Technological Advancements in Affinity Proteomics

From planar antibody microarrays towards a solution-based platform

Brofelth, Mattias

2018

Document Version:
Publisher's PDF, also known as Version of record

Technological Advancements in Affinity Proteomics

From planar antibody microarrays towards a solution-based platform

Mattias Brofelth

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Engineering, Lund University, Sweden.
To be defended at Hörsalen, Medicon Village, Lund,
Friday 23 February 2018 at 13:15

Faculty opponent
Prof. Amelie Eriksson Karlström
School of Biotechnology, KTH - Royal Institute of Technology,
Stockholm, Sweden

Organization LUND UNIVERSITY	Document name: Doctoral dissertation
Department of Immunotechnology Medicon Village (Building 406) SE-22381 LUND	Date of issue: 2018-01-29
Author: Mattias Brofelth	Sponsoring organization

Title and subtitle: Technological Advancements in Affinity Proteomics - From planar antibody microarrays towards a solution-based platform

Abstract: Proteomics has the potential to deliver disease-associated biomarkers that could provide diagnostic, prognostic and predictive information to enable precision medicine. Affinity proteomics, most commonly based on antibodies and their ability to specifically capture target proteins, has emerged has a valuable tool in biomarker discovery. Our group has developed a recombinant antibody microarray platform that can be used for protein expression profiling of serum samples to define multiplex biomarker signatures. This thesis is focused on antibody engineering and assay development to further improve the current microarray platform and also present proof-of-concept for a novel solution-based platform.

In Paper I we evaluated the novel detection reagent PID in search for increased signal-to-noise ratio and improved sensitivity of the microarray assay. PID is a fluorophore-packed nanoparticle and was here used to replace the currently employed single fluorophore molecule. The result showed that it was possible to use PID as a detection reagent in our assay and even higher signals were achieved, although accompanied by a heterogeneous background that will require further optimization.

In Paper II and III we explored the Dock'n'Flash method for site-specific conjugation to enable oriented immobilization or functionalization of scFvs. Immobilizing the scFvs in an oriented configuration on the slide could lead to increased sensitivity and performance of the microarray assay. Functionalization could enable novel scFv applications. The scFvs were equipped with the unnatural amino acid *p*Bpa and photocrosslinked to beta-cyclodextrin on a coated slide or in solution. Proof-of-concept was demonstrated for one scFv in Paper II and the study was expanded and the *p*Bpa position optimized in Paper III.

In Paper IV we sought to overcome some inherent limitations associated with planar microarrays for global serum profiling by developing the solution-based MIAS platform. In MIAS, proteins were displayed on beads and quantified via DNA-barcoded scFvs using next generation sequencing (NGS). Sortase A was used to site-specifically conjugate the oligonucleotide barcode to scFvs. Proof-of-concept for the assay steps was demonstrated using barcoded scFvs targeting three different serum proteins.

In summary, the work presented in this thesis can be used to improve the performance of our current antibody microarray platform and also provides the first steps towards a novel solution-based platform. This could in turn enable improved development of disease-associated protein biomarkers.

Key words: Affinity proteomics, antibody microarrays, scFv, biomarkers, MIAS

Classification system and/or index terms (if any)

Supplementary bibliographical information		Language: English
ISSN and key title		ISBN 978-91-7753-563-8 (Print) ISBN 978-91-7753-564-5 (PDF)
Recipient's notes	Number of pages	Price
	Security classification	

Signature _____ Date _____

Technological Advancements in Affinity Proteomics

From planar antibody microarrays towards a solution-based platform

Mattias Brofelth

LUND
UNIVERSITY

Coverphoto by Mattias Brofelth

Faculty of Engineering
Department of Immunotechnology

ISBN 978-91-7753-563-8 (Print)
ISBN 978-91-7753-564-5 (PDF)

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

Contents

Original papers

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

I. M. Brofelth, E. Futaya, R. Erhnström, Y. Motokui, T. Shiraishi, C. Wingren. *Evaluation of a novel fluorescent nanoparticle for detection of planar multiplexed recombinant antibody microarrays.* Manuscript.

II. L. Petersson, L.W. Städe, M. Brofelth, S. Gartner, E. Fors, M. Sandgren, J. Vallkil, N. Olsson, K.L. Larsen, C. Borrebaeck, L. Duroux, C. Wingren. *Molecular design of recombinant scFv antibodies for site-specific photocoupling to beta-cyclodextrin in solution and onto solid support.* Biochim. Biophys. Acta (2014)

III. M. Brofelth, L.W. Städe, A. I. Ekstrand, L. P. Edfeldt , R. Kovačič, T. T. Nielsen, K. L. Larsen, L. Duroux. C. Wingren. *Site-specific photocoupling of pBpa mutated scFv antibodies for use in affinity proteomics.* Biochim. Biophys. Acta (2017)

IV. A. I. Ekstrand*, M. Brofelth*, T. Törngren, R. Jansson, S. Gour, M. Hedhammar, C. Borrebaeck, A. Kvist, C. Wingren. *MIAS – A Multiplex Immuno-Assay for sensitive protein expression profiling.* Manuscript.

* Shared first authors.

My contribution to the papers

Paper I — I took part in planning and performed all the experiments, analyzed the data and had a major role in writing the manuscript.

Paper II — I took part in planning and performed some of the experiments, analyzed the corresponding data and participated in writing the manuscript.

Paper III — I took part in planning and performed a large part of the experiments, analyzed the data and had a major role in writing the manuscript.

Paper IV — I took part in planning and performed a large part of the experiments, analyzed the corresponding data and had a major role in writing the manuscript.

Abbreviations

CDR	Complementarity determining region
Fab	Antigen-binding fragment
Fc	Fragment crystallizable
Ig	Immunoglobulin
LoD	Limit of detection
mAb	Monoclonal antibody
MIAS	Multiplex Immuno-Assay in Solution
MS	Mass spectrometry
NGS	Next-generation sequencing
pAb	Polyclonal antibody
*p*Bpa	*p*-benzoyl-L-phenylalanine
PCR	Polymerase chain reaction
PID	Phosphor-Integrated Dot
rAb	Recombinant antibody
scFv	single-chain Fragment variable
SrtA	Sortase A
UAA	Unnatural amino acid
VH	Variable domain, heavy chain
VL	Variable domain, light chain

1. Introduction

Our knowledge about life at the molecular level continues to grow and enables us to better understand the mechanisms behind health and disease. Groundbreaking discoveries were made in the middle of the twentieth century with the elucidation of the DNA double helix structure and how the genetic code is linked to RNA and proteins (Crick et al., 1961; Nirenberg and Matthaei, 1961; Watson and Crick, 1953). This opened up for the rapid development of the branch of science that would later be called the life sciences. The realization that DNA was the carrier of our hereditary information lead to intense studies in the field of *genomics*. The Human Genome Project (HGP) started in 1990 and together with Celera Genomics the complete sequence of the whole human genome was presented in 2003 (International Human Genome Sequencing, 2004). More than 20,000 protein-coding genes were found and hopes were that studying abnormalities in these genes would explain the cause of different diseases. Indeed, some genetic markers have been found but it has also become evident that genomics alone cannot explain the complex biology that is present at a given point in time (Graves and Haystead, 2002). More has been discovered about how genes are regulated and RNA has been found to not merely be an intermediate between DNA and proteins. As proteins carries out most of life's functions, the study of proteins - *proteomics* - has attracted increased attention in order to understand the present status in our bodies. A research area within proteomics aims to explore the proteome in search for measureable characteristics that can improve disease treatment, so called disease-associated protein *biomarkers*. There are many diseases, not least cancers, where *precision medicine* enabled by earlier detection and more precise diagnosis could save lives and reduce the suffering of many. As the proteome is enormously complex in terms of abundance, variety and interactions, it provides both opportunities as well as challenges for biomarker discovery. One way to tackle the complexity has been through *affinity-based proteomics*, where *antibodies* have been the main workhorse. With their natural binding abilities, antibodies can fish out specific proteins to allow measurements and detailed studies. This has been the main advantage over *mass spectrometry*, the other major proteomic technology.

Our research group is working with technology development and application of antibody-based methods with the aim to discover candidate biomarkers. The search is mainly focused within the blood proteome, due to the rich source of potential protein biomarkers it provides and because the sample format is easily accessible. Among the proteins that are being transported in the blood are many regulators of the immune system (Anderson and Anderson, 2002). As the immune system is constantly responding to threats and deviations in our body, our hypothesis is that immuno-regulatory proteins that specifically respond to a certain disease can be used as biomarkers to detect even very early signs of disease. To ensure specificity and sensitivity, the collective response of many proteins are simultaneously measured in so called *immunosignatures*. A number of candidate biomarker signatures for various cancers and autoimmune diseases have already been identified using our in-house developed *antibody microarrays* (Borrebaeck et al., 2014; Delfani et al., 2016; Wingren et al., 2012). The microarrays are produced by printing microscopic droplets of *single-chain fragment variable (scFv) antibodies* onto a solid surface. scFv is a recombinant antibody format that only contains the variable antigen-binding domains and can be engineered for improved performance and new functions (Borrebaeck and Wingren, 2011).

The work presented in this thesis is focused on antibody engineering and further technological development in antibody-based proteomics to enable improved and expanded protein expression profiling. Although the current antibody microarray platform allows simultaneous detection of a couple of hundred target proteins, further expansion in multiplexity is limited by physical and practical restrictions. Blood contains thousands of proteins and increased proteome coverage could potentially allow even better biomarker discovery. Improved detection and sensitivity could also allow better measurements of low-abundant proteins. Furthermore, there are certain assay steps that are laborious and difficult to automate which would be beneficial to circumvent in a high-throughput setting. These reasons motivated the continued developmental efforts presented in **Paper I-IV** upon which this thesis is based.

In **Paper I** we evaluated a novel detection reagent in search for increased signal-to-noise ratio and improved sensitivity of the microarray assay. In the current set-up, biotinylated target proteins are captured by the antibodies and subsequently detected using the fluorophore Alexa-647 coupled to streptavidin. In this study, the Alexa-647 molecule was replaced by a novel fluorophore-packed nanoparticle called PID which was for the first time used in a microarray assay. The results

showed that it was possible to use PID in microarrays and higher signals were found in comparison to Alexa-647. The nanoparticle was however accompanied by a heterogeneous background issues that will require further optimization.

In **Paper II** we explored the novel engineering method called Dock'n'Flash for site-specific conjugation of our scFvs. Such conjugation could be used to immobilize the scFvs in an oriented configuration on the slide that could lead to increased sensitivity and performance of the microarray assay. Further, it could also enable functionalization of the scFvs. Dock'n'Flash is based on covalent photocrosslinking of the unnatural amino acid *p*Bpa to the ligand beta-cyclodextrin. *p*Bpa was for the first time incorporated in a scFv which allowed covalent site-specific conjugation to a microarray slide coated with beta-cyclodextrin. In addition, conjugation was also demonstrated to beta-cyclodextrin in solution which opened up for site-specific functionalization of scFvs.

The Dock'n'Flash method was further evaluated in **Paper III** with the aim to optimize the site for *p*Bpa incorporation in our scFvs. Four different scFv were used as model clones to evaluate 13 different *p*Bpa sites and their individual performance were compared to the corresponding wild-type antibodies. Conjugation to beta-cyclodextrin was performed both in solution and to a novel coated surface. A position in the C-terminal affinity tag called T7 was selected as the best candidate. The results confirmed that Dock'n'Flash could be used for oriented immobilization and that the method could enable future scFv functionalization studies.

In **Paper IV** we sought to overcome some inherent limitations with planar microarrays by developing the solution-based platform Multiplex Immuno-Assay in Solution (MIAS). MIAS was designed to potentially enable higher sensitivity, multiplexity and sample throughput as well as allow a higher degree of automation and direct digital read-out. In MIAS, the biotinylated sample proteins are first captured to magnetic beads coated with streptavidin. Next, scFvs site-specifically conjugated to DNA-barcodes are added and allowed to bind, after which unbound scFvs are washed away. The target proteins can then be quantified via the barcode by next generation sequencing (NGS). Proof-of-concept for the assay steps was demonstrated using antibodies targeting three different proteins.

2. Disease-associated biomarkers

2.1 Biomarkers for precision medicine

The U.S. National Institutes of Health (NIH) defined a biomarker as (Biomarkers Definitions Working Group, 2001):

> "A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."

Disease-associated biomarkers generally refer to biological substances, such as genes or proteins, that can be correlated to a certain disease or disease state. The aim with identifying such biomarkers is to aid clinicians in making informed decisions on how to best treat a patient.

Biomarkers can be classified into three main categories: *diagnostic*, *prognostic* and *predictive* biomarkers (Frantzi et al., 2014). The diagnostic group include biomarkers that can be used to screen asymptomatic populations for early signs of disease as well as biomarkers that can precisely divide patients into disease subgroups. Prognostic biomarkers are useful to give a prognosis for the disease progression and the likelihood of reoccurrence and survival. Predictive biomarkers can predict the response to treatment and thereby guide the choice of therapy.

Finding such biomarkers would make an important contribution to enable tailoring of the treatment to an individual patient's profile, so called *precision medicine*. The ultimate goal with personalized medicine is to give the right drug, for the right patient, at the right dose and time (Sadee and Dai, 2005). Earlier and more precise diagnosis would furthermore minimize "trial-and-error" medication and lead to less over-treatment and side-effects as well as quicker recovery (Mathur and Sutton, 2017).

An example where early detection by biomarkers could have a major impact is in pancreatic cancer. Today, pancreatic ductal adeno-carcinoma (PDAC) is most commonly diagnosed at a late stage where the 5-year survival rate is under 5% (Siegel et al., 2017). This number can however be increased to around 50% if the tumor is detected and removed at an early stage (Furukawa et al., 1996; Matsuno et al., 2004). PDAC is one of the diseases that has been studied with our antibody microarrays and for which candidate immunosignatures have been found (Wingren et al., 2012; Gerdtsson et al., 2015; Gerdtsson et al., 2016).

2.2 Protein biomarker discovery

As touched upon in the introduction, searching for biomarkers in proteomes has many challenges due to the sample complexity. As each of the around 20,000 protein-coding genes can result in many different proteoforms, owing to alterations such as alternative splicing and post-translational modifications, the human proteome is estimated to consist of millions of different proteins (Smith et al., 2013; Ponomarenko et al., 2016). The number of proteforms varies between and within cells and tissue types, from person to person and over time. The protein abundance also spans over a wide dynamic range with more than 10 orders of magnitude in plasma (Anderson and Anderson, 2002). Furthermore, proteins are difficult to study as they are limited in quantity and cannot be amplified in the same manner as genes (Tyers and Mann, 2003). The twenty natural amino acids that make up proteins also have various chemical properties and can be combined and folded into complex structures in numerous ways. Moreover, careful sample handling is important as proteins risk denaturation and degradation after sampling.

At the same time, the complexity is what makes the proteome attractive for biomarker discovery as the protein levels are dynamic and change in response to disease. The challenge to manage the complexity has not stopped ambitious efforts to explore the proteome. The Human Proteome Project (HPP) was launched in 2010 by the Human Proteome Organization (HUPO) with the initial aim to create a database with at least one representative protein from every human gene (HUPO, 2010). This database is intended to include a parts list, a distribution

atlas, and a pathway and network map to create a backbone for further proteome research.

Fundamental to the exploration of the proteome in search for disease-associated biomarkers is the development of proteomic assays, which is the focus of the work presented in this thesis (**Paper I-IV**). Different technological platforms that are commonly used in assay development will be presented in **Section 2.2.2** and **Chapter 4**. It should however be remembered that the technology is not the only consideration for a successful biomarker study. There have already been many candidate biomarkers reported in discovery studies, but few protein biomarkers have made it all the way to clinical implementation (Anderson et al., 2013). There are likely many reasons for the low success rate but the lack of a clear aim, i.e. intended use, has been identified as an important factor. The aim should be to answer a defined and relevant clinical question and should be the set already from the beginning so that an appropriate roadmap to reach this aim can be designed (Pavlou et al., 2013; Borrebaeck, 2017).

With a clear aim in mind, the development pipeline of biomarkers for clinical use can be divided into three main phases; *discovery*, *verification* and *validation* (**Figure 1**) (Frantzi et al., 2014). The discovery phase generally begins with a broad search to find differentially expressed proteins. The target proteins can include proteins that are already known to be associated with a certain disease, but a non-targeted identification by comparing cohorts is common. Candidate biomarkers that are identified in the discovery phase are then transferred into the verification phase, also known as pre-validation. A reduced number of targeted proteins often allows an increase in the number of analyzed samples. The verified proteins are finally evaluated for clinical utility in the validation phase and parameters such as assay sensitivity and specificity are defined. The first validation step (and also the previous discovery and verification phases) is usually performed on retrospective samples stored in biobanks. A second validation step is then performed as a prospective study to show clinical utility (Pavlou et al., 2013).

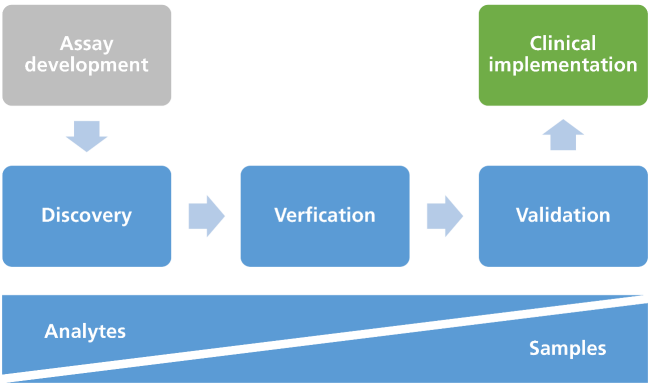

Figure 1. Discover, verification and validation are the three main phases of the biomarker development pipeline. Commonly, the discovery phase involves targeting a large number of proteins to identify those associated to the disease. Later in the pipeline, selected candidate biomarkers need to be verified and validated which requires a large number of samples.

2.2.2 Analytical platforms for proteomics

A key element in the search for biomarkers is the employed analytical platform. The platform must not necessary be the same throughout the development phases. During the discovery phase, the technology needs to be able to simultaneously analyze a large number of proteins (high *multiplexity*) and also detect low-abundant targets (high *sensitivity*) to cover as many potential biomarkers as possible. Once the candidate biomarker (signature) is found, the need for multiplexity is reduced, which in some cases can open up for other technological solutions that might be better suited for implementation in routine clinical practice (Fuzery et al., 2013). Requirements on high *sample throughput* is generally increased in the later developmental phases, although as discussed earlier, also important in the discovery phase.

2.2.2.1 Mass spectrometry

Protein biomarker discovery has mainly been pursued in two separate branches, *mass spectrometry (MS)* and *affinity assays*. In MS, a mass spectrometer is used to separate ions based on their mass-to-charge ratio (m/z) and consist of three main steps (Aebersold and Mann, 2003). The first step involves ionizing the proteins to later enable separation in a magnetic field. Ionization is mainly achieved using matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988) or electrospray ionization (ESI) (Fenn et al., 1989), where ESI is commonly

preferred for complex samples. The second step is the separation of ions based on their charge in a magnetic field in a mass analyzer. Finally, the ions are measured in a detector and the signal is correlated to the mass. By including a fragmentation step and a second mass analyzer, so called *tandem mass spectrometry (MS/MS)* is achieved that enables amino acid sequencing and more detailed analysis. The most widely employed MS approach is called *shotgun proteomics*. Here, the sample is first prepared by digesting the proteins into peptides with the enzyme trypsin and separated by liquid chromatography (LC) before MS/MS analysis.

The main advantages of MS are the specificity in identification and quantification of proteins (including modifications), and that MS can be performed without prior knowledge of the target (Aebersold and Mann, 2003). Drawbacks of MS for biomarker discovery has been the limited sensitivity, dynamic range and reproducibility when analyzing complex samples such as serum. Low-abundant proteins can be difficult to measure as they are masked by more abundant proteins. Some of these shortcomings have been addressed by reducing the complexity with pre-fractionation of the sample, although this can introduce bias and increase the time of the analysis. Targeted strategies has been developed such as *selected reaction monitoring (SRM)* (also called MRM - multiple reaction monitoring) (Lange et al., 2008). In SRM, only selected ions are allowed to enter the second mass analyzer which enables higher sensitivity. Still, protein concentrations below ng/ml are outside the dynamic range and the multiplexity is limited (Geyer et al., 2017). Another approach to achieve a wider proteome coverage has been developed based on *data independent acquisition (DIA)*, wherein the *Sequential Window Acquisition of All Theoretical Mass Spectra (SWATH-MS)* platform has gained most visibility (Anjo et al., 2017; Gillet et al., 2012). SWATH does not limit the number of peptides being analyzed, which is why quantification is possible over a wider dynamic range compared to shotgun MS.

2.2.2.2 Affinity Proteomics

The second technology branch for protein discovery is the affinity-based assays, the focus in this thesis (**Paper I-IV**). Originally, these assays were constructed using antibodies or antibody fragments and therefore called *immunoassays*. After development of novel types of binders, all assays are now included under the broader term *affinity proteomics*. Affinity-based assays have historically had an advantage in sensitivity compared to MS, due to the affinity binder's ability to fish out even low abundant proteins from complex samples. Affinity proteomics

and combinations of the both branches will later be discussed in more detail in **Chapter 4**.

2.2.2.3 Assay performance

A number of characteristics must be optimized in order to achieve a robust analytical assay (**Figure 2**). A fundamental property of an analytical instrument is to have a high enough *resolution* to enable detection of small variations in the sample. It should be noted that the term resolution is also used in related contexts and can for example mean the depth of proteome coverage in MS or the pixel size of a laser scanning. Furthermore, sufficient *precision*, the closeness of agreement between independent test results, must be accomplished within (repeatability) and between (reproducibility) runs (ISO 3534-1) (Fuzery et al., 2013). Precision needs to be maintained over time and also between different operators and laboratories. Another analytical parameter is *accuracy*, closeness of a measurement to the true value (reference value) (ISO 3534-1). In repeated measurements, *trueness* is used as the accuracy of the mean. Determining the accuracy/trueness is however often difficult in proteomics as reference values are not existing for all analytes. This means that the size of the bias (systematic error) is unknown. Although highly desired, the accuracy/trueness can however be less critical compared to precision if the aim is to find variations (biomarkers) and not determine the true analyte concentration.

In our antibody microarrays, the captured proteins are labeled with a detection reagent that generates a signal which is measured in Relative Fluorescence Units (RFU) and corresponds to the relative concentration of the analyte. The noise level of the measurement provides the lower limit, while the upper limit is reached when the detector becomes saturated (**Figure 3**). In affinity assays, an upper limit can also be reached if all binding-sites becomes saturated with target proteins. The lower limit is critical when determining the assay *sensitivity*. The limit of detection (LoD) is usually defined as the lowest concentration at which an analyte can *reliably* be detected (Fuzery et al., 2013). LoD is commonly estimated by repeated measurements of a blank sample, determining the mean and standard deviation (SD), and calculating LoD as the mean of the blank + 2 SD (Armbruster and Pry, 2008). Being able to reliably distinguish a signal from noise might however not be a measurement within the detection range where the analysis meets other requirements, such as precision. Therefore, a linear dynamic range should be

defined between a limit of quantification (LoQ) and a limit of linearity (LoL) where the measurements are performed according to all requirements (**Figure 3**).

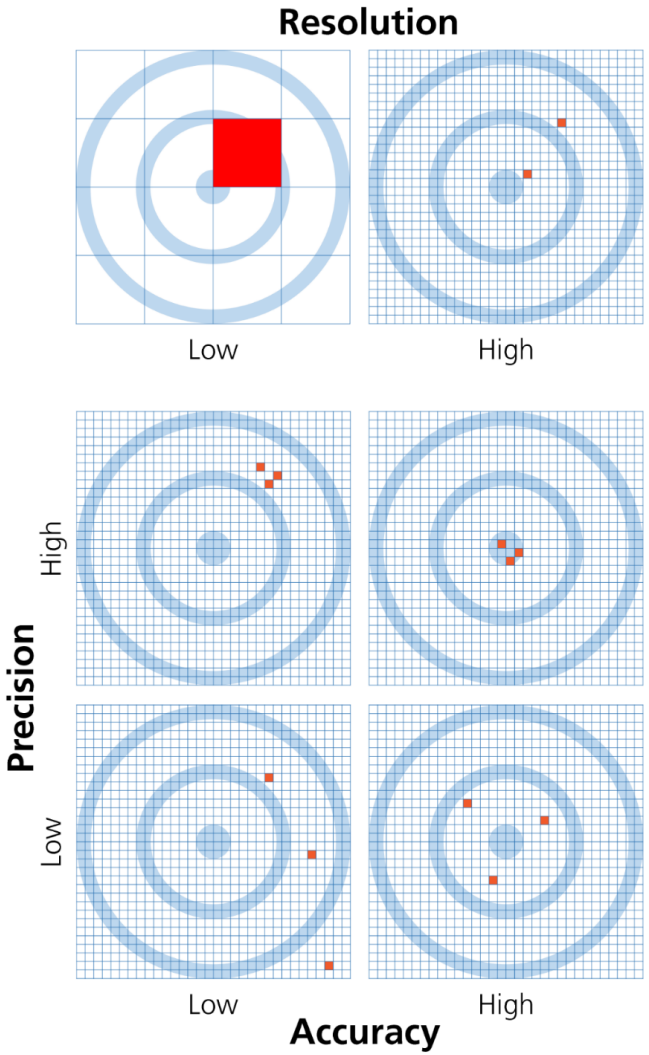

Figure 2. A robust analytical assay should have high resolution, precision and accuracy.

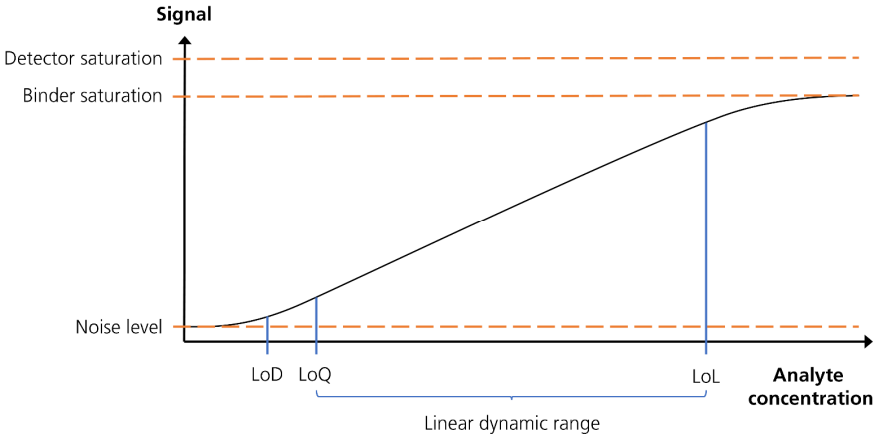

Figure 3. A microarray analysis is limited by the noise level and saturation of either the binder or the detector. The limit of detection (LoD) is the lowest concentration that can reliably be detected and the linear dynamic range is defined between the limit of quantificaiton (LoQ) and the limit of linearity (LoL).

The LoQ is thereby often defined as the *functional* assay sensitivity. The studies in **Paper I-III** were largely focused on further development of steps in the microarray assay that could lead to even better sensitivity. In **Paper I**, a novel detection reagent was evaluated to increase the signal per bound target protein and in **Paper II-III** we tested a method for oriented immobilization of the scFvs that could enable increased target capture per spot.

2.3 Biomarker panels and immunosignatures

A diagnostic test needs to answer the clinical question with high *specificity* and *sensitivity*. In this context, specificity refers to the ability to correctly identify healthy individuals (true negatives) and sensitivity refers to correctly identifying diseased individuals (true positives). However, finding a single biomarker that is both specific and sensitive for just one disease is not likely to be found (Brody et al., 2010). Using multiplex *biomarker panels*, the combination of how several disease-associated biomarkers change in response to a disease can provide better discriminatory power. So far, only one such panel of protein biomarkers, OVA1, has received clearance by the U.S. Food and Drug Administration (Borrebaeck, 2017; Fung, 2010). OVA1 is a blood-based test that measures five proteins and is used for assessment of ovarian cancer risk in women.

Blood, either as *plasma* or *serum*, is a commonly used sample format for biomarker discovery studies and is also the focus in our research group. More than 3500 proteins have been identified in plasma (Farrah et al., 2014), thus providing a rich source of potential biomarkers. The blood proteome is dynamic with proteins originating from various locations in the body. Among the proteins found in blood are signaling molecules of the immune system, such as cytokines, chemokines and growth factors, which are highly interesting targets for biomarker discovery. As the immune system is constantly screening our bodies for even the smallest deviations, it is likely that it will start signaling a response to disease already at an early development stage. That the immune system is responding during cancer progression has since long been suggested (Burnet, 1957), and the need for a tumor to evade the immune system is one of the hallmarks of cancer (Hanahan and Weinberg, 2011).

Targeting mainly immunoregulatory proteins, our research group have used antibody microarrays to identify differentially expressed blood proteins and defined multiplex biomarker panels called *immunosignatures*. Candidate biomarker signatures have been deciphered for different cancers, autoimmune disease and pre-eclampsia, including *B-cell lymphoma* (Pauly et al., 2016; Pauly et al., 2014), *breast cancer* (Skoog et al., 2017; Olsson et al., 2013; Carlsson et al., 2011a; Carlsson et al., 2008), *glioblastoma* (Carlsson et al., 2010), *H. pylori-induced gastric adenocarcinoma* (Ellmark et al., 2006), *pancreatic cancer* (Gerdtsson et al., 2016; Gerdtsson et al., 2015; Sandstrom et al., 2012; Wingren et al., 2012; Ingvarsson et al., 2008), *prostate cancer* (Nordstrom et al., 2014) and *systemic lupus erythematosus (SLE)* (Delfani et al., 2017; Petersson et al., 2014c; Carlsson et al., 2011b), *systemic sclerosis* (Carlsson et al., 2011b) and *pre-eclampsia* (Centlow et al., 2011).

3. Affinity binders in proteomics

3.1 Antibodies

Antibodies, also called immunoglobulins (Ig), were the first binders to be used in affinity proteomics and are still today the most widely employed binder in research and clinical assays, including the set-ups used in this thesis (**Paper I-IV**) (Yalow and Berson, 1959). They have enabled detailed studies and detection of even low abundant proteins in complex samples, something MS have been struggling to achieve. The key to the success of antibodies are their natural ability to specifically bind targets with high affinity, owing to their modular structure that allow generation of affinity towards almost any target.

In our bodies, antibodies have a vital role in both the maintenance of internal homeostasis and the defense against external threats such as bacteria and other pathogens (Murphy et al., 2007). They are produced by B-cells and either secreted in soluble form into the circulation or bound to the membrane as B-cell receptors. As part of the adaptive response of the immune system, antibodies constantly survey the body in search of foreign molecules, so called *antigens*. Once a deviating structure is found, the antibodies have two main functions to neutralize the antigen; antigen binding and recruitment of other actors in the immune system.

The antibody molecule is made up of four separate subunits; two longer *heavy chains* and two shorter *light chains*, each folded into several discrete domains (**Figure 4**) (Murphy et al., 2007). Together they resemble a Y-shaped form with separate parts of the structure corresponding to the two functions. The antigen-binding parts of the structure are located in the "arms" of the antibody, called the *antigen binding fragments (Fabs)*. At the tip of each arm, the *variable domains* of both the heavy (VH) and light (VL) chains form the antigen binding "hands". Particularly important for the antigen binding are six exposed loops called the *complementarity-determining regions (CDRs)* that together form the *paratope,*

which is matching the *epitope* on the antigen both in terms of structure and chemical properties. The CDRs are hypervariable regions of the amino acid sequence while the rest of the domain *framework* forms a scaffold that is relatively constant between different antibodies. The variability of the CDRs is created by a multistep process in the B-cell, a process that generate enormous diversity and lead to the development of antibodies with high *specificity* and *affinity* for the antigen (Murphy et al., 2007).

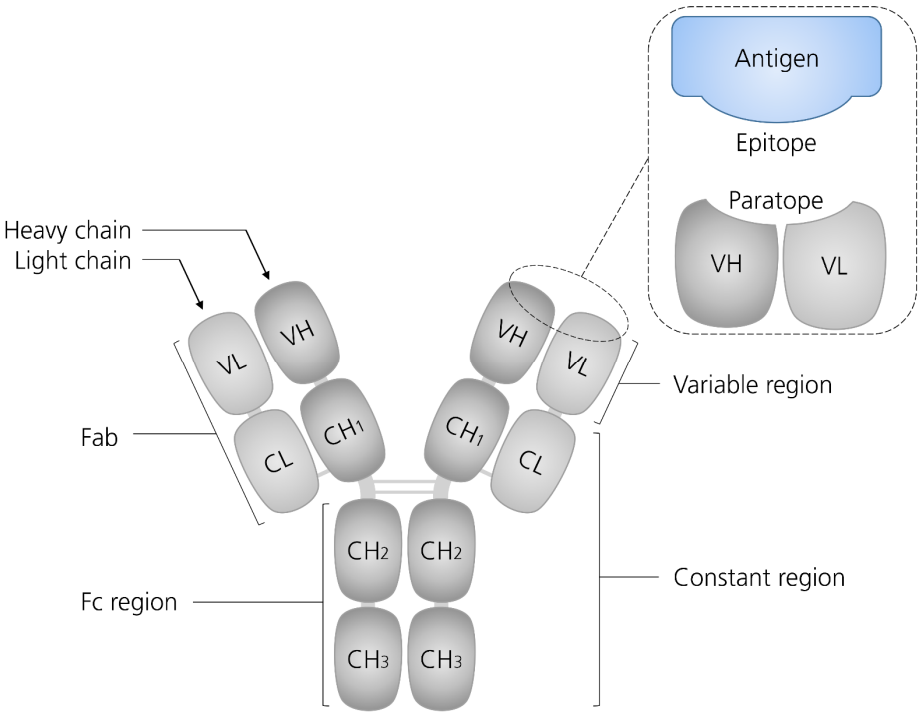

Figure 4. Schematic representation of a full-length IgG antibody.

The "legs" of the antibody Y is called the *fragment crystallizable (Fc)* region and is used by the antibody to recruit and activate other components of the immune system to aid in the neutralization of a threat. The Fc region is made up of the C-terminal ends of the heavy chains and depending on which of the five different heavy chains (α, γ, δ, ε, and μ) that is used determine the antibody *isotype*; IgA, IgG, IgD, IgE or IgM. In contrast, there are only two types of human light chains, *lambda (λ)* and *kappa (κ)*, which have no known functional difference and can both be combined with any of the heavy chain types (Murphy et al., 2007)

3.1.1 Antibody production and formats

The natural ability of B-cells to generate antibodies after antigen stimulation has been utilized to produce antibodies for research purposes. By immunizing animals with a specific antigen, *polyclonal antibodies (pAbs)* can be extracted from the blood. pAbs are a mixture of different antibodies that potentially target different epitopes on the antigen. This can in some settings be an advantage as the antigen binding becomes less dependent on a single epitope and performance of a single antibody. The use of pAbs is however limited by the obtained amount from one extraction and repeated immunizations comes with batch-to-batch variations (Hust et al., 2011). Further, there is a risk of cross-reactivity as the extraction can contain also less specific or even non-target specific antibodies. Animal use is also an obvious drawback. Great progress was made in this area after the introduction of the *hybridoma technology* (Kohler and Milstein, 1975). By fusing a B-cell with a myeloma cell and thereby creating an immortalized cell line, *monoclonal antibodies (mAbs)* could now be produced. As all mAbs from one cell line are identical, monospecific antibodies are produced in a reproducible and renewable manner without the need for repeated animal use.

Advancements in gene technology have further revolutionized how antibodies can be designed and produced. Isolation of the genes coding for the antibody allow the creation of *recombinant antibodies (rAbs)* and enables construction of other antibody formats beyond the conventional full-size antibody (Borrebaeck and Wingren, 2011). Various constructs have been designed over the years, such as *Fab fragments*, *single-chain fragments variable (scFv)*, *diabodies* (dimeric scFvs), and *single-domain antibody* (*sdAb*, also called Nanobodies) (Holliger and Hudson, 2005). Furthermore, rAb genes can be introduced in bacteria, yeast, mammalian and other cell lines, which opens up for fast, cost-effective and animal-free production in large scale. Diversity of rAbs can be generated by creating either *in vivo* or *synthetic antibody libraries* (Marks et al., 1991; Lee et al., 2004; Barbas et al., 1991). *In vivo* libraries are created by isolating the variable parts from large pools of B-cell derived genes and inserting them in the desired format. Synthetic libraries are instead created by modifying the CDRs using for example site-directed or random mutagenesis. Selection of the best performing antibodies is then accomplished using display technologies, such as *phage display* (McCafferty et al., 1990). In phage display, the antibody gene is introduced in a filamentous phage, commonly the bacteriophage M13, and fused with one of the coat proteins to allow expression on the phage surface. This provides a link between the

genotype and phenotype so that the antibody gene from a selected phage can be retrieved.

scFv is the format used by our group and in the work presented in this thesis (**Paper I-IV**). They were selected from two recombinant libraries where each library had a universal framework and only the CDRs were unique to each antibody clone (Sall et al., 2016b; Soderlind et al., 2000). The scFv format was chosen due to several reasons, such as the engineering options and the compatibility with phage display selection and large-scale animal-free production in *E. coli*. The scFv is smaller (~25 kDa) compared to full-length IgG (~150 kDa) and is also expressed from a single gene (VH and VL domains are connected with a flexible amino acid linker), which facilitates expression in phages and bacteria.

A common genetic modification of rAbs is the addition of peptide sequences with certain functions, so called *fusion tags* (Malhotra, 2009). The tags are usually fused by recombinant DNA technology to the gene in the protein C-terminal. The position is chosen for accessibility as well as to avoid disrupting the protein fold and interfering with the antigen-binding site, especially as many tags are long and bulky (Shen et al., 2005). Most recombinant proteins are fused with an *affinity tag* to enable selection or purification. These tags either have affinity for a ligand or act as a binding site for a secondary antibody (Kimple et al., 2013). Some of the most frequently used affinity tags (with their amino acid sequence or length) are Protein A (280 aa), Poly-histidin/6xHis (6 aa: HHHHHH), GST - Glutathione S-transferase (211), FLAG (8 aa: DYKDDDDK), and MBP - Maltose Binding Protein (396 aa). The scFv antibodies used in this thesis (**Paper I-IV**) were purified using a C-terminal 6xHis-tag that allows separation with Ni^{2+} in immobilized metal affinity chromatography (IMAC). Some affinity tags and other fusion tags can also be used for immobilization or functionalization, which will be discussed in **Section 3.1.2.1**.

3.1.2 Antibody engineering

When designing rAbs, a number of engineering options are available to tailor the properties of the binders for high performance in the intended application. As discussed above, the choice of antibody format is a fundamental step. The format must be compatible with the display technology, production host, storage and assay conditions to allow selection, production and final use in large scale. Here,

the use of rAbs can be advantageous as it allows a single universal framework to be selected that is known to be compatible with all the steps mentioned above (Borrebaeck and Wingren, 2009). For pAbs and mAbs, this control is not possible which results in a range of differently performing antibodies due to the generation in different B-cells. As an example, studies have shown that up to 95% of commercially available antibodies are not directly functional in microarray assays as they were developed for other purposes (Haab et al., 2001; MacBeath, 2002). The antibody microarray production and assay involve a number of harsh conditions which the antibodies need to withstand and remain functional in, including printing and adsorption to the slide surface, dry storage and rewetting for final assay. A stable model framework should be selected to begin with but further stability enhancement of rAbs can be achieved either by *directed evolution* or *rational design* (Worn and Pluckthun, 2001). The former uses random mutations and can be achieved by error-prone PCR, while the later relies on detailed knowledge about the structure to make site-specific mutations. Using iterative rounds of mutagenesis and applying a selection pressure (usually elevated temperature or denaturing chemicals such as guanidine hydrochloride), more stable structures can be selected.

3.1.2.1 Conjugation strategies for immobilization and functionalization

Beyond affinity tags as previously discussed, there are also other modifications that can add new features to the rAb. Such modifications can facilitate the selection and purification as well as enable and improve the final assay steps. Sought-after engineering solutions often involve means for *immobilization* or *functionalization* of antibodies. Many immunoassays, including microarrays, rely on immobilization of the antibodies to a surface (Hu et al., 2013). In our current antibody microarrays, the antibodies are bound to the surface through non-covalent adsorption which is a simple and common strategy. However, this strategy results in random 3D orientation of the antibodies with a risk that a large portion of the antibodies become inactive on the surface due to steric obstruction of the antigen-binding site. When pushing the limits through miniaturization, the number of actively binding antibodies become increasingly important to achieve a high sensitivity. By controlling the orientation of immobilized antibodies, up to a 200-fold increase in sensitivity has been reported (Cho et al., 2007; Trilling et al., 2013b).

Some general considerations apply when choosing a suitable immobilization strategy for scFv antibodies, the antibody format we use in our microarray platform and in **Paper I-IV**. Care must be taken to avoid interference with the antigen-binding site or disruption of the protein fold, as this can lead to reduced binding capability. As the scFv format only contain the variable domains, some methods for antibody immobilization are not directly applicable. Strategies based on oxidized glycochains or reduced inter-molecular disulfide bonds are not possible as these features are only present in the constant region (Makaraviciute and Ramanaviciene, 2013). Nor are other Fc binding strategies such as Protein A, Protein G or Fc-binding mAbs applicable for the same reason (Lu et al., 1996).

A convenient solution would be to use the poly-histidine affinity tag already present in the scFv design. Binding of 6xHis-tagged scFvs to surfaces modified with nickel-nitriloacetic acid (Ni-NTA) have been demonstrated (Baio et al., 2011; Lo et al., 2009), however the binding is weak and reversible (Trilling et al., 2013a). Single and double 6xHis-tags have also been tested for our scFvs for binding to Ni-NTA surfaces (Steinhauer et al., 2006; Wingren et al., 2005). Although initially promising, the experiments were discontinued due to issues with heterogeneous coating of the Ni-NTA surfaces and elevated background noise (Wingren et al, unpublished observations).

A novel immobilization procedure based on fusion of scFvs with partial spider silk has also been tested in our group (Thatikonda et al., 2016). The stickiness of the silk was used to guide the fusion protein to attach with the silk part to the surface and thereby orient the scFvs in the opposite direction, which resulted in stronger spot signals. Spotting was here performed either manually or by using a dip-pen contact printer. It remains to be evaluated if the silk-fusion can be used in non-contact printing without clogging the thin nozzles and also if the introduction of the fusion can be expanded beyond the two clones tested in this study.

Various methods have also been suggested for controlled orientation of antibodies via functionalization with a secondary molecule, such as biotin-streptavidin systems, heterobifunctional linkers and DNA-directed methods to mention a few (Liu and Yu, 2016; Niemeyer et al., 1999; Welch et al., 2017). However, these functional molecules first need to be conjugated to the antibody. Conventional conjugation strategies often involve chemical crosslinking to reactive functional groups such as primary amines or sulfhydryls present on sterically available amino

acid side chains or protein terminals (Boutureira and Bernardes, 2015). One such strategy was used in **Paper IV** to first demonstrate the use of oligo-conjugated scFv antibodies. There we used the heterobifunctional crosslinker sulfo-SMCC containing a N-hydroxysuccinimide (NHS) ester that reacts with primary amines on lysines or the N-terminal. The other end of the crosslinker contained a maleimidie group that reacts with the thiol-modified oligonucleotide. A drawback with conventional conjugation strategies is the lack of control over the conjugation instead. Several conjugation sites can be available on the antibody and even appear in the antigen-binding site, and the number of sites can differ between clones and even be absent in some (Welch et al., 2017). There is consequently a risk that the conjugation can give a heterogeneous result and even inhibit the antigen-binding, hence similar problems to the issue with direct surface adsorption. Preferably, the immobilization should instead be site-specific with only a single conjugation site (1:1 molar ratio) per scFv. Such strategies were explored in **Paper II-IV** but I will first mention a couple of other available options.

Säll et al used *in vitro* biotinylation to couple scFvs to streptavidin-coated beads (Sall et al., 2016a). This was achieved by adding a C-terminal biotin acceptor domain (BAD), also known as the Avi-tag (Cull and Schatz, 2000), and expressing the scFvs in an *E. coli* strain also co-expressing the biotin ligase BirA. This approach could potentially also be used in our antibody microarrays but is avoided as the biotin-streptavidin system is already used for sample labeling (Petersson et al., 2014d). Another fusion tag that can be used for oriented protein immobilization is the tetracysteine (TC) motif (CCXXCC) which has high affinity for biarsenical dyes such as FlAsH or CrasH (Griffin et al., 1998; Schulte-Zweckel et al., 2016). Oriented microarrays TC motifs have been demonstrated for other proteins but not yet antibodies (Schulte-Zweckel et al., 2014). Adding an engineered glycan moiety to scFvs has been demonstrated for oriented immobilization to amine-functionalized beads (Hu et al., 2013). This method also requires incorporation of N-glycosylation machinery as it is naturally absent in *E. coli*. Risk of functional loss has however been discussed when using periodate oxidation for the conjugation to oligosaccharides (Abraham et al., 1991).

In the next two sections, I will discuss the conjugation strategies used in this thesis. They are based on the use of unnatural amino acids (**Paper II, Paper III**) and the enzyme Sortase A (**Paper IV**). These strategies can be used for immobilization but can also be applied in functionalization of rAbs for other purposes, such as DNA-

conjugation for detection by sequencing as in **Paper IV**. Generally, the same strategies for immobilization as discussed above can be used for functionalization.

3.1.2.2 Expanding the genetic code of E. coli

With only two rare exceptions (Ambrogelly et al., 2007), all proteins are created from a limited set of 20 natural amino acids. mRNA transcribed from DNA carries the genetic instructions on how the protein should be constructed by the translation machinery of the cell. Codons, triplets of nucleotides, are recognized by corresponding anticodons on tRNAs which carries the encoded amino acid. The correct amino acid has previously been loaded to the tRNA by aminoacyl-tRNA synthetase (aaRS) and the protein is finally assembled in the ribosome where tRNAs deliver the amino acids in the order coded by the mRNA. Schultz and coworkers used this knowledge to design and incorporate also *unnatural amino acids (UAAs)* into proteins and thereby expanding the genetic code (Noren et al., 1989; Wang et al., 2001). UAAs are added to the protein using orthogonal tRNA/aaRS pairs (O-tRNAs and O-aaRS), where the O-aaRS loads the UAA to the O-tRNA. The anticodon of the O-tRNA is directed towards a stop codon in the natural genetic code, commonly the Amber stop codon UAG. In this way, the UAA is incorporated at the stop codon where otherwise the translation would be terminated (**Figure 5**) (Chin, 2017).

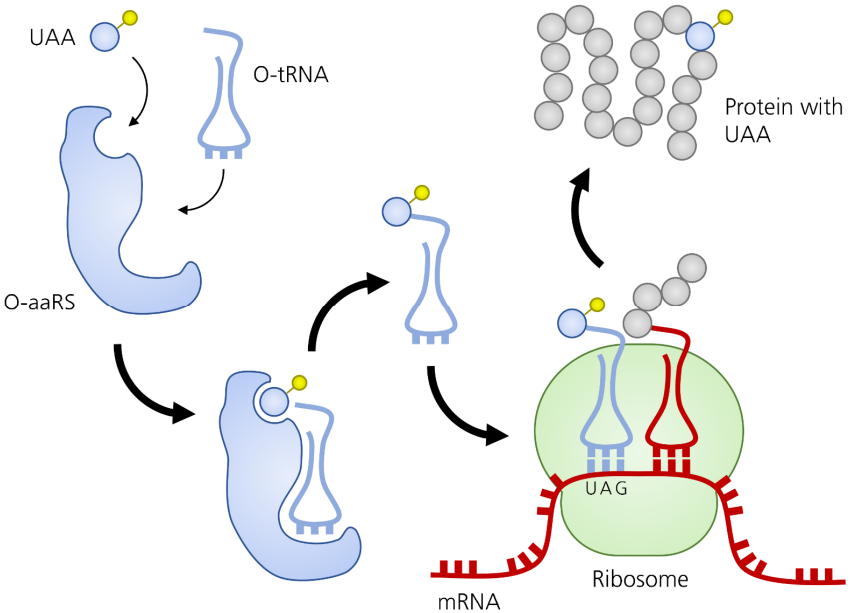

Figure 5. Unnatural amino acids (UAAs) loaded on orthogonal tRNA (O-tRNA) by orthogonal aminoacyl-tRNA synthetase (O-aaRS) and incorporated into proteins at the stop codon UAG introduced in the mRNA.

UAAs have opened up for a new dimension of protein modifications, including means to do site-specific conjugations (Davis and Chin, 2012). Using recombinant DNA technology, the UAA can be positioned in the protein at choice by introducing the stop codon at the corresponding position in the coding sequence, either as a new insertion or a substitution of an existing amino acid. In this way, the incorporation can be done with minimal impact on the protein fold and the position can be chosen to avoid interference with active sites, such as the antigen-binding site of antibodies.

Over 70 different UAAs have been created with a wide range of chemical properties (Liu and Schultz, 2010). Among them is the *p*-benzoyl-L-phenylalanine (*p*Bpa), a chemically stable UAA that upon exposure to near-UV light can be used for photocrosslinking (Chin et al., 2002). When excited by wavelengths around 350-365 nm, a radical is formed that reacts with carbon-hydrogen bonds (C—H) in its proximity and forms a covalent bond (Chin and Schultz, 2002). If no such C-H bond is available, the exited state readily relaxes after terminated exposure. *p*Bpa has for example been used to map proteins interactions (Forne et al., 2012) and for site-specific conjugation of Fc binding Z domains to full-length IgG (Perols and Karlstrom, 2014)

In **Paper II** and **Paper III,** we explored the incorporation of *p*Bpa in scFvs and application of the method *Dock'n'Flash* for site-specific, 1:1 and covalent conjugation. Dock'n'Flash is based on the interaction between *p*Bpa and the seven-membered sugar ring *beta-cyclodextrin (β-CD)* and subsequent photocrosslinking (Jensen et al., 2010). The sidechain of *p*Bpa is attracted to the center of the β-CD ring (the "docking") and binds to C-H bonds upon near-UV irradiation (the "flash"). The wavelengths used are not protein damaging (Chin and Schultz, 2002), why the photocrosslinking is a simple and biocompatible method. In the **Paper II-III**, we show that the amber stop codon could be introduced at several positions in several different scFv clones and that plasmids for mutated scFvs and O-tRNA/O-aaRS pair could be co-transformed and expressed in *E. coli*. Further, using this approach we could demonstrate conjugation of purified *p*Bpa-scFvs to β-CD coated slides for an oriented immobilization. In addition, conjugation to β-CD in solution was also shown, which could be used for scFv functionalization via modified -CD.

3.1.2.3 Conjugation using Sortase A

Sortases are cysteine transpeptidase enzymes that primarily gram-positive bacteria use to covalently attach proteins to their cell wall and to assemble pili (Jacobitz et al., 2017). They specifically recognize the sorting motif LPXTG, where X can be any amino acid, which is cleaved after the threonine and catalyze the formation of a new covalent amide bond to N-terminal oligoglycines on a secondary protein (Chen et al., 2016). This motif-specific conjugation has been applied in research where especially *Sortase A (SrtA)* from *S. aureus* has been widely adopted for site-specific protein engineering. SrtA has been used in for example surface immobilization of unstable membrane proteins (Ito et al., 2010), functionalization of Affibody molecules with peptide nucleic acid (PNA) probes (Westerlund et al., 2015) and generate antibody-drug conjugates (Beerli et al., 2015).

In **Paper IV** we used SrtA to site-specifically and covalently conjugate scFvs with oligonucleotide sequences at 1:1 molar ratio. The SrtA recognition motif LPETG was introduced as a fusion tag prior to the C-terminal 6xHis tag of the scFvs. SrtA was then used to mediate the conjugation of oligos with a tri-glycine (GGG) modification in the 5'-end (**Figure 6**). This conjugation enabled the proof-of-concept demonstration of the Multiplex Immuno-Assay in Solution (MIAS) platform, further explained in **Section 4.3.1**.

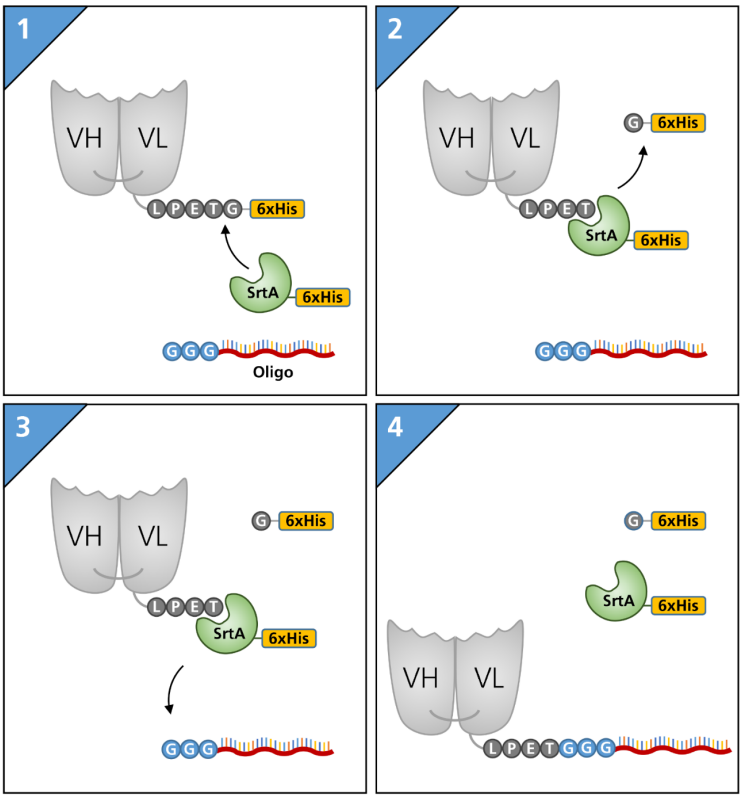

Figure 6. Sortase A (SrtA) mediated conjugation of scFv with the Sortase recognition motif LPETG to an tri-glycine (GGG) modified oligonucleotide.

3.2 Alternative affinity binders

Antibodies have been studied for over a century and the conventional full-length antibodies are still the most commonly used affinity binders in research and used for various applications (Greenspan, 2017; Helma et al., 2015). However, just as rAbs were developed in response to a demand for binders with enhanced quality, modular possibilities and animal-independent production; several affinity binders based on other molecular scaffolds have also been developed. I will briefly mention some available alternative affinity binders that are not derived from the antibody molecule. Alternative binders generally possess similar advantages in terms of the ability to tailor compatibility and function for high performance in the intended use, as previously discussed for rAbs.

Affibodies are small affinity proteins developed from the immunoglobulin binding Z domain of Protein A from *Staphylococcus aureus* (Nord et al., 1997). Consisting of only three helices in a single domain, the affibodies are small with a size of only around 7 kDa, compared to 150 kDa of IgG and 25 kDa for scFvs (Frejd and Kim, 2017). By randomizing amino acids on two of the helices, libraries with more than 10^{10} clones can be created from where affibodies with high specificity and affinity can be selected (Feldwisch and Tolmachev, 2012). Affibodies have also been shown to be compatible with application in microarrays (Renberg et al., 2007; Renberg et al., 2005). Another group of natural binding proteins are repeat proteins, such as the ankyrin repeat proteins from which *Designed ankyrin repeat proteins (DARPins)* have been developed (Bork, 1993; Forrer et al., 2003). DARPins are assembled as tandem arrays with two to four repeated motifs forming a binding site which is flanked by terminal capping repeats, with a size of around 15-18 kDa (Pluckthun, 2015). Further, also non-protein scaffolds have been constructed for affinity binding. *Aptamers* are short, single-stranded DNA or RNA oligonucleotides with the ability to fold into complex and stable structures (Ellington and Szostak, 1990; Robertson and Joyce, 1990; Tuerk and Gold, 1990). The size of aptamers can vary from around the size of scFvs down to only a handful of kilodaltons. In addition, aptamers are chemically synthesized which circumvent the need for a production host (Dunn et al., 2017). Advantages of aptamers include their small size, ease of production, flexibility and stability, while a challenge is the limited chemical diversity of nucleotides compared to amino acids (Groff et al., 2015). These are only a few examples from the plethora of novel affinity binders that has been developed (Helma et al., 2015; Yu et al., 2017).

3.3 Validation of affinity binders

To conclude this chapter, a large number of various antibody formats and also alternative affinity binders are available that can be used for proteomic studies. Independent of the choice of binder, an important aspect is to validate the specificity and function of the binders in the final application. As previously discussed, many "off-the-shelf" antibodies are not developed for microarray and lose their activity on the slide (MacBeath, 2002; Haab et al., 2001). A study within the *Human Protein Atlas (HPA*) showed that out of 5436 commercially available pAbs and mAbs, only 49% could be validated in Western blotting and

Immunohistochemistry (IHC) on tissue microarrays (Berglund et al., 2008). This highlights the fact that the function of an antibody is context dependent. The *International Working Group for Antibody Validation (IWGAV)* recently proposed five "pillars" for validating antibodies in order to generate high quality and consistent data; i) genetic strategies (such as use of samples with target knockdown), ii) orthogonal strategies, iii) independent antibody strategies, iv) expression of tagged proteins (for parallel detection), and v) immunocapture followed by mass spectrometry (Uhlen et al., 2016).

In comparison to antibodies from immunized animals, recombinant binders have some advantage in terms of batch-to-batch reproducibility and options for application-tailored performance. Binders selected for high affinity and specificity in solution does however not guarantee retained function in another context, such as immobilized in microarrays. The scFvs used by our group and the studies presented in this thesis (**Paper I-IV**) have nevertheless shown to perform well in antibody microarrays. Stringent phage-display selection promotes a high specificity (Sall et al., 2016b; Soderlind et al., 2000) and several clones have also been validated using samples with known target concentrations, spiked or depleted samples, and multiple clones for the same target, as well as orthogonal methods such as MS, ELISA, Meso Scale Discovery (MSD) and cytometric bead assay (Wingren et al., 2007; Pauly et al., 2013; Kristensson et al., 2012; Ingvarsson et al., 2008; Ingvarsson et al., 2007; Dexlin-Mellby et al., 2010; Carlsson et al., 2011b). After modifications such as the ones applied in **Paper II-IV**, the validation has to be repeated to ensure retained specificity.

4. Affinity proteomics

4.1 Antibody-based technologies

As described in the previous chapters, proteomics has historically been pursued in two main technology branches; mass spectrometry and affinity-based assays. In this chapter, I will focus on the later and especially on antibody-based technologies, also known as immuno-assays, and their application in biomarker discovery as this is the technology used in **Paper I-IV**. As discussed in **Chapter 2**, successful protein biomarker discovery in blood samples will require the identification of multiplex biomarker panels, such as the immunosignatures defined by our group. This requires a technology platform with high multiplexity to study many proteins in parallel, high sensitivity to also target highly interesting low-abundant targets and high sample throughput to allow rapid generation of large amount of data for a statistically powerful analysis (Fuzery et al., 2013). Fundamental demands on the assay performance are to allow analysis at high resolution and precision over a wide linear dynamic range, preferably as accurate as possible. These requirements limit the number of available platforms and motivates further technological development, which is the focus in the work presented in this thesis (**Paper I-IV**).

One of the key advantages of immunoassays compared to MS has been the high sensitivity. The high affinity of antibodies has for over 50 years been used in Enzyme-linked immunosorbent assays (ELISAs) which has become a standard method in research and also implemented as a clinical routine assay (Tighe et al., 2015). A standard ELISA is however not suitable for multiplex detection, which also applies to other low-plex ultrasensitive assays such as SMC (Singulex Erenna) (Todd et al., 2007), Immuno-PCR (Niemeyer et al., 2007) and Digital ELISA (Simoa) (Rissin et al., 2010). Instead, a number of different methods have been developed for simultaneous analysis of large number of proteins to enable biomarker discovery. They can be divided in two main groups, *planar protein*

microarrays (used in **Paper I-III**) and *solution-based platforms* (**Paper IV**), which will be discussed in **Section 4.2 and 4.3**.

Technologies that combine the both branches of proteomics (MS and affinity assays) have also been developed. Hybrid technologies utilize the sensitivity of affinity assays to circumvent the sample complexity and confirm the identity of the captured protein in subsequent MS analysis (Weiss et al., 2014). Different approaches have been suggested such as the *Mass Spectrometric ImmunoAssays (MSIA)* (Nelson et al., 1995) and *Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA)* (Anderson et al., 2004). Work within our department have resulted in two hybrid technologies, the *AFFIRM (AFFInity sRM)* (Sall et al., 2014) and the *Global Proteome Survey (GPS)* (Sall et al., 2016a; Olsson et al., 2011; Olsson et al., 2012; Olsson et al., 2013). AFFIRM use scFvs coupled to magnetic beads for targeted protein enrichment followed by SRM-MS. The GPS platform is instead developed for global proteome discovery and use *Context Independent Motif Specific (CIMS)* scFvs to target short amino acid sequences generated by enzymatic digestion of the sample and subsequent LC-MS-MS detection. As the same motif will be generated in peptides from many different proteins, only 100 CIMS scFvs could in theory cover almost 50% of the nonredundant human proteome and thereby reduce the need for developing large amounts of binders (Sall et al., 2016a). A similar motif-approach is Triple X Proteomics (TXP) which instead is based on enrichment by pAbs (Poetz et al., 2009).

4.2 Planar protein microarrays

The history of microarrays dates back to the late 80s when Ekins and colleagues first described the use of "microspots" of antibodies for multiplex protein analysis (Ekins, 1989). The microarray concept was then adopted and for many years largely driven by the studies of the human genome and the development of DNA microarrays (Angenendt, 2005). However, after realizing that genomic studies would not answer all clinical question, a growing interest for protein microarrays was sparked. Knowledge gained from development of DNA microarrays were inherited and transformed into the early protein microarrays, including instrumentation, detection reagents and bioinformatics (Haab et al., 2001; Hall et al., 2007; MacBeath and Schreiber, 2000; Tyers and Mann, 2003).

Antibody microarrays can be created as *forward arrays* by spotting and immobilizing a capture antibody on a planar solid surface using a high-precision robotic printer (MacBeath and Schreiber, 2000; Sanchez-Carbayo, 2011; Wingren and Borrebaeck, 2006). This allows utilization of the antibody's ability to specifically fish out the target from a complex sample and non-specific content can be removed by washing. Forward arrays can also be constructed with sandwich antibody pairs, where a second detection antibody directed towards the same target is used. Sandwich arrays promotes the specificity of the detection, but are more complicated to implement for multiplex analysis as development of high-performing antibody pairs is needed. Detection is either achieved by direct labeling of the target or using a second labeled detection antibody directed towards a sample tag. The amount of bound target is finally quantified, commonly using a laser scanner. The antibody microarrays develop by our group are forward arrays and technological development presented in the papers of this thesis (**Paper I-IV**) is mainly focused on the immobilization and detection steps. There are also *reverse phase protein arrays (RPPA)* where instead a minute amount of sample is spotted on the surface after which the antibodies are added (Sanchez-Carbayo, 2011).

Antibody microarrays have been developed based on various antibody formats and in numerous assay set-ups and have become an established method for proteomic studies (Borrebaeck and Wingren, 2011; Borrebaeck and Wingren, 2014; Haab, 2006; Sanchez-Carbayo, 2011). In the following section, I will provide an overview of the current antibody microarray platform used in our group and which we in **Paper I-III** tried to further develop in terms of antibody immobilization and target detection.

4.2.1 Our current antibody microarray platform

The in-house designed recombinant antibody microarray platform used in **Paper I-III** has since the introduction in the early 2000s developed into a state-of-the-art analytical platform for large-scale protein expression profiling (Delfani et al., 2016; Steinhauer et al., 2002). Several of the key technological parameters have been optimized and the current version includes over 350 unique scFvs printed in high-density microarrays (Borrebaeck and Wingren, 2009; Delfani et al., 2016). As specified in **Section 2.3**, the platform has been used to define numerous candidate biomarker signatures in various cancers and autoimmune disease. In parallel with continued academic research and development, some of the

promising results are currently under validation for clinical implementation by the company Immunovia AB (www.immunovia.com).

A conceptual overview of the microarray platform is presented in **Figure 7**. scFvs selected from large in-house recombinant antibody libraries are produced in *E. coli* and purified using a C-terminal 6xHis affinity tag (Sall et al., 2016b; Soderlind et al., 2000). Purified scFvs are diluted to suitable on-chip concentrations and printed onto planar Black Polymer MaxiSorp microarray slides using a Scienion SciFlexarrayer S11 non-contact printer. The printer generates droplets with a volume of around 330 pl each and arrange them in a predefined array of spots on the slide. A single droplet creates a spot with a diameter of around 140 μm. Fourteen identical subarrays, each containing all the specificities in at least triplicates, are printed on the same slide which enables parallel analysis of separate samples in each subarray. In addition to the scFv spots, each subarray also contains biotinylated BSA spots as positive controls and PBS spots as negative controls. After printing, the spots are allowed to dry and the slide is stored in controlled conditions until use.

In the current set-up, the printed scFvs are immobilized to the slide surface by random non-covalent adsorption. This might have the effect that some of the scFvs are adsorbed with the binding-sites unavailable for the target. If instead the scFvs were immobilized in a oriented fashion, antibody activity could be increased which in turn could lead to improved assay sensitivity (Cho et al., 2007; Steinhauer et al., 2002; Welch et al., 2017). This was one of the reasons that motivated the studies in **Paper II** and **Paper III** where the Dock'n'Flash method was used to photocrosslink scFvs equipped with the unnatural amino acid *p*Bpa to slides coated with the ligand β-cyclodextrin.

For future studies, the Sortase A mediated site-specific conjugation strategy applied in **Paper IV** could in theory also be used for scFv orientation to a surface with a coating containing available glycines. Furthermore, scFvs functionalized with single-stranded oligonucleotides either by Sortase or Dock'n'Flash could also open up for orientation to slides coated with complementary oligonucleotides. DNA-directed immobilization (DDI) strategies have previously been suggested but have been impaired by the lack of site-directed conjugation strategies (Niemeyer et al., 1999).

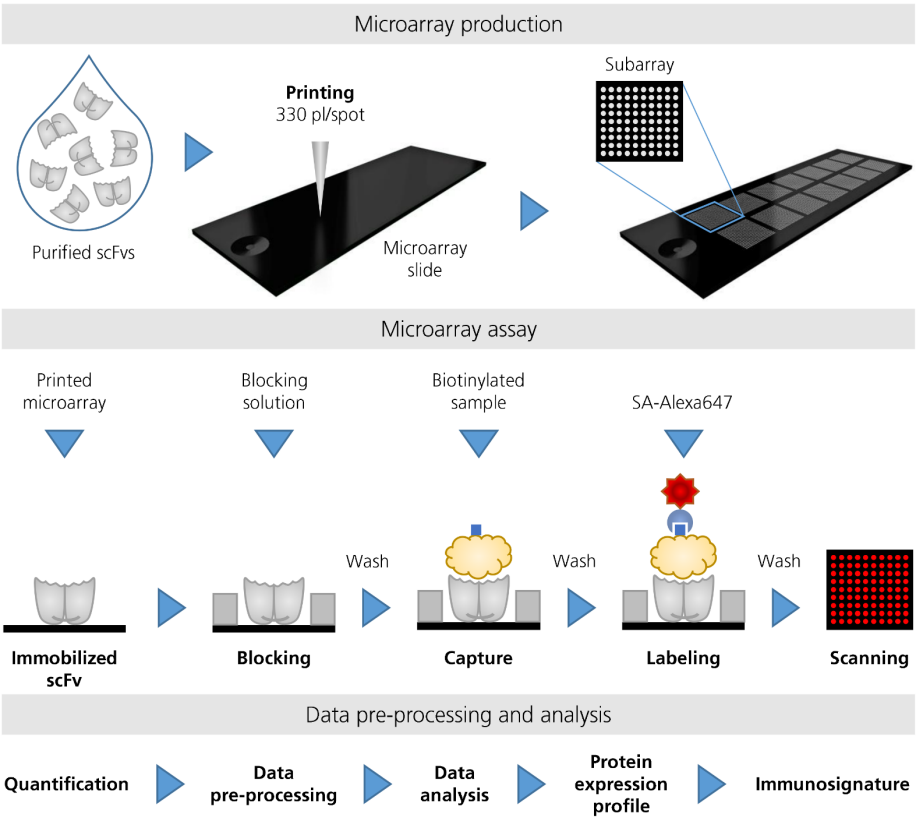

Figure 7. Conceptual overview of our recombinant antibody microarray platform.

A microarray assay is initiated by creating separate reaction wells for each subarray, in our case using a slide holder and a gasket. Naked slide surface is first blocked using a milk protein solution to minimize non-specific background binding. After washing away excess blocking solution, a biotinylated and diluted serum sample is added and incubated to allow target capture by the scFvs. The subarrays are then washed before addition of Alexa647-labelled streptavidin (SA-Alexa) for fluorescent detection. After final wash, the slide is dried before immediate imaging in a laser scanner.

The detection step was the focus of **Paper I** where a novel fluorophore-packed nanoparticle entitled Phosphor-Integrated Dot (PID) was for the first time used in an antibody microarray application as a detection reagent. The hypothesis was that a simple replacement of SA-Alexa with PID would generate stronger

fluorescent signals from the bound targets. This could potentially raise signals from low-abundant proteins above the noise level to allow or improve the detection. Indeed, stronger signals were in general detected from the arrays with PID but they were accompanied with increased and heterogeneous background. The background issues could partly be reduced by optimization, but further optimization would be required for the implementation of PID. As this type of PID were around 30 times larger compared to the scFvs, smaller sized nanoparticles would also be interesting to evaluate. Smaller sized nanoparticles would to lesser extent block the target binding of adjacent scFvs.

After scanning, the signals from the spots are quantified and the background subtracted. The data is pre-processed and normalized to avoid bias by different days of analysis or array-to-array variation. The data is analyzed by machine learning algorithms and the out-put is a protein expression profile. By comparing the profiles from large amount of samples, differentially expressed proteins can be condensed to an immunosignature that best discriminates between different sample cohorts.

4.3 Solution-based platforms

In parallel to planar microarrays, multiplex platforms based on *microbeads* have also been developed where the analysis is carried out in solution. Common solution-based assays are the *suspension bead arrays* where affinity binders are immobilized on the beads and allowed to incubate with a sample followed by quantification of the captured targets (Schwenk and Nilsson, 2011; Schwenk et al., 2008). Multiplex analysis of a sample can be performed by first coating sets of beads with binders of only one specificity per set. Quantification is the commonly made using flow cytometry where both the bead and the labeled target is detected. The quantified signal and target identity is connected by using a unique bead-set per binder specificity, similar to how the known position of a spot connects the signal to the identity in planar microarrays. Unique beads can be created using either different sized beads or beads filled with different fluorescent dyes. The multiplexity of suspension arrays is thereby limited by the number of unique beads. However, by mixing different ratios of several fluorescent dyes inside the beads, spectrally unique beads can be created that today allow up to 500-plex analysis (Fulton et al., 1997; Schwenk and Nilsson, 2011).

The use of in-solution assays can solve some bottlenecks of planar microarrays when it comes to increasing the multiplexity towards global serum protein profiling. No high-precision printing robot is needed as the spotting step is circumvented which also makes the assay more flexible and scalable. For planar arrays, increased number of spots would also result in a larger array footprint that could limit the sample throughput. Reducing the spots size is possible but difficult and not yet compatible with higher multiplexity (Petersson et al., 2014a; Petersson et al., 2014b; Petersson et al., 2014c). The microarray slides must also be carefully handled as they are sensitive to dust and scratches. The natural environment for affinity binders is also in a solution, why the dried state on slides can be detrimental for many binders. Although our recombinant scFvs are optimized to withstand this handling, they might perform even better in solution. To take advantage of the opportunities with solution-based assay, in **Paper IV** we developed the novel platform Multiplex Immuno-Assay in solution (MIAS) which is further explained in the next section.

4.3.1 Multiplex Immuno-Assay in Solution

Our current planar antibody microarray technology provides a state-of-the-art protein expression profiling platform with capacity to simultaneous analyze the response from over 350 scFvs. Although the current platform already has delivered several candidate biomarker signatures for various diseases, there might still be many serum proteins not covered by the discovery platform that potentially could contribute to even better signatures. A large increase in multiplexity would however be challenging to practically handle with planar microarrays. Switching to a solutions-based assay can have some advantages when aiming for global serum profiling, as discussed in the section above.

In **Paper IV** we developed the novel platform Multiplex Immuno-Assay in solution (MIAS) which was designed for large-scale protein expression profiling. The MIAS concept uses streptavidin-coated magnetic beads to immobilize and display biotinylated proteins of a sample followed by target quantification using barcoded scFvs and massive parallel sequencing by NGS (**Figure 8**).

Key to the MIAS concept is the site-specific 1:1 functionalization of scFvs with oligonucleotides, mediated by Sortase A as described in **Section 3.1.2.3**. Using antibody engineering, the SrtA recognition motif LPETG was introduced in scFvs

clones which enabled site-specific conjugation to tri-glycine modified oligonucleotides at 1:1 ratio. This in turn enabled highly sensitive quantification using NGS which provides a direct digital read-out.

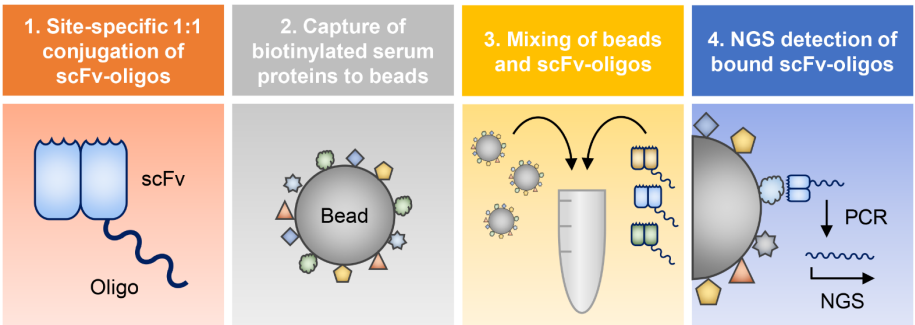

Figure 8. Overview of the Multiplex Immuno-Assay in solution (MIAS) concept.

The oligonucleotides were designed to include different parts to allow identification and quantification (**Figure 9**). A 6 bp barcode sequence represent the specificity of the scFv and an 8 bp random sequence makes each barcode unique to allow quantification. The oligonucleotides also contain two separate adaptor sequences that allow addition of two primers by adaptor PCR. A specific index primer is added to all barcodes from the same sample to connect the quantified target to the right sample. Finally, a universal primer (Primer 1) is added in the 5' end to signal where the sequencing should start in the NGS. As all the information is acquired in this way, coded beads are not needed and thereby the multiplexity is not limited by the availability of unique beads. The multiplexity is determined by the possible combinations in the barcode and is dramatically increased per extended nucleotide if needed.

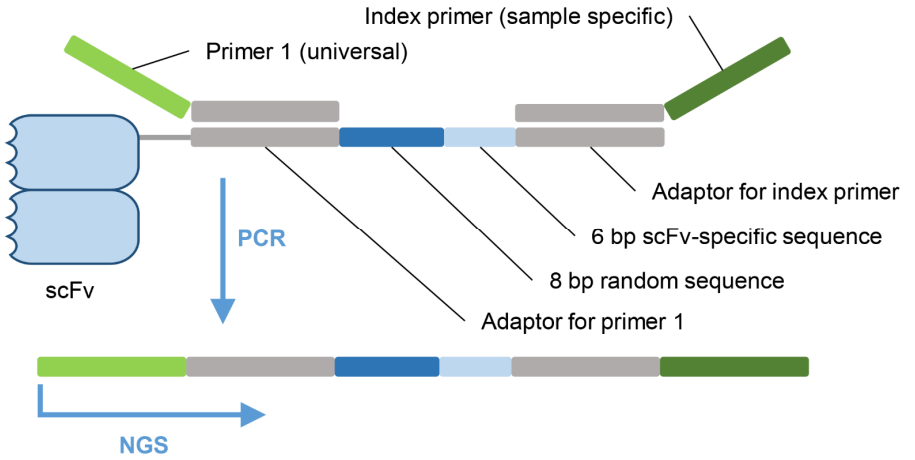

Figure 9. Oligonucleotide design for quantification in the MIAS platform.

Proof-of-concept for MIAS was demonstrated in four major steps. Firstly, the immobilization of sample proteins to the magnetic beads was optimized. Secondly, the assay steps including NGS detection of scFv-oligos was verified using standard SMCC conjugation and thiol-modified oligos. Thirdly, three scFv clones modified with the LPETG motif were generated and conjugated to oligos using SrtA. Finally, the complete assay was run with oligos site-specifically conjugated to the scFv-LPETG mutants and successfully detected by NGS. The proof-of-concept encourage further development of the MIAS platform in future studies.

One of the strengths of MIAS is the use of scFvs. As it is a recombinant antibody format, the use of scFvs enables engineering options and the animal-free production in bacteria provide a renewable source of tailor-made binders with high affinity. scFvs also allow the covalent site-specific 1:1 conjugation to oligos. Altogether, this makes MIAS a unique platform compared to other methods such as ProteinSeq which are based on pAbs and conventional conjugation strategies (Darmanis et al., 2011).

5. Concluding remarks

In 2005, a couple of years after the sequencing of the human genome was completed, Baak and colleagues described the level of complexity of proteomics compared to genomics (Baak et al., 2005):

> "…we could say that we now have the letters (the sequence) and have thus far found a few sentences (genes that we know of), but we have only merely begun reading the chapter contents (how genes may be transcribed), while the books (the proteins and the metabolites) will keep human mankind reading for many centuries to come."

The complexity of the proteome owes to the vast number of proteoforms that each coding gene can give rise to, its dynamic nature and the large variation in chemical properties that allow formation of intricate structures and interactions, as discussed in **Chapter 2**. Although this makes the proteome challenging to study, we have already begun to see the possibilities that proteomics can provide for understanding the molecular processes behind human diseases. Proteomic research, such as the work conducted within our group, have resulted in candidate biomarker signatures that can provide invaluable information to improved disease treatment. Precise diagnostic, prognostic and predicative biomarkers are the first steps towards implementation of precision medicine. So far, there has however been a very limited transfer of research findings into clinical implementation which points out the need for continued and improved research and development.

The work within this thesis revolves around engineering of affinity-binders (**Chapter 3**) and technological development within affinity-based proteomics (**Chapter 4**). In four original papers, we have focused on improving current planar antibody microarray platform by a novel detection reagent (**Paper I**) and oriented immobilization of the scFvs (**Paper II-III**) and also demonstrated proof-of-concept for the novel MIAS platform (**Paper IV**). The ultimate goal was to facilitate large-scale protein expression profiling of blood samples to decipher disease-associated biomarker panels.

Paper I was focused on the employed detection reagent in pursuit of increased sensitivity of the microarray assay. PID, a novel fluorophore-packed nanoparticle, was for the first time used in antibody microarrays and replaced the Alexa-647 used in the current set-up. We showed that it was possible to use PID in the microarray assay and that the resulting signals were stronger compared to Alexa-647. However, PID gave increased and heterogeneous background that, although it could partly be reduced, will require further optimization.

Paper I-II aimed to find means for site-specific conjugation of our scFvs to enable oriented immobilization or functionalization. To achieve this, the Dock'n'Flash method was for the first time used for antibodies where the incorporation of the unnatural amino acid *p*Bpa in scFvs allowed photocrosslinking to a novel slide coated with the ligand β-cyclodextrin. Proof-of-concept was shown in **Paper I** where a first *p*Bpa mutant scFv was generated and conjugated to β-cyclodextrin coated on a slide and also in solution. The promising results were followed up in **Paper II** where *p*Bpa incorporation was expanded to in total 4 different scFv clones and 13 different positions. The position T7 in the affinity tag was chosen as the preferred site for *p*Bpa. Taken together, the conclusion from both studies is that site-specific conjugation is possible via Dock'n'Flash. This opens up for future studies to evaluate if the oriented scFvs indeed provide increased sensitivity and performance compared to random adsorption. An interesting future study would be to combine oriented immobilization from **Paper II-III** with an optimized version of PID nanoparticles from **Paper I** as detection reagent. Together, they might produce far better signal-to-noise ratios and increase the sensitivity to enable better measurements of also very low-abundant targets. In addition, further studies can investigate scFv functionalization using modified β-cyclodextrin.

Besides oriented immobilization of scFvs, Dock'n'Flash was also used both in **Paper I** and **Paper II** for covalent conjugation to β-cyclodextrin in solution. This could pave the way for site-directed functionalization of recombinant proteins via β-cyclodextrin. The use of unnatural amino acids provides a rather unique possibility for site-directed conjugation as they in many cases involve substitution of only a single amino acid. This reduces the risk of significant interference with the delicate protein fold and provides more freedom to position the conjugation site. Many other strategies involve addition of multiresidue fusion tags which usually restrict the positioning to the protein terminals. However, a remaining obstacle for efficient functionalization using Dock'n'Flash is the conjugation

yield. Judging from MALDI-TOF results in both **Paper I** and **Paper II**, only moderate yields were achieved. Pure conjugated product can be a requirement depending on the application. This could be solved by purification, although it can be challenging and adds an additional step.

In **Paper IV** we sought to overcome some inherent limitations with planar microarrays by developing the solution-based MIAS platform. In MIAS, biotinylated samples are displayed on magnetic beads and target proteins are subsequently quantified using DNA-barcoded scFvs and massive parallel sequencing by NGS. This abandon the need for printing, the assay is not limited by slide surface and the scFvs are kept in solution. This allow for even higher multiplexity which will be needed when aiming for global serum profiling and at the same time allow high sensitivity and high sample throughput. Key to the MIAS platform was the scFv-oligo conjugates that was covalently and site-specifically generated using Sortase A at 1:1 ratio. Proof-of-concept was demonstrated using three Sortase-conjugated scFv-oligos which all could be detected and quantified using NGS.

The results in **Paper IV** encourage future studies and much work remains to achieve a robust platform. For comparison, the microarray platform has matured over many years of refinement. Knowledge gained from the in-house development combined with general progress in the field can however be transferred into the MIAS platform to speed up the development. Each assay step can likely be further optimized, including the Sortase-mediated conjugation where for example other enzyme variants and different reaction conditions can be evaluated. In the end, assay performance needs to be demonstrated in terms of analytical parameters as discussed in **Chapter 2**.

After modifications of the scFvs, such as the conjugations in **Paper II-IV**, the antibodies also need to be revalidated as discussed in **Section 3.3**. Modifying a protein and conjugation always come with a risk of affecting the fold and stability of the protein. In the case of antibodies, this can lead to loss of target affinity or cause unspecific binding. In **Paper II-III**, a first validation confirmed that the on-chip activity and affinity of the *p*Bpa mutant scFvs were retained compared to the wild-type.

In summary, proteomics has the potential to revolutionize our understanding of human biology and provide efficient tools to enable precision medicine. Affinity proteomics is one such tool that can be used to overcome the complexity of the proteome and decipher its valuable information. I hope that the work presented in this thesis will contribute to the technological development needed to further identify the pieces of the large proteome puzzle.

Populärvetenskaplig sammanfattning

Vår kunskap om livets minsta beståndsdelar blir allt mer detaljerad och därmed kan vi bättre förstå bakomliggande orsaker som på molekylär nivå påverkar vår hälsa. Avgörande steg för den moderna molekylärbiologiska forskningen togs i och med insikten om att DNA-molekylens komponenter och uppbyggnad kan lagra information som sedan kan överföras till RNA och översättas till proteiner. Fortsatt forskning har lett fram till 2000-talets kartläggning av människans fullständiga DNA-uppsättning, även kallat sekvenseringen av det mänskliga genomet. Studier av mutationer i en människas genom har sedan kunnat kopplas till sjukdomar så som cancer. Forskningen har dock också konstaterat att en förändring i vårt genom som skulle kunna orsaka sjukdom inte nödvändigtvis alltid kommer till uttryck. Detta kan exempelvis förklaras av att alla gener i vårt genom inte är ständigt aktiva och då i stunden inte bidrar till uppsättningen av våra proteiner, det så kallade proteomet. Den insikten har lett fram till ett ökat intresse för att mäta och förstå de proteiner som vid ett givet tillfälle faktiskt är närvarande, en disciplin inom forskningen som fått namnet proteomik.

Vår forskargrupp är fokuserad på att utveckla proteomisk teknik som möjliggör bättre studier av proteiner och att med hjälp av den tekniken hitta de proteiner som förändras när vi blir sjuka. Sådana proteiner kallas för *sjukdomsspecifika biomarkörer*. Förhoppningen är att hitta biomarkörer som exempelvis kan leda till tidigt upptäckt av en växande cancertumör långt innan symptom visar sig. Tidig upptäckt är i många fall avgörande för en lyckad behandling och cancern kan stoppas innan den vuxit sig stor och börjat sprida sig vidare i kroppen. Biomarkörer spås bli viktiga för utvecklingen av mer precis och personanpassad behandling av sjukdomar. De skulle kunna bidra till att fastställa en mer noggrann diagnos, förutse patientens prognos och föreslå bästa behandling samt följa upp hur patienten svarar på vald behandling. Allt detta skräddarsytt utifrån patientens unika förutsättningar.

Tekniken vi arbetar med kallas för *antibody microarrays* (antikropps-mikromatriser) och baserar sig på antikroppars särskilda förmåga att fånga in utvalda proteiner. I vår kropp har antikropparna en viktig funktion i vårt immunförsvar där de används för att möjliggöra upptäckt och eliminering av främmande objekt. Till våra mikromatriser har vi tagit fram hundratals specialdesignade antikroppar, så kallade *single-chain fragment variable (scFv)* antikroppar, som var för sig fångar in ett särskilt protein. Mikroskopiska droppar (omkring 300 pikoliter) med dessa antikroppar placeras i ett ordnat mönster (mikromatris) på en plastyta med hjälp av en precisionsrobot. Därefter tillsätts prov varifrån antikropparna fångar in sina respektive proteiner och övriga proteiner som inte bundit tvättas sedan bort. Alla proteiner i provet är på förhand märkta vilket gör det möjligt att mäta antalet proteiner som fångats in av antikropparna med hjälp med en fluorescerande molekyl och en laserscanner. Sjukdomsspecifika biomarkörer kan då hittas genom att identifiera de proteiner vars nivåer skiljer mellan prover från friska personer och patienter med en viss sjukdom.

Vi har valt att studera proteiner som finns i vårt blod. Blodprover är lätta att ta utan komplicerade ingrepp och de innehåller också ett stort antal intressanta proteiner. Antikroppar har särskilt tagits fram för att fånga in proteiner som reglerar vårt immunförsvar. Eftersom immunförsvaret ständigt övervakar och reagerar på minsta förändring i vår kropp tror vi att det här går att hitta biomarkörer som indikerar avvikelser redan tidigt i ett sjukdomsförlopp. För att ytterligare stärka tillförlitligheten att dessa biomarkörer verkligen är kopplade till en viss sjukdom använder vi oss av den kombinerade information från hur flera proteiner varierar, så kallade *biomarkörs-signaturer*.

Min avhandling består av fyra artiklar där vi på olika sätt har arbetat med vidareutveckling av tekniker för att möjliggöra ännu bättre analys av proteiner i blodet. Även om antibody microarrays redan är en mycket användbar teknik har den i dagens utförande några tekniska egenskaper och även en del arbetskrävande delsteg som begränsar möjligheten att ytterligare utöka antalet antikroppar.

I den första artikeln (**Paper I**) testar vi en ny typ av märknings-reagens till våra microarrayer där ett stort antal fluorescerande molekyler packats samman i en *nanopartikel*. Syftet var att höja signalen från varje bundet protein och därmed på ett enkelt sätt åstadkomma en känsligare detektion.

I den andra och tredje artikeln (**Paper II-III**) utforskar vi möjligheten att höja känsligheten genom att styra hur antikroppar fäster till ytan och på så sätt göra en större andel av antikropparna till aktiva proteinbindare. Istället för att som i dagens microarrays låta antikropparna slumpvis fästa till ytan och riskera att några fäster med den proteinbindanden delen otillgänglig för provet, så ville vi här rikta varje antikropp rättvänd. Detta testas med hjälp av en metod kallad *Dock'n'Flash* där den onaturliga aminosyran *pBpa* introduceras i antikroppen på motsatt sida av den proteinbindande delen. pBpa binder söker sedan upp en motpart kallad *β-cyclodextrin* som vi täckt ytan med och genom att lysa med ultraviolett ljus av en viss våglängd så kopplas de båda parterna samman. Att kopplingen görs på en kontrollerad plats i antikroppens struktur och enbart där gör att Dock'n'Flash även skulle kunna användas för att koppla antikroppar till andra molekyler och på så sätt lägga till funktioner som inte går att åstadkomma på naturligt väg.

I den fjärde artikeln (**Paper IV**) utvecklar vi en ny teknik kallad *Multiplex Immuno-Assay in Solution* (*MIAS)* för att möjliggöra mätning av ännu fler proteiner parallellt. I MIAS utförs analysen i lösning och proteinerna som ska studeras kopplas till magnetiska kulor. Därefter tillsätts antikroppar som med hjälp av enzymet *Sortas* märkts med en kort DNA-sekvens vilken fungerar som en unik streckkod för varje antikropp. Antalet proteiner kan då mätas genom att läsa av streckkoden på de antikroppar som har bundit in. Avläsningen sker med hjälp av en metod som kallas *next generation sequencing (NGS)* som är så känslig att även minimala skillnader mellan prover kan detekteras.

Acknowledgements

Denna avhandling och resan som lett fram till den hade inte varit möjlig utan hjälp från en lång rad människor genom åren och jag vill rikta ett stort tack till er alla! Ett särskilt tack till handledarna samt Anna S, Catja F, Corinna R, Kristoffer S, Malin B, Mats O, Petter S och Payam D för kommentarer och korrekturläsning av olika delar av kappan. Jag är också innerligt tacksam för det extra stöd och den värme som visades mig i samband med att min mamma gick bort våren 2016. Det lyste upp i en annars väldigt mörk tid.

Huvudhandledare Christer – Tack för att du tog mig ombord via projektanställningen och sedan vidare som doktorand. Jag mötte dig i kurser redan under ingenjörsutbildningen där du lärde ut med en smittsam entusiasm och inlevelse. Jag är säker på att du kommer inspirera många elever till fortsatt lärande och nyfikenhet i ditt nya jobb.

Biträdande handledare Anna – Du har varit ett enormt viktigt stöd och bollplank genom åren. Du bjuder alltid på dig själv var sig det är genom att dela med dig labberfarenhet, kloka råd om livet eller roliga historier på fikarasten.

Biträdande handledare Carl – Det är fantastiskt inspirerande att få vara en del av vad du varit med och byggt upp som sträcker sig långt utanför institutionens väggar.

Alla medlemmar av chipgruppen genom åren! Ni har varit enormt viktiga och jag har lärt mig så mycket av er och haft fantastiskt kul tillsammans med er både på och utanför jobbet. Särskilt ni ”Chipster” – Anna IE, Anna S/Z, Payam och Petter – som var med nästan hela vägen och jag lärt känna lite närmre: ni är fantastiska! Extra tack till Linn P som guidade mig in i världen av ljusmutanter och nanoarrayer samt för alla bra samtal och galna halloweenfester genom åren!

Alla härliga kollegor på IT, nuvarande som utflugna samt de som blev kvar på BMC! Tillsammans skapar ni en fantastisk arbetsplats med spännande forskning i framkant. Jag är väldigt tacksam för alla som kommit med input i mina projekt och stöttat i svåra stunder samt all hjälp med det praktiska runt omkring forskningen. Tack också för allt kul på fikarasterna, utflykter, spexinspelningar, disputationsfester mm!

To all the great people I have come in contact with in my projects, at Immunovia, the Dock'n'Flash team, My Hedhammars group and Konica Minolta: Thank you for fruitful collaborations!

Carl-Magnus – stort tack för att du så generöst delade med dig av dina erfarenheter och ditt nätverk under året som min mentor.

Till alla underbara vänner utanför jobbet: Halmstadgänget, LTH-vänner, Korridorsfolket, Innebandylirare mfl; utan er hade livet varit tråkigt! Tack för allt kul och för fin vänskap!

Jag har också haft förmånen att växa upp med så många fantastiska personer och goda förebilder i släkten vilket jag är väldigt tacksam för. Många härliga minnen har jag från glada stunder på Barrstigen, i stugorna och på stranden i Frösakull.

Malin och Pappa - ni är bäst! Ni betyder så enormt mycket för mig och ni har alltid stöttat i allt jag tagit mig för. Särskilt tack för all service och alla matlådor som gjort att jag kunnat fokusera på skrivandet nu på slutet och samtidigt kunna njuta av god mat!

Mamma – Du är för evigt saknad! Även om du inte fick se mig gå i mål med doktorandtiden så finns du med mig i allt jag gör.

References

Abraham R, Moller D, Gabel D, et al. (1991) The influence of periodate oxidation on monoclonal antibody avidity and immunoreactivity. *J Immunol Methods* 144: 77-86.

Aebersold R and Mann M. (2003) Mass spectrometry-based proteomics. *Nature* 422: 198-207.

Ambrogelly A, Palioura S and Soll D. (2007) Natural expansion of the genetic code. *Nat Chem Biol* 3: 29-35.

Anderson NL and Anderson NG. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845-867.

Anderson NL, Anderson NG, Haines LR, et al. (2004) Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res* 3: 235-244.

Angenendt P. (2005) Progress in protein and antibody microarray technology. *Drug Discov Today* 10: 503-511.

Anjo SI, Santa C and Manadas B. (2017) SWATH-MS as a tool for biomarker discovery: From basic research to clinical applications. *Proteomics* 17.

Armbruster DA and Pry T. (2008) Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev* 29 Suppl 1: S49-52.

Baak JP, Janssen EA, Soreide K, et al. (2005) Genomics and proteomics--the way forward. *Ann Oncol* 16 Suppl 2: ii30-44.

Baio JE, Cheng F, Ratner DM, et al. (2011) Probing orientation of immobilized humanized anti-lysozyme variable fragment by time-of-flight secondary-ion mass spectrometry. *J Biomed Mater Res A* 97: 1-7.

Barbas CF, 3rd, Kang AS, Lerner RA, et al. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proc Natl Acad Sci U S A* 88: 7978-7982.

Beerli RR, Hell T, Merkel AS, et al. (2015) Sortase Enzyme-Mediated Generation of Site-Specifically Conjugated Antibody Drug Conjugates with High In Vitro and In Vivo Potency. *PLoS One* 10: e0131177.

Berglund L, Bjorling E, Oksvold P, et al. (2008) A genecentric Human Protein Atlas for expression profiles based on antibodies. *Mol Cell Proteomics* 7: 2019-2027.

Biomarkers Definitions Working Group. (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 69: 89-95.

Bork P. (1993) Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins* 17: 363-374.

Borrebaeck CA. (2017) Precision diagnostics: moving towards protein biomarker signatures of clinical utility in cancer. *Nat Rev Cancer* 17: 199-204.

Borrebaeck CA, Sturfelt G and Wingren C. (2014) Recombinant antibody microarray for profiling the serum proteome of SLE. *Methods Mol Biol* 1134: 67-78.

Borrebaeck CA and Wingren C. (2009) Design of high-density antibody microarrays for disease proteomics: key technological issues. *J Proteomics* 72: 928-935.

Borrebaeck CA and Wingren C. (2011) Recombinant antibodies for the generation of antibody arrays. *Methods Mol Biol* 785: 247-262.

Borrebaeck CA and Wingren C. (2014) Antibody array generation and use. *Methods Mol Biol* 1131: 563-571.

Boutureira O and Bernardes GJ. (2015) Advances in chemical protein modification. *Chem Rev* 115: 2174-2195.

Brody EN, Gold L, Lawn RM, et al. (2010) High-content affinity-based proteomics: unlocking protein biomarker discovery. *Expert Rev Mol Diagn* 10: 1013-1022.

Burnet M. (1957) Cancer; a biological approach. I. The processes of control. *Br Med J* 1: 779-786.

Carlsson A, Persson O, Ingvarsson J, et al. (2010) Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients. *Proteomics Clin Appl* 4: 591-602.

Carlsson A, Wingren C, Ingvarsson J, et al. (2008) Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays. *Eur J Cancer* 44: 472-480.

Carlsson A, Wingren C, Kristensson M, et al. (2011a) Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases. *Proc Natl Acad Sci U S A* 108: 14252-14257.

Carlsson A, Wuttge DM, Ingvarsson J, et al. (2011b) Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. *Mol Cell Proteomics* 10: M110 005033.

Centlow M, Wingren C, Borrebaeck C, et al. (2011) Differential gene expression analysis of placentas with increased vascular resistance and pre-eclampsia using whole-genome microarrays. *J Pregnancy* 2011: 472354.

Chen L, Cohen J, Song X, et al. (2016) Improved variants of SrtA for site-specific conjugation on antibodies and proteins with high efficiency. *Sci Rep* 6: 31899.

Chin JW. (2017) Expanding and reprogramming the genetic code. *Nature* 550: 53-60.

Chin JW, Martin AB, King DS, et al. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* 99: 11020-11024.

Chin JW and Schultz PG. (2002) In vivo photocrosslinking with unnatural amino Acid mutagenesis. *Chembiochem* 3: 1135-1137.

Cho IH, Paek EH, Lee H, et al. (2007) Site-directed biotinylation of antibodies for controlled immobilization on solid surfaces. *Anal Biochem* 365: 14-23.

Crick FH, Brenner S, Watstobi.Rj, et al. (1961) General Nature of Genetic Code for Proteins. *Nature* 192: 1227-&.

Cull MG and Schatz PJ. (2000) Biotinylation of proteins in vivo and in vitro using small peptide tags. *Methods Enzymol* 326: 430-440.

Darmanis S, Nong RY, Vanelid J, et al. (2011) ProteinSeq: high-performance proteomic analyses by proximity ligation and next generation sequencing. *PLoS One* 6: e25583.

Davis L and Chin JW. (2012) Designer proteins: applications of genetic code expansion in cell biology. *Nat Rev Mol Cell Biol* 13: 168-182.

Delfani P, Dexlin Mellby L, Nordstrom M, et al. (2016) Technical Advances of the Recombinant Antibody Microarray Technology Platform for Clinical Immunoproteomics. *PLoS One* 11: e0159138.

Delfani P, Sturfelt G, Gullstrand B, et al. (2017) Deciphering systemic lupus erythematosus-associated serum biomarkers reflecting apoptosis and disease activity. *Lupus* 26: 373-387.

Dexlin-Mellby L, Sandstrom A, Centlow M, et al. (2010) Tissue proteome profiling of preeclamptic placenta using recombinant antibody microarrays. *Proteomics Clin Appl* 4: 794-807.

Dunn MR, Jimenez RM and Chaput JC. (2017) Analysis of aptamer discovery and technology. *Nature Reviews Chemistry* 1: 0076.

Ekins RP. (1989) Multi-analyte immunoassay. *J Pharm Biomed Anal* 7: 155-168.

Ellington AD and Szostak JW. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346: 818-822.

Ellmark P, Ingvarsson J, Carlsson A, et al. (2006) Identification of protein expression signatures associated with Helicobacter pylori infection and gastric adenocarcinoma using recombinant antibody microarrays. *Mol Cell Proteomics* 5: 1638-1646.

Farrah T, Deutsch EW, Omenn GS, et al. (2014) State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. *J Proteome Res* 13: 60-75.

Feldwisch J and Tolmachev V. (2012) Engineering of affibody molecules for therapy and diagnostics. *Methods Mol Biol* 899: 103-126.

Fenn JB, Mann M, Meng CK, et al. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246: 64-71.

Forne I, Ludwigsen J, Imhof A, et al. (2012) Probing the conformation of the ISWI ATPase domain with genetically encoded photoreactive crosslinkers and mass spectrometry. *Mol Cell Proteomics* 11: M111 012088.

Forrer P, Stumpp MT, Binz HK, et al. (2003) A novel strategy to design binding molecules harnessing the modular nature of repeat proteins. *FEBS Lett* 539: 2-6.

Frantzi M, Bhat A and Latosinska A. (2014) Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. *Clin Transl Med* 3: 7.

Frejd FY and Kim KT. (2017) Affibody molecules as engineered protein drugs. *Exp Mol Med* 49: e306.

Fulton RJ, McDade RL, Smith PL, et al. (1997) Advanced multiplexed analysis with the FlowMetrix system. *Clin Chem* 43: 1749-1756.

Fung ET. (2010) A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. *Clin Chem* 56: 327-329.

Furukawa H, Okada S, Saisho H, et al. (1996) Clinicopathologic features of small pancreatic adenocarcinoma. A collective study. *Cancer* 78: 986-990.

Fuzery AK, Levin J, Chan MM, et al. (2013) Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteomics* 10: 13.

Gerdtsson AS, Malats N, Sall A, et al. (2015) A Multicenter Trial Defining a Serum Protein Signature Associated with Pancreatic Ductal Adenocarcinoma. *Int J Proteomics* 2015: 587250.

Gerdtsson AS, Wingren C, Persson H, et al. (2016) Plasma protein profiling in a stage defined pancreatic cancer cohort - Implications for early diagnosis. *Mol Oncol* 10: 1305-1316.

Geyer PE, Holdt LM, Teupser D, et al. (2017) Revisiting biomarker discovery by plasma proteomics. *Mol Syst Biol* 13: 942.

Gillet LC, Navarro P, Tate S, et al. (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 11: O111 016717.

Graves PR and Haystead TA. (2002) Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev* 66: 39-63; table of contents.

Greenspan NS. (2017) Celebrating More Than a Century of Research on Antibodies: Affirmation Through Negation via Complex Formation. *Pathog Immun* 2: 60-65.

Griffin BA, Adams SR and Tsien RY. (1998) Specific covalent labeling of recombinant protein molecules inside live cells. *Science* 281: 269-272.

Groff K, Brown J and Clippinger AJ. (2015) Modern affinity reagents: Recombinant antibodies and aptamers. *Biotechnol Adv* 33: 1787-1798.

Haab BB. (2006) Applications of antibody array platforms. *Curr Opin Biotechnol* 17: 415-421.

Haab BB, Dunham MJ and Brown PO. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol* 2: RESEARCH0004.

Hall DA, Ptacek J and Snyder M. (2007) Protein microarray technology. *Mech Ageing Dev* 128: 161-167.

Hanahan D and Weinberg RA. (2011) Hallmarks of cancer: the next generation. *Cell* 144: 646-674.

Helma J, Cardoso MC, Muyldermans S, et al. (2015) Nanobodies and recombinant binders in cell biology. *J Cell Biol* 209: 633-644.

Holliger P and Hudson PJ. (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotechnol* 23: 1126-1136.

Hu X, Hortiguela MJ, Robin S, et al. (2013) Covalent and oriented immobilization of scFv antibody fragments via an engineered glycan moiety. *Biomacromolecules* 14: 153-159.

HUPO. (2010) A gene-centric human proteome project: HUPO - the Human Proteome organization. *Mol Cell Proteomics* 9: 427-429.

Hust M, Meyer T, Voedisch B, et al. (2011) A human scFv antibody generation pipeline for proteome research. *J Biotechnol* 152: 159-170.

Ingvarsson J, Larsson A, Sjoholm AG, et al. (2007) Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins. *J Proteome Res* 6: 3527-3536.

Ingvarsson J, Wingren C, Carlsson A, et al. (2008) Detection of pancreatic cancer using antibody microarray-based serum protein profiling. *Proteomics* 8: 2211-2219.

International Human Genome Sequencing C. (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.

Ito T, Sadamoto R, Naruchi K, et al. (2010) Highly oriented recombinant glycosyltransferases: site-specific immobilization of unstable membrane proteins by using Staphylococcus aureus sortase A. *Biochemistry* 49: 2604-2614.

Jacobitz AW, Kattke MD, Wereszczynski J, et al. (2017) Sortase Transpeptidases: Structural Biology and Catalytic Mechanism. *Adv Protein Chem Struct Biol* 109: 223-264.

Jensen RL, Stade LW, Wimmer R, et al. (2010) Direct site-directed photocoupling of proteins onto surfaces coated with beta-cyclodextrins. *Langmuir* 26: 11597-11604.

Karas M and Hillenkamp F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60: 2299-2301.

Kimple ME, Brill AL and Pasker RL. (2013) Overview of affinity tags for protein purification. *Curr Protoc Protein Sci* 73: Unit 9 9.

Kohler G and Milstein C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256: 495-497.

Kristensson M, Olsson K, Carlson J, et al. (2012) Design of recombinant antibody microarrays for urinary proteomics. *Proteomics Clin Appl* 6: 291-296.

Lange V, Picotti P, Domon B, et al. (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* 4: 222.

Lee CV, Liang WC, Dennis MS, et al. (2004) High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. *J Mol Biol* 340: 1073-1093.

Liu CC and Schultz PG. (2010) Adding new chemistries to the genetic code. *Annu Rev Biochem* 79: 413-444.

Liu YS and Yu J. (2016) Oriented immobilization of proteins on solid supports for use in biosensors and biochips: a review. *Microchimica Acta* 183: 1-19.

Lo YS, Nam DH, So HM, et al. (2009) Oriented immobilization of antibody fragments on Ni-decorated single-walled carbon nanotube devices. *ACS Nano* 3: 3649-3655.

Lu B, Smyth MR and O'Kennedy R. (1996) Oriented immobilization of antibodies and its applications in immunoassays and immunosensors. *Analyst* 121: 29R-32R.

MacBeath G. (2002) Protein microarrays and proteomics. *Nat Genet* 32 Suppl: 526-532.

MacBeath G and Schreiber SL. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* 289: 1760-1763.

Makaraviciute A and Ramanaviciene A. (2013) Site-directed antibody immobilization techniques for immunosensors. *Biosens Bioelectron* 50: 460-471.

Malhotra A. (2009) Tagging for protein expression. *Methods Enzymol* 463: 239-258.

Marks JD, Hoogenboom HR, Bonnert TP, et al. (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol* 222: 581-597.

Mathur S and Sutton J. (2017) Personalized medicine could transform healthcare. *Biomed Rep* 7: 3-5.

Matsuno S, Egawa S, Fukuyama S, et al. (2004) Pancreatic Cancer Registry in Japan: 20 years of experience. *Pancreas* 28: 219-230.

McCafferty J, Griffiths AD, Winter G, et al. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* 348: 552-554.

Murphy K, Travers P and Walport M. (2007) *Janeway's Immunobiology,* New York: Garland Science.

Nelson RW, Krone JR, Bieber AL, et al. (1995) Mass spectrometric immunoassay. *Anal Chem* 67: 1153-1158.

Niemeyer CM, Adler M and Wacker R. (2007) Detecting antigens by quantitative immuno-PCR. *Nat Protoc* 2: 1918-1930.

Niemeyer CM, Boldt L, Ceyhan B, et al. (1999) DNA-Directed immobilization: efficient, reversible, and site-selective surface binding of proteins by means of covalent DNA-streptavidin conjugates. *Anal Biochem* 268: 54-63.

Nirenberg M and Matthaei JH. (1961) Dependence of Cell-Free Protein Synthesis in E Coli Upon Naturally Occurring or Synthetic Polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America* 47: 1588-&.

Nord K, Gunneriusson E, Ringdahl J, et al. (1997) Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain. *Nat Biotechnol* 15: 772-777.

Nordstrom M, Wingren C, Rose C, et al. (2014) Identification of plasma protein profiles associated with risk groups of prostate cancer patients. *Proteomics Clin Appl* 8: 951-962.

Noren CJ, Anthony-Cahill SJ, Griffith MC, et al. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244: 182-188.

Olsson N, Carlsson P, James P, et al. (2013) Grading breast cancer tissues using molecular portraits. *Mol Cell Proteomics* 12: 3612-3623.

Olsson N, James P, Borrebaeck CAK, et al. (2012) Quantitative Proteomics Targeting Classes of Motif-containing Peptides Using Immunoaffinity-based Mass Spectrometry. *Molecular & Cellular Proteomics* 11: 342-354.

Olsson N, Wingren C, Mattsson M, et al. (2011) Proteomic analysis and discovery using affinity proteomics and mass spectrometry. *Mol Cell Proteomics* 10: M110 003962.

Pauly F, Dexlin-Mellby L, Ek S, et al. (2013) Protein expression profiling of formalin-fixed paraffin-embedded tissue using recombinant antibody microarrays. *J Proteome Res* 12: 5943-5953.

Pauly F, Fjorden K, Leppa S, et al. (2016) Plasma immunoprofiling of patients with high-risk diffuse large B-cell lymphoma: a Nordic Lymphoma Group study. *Blood Cancer J* 6: e501.

Pauly F, Smedby KE, Jerkeman M, et al. (2014) Identification of B-cell lymphoma subsets by plasma protein profiling using recombinant antibody microarrays. *Leuk Res* 38: 682-690.

Pavlou MP, Diamandis EP and Blasutig IM. (2013) The long journey of cancer biomarkers from the bench to the clinic. *Clin Chem* 59: 147-157.

Perols A and Karlstrom AE. (2014) Site-specific photoconjugation of antibodies using chemically synthesized IgG-binding domains. *Bioconjug Chem* 25: 481-488.

Petersson L, Berthet Duroure N, Auger A, et al. (2014a) Generation of miniaturized planar ecombinant antibody arrays using a microcantilever-based printer. *Nanotechnology* 25: 275104.

Petersson L, Coen M, Amro NA, et al. (2014b) Miniaturization of multiplexed planar recombinant antibody arrays for serum protein profiling. *Bioanalysis* 6: 1175-1185.

Petersson L, Dexlin-Mellby L, Bengtsson AA, et al. (2014c) Multiplexing of miniaturized planar antibody arrays for serum protein profiling--a biomarker discovery in SLE nephritis. *Lab Chip* 14: 1931-1942.

Petersson L, Stade LW, Brofelth M, et al. (2014d) Molecular design of recombinant scFv antibodies for site-specific photocoupling to beta-cyclodextrin in solution and onto solid support. *Biochim Biophys Acta* 1844: 2164-2173.

Pluckthun A. (2015) Designed ankyrin repeat proteins (DARPins): binding proteins for research, diagnostics, and therapy. *Annu Rev Pharmacol Toxicol* 55: 489-511.

Poetz O, Hoeppe S, Templin MF, et al. (2009) Proteome wide screening using peptide affinity capture. *Proteomics* 9: 1518-1523.

Ponomarenko EA, Poverennaya EV, Ilgisonis EV, et al. (2016) The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem* 2016: 7436849.

Renberg B, Nordin J, Merca A, et al. (2007) Affibody molecules in protein capture microarrays: evaluation of multidomain ligands and different detection formats. *J Proteome Res* 6: 171-179.

Renberg B, Shiroyama I, Engfeldt T, et al. (2005) Affibody protein capture microarrays: synthesis and evaluation of random and directed immobilization of affibody molecules. *Anal Biochem* 341: 334-343.

Rissin DM, Kan CW, Campbell TG, et al. (2010) Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nat Biotechnol* 28: 595-599.

Robertson DL and Joyce GF. (1990) Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* 344: 467-468.

Sadee W and Dai Z. (2005) Pharmacogenetics/genomics and personalized medicine. *Hum Mol Genet* 14 Spec No. 2: R207-214.

Sall A, Carlsson F, Olsson N, et al. (2014) AFFIRM--a multiplexed immunoaffinity platform that combines recombinant antibody fragments and LC-SRM analysis. *J Proteome Res* 13: 5837-5847.

Sall A, Persson H, Ohlin M, et al. (2016a) Advancing the global proteome survey platform by using an oriented single chain antibody fragment immobilization approach. *N Biotechnol* 33: 503-513.

Sall A, Walle M, Wingren C, et al. (2016b) Generation and analyses of human synthetic antibody libraries and their application for protein microarrays. *Protein Eng Des Sel* 29: 427-437.

Sanchez-Carbayo M. (2011) Antibody microarrays as tools for biomarker discovery. *Methods Mol Biol* 785: 159-182.

Sandstrom A, Andersson R, Segersvard R, et al. (2012) Serum proteome profiling of pancreatitis using recombinant antibody microarrays reveals disease-associated biomarker signatures. *Proteomics Clin Appl* 6: 486-496.

Schulte-Zweckel J, Rosi F, Sreenu D, et al. (2014) Site-specific, reversible and fluorescent immobilization of proteins on CrAsH-modified surfaces for microarray analytics. *Chem Commun (Camb)* 50: 12761-12764.

Schulte-Zweckel J, Rosi F, Sreenu D, et al. (2016) High Affinity Immobilization of Proteins Using the CrAsH/TC Tag. *Molecules* 21.

Schwenk JM, Gry M, Rimini R, et al. (2008) Antibody suspension bead arrays within serum proteomics. *J Proteome Res* 7: 3168-3179.

Schwenk JM and Nilsson P. (2011) Antibody suspension bead arrays. *Methods Mol Biol* 723: 29-36.

Shen Z, Mernaugh RL, Yan H, et al. (2005) Engineered recombinant single-chain fragment variable antibody for immunosensors. *Anal Chem* 77: 6834-6842.

Siegel RL, Miller KD and Jemal A. (2017) Cancer Statistics, 2017. *CA Cancer J Clin* 67: 7-30.

Skoog P, Ohlsson M, Ferno M, et al. (2017) Tumor tissue protein signatures reflect histological grade of breast cancer. *PLoS One* 12: e0179775.

Smith LM, Kelleher NL and Consortium for Top Down P. (2013) Proteoform: a single term describing protein complexity. *Nat Methods* 10: 186-187.

Soderlind E, Strandberg L, Jirholt P, et al. (2000) Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nature Biotechnology* 18: 852-856.

Steinhauer C, Wingren C, Hager AC, et al. (2002) Single framework recombinant antibody fragments designed for protein chip applications. *Biotechniques* Suppl: 38-45.

Steinhauer C, Wingren C, Khan F, et al. (2006) Improved affinity coupling for antibody microarrays: engineering of double-(His)6-tagged single framework recombinant antibody fragments. *Proteomics* 6: 4227-4234.

Thatikonda N, Delfani P, Jansson R, et al. (2016) Genetic fusion of single-chain variable fragments to partial spider silk improves target detection in micro- and nanoarrays. *Biotechnol J* 11: 437-448.

Tighe PJ, Ryder RR, Todd I, et al. (2015) ELISA in the multiplex era: potentials and pitfalls. *Proteomics Clin Appl* 9: 406-422.

Todd J, Freese B, Lu A, et al. (2007) Ultrasensitive flow-based immunoassays using single-molecule counting. *Clin Chem* 53: 1990-1995.

Trilling AK, Beekwilder J and Zuilhof H. (2013a) Antibody orientation on biosensor surfaces: a minireview. *Analyst* 138: 1619-1627.

Trilling AK, Harmsen MM, Ruigrok VJ, et al. (2013b) The effect of uniform capture molecule orientation on biosensor sensitivity: dependence on analyte properties. *Biosens Bioelectron* 40: 219-226.

Tuerk C and Gold L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249: 505-510.

Tyers M and Mann M. (2003) From genomics to proteomics. *Nature* 422: 193-197.

Uhlen M, Bandrowski A, Carr S, et al. (2016) A proposal for validation of antibodies. *Nat Methods* 13: 823-827.

Wang L, Brock A, Herberich B, et al. (2001) Expanding the genetic code of Escherichia coli. *Science* 292: 498-500.

Watson JD and Crick FHC. (1953) Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* 171: 737-738.

Weiss F, van den Berg BH, Planatscher H, et al. (2014) Catch and measure-mass spectrometry-based immunoassays in biomarker research. *Biochim Biophys Acta* 1844: 927-932.

Welch NG, Scoble JA, Muir BW, et al. (2017) Orientation and characterization of immobilized antibodies for improved immunoassays (Review). *Biointerphases* 12: 02D301.

Westerlund K, Honarvar H, Tolmachev V, et al. (2015) Design, Preparation, and Characterization of PNA-Based Hybridization Probes for Affibody-Molecule-Mediated Pretargeting. *Bioconjug Chem* 26: 1724-1736.

Wingren C and Borrebaeck CA. (2006) Antibody microarrays: current status and key technological advances. *OMICS* 10: 411-427.

Wingren C, Ingvarsson J, Dexlin L, et al. (2007) Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support. *Proteomics* 7: 3055-3065.

Wingren C, Sandstrom A, Segersvard R, et al. (2012) Identification of serum biomarker signatures associated with pancreatic cancer. *Cancer Res* 72: 2481-2490.

Wingren C, Steinhauer C, Ingvarsson J, et al. (2005) Microarrays based on affinity-tagged single-chain Fv antibodies: sensitive detection of analyte in complex proteomes. *Proteomics* 5: 1281-1291.

Worn A and Pluckthun A. (2001) Stability engineering of antibody single-chain Fv fragments. *J Mol Biol* 305: 989-1010.

Yalow RS and Berson SA. (1959) Assay of plasma insulin in human subjects by immunological methods. *Nature* 184 (Suppl 21): 1648-1649.

Yu X, Yang YP, Dikici E, et al. (2017) Beyond Antibodies as Binding Partners: The Role of Antibody Mimetics in Bioanalysis. *Annu Rev Anal Chem (Palo Alto Calif)* 10: 293-320.