and repair: Relevance to mechanisms of neurodegeneration. *J. Neuropathol. Exp. Neurol.* **67**, 377–387 (2008).
155. Filges, S., Mouhanna, P. & Ståhlberg, A. Digital Quantification of Chemical Oligonucleotide Synthesis Errors. *Clin. Chem.* **67**, 1384–1394 (2021).
156. LeProust, E. M. *et al.* Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. *Nucleic Acids Res.* **38**, 2522–2540 (2010).
157. Barth, H. G. & Boyes, B. E. Size Exclusion Chromatography. *Anal. Chem.* **62**, 268–303 (1990).
158. Xiao, W. & Oefner, P. J. Denaturing High-Performance Liquid Chromatography : A Review. **474**, 439–474 (2001).

159. Monserud, J. H. & Schwartz, D. K. Effects of Molecular Size and Surface Hydrophobicity on Oligonucleotide Interfacial Dynamics. *Biomacromolecules* **13**, 4002–4011 (2012).
160. Walton, H. Ion Exchange Chromatography. *Anal. Chem.* **36**, 51–55 (1964).
161. Cook, K. & Thayer, J. Advantages of ion-exchange chromatography for oligonucleotide analysis. *Bioanalysis* **3**, 1109–1120 (2011).
162. Albright, L. M. & Slatko, B. E. Denaturing Polyacrylamide Gel Electrophoresis. *Curr. Protoc. Nucleic Acid Chem.* **00**, 1–5 (2000).
163. AGARWAL, K. L. *et al.* Total Synthesis of the Gene for an Alanine Transfer Ribonucleic Acid from Yeast. *Nature* **227**, 27–34 (1970).
164. Au, L. C., Yang, F. Y., Yang, W. J., Lo, S. H. & Kao, C. F. Gene synthesis by a LCR-based approach: High-level production of leptin-L54 using synthetic gene in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **248**, 200–203 (1998).
165. Hughes, R. A. & Ellington, A. D. Synthetic DNA synthesis and assembly: Putting the synthetic in synthetic biology. *Cold Spring Harb. Perspect. Biol.* **9**, (2017).
166. Stemmer, W. P. C., Cramer, A., Ha, K. D., Brennan, T. M. & Heyneker, H. L. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**, 49–53 (1995).
167. Wittig, I., Braun, H. P. & Schägger, H. Blue native PAGE. *Nat. Protoc.* **1**, 418–428 (2006).
168. Wittig, I. & Schägger, H. Native electrophoretic techniques to identify protein-protein interactions. *Proteomics* **9**, 5214–5223 (2009).
169. Vasilescu, J. & Figeys, D. Mapping protein-protein interactions by mass spectrometry. *Curr. Opin. Biotechnol.* **17**, 394–399 (2006).
170. Free, R. B., Hazelwood, L. A. & Sibley, D. R. Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy. *Curr. Protoc. Neurosci.* 1–19 (2009). doi:10.1002/0471142301.ns0528s46
171. Louche, A., Salcedo, S. P. & Bigot, S. Protein-Protein Interactions: Pull-Down Assays. in 247–255 (2017). doi:10.1007/978-1-4939-7033-9\_20
172. Tomizaki, K. Y., Usui, K. & Mihara, H. Protein-protein interactions and selection: Array-based techniques for screening disease-associated biomarkers in predictive/early diagnosis. *FEBS J.* **277**, 1996–2005 (2010).
173. Fields, S. & Song, O. K. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).
174. Ma, Y., Nagamune, T. & Kawahara, M. Split focal adhesion kinase for probing protein-protein interactions. *Biochem. Eng. J.* **90**, 272–278 (2014).
175. Martell, J. D. *et al.* A split horseradish peroxidase for the detection of intercellular protein-protein interactions and sensitive visualization of synapses. *Nat. Biotechnol.* **34**, 774–780 (2016).
176. Paulmurugan, R. & Gambhir, S. S. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Anal. Chem.* **75**, 1584–1589 (2003).
177. Kerppola, T. K. Bimolecular Fluorescence Complementation (BiFC) Analysis as a Probe of Protein Interactions in Living Cells. *Annu. Rev. Biophys.* **37**, 465–487 (2008).

178. Van Crielinge, W. & Beyaert, R. Yeast two-hybrid: State of the art. *Biol. Proced. Online* **2**, 1–38 (1999).
179. Shroff, H. *et al.* Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. *Proc. Natl. Acad. Sci.* **104**, 20308–20313 (2007).
180. Liu, Z. *et al.* Super-resolution imaging and tracking of protein–protein interactions in sub-diffraction cellular space. *Nat. Commun.* **5**, 4443 (2014).
181. Sorkin, A., McClure, M., Huang, F. & Carter, R. Interaction of EGF receptor and Grb2 in living cells visualized by fluorescence resonance energy transfer (FRET) microscopy. *Curr. Biol.* **10**, 1395–1398 (2000).
182. Clegg, R. M. Fluorescence resonance energy transfer. *Curr. Opin. Biotechnol.* **6**, 103–110 (1995).
183. Fredriksson, S. *et al.* Protein detection using proximity-dependent DNA ligation assays. *Nat. Biotechnol.* **20**, 473–477 (2002).
184. Raykova, D. *et al.* A method for Boolean analysis of protein interactions at a molecular level. *Nat. Commun.* **13**, 1–17 (2022).
185. Pellegrino, C. *et al.* Impact of Ligand Size and Conjugation Chemistry on the Performance of Universal Chimeric Antigen Receptor T-Cells for Tumor Killing. *Bioconjug. Chem.* **31**, 1775–1783 (2020).
186. Sadiki, A. *et al.* Site-specific Bioconjugation and Convergent Click Chemistry Enhances Antibody–Chromophore Conjugate Binding Efficiency. *Photochem. Photobiol.* **96**, 596–603 (2020).
187. Kazlauskas, A. & Cooper, J. A. Autophosphorylation of the PDGF receptor in the kinase insert region regulates interactions with cell proteins. *Cell* **58**, 1121–1133 (1989).
188. Heldin, C.-H. & Westermark, B. Mechanism of Action and In Vivo Role of Platelet-Derived Growth Factor. *Physiol. Rev.* **79**, 1283–1316 (1999).
189. Tsuneoka, Y. & Funato, H. Modified in situ Hybridization Chain Reaction Using Short Hairpin DNAs. *Front. Mol. Neurosci.* **13**, 1–14 (2020).
190. Goddard, N. L., Bonnet, G., Krichevsky, O. & Libchaber, A. Sequence dependent rigidity of single stranded DNA. *Phys. Rev. Lett.* **85**, 2400–2403 (2000).

# Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations  
from the Faculty of Pharmacy 326*

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](http://publications.uu.se)  
urn:nbn:se:uu:diva-496889



ACTA  
UNIVERSITATIS  
UPSALIENSIS  
UPPSALA  
2023