LUND UNIVERSITY

Biomarker Discovery in Cancer and Autoimmunity using an Affinity Proteomics Platform - a Tool for Personalized Medicine

Nordström, Malin

2013

Biomarker Discovery in Cancer and Autoimmunity using an Affinity Proteomics Platform

- a Tool for Personalized Medicine

Malin Nordström

LUND
UNIVERSITY

Department of Immunotechnology,
Lund University 2013

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. Kristensson, M., Olsson, K., Carlson, J., Wullt, B., Sturfelt, G., Borrebaeck, CAK., Wingren, C.,
Design of recombinant antibody microarrays for urinary proteomics *Proteomics Clin. Appl. 2012 Jun;6(5-6):291-6*

II. Nordström, M., Vallkil, J., Borrebaeck, CAK., and Wingren, C.,
Stability engineering of recombinant antibodies for microarray applications *Manuscript*

III. Nordström, M., Stål Hallengren, C., Mårtensson, S., Bengtsson, A., Sturfelt, G., Borrebaeck, CAK. and Wingren, C.,
Serum and urine biomarker signatures reflecting disease activity in systemic lupus erythematosus revealed by affinity proteomics *Manuscript*

IV. Nordström, M., Wingren, C., Rose, C., Bjartell, A., Becker, C., Lilja, H., and Borrebaeck, CAK.,
Identification of Plasma Protein Profiles Associated with Prostate Cancer Risk Groups. *Manuscript submitted for publication.*

My contribution to the papers

Paper I. I planned experiments and performed experiments together with KO. I participated in writing the manuscript.

Paper II. I performed experiments together with JV. I analyzed the data and participated in writing the manuscript.

Paper III. I planned experiments and performed experiments together with SM. I analyzed the data and participated in writing the manuscript.

Paper IV. I planned and performed all experiments. I analyzed the data and was main responsible for writing the manuscript.

I have also contributed to the following scientific papers, not included in this thesis:

i) Gustavsson E, Ek S, Steen J, Kristensson M, Algenäs C, Uhlén M, Wingren C, Ottosson J, Hober S, Borrebaeck CAK.
Surrogate antigens as targets for proteome-wide binder selection. *N Biotechnol. 2011 Jul;28(4):302-11*

ii) Carlsson A, Wingren C, Kristensson M, Rose C, Fernö M, Olsson H, Jernström H, Ek S, Gustavsson E, Ingvar C, Ohlsson M, Peterson C, Borrebaeck CAK
Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases *Proc Natl Acad Sci U S A. 2011 Aug 23;108(34):14252-7*

Abbreviations

%fPSA	ratio between free and total serum prostate specific
AUC	area under the curve
CDR	complementary determining regions
CML	chronic myeloid leukemia
EGFR	epidermal growth factor receptor
ELISA	enzyme-linked immunosorbant assay
Fab	fragment antigen-binding
FDA	US food and drug administration
FFPE	formalin-fixed and paraffin-embedded
FW	framework
GdmCl	guanidine hydrochloride
Ig	immunoglobulin
mAb	monoclonal antibodies
MRM	multiplexed reaction monitoring
MS	mass spectrometry
NSCLC	non-small cell lung cancer
pAb	polyclonal antibodies,
PCR	polymerase chain reaction
PSA	prostate specific antigen
PTM	post translational modification
RCA	rolling circle amplification
ROC	receiver operating characteristics
RPPM	reverse-phase protein microarrays
S/N	signal-to-noise ratio
scFv	single chain Fragment variable
SLE	systemic lupus erythematosus
SLEDAI	SLE disease activity index
SVM	support vector machine
tPSA	total serum PSA
VH	variable domain of immunoglobulin heavy chain
VL	variable domain of immunoglobulin light chain
Tm	melting temperature
TSA	tyramide signal amplification

1. Introduction

Medicine has always been personal, and aimed at giving each patient optimal and individualized treatment. The term "**personalized medicine**" has in the last decades been referred to the tailoring of medical treatment based on individual characteristics of each patient, giving the "right treatment to the right person at the right time" (Bates 2010). Traditionally, these characteristics have been solely of a clinical and demographic nature, such as performance status and age of the patient. However, in recent years, genetic and protein biomarkers has emerged and now enable more detailed decoding of personal differences that can be used for even more specific treatment selection (Bates 2010, Mehta, Jain et al. 2011). Two major indications, with large unmet clinical needs demanding individualized management of patients, are cancer and autoimmune disorders (Rovin, McKinley et al. 2009, Ross 2011).

Most people in developed countries are today affected by **cancer** in one way or another. One out of three people will be diagnosed with cancer in their life-time, and this is a number expected to increase to 50%, due to an aging population and life-style choices (Stein and Colditz 2004). Great hopes are set on the field of personalized medicine for providing e.g. early and accurate diagnostics, classification of tumors into distinct molecular subtypes, each with a corresponding treatment, and monitoring of disease relapse. Detecting tumors in an early stage improves the odds of successful treatment, and treatment selection based on molecular subtypes has been shown to be essential for the efficacy of a number of treatment regimens (e.g. therapeutic agents imatinib in chronic lymphoid leukemia and trastuzumab in HER-2 positive breast cancers) (Joske 2008, Ross, Slodkowska et al. 2009). **Autoimmune diseases** are often chronic and systemic disorders, characterized by diverse manifestations, motivating individualized management of patients for optimal prognosis (Maecker, Lindstrom et al. 2012). The benefit of personalizing the treatment lies not only in treating the right patients, but also in sparing those who would not need or respond to the treatment. Current treatment regimens for cancer and

autoimmune diseases are often associated with severe side-effects and the severity of the disease will be a major factor for deciding how much side-effects can be tolerable.

Aiming for detection of novel gene and protein markers for diagnosis and prognosis of disease, numerous **biomarker discovery** studies have been reported recently, while the clinical utility of these markers remain to be proven (Boschetti, Chung et al. 2012). Clinical demands on biomarkers include their ability to answer a clinical question with high specificity and sensitivity, and that they can be reliably measured in an accessible sample format (Sanchez-Carbayo 2011). Protein biomarkers are an attractive solution to these demands, as proteins are the actual executor of most cellular events and are available in body fluids for minimal invasive sampling.

Proteomic techniques are powerful discovery tools, targeting up to thousands of proteins in a single experiment. In this context, **affinity proteomics**, with antibody arrays in particular, has positioned itself as a sensitive, multiplex and high-throughput tool for biomarker discovery (Stoevesandt and Taussig 2012). Our group has in the last decade developed and implemented an affinity proteomic platform, where recombinant antibodies are printed onto a solid surface, creating an array of binder molecules (Wingren, Ingvarsson et al. 2007, Borrebaeck and Wingren 2009). The analyzed sample is labeled and added to this antibody array, and bound proteins are detected in a scanner. By comparing the detected protein patterns in samples of different disease status, disease related protein signatures can be identified. Key features for the assay is the on-chip performance of the affinity probes, and optimized protocols for analysis of all relevant clinical sample formats.

The aim of this thesis has been to further optimize key features of our affinity proteomics platform, recombinant antibody microarrays, and to apply the platform in clinical studies. This work is based on four original papers, where **paper I and II** address technology development of the platform, while in **paper III and IV** the optimized platform is applied in clinical studies. Large efforts have been devoted to optimizing all different parameters including choice of surface, printing parameters, detection system and choice of probes. The analyzed sample formats include serum/plasma, tissue extracts, cell lysates, intact cells, and I have in **paper I** extended our platform and re-optimized the set-up for urine analysis. The on-chip stability of the affinity probes is a key

feature for a robust and reproducible array set-up, and has this been evaluated and further optimized in **paper II**.

The optimized affinity proteomic platform has then been applied in two clinical studies targeting prostate cancer and the autoimmune disorder systemic lupus erythematosus (SLE). In **paper III**, I have analyzed serum and urine samples from patients with the most severe manifestation of SLE, SLE nephritis. Candidate protein biomarker signatures associated with disease activity has been identified. This data is a first step towards monitoring and ultimately predicting flares, which would enable individualized management and therapy selection of SLE nephritis patients. In **paper IV**, I have analyzed plasma samples from potential prostate cancer patients. The data showed that we have successfully identified biomarkers that could be used for stratification of patient risk groups. Of note, heterogeneous patient groups could be stratified into groups of high or low risk of having prostate cancer. Thus, we showed that our affinity proteomics platform could be used for identification of biomarker signatures for decision basis in the selection of patients for biopsy testing.

2. Biomarkers in personalized medicine

A disease biomarker is virtually anything that can be used as an indicator of disease, but the term has predominantly been used for genes or proteins that can be detected in tissue or body fluids and reflect a disease status. In order to pursue personalized medicine, access to well-defined biomarkers will be a prerequisite for correct and effective decision making in diagnosis, prognosis and treatment decision (Mehta, Jain et al. 2011). An ideal biomarker would be a single molecule, easily detected in a patient with a certain disease, but not at all detected in a healthy person. In reality, these kinds of magic bullets rarely exist, forcing us to study more complex patterns of genes or proteins (Wallstrom, Anderson et al. 2013). The performance of biomarkers is usually evaluated in terms of sensitivity and specificity, where sensitivity is the ability to detect disease where the disease is truly present, and specificity is the ability to accurately recognize absence of disease.

In this chapter, I will exemplify important gene and protein markers that substantially have influenced over-all survival and quality of life for thousands of patients in a variety of diseases, and I then focus on the role of biomarkers and personalized medicine in prostate cancer and SLE. Finally, I will address some of the challenges scientists are faced with when pursuing biomarker discovery.

2.1 Gene and protein biomarkers

The mapping of the human genome at the turn of the century has enabled large scale studies of genetic profiles, as well as identification of mutations and altered expression profiles. This has resulted in discovery of individual genes or

gene profiles associated with different diseases or response to treatment. Proteins are more complex than DNA both in structure and composition, placing higher demands on the techniques used (Phizicky, Bastiaens et al. 2003). On the other hand, proteins hold great promise in harboring more information on current disease status, as they are the actual executor of molecular events. Gene and protein markers often provide complementary information, and will continue to play important roles in personalized medicine, independent of each other or used in combination.

2.1.1 Genetic and gene expression biomarkers

Genetic biomarkers have so far predominantly been identified in oncology, where mutations and translocations can e.g. inactivate tumor suppressors, or result in fusion proteins with oncogenic properties. An early example of gene based personalized medicine is the identification of the Philadelphia chromosome in chronic myeloid leukemia (CML). A reciprocal translocation between chromosomes 9 and 22 (Rowley 1973), known as the Philadelphia chromosome, is responsible for the fusion protein BCR-ABL which induces the myeloproliferative disorder typical of CML. Presence of the Philadelphia chromosome identifies CML with 100% specificity among other leukemia, and these patients can effectively be treated with tyrosine-kinase inhibitor imatinib (Gleevec/Glivec) targeting BCR-ABL (Joske 2008). A more recent example is the use of gefitinib (Iressa) in non-small cell lung cancer (NSCLC) (Paez, Janne et al. 2004). Iressa was first approved for treatment of NSCLC, but withdrawn due to disappointing results in phase II studies. Further retrospective studies showed association between epidermal growth factor receptor (EGFR) mutation status and response to Iressa treatment, and Iressa was in 2010 again approved for treatment, this time for the subset of NSCLC patients with confirmed EGFR mutations. The Philadelphia chromosome and the mutated EGFR are examples of gene-based companion diagnostics, gene biomarkers crucial for the employment of the corresponding therapy.

Extensive work in gene expression profiling has resulted in identification of mRNA signatures associated with different sub-sets of breast cancer. In 2002, van't Veer and colleagues presented a gene expression profile for prediction of clinical outcome (short interval to distant metastasis) of breast cancer patients (van 't Veer, Dai et al. 2002). After optimization and validation, a 70-gene signature (MammaPrint®) was in 2007 approved by US food and drug

administration (FDA) as the first diagnostic microarray test (Cardoso, Van't Veer et al. 2008). Similarly, in 2004 Paik et al. identified a 21 gene polymerase chain reaction (PCR) panel (Oncotype DX) that predicts disease relapse in a subset of breast cancer patients receiving endocrine therapy (tamoxifen) (Paik, Shak et al. 2004). Also, the feasibility of using DNA array data for stratification of breast cancer patients into subgroups has been elegantly demonstrated by the Börresen-Döle group (Sorlie, Perou et al. 2001). Gene expression patterns derived from cDNA microarrays were used for unsupervised clustering of breast cancer patients and the obtained cluster groups correlated to the clinical subgroups, which include basal like, ERBB2 positive, normal breast like and luminal breast cancer, with high accuracy.

2.1.2 Protein biomarkers

The notion that mRNA levels on many occasions do not correlate with protein levels (Gygi, Rochon et al. 1999) has fueled the interest of identifying protein and protein profiles as markers of disease (Liang and Chan 2007). Protein biomarkers can be detected in tissue samples using antibody probes, or as circulating proteins in serum or other body fluids. The human epidermal growth factor receptor (HER2) is a trans-membrane tyrosine kinase receptor up-regulated in 10-34% of invasive breast cancers (Schechter, Stern et al. 1984), and is today routinely used both as a tissue biomarker for classification of aggressive cancers and as an effective drug target. The monoclonal antibody Herceptin (trastuzumab) targets HER2 and is solely administered to HER2-positive patients, most likely to respond to the treatment. Herceptin is associated with substantial risk of cardio toxicity (Telli, Hunt et al. 2007), why sparing HER2-negative patients from this therapy improves their quality of life (Ross, Slodkowska et al. 2009).

Detecting circulating protein biomarkers is an attractive approach, due to their less invasive sampling procedures. The use of serum prostate specific antigen (PSA) for assessment of risk of prostate cancer has revolutionized care of prostate cancer patients, and will be further discussed in section 2.2. Several circulating glycoproteins have been proposed as tumor markers (Chatterjee and Zetter 2005). Elevated levels of CA19-9 (sialylated Lewis (a) antigen) were initially detected in colorectal cancer cell lines(Koprowski, Steplewski et al. 1979). Since then, several studies have shown correlation between increased serum levels of CA19-9 and pancreatic cancer (Goonetilleke and Siriwardena

2007). However, due to insufficient specificity (68–91%) and sensitivity (70–90%) of the test, CA19-9 is not recommended as a diagnostic biomarker. Possible causes for false positives include elevated levels due to jaundice, and the low sensitivity can in part be explained by that certain people are lewis-negative (von Rosen, Linder et al. 1993). In pancreatic cancer patients that do have a verified CA19-9 secretion, the marker can be used for monitoring of response to treatment and of disease recurrence (Goonetilleke and Siriwardena 2007). Glycoprotein mucin 16, also known as CA-125, is used as a marker for detection of ovarian cancer with a sensitivity of 80-90 % (Canney, Moore et al. 1984). The specificity is, however, more modest, as CA-125 can be elevated in other cancers and benign states, while usually in lower levels. Circulating protein biomarkers also have the capability of identifying more acute events, as Troponin T detecting myocardial infarctions (Mair, Artner-Dworzak et al. 1991) and C-reactive protein as a marker of inflammation (Tillett and Francis 1930, Ridker 2009).

Using a single protein biomarker would obviously be the most practical choice for point-of-care applications. However, due to the complexity of many diseases such as cancer and auto-immune diseases, physicians will most likely have to rely on multiplex marker signatures (Chatterjee and Zetter 2005, Liang and Chan 2007, Wallstrom, Anderson et al. 2013). This applies especially for markers for early detection, where the probed population constitutes of a group of vast heterogeneity in individual pathophysiology, as exemplified with CA19-9 above. Multiplex markers can be obtained either by combining different known markers (Cordero, De Chiara et al. 2008, Bansal and Sullivan Pepe 2013), or by designing discovery studies for identification of complex patterns, and the latter approach has been the focus of this thesis.

I will next turn to exemplifying current diagnostic procedures and challenges in prostate cancer and SLE.

2.2 Personalized medicine in prostate cancer

Prostate cancer is currently the most frequently diagnosed cancer among men in developed countries (Ferlay, Shin et al. 2010), and for improved prognosis individualized management of these patients is required. In the process of

diagnosing prostate cancer, the physicians are faced with two major challenges: First, who is at risk of having prostate cancer and should be selected for biopsy testing?, and second, once a malignancy is detected, what treatment alternative should be chosen?

The first challenge was revolutionized by the introduction of PSA testing, resulting in an increased number of early diagnosed cases (Parekh, Ankerst et al. 2007, Shariat, Semjonow et al. 2011). Elevated total serum PSA (tPSA) is associated with prostate cancer, as the malignant prostate usually leaks PSA to much larger extent than the healthy prostate. There is, however, also a significant leakage of PSA from a prostate of benign enlargement (BPH), which is a common complication among aging men. Therefore, PSA testing has dramatically increased the number of unnecessary biopsies, causing a major burden on both well-being of individual patients and national health economics.

In order to improve PSA's specificity for malignant disease, the ratio between free and tPSA (%fPSA) can be assessed (Lilja, Christensson et al. 1991, Catalona, Partin et al. 1998). PSA circulates in the blood stream, both free as well as complex bound. The free, non-complexed form has shown more frequent in leakage from a prostate of benign enlargement, why men with %fPSA above 15-20% is usually spared from biopsy testing. Still, men subjected to biopsy testing are a very heterogeneous group (Parekh, Ankerst et al. 2007), why further stratification of this patient cohort is essential, and was explored in **paper IV**.

Turning to the second challenge of treatment selection, it should be noted that detection of malignant tissue might not always motivate heavy treatment: For instance, 25-35% of young men have indolent tumors in prostatic tissue that, in most cases, will not progress into aggressive tumors (autopsy finding on men with other cause of death (Sakr, Haas et al. 1993)). For classification of detected tumors, and treatment selection, factors to consider include grading and staging of the tumor and demographic factors, such as patient age. The grading of the tumor is based on the histological assessment of a biopsy specimen and presented as a Gleason score, where a high score represents poorly differentiated prostate gland cells and a high risk of metastasis (Gleason and Mellinger 1974). The staging communicates if the tumor is spread to lymph nodes or further metastasized, usually using the Tumor, Lymph Node, and

Metastasis staging system (Cheng, Montironi et al. 2012). As a basis for treatment selection, these factors are compiled into classification systems (D'Amico, Desjardin et al. 1998) or more complex predictive algorithms, known as nomograms (Katz, Efstathiou et al. 2010). Therapy options include prostatectomy and hormonal treatment, both associated with severe side-effects as impotence and incontinence. Active surveillance is a treatment option of indolent cancers, especially among elderly patients. Still, the difficulty of distinguishing indolent from aggressive tumors remains and motivates the need for improvement of classification systems.

2.3 Personalized medicine in systemic lupus erythematosus (SLE)

Systemic lupus erythematosus (SLE) is a chronic, autoimmune disorder characterized by the formation of autoantibodies and immune complexes, leading to a plethora of different clinical presentations and manifestations, ranging from rashes to glomerulonephritis (Tsokos 2011). The diagnosis of SLE include 11 classification criteria, and patients displaying four or more of these criteria are diagnosed with a specificity of 95% and a sensitivity of 85 % (Maidhof and Hilas 2012). Although certain clinical presentations are common for many SLE patients, the disease is to great extent characterized by a unique set of identifiers and autoantibody repertoires for each patient, requiring an individualized approach in treatment decision (Agmon-Levin, Mosca et al. 2012). In 2011, FDA approved the monoclonal antibody belimumab for treatment of SLE patients, as the first novel therapy in SLE for 56 years (Chugh and Kalra 2013). Only around 30% of the patients benefit from belimumab treatment, and patients with severe manifestations as kidney involvement were not included in the clinical trials. Further studies are required to evaluate which sub-populations would benefit most from belimumab treatment, in order to more accurately decide who is eligible for therapy.

The underlying disease etiology of SLE is still largely unknown, but the heterogeneity of symptoms has led to the suggestion that SLE is actually a variety of different diseases with diverse pathogenic mechanism (Agmon-Levin, Mosca et al. 2012). This notion motivates studies of stratification of SLE into

different sub-diseases, which has primarily been taken on using genetic studies in the last decade. For instance, mapping of SLE genes into pathogenetic pathways has revealed that a subgroup of patients with an activated interferon-α (IFN-α) pathway were associated with distinct serologic features (low complement, high α-dsDNA) (Kirou, Lee et al. 2005).

SLE patients go through periods of active disease (flares) and periods of inactive disease (remission) (Tsokos 2011). The disease itself is chronic, but the flares can be reduced using effective treatment regimens. SLE disease activity is currently assessed using activity indices, for instance SLE disease activity index 2000 (SLEDAI-2K), covering systemic symptoms, and renal SLEDAI, pin-pointing renal involvement. Albeit useful, the SLEDAI-2K index requires observation of 24 different clinical parameters observed over a longer (> 10 days) time period, which could delay treatment. Therefore, molecular biomarkers for monitoring, or ultimately, predicting flares could improve quality of life for SLE patients (Gibson, Banha et al. 2010). Markers of disease activity used in clinics today include complement protein C3 and auto-antibodies directed against complement protein C1q, but their accuracy is unfortunately limited, why additional markers are highly warranted (Rovin and Zhang 2009). Also, the heterogeneity of the disease motivates the need to study multiplex panels of biomarkers (Wallstrom, Anderson et al. 2013), which has been pursued in **paper III**.

2.4 Challenges in biomarker discovery

Pursuing protein biomarker discovery is faced with a number of challenges. Recently evolved proteomic techniques have reported numerous candidate biomarkers (Hu, Loo et al. 2006, Lescuyer, Hochstrasser et al. 2007), while the transition into clinical application of these potential markers has been much more modest (Anderson, Ptolemy et al. 2013). The reasons for this discrepancy could be several, and I will here focus on the impact of study design, sample format and requirements on the techniques used.

2.4.1 Study design

The route of biomarker development, from raising a valid clinical question to implementation in clinical practice, has proven to be long and difficult. The starting-point of all biomarker discovery studies should include addressing an unmet clinical need, why close collaborations between scientists and practicing physicians is essential. It has even been proposed that national health institutes ought to be involved in prioritizing important clinical questions by their impact on overall healthcare (Anderson, Ptolemy et al. 2013). Once the relevant clinical question is formulated, the optimal study design is to be chosen.

Biomarker discovery studies can be performed as case-control studies where one group of patients are compared to a control group, or longitudinal cohort studies, where patients are followed and sampled over a period of time (Mann 2003). A case-control study design is attractive due to its relative speed and cost-effectiveness, while hampered by difficulties in the selection of, and access to, representative cases and controls. Cases might be few and time-consuming to collect in sufficient number, and the controls should be absent of the disease that they control for, but in all other aspects be comparable to the cases. Case-control studies are faced with a substantial risk of identifying candidate markers reflecting differences related to the particular patient cohort and not to the disease *per se*, which could be a reason for many candidate marker not transforming into clinical practice.

Longitudinal studies are performed either retrospectively, where previously collected samples are analyzed at one time-point and related to the present clinical outcome of the patient, or prospectively where the cases are followed over time and samples are collected at different occasions (Euser, Zoccali et al. 2009). The retrospective study is faster and more convenient, but relies on the relevant samples or data being collected. The prospective study can take several years to follow up, but is more likely to provide markers of clinical utility (Euser, Zoccali et al. 2009, Brennan, O'Connor et al. 2010).

The process of bringing candidate biomarker signatures into clinical implementation has turned out to be very challenging, and a successful discovery study is followed be several validation phases (Rifai, Gillette et al. 2006, Puntmann 2009) (Figure 1). In the initial discovery phase, a candidate biomarker panel, sometimes encompassing hundreds of different markers, is

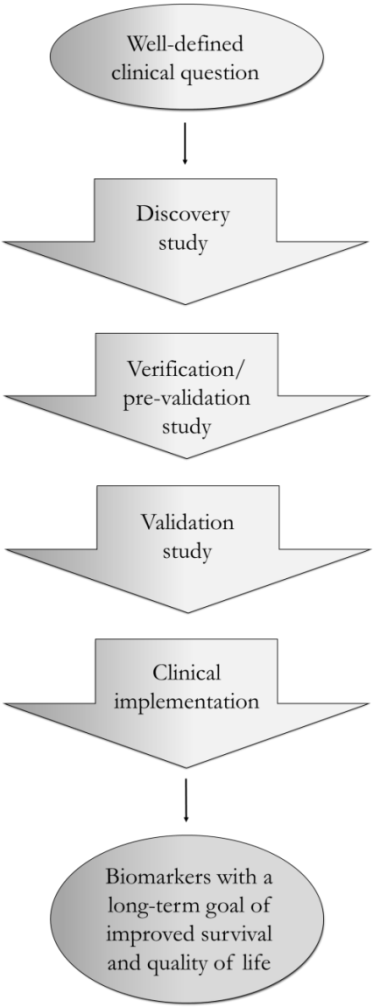

Figure 1. All biomarker studies ought to start with a well-defined clinical need. The biomarker discovery study is then followed by validation studies and finally introduction into a clinical setting.

identified. In a second step, denoted pre-validation or verification, these candidate panels are condensed and then validated in a second independent data-set. Third, the condensed biomarker panel is validated in a large independent population, using the analysis platform attended for its clinical application (e.g. an immunoassay). The large number of samples needed in the validation studies can be demanding to access, and has often become a key bottleneck. Finally, after approval from regulatory authorities (e.g. FDA for the US market) the validated biomarker(s) can be introduced into a clinical setting, and the long-term clinical utility, e.g. improved survival, can be assessed. The final step of introducing a biomarker into the clinic is strictly controlled by regulatory authorities. However, the process of taking the candidate through the proceeding pre-validation and verification have fewer guidelines, in contrast to the drug discovery pipe-line where each phase is carefully regulated (Anderson, Ptolemy et al. 2013). Also, the discovery phase is usually performed in academia, while the point-of-care assay is developed in a commercial/industry setting, and the transition between the two demands new routes of financing of projects etc. (Mischak, Ioannidis et al. 2012).

Taken together, formulation of a clinical question, choice of study design and strategy for validation studies are all crucial factors in the route of developing and implementing biomarkers. In addition, the patient subgroup identified by the marker requires an available

treatment option, in order to make the biomarker attractive for the clinic application. It is, however, not rare that the discovery of a marker subsequently has led to discovery of drug target(s), as in the example of the Philadelphia chromosome above.

2.4.2 Samples for biomarker discovery

The outcome of a biomarker discovery study relies to great extent on the nature and quality of the analyzed biological sample, usually a tissue specimen or a biological fluid, such as serum or urine. The choice of sample format involves both demands from clinic and from the chosen analysis platform, and the latter will be discussed further in section 4.2.1. From the clinician's and patient's point of view, the sample should preferably be obtained through non-invasive, convenient, and cost-effective sampling, and only require simple protocols for handling and storage.

Sample formats

Tissue is a valuable sample format, used for histological diagnosis of many indications including cancers and renal disease. Tissue samples can, however, only be obtained through invasive sampling i.e. biopsies or tissue removed by surgery. In addition, for samples obtained during surgery, standard protocols regarding timing of handling can be difficult to implement. Tissue samples can be stored as either unfixed and freshly frozen or formalin-fixed and paraffin-embedded (FFPE) (Grantzdorffer, Yumlu et al. 2010). The freshly frozen samples are better suited for protein extraction, while demanding more stringent handling protocols why samples often need to be discarded after a single analysis. In contrast, FFPE samples are more conveniently handled and stored, and are robust enough to be used in many different studies. However, due to protein-crosslinking in the formalin fixation, the protein extraction protocols have traditionally been far more complex than for frozen tissue (Grantzdorffer, Yumlu et al. 2010). However, using FFPE material in proteomic studies has recently gained interest due to the vast FFPE collections available, together with the increasing demands on large sample cohorts for proteomic studies. New improved protocols have been developed, for instance Pauly et al. (manuscript in preparation) have optimized a protocol for analysis of FFPE samples using recombinant antibody microarrays.

Attracted by the minimally invasive sampling procedures, several biomarker initiatives are instead turning to searching for protein biomarkers in body fluids. Serum and plasma are the most frequently used body fluids for biomarker discovery, and it has in several studies been demonstrated that their protein levels reflect both physiological and pathological states that can be used for disease diagnosis and prognosis (Anderson and Anderson 2002, Thadikkaran, Siegenthaler et al. 2005). Serum is obtained from withdrawn blood after removal of blood cells, as well as coagulation factors, through clotting and centrifugation. Plasma, on the other hand, is prevented from clotting by addition of an anticoagulant (EDTA, sodium citrate or heparin). Studies on systematic variation in protein abundances of serum and plasma samples have indeed shown variation between different sample preparations, but also dependence on the technique used for analysis and individual protein of interest (Haab, Geierstanger et al. 2005). For instance, cytokines appeared to be most stable in EDTA-plasma, which could be explained by EDTA's protease inhibitory properties (Haab, Geierstanger et al. 2005). Most importantly, in a single biomarker study, all included blood samples need to be collected using the same sample preparation method.

Urine has been utilized in clinical testing for centuries, including assessment of albumin concentration as a measure of kidney disease (Guh 2010). Urine is readily available and non-invasive in sampling and has attracted interest in clinical proteomics as a valuable source of both renal and systemic biomarkers. More than 1500 unique proteins have been identified in healthy urine (Adachi, Kumar et al. 2006), and the urinary proteome of various physiological and pathological conditions is estimated to comprise more than 5000 proteins (Coon, Zurbig et al. 2008). The majority of urinary proteins are indeed of renal origin (70%), while 30% of the proteins are filtered through the glomerulus (Decramer, Gonzalez de Peredo et al. 2008), and can provide insights into mechanisms of indications originating outside the urinary tract system, such as cancer and autoimmune conditions (Voss, Goo et al. 2011).

The physiological composition of urine is effected by diet and exercise why patients usually need to follow more strict guidelines before sample collection. Also, the timing of sampling (e.g. first morning, second morning or 24 hour sample collection) needs to be standardized (Voss, Goo et al. 2011). Examples of other body fluids used in proteomics experiments include cerebrospinal fluid, saliva and tear fluid (Hu, Loo et al. 2006). Cerebrospinal fluid is the primary

sample for central nervous system disorders, and is collected by lumbar puncture, aspiration of fluid from the lower spine. Saliva and tear fluid are minimally invasive sample formats, which have also gained interest in proteomics.

Pre-analytical processing of samples

All of the above described sample formats need to be collected, handled and stored following strict standard operation procedures (SOP) in order to avoid pre-analytical sources of data bias. Even small differences in processing of samples could have dramatic effects on analytical reliability and study outcome (Tuck, Chan et al. 2009). Pre-analytical bias between cases and controls could result in false positive results, and processing variations within the sample groups of cases and controls could potentially mask disease related differences (false negatives). This is especially crucial for samples collected from different sites, where indeed site-to-site normalization of data often is required. Standard operating procedures for standardizing of sample collection have to take into account e.g. type of additives, sample processing temperature and time, as well as hemolysis of samples. In the subsequent sample processing, special caution should be observed for freeze-thaw cycles of samples, where cytokines have been shown particularly vulnerable (Thavasu, Longhurst et al. 1992).

Biobanking

Access to well-defined, high-quality biospecimens has been identified as a major limiting factor in the development of biomarkers (LaBaer 2012). The organizing of large sample collections in biobanks will be a prerequisite for running large-scale discovery and validation studies needed for identification and approval of biomarkers (Schrohl, Wurtz et al. 2008, Hewitt 2011, Marko-Varga, Vegvari et al. 2012). Biobanking methodology is now a fast developing research field, and several networks for organization of biobanks on national and international level are now being established. These networks will facilitate both cataloging and availability of samples, and the complex infrastructure needed for organization and storage of thousands of samples. One such network is the European collaboration BBMRI (Biobanking and Biomolecular Resources Research Infrastructure) with branches in several European countries and encompassing 30 scientific partners and 24 funding organizations (bbmri.eu). An obstacle in fruitful employment of biobanks is the lack of collaboration between public sector biobanks and pharmaceutical companies. Concerns of commercial use of patients samples as well as intellectual property issues has

been pointed out as explanations for this, as well as lack of proper quality assurance in public biobanks (Schrohl, Wurtz et al. 2008, Hewitt 2011, Marko-Varga, Vegvari et al. 2012).

The issue of ethics and data protection is central in all biobanking initiatives. All collection of biospecimens from humans needs to be accompanied by an informed consent from the donor, and the consent must include a specification of the purpose of the collection. This causes a problem for creating large biobanks, where the specific application of each sample will not be known at forehand. For this reason the Swedish Data Inspection Board has stopped the Lifegene project (www.lifegene.se), a large-scale biobanking collaboration between six Swedish universities. This project is now on hold waiting for further legal investigation.

2.4.3 Technological requirements

Protein biomarker discovery requires technologies capable of detecting molecular differences between samples of different disease statuses. In large-scale proteomic approaches, the chosen technology platform will need to be multiplexed, and target many proteins simultaneously, while using minute volumes of sample. In addition, working with complex sample formats as serum, the platform should target a wide range of proteins, ranging from low abundant cytokines to high-abundant complement factors. Also, in order to analyze large sample sets in a reasonable time frame, a high-throughput platform is required.

Initially, proteomic biomarker discovery has been pursued using protein separation techniques, as 2D gels and liquid chromatography, in combination with a mass spectrometry (MS) read-out (Hanash 2003, Hu, Loo et al. 2006). The results from discovery studies have been promising, with hundreds of candidate biomarker and biomarker signatures. Unfortunately, the translation of candidate markers into clinical utility has not been equally successful. Also, biomarker discovery studies of a given disease conducted by different research groups have often resulted in quite different panels of markers (Boschetti, Chung et al. 2012).

The technological explanations for this discrepancy can be several (Kingsmore 2006, Boschetti, Chung et al. 2012). First, the sensitivity of MS-based

techniques is significantly hampered by high-abundant proteins masking low-abundant proteins. To circumvent this, samples can be fractionated, usually through albumin removal. This action will allow targeting of proteins of lower concentration, but at the same time the introduced pre-treatment might influence reproducibility of the platform. A recent advancement, multiplexed reaction monitoring (MRM) has indeed increased the sensitivity of the MS platform, but the read-out is instead focused to a narrow pre-defined mass interval, significantly limiting its utility as a discovery tool. Also, MS-based techniques can be limited by their dependence on database searches, as a potential source of false negatives. Further, certain proteins are more difficult to analyze than others, due to their inability of displaying peptides of sufficient number and quality for MS identification.

Affinity proteomics has arisen as an alternative tool for biomarker discovery, and will be carefully reviewed in chapter 3.

3. Affinity proteomics

The use of affinity probes for protein analysis is well established in biomedical research (Brennan, O'Connor et al. 2010). The intrinsic ability of antibodies to specifically recognize proteins has given them a natural position as the most frequently used affinity probe. Antibodies are a cornerstone in widely used immune assays, like enzyme-linked immunosorbant assay (ELISA) and immunohistochemistry, and now also in the more systematic screenings of the proteome, known as proteomics.

In affinity proteomics, the proteome is explored by utilizing affinity probes targeting each studied protein, and by coupling the probes to a read-out, usually fluorescence or MS (Stoevesandt and Taussig 2007). Recent advancement in affinity proteomics has been facilitated by the development of new technologies in i) miniaturization, e.g. printing robotics and bead assays allowing for multiplexing of assays ii) automation, allowing for high-throughput handling of samples, and iii) recombinant techniques allowing for new strategies of obtaining numerous high-performing binders.

The availability of high-performing binders in a sufficient number will be crucial for large-scale surveys of proteomes, and has so far been a limiting factor for global, untargeted approaches using affinity proteomics. Annotating the entire human proteome will require at least 20 000 unique binders, just to target each non-redundant gene product, and at least 10 times more in order to cover splice products and post translational modifications (PTMs) [ensemble.org, (Clamp, Fry et al. 2007, Stoevesandt and Taussig 2012). For this purpose, several national and international initiatives have been taken on for identification and evaluation of optimal binders. For instance, the Affinomics program, an EU granted collaboration between 20 European research groups, aims at generating large-scale resources of validated affinity reagents (Stoevesandt and Taussig 2012). Binders targeting 1000 proteins will be made over the course of the program, and binders directed against protein kinases, SH2-domain containing

proteins, protein tyrosine phosphatases, and candidate cancer biomarkers are prioritized. Also, the Stockholm-based human proteome resource project aims at raising affinity-purified polyclonal antibodies against all non-redundant human proteins (Uhlen, Oksvold et al. 2010). The project has today gathered more than 17000 antibodies, targeting proteins from more than 14000 human genes (proteinatlas.com).

An alternative strategy for raising affinity reagents targeting entire proteomes has recently been developed by our group (Olsson, Wingren et al. 2011) and others (Hoeppe, Schreiber et al. 2011), in efforts combining affinity proteomics with an MS-based readout. By using antibodies directed against C- or N-terminal short motifs composed of about 4 to 6 amino acids shared among several proteins, instead of single proteins, the number of affinity reagents needed to probe the human proteome can be substantially reduced. In other words, instead of using one antibody per protein, one such motif-specific antibody could target 10 to 200 proteins.

In this chapter I will cover the most commonly used affinity reagents, demands on the chosen reagents, and different applications for affinity reagents.

3.1 Choice of affinity probes

Traditionally, full-length immunoglobulins (Igs) obtained from either immunization of animals (polyclonal antibodies, pAb) or hybridoma technology (monoclonal antibodies, mAb), have been used as affinity reagent and are still the primary choice in assays where an intact constant region is required e.g. for detection. However, the use of full-length antibodies has raised concerns regarding specificity and functionality in certain assays, why other probe formats also needs to be considered.

Advancement in recombinant technology in the last decades has allowed for the development of a wide range of alternative binders, where the protein scaffold often is based on antibodies or other natural molecules. The fragment antigen-binding (Fab) and single chain Fragment variable (scFv) are fragments derived from Igs variable region, retaining the specific binding ability of the Ig, while significantly smaller and more simple in structure. Fabs consist of one constant

and one variable domain from each of the heavy and light chain of the antibody, while the scFv consist of only the variable domains of the heavy (VH) and light (VL) chain of the Ig, linked by a recombinant polypeptide linker allowing for expression of both domains as one single chain. Scaffolds based on entities other than Ig include i) alpha-helical receptor domains derived from staphylococcal protein A, where diversity was introduced through randomizing of 13 solvent-accessible surface residues (Affibodies, 6 kDa) (Nord, Gunneriusson et al. 1997), ii) repeat proteins derived from ankyrin adaptor proteins, usually composed of 4-5 repeat motifs (DARPins, 14-18 kDa) (Binz, Stumpp et al. 2003), and iii) single- or double-stranded oligonucleotides, which fold upon associating with their ligands (aptamers, ~10-20 kDa) (Ellington and Szostak 1990, Tuerk and Gold 1990). Despite promising results and the advancement among alternative scaffolds, binders based on Igs are still most commonly used in affinity based assays.

All of the above described novel recombinant binders are of substantially smaller size than full-length Igs (~6-30 kDa versus 150 kDa for IgG), and their function is independent of complex structures, such as the glycosylation of the Ig constant region. These factors together allow for in vitro production of recombinant fragments, as well as display of fragments in various display systems as bacteriophages, ribosome- and yeast-display. This, in turn, enables the design and construction of combinatorial libraries constituting of vast members of binders (Barbas, Bain et al. 1992, Hoogenboom and Winter 1992), from which desired specificities can be selected. These libraries provide a renewable probe source for virtually any binder, even including toxins and self-antigens (Griffiths, Malmqvist et al. 1993, Kasman, Lukowiak et al. 1998).

The primary requirement of all binding probes is the specific identification and high affinity binding of the intended target protein. The term *specificity*, in this context, describes the ability of the probe to single out target proteins in a complex sample, while a probe's *affinity* describes the strength of binding to its target. However, for practical reasons, the probes also need to be easily accessible and renewable, and meet different demands of the assay, including detection system and physical properties (e.g. stability).

3.1.1 Probe specificity

All immune assays are dependent on access to binders with high specificity and affinity. Unfortunately, many commercially available antibodies have not lived up to this requirement, and also suffer from insufficient characterization and/or documentation (Stoevesandt and Taussig 2007, Brennan, O'Connor et al. 2010). Also, probes that are specific in one assay might cross-react or not recognize its target in another. For instance, antibodies specifically targeting the epitope of a native proteins (e.g. in ELISA) could fail to recognize its denatured counterpart (e.g. in western blots). In addition, analysis of more complex samples, such as serum/plasma also place higher demands on probe specificity, and targeting low-abundant analytes as cytokines calls for binders of high affinity. Consequently, there is a need for well-characterized high-performing binders, developed with intended assay in mind (Stoevesandt and Taussig 2007, Brennan, O'Connor et al. 2010, Stoevesandt and Taussig 2012).

To ensure sufficient specificity, the affinity probes can be evaluated using spiking and blocking experiments as well as capture assays in combination with MS-based detection. High-throughput validation of antibodies can preferably be performed using microarray-based screening, using protein and peptide arrays (Lueking, Horn et al. 1999, Poetz, Ostendorp et al. 2005).

For affinity reagents obtained through library panning, the selection pressure and screening strategies will influence the properties of obtained binders, and stringent protocols will result in binders of high specificity and affinity (Hoogenboom and Winter 1992). Another advantage of working with recombinant reagents is that the obtained binders can be further engineered for increased specificity and affinity, using site-directed and/or evolutionary approaches (von Schantz, Gullfot et al. 2009). Still, before introduction into its intended application the selected binders always need to be carefully characterized with regard to specificity and functionality.

3.1.2 Physical demands on probes

Each technology poses its specific physical demands on the reagents used. Probes used in *in vivo* application require sufficient half-lives to reach its target, and reagents used in *in vitro* assays need to be compatible with buffers used. Affinity probes used on planar microarrays are subjected to particularly harsh

treatment, as they are dispensed onto a solid support and then allowed to dry out. Many scaffolds/probe formats cannot sustain such treatment but would denature and loose its binding properties. In fact, early microarray studies showed that more than 90% of evaluated probes (mainly mAbs and pAbs) did not retain its binding properties when dispensed on-chip (Haab, Dunham et al. 2001, MacBeath 2002, Mitchell 2002), which would demand huge laborious efforts and resources in order to identify binders suitable for on-chip applications. One solution to this problem is to work with binders that all share a common framework (FW), known to be stable on-chip (Borrebaeck and Ohlin 2002). Another advantage of using a common master FW is the compatibility with assay buffers: In multiplex affinity assays, all binding events will take place in a single reaction chamber. This means that all antibody-antigen pairs will be subjected to the same assay conditions, e.g. choice of buffer, temperature, incubation time etc.. Using affinity reagents with a common FW increases the likelihood of finding assay conditions that suits all included reagents. Similar to the protein engineering for improved specificity and affinity, recombinant affinity probes can be engineered on molecular level for improvement of physical properties e.g. increased stability (Worn and Pluckthun 2001), which has been explored in **paper II** and further discussed in chapter 4.

3.2 Assay formats

Traditional techniques utilizing the unique properties of affinity reagents include ELISA, western blots, immunohistochemistry, and immunopercipitation. ELISA is still regarded as the gold standard, and has had recent improvement in sensitivity due to novel detection systems, often utilizing DNA based amplification, including PCR and rolling circle amplification (RCA). However, simultaneous analysis of multiple proteins in the ELISA format would be laborious and consume sample volumes far beyond what is usually available. Emerging assays for multiplexed protein analysis using affinity reagents include printed planar arrays, suspension bead assays and affinity assays coupled to MS (Anderson, Anderson et al. 2004, Kingsmore 2006, Schwenk and Nilsson 2011). I will here focus on planar arrays.

Advantages of scaling down the assay from macro format (e.g. ELISA) to micro format (arrays) include i) minute volumes of sample and reagent required (μL scale) and consequently lower cost of assays, ii) reduced reaction times due to short diffusion distances and, iii) improved signal-to-noise ratios as a result of miniaturized immunoassays following the ambient analyte theory, as described by Ekins. (Ekins 1998). Promising proof-of-principle studies in late 1990' by Snyder's (Zhu, Klemic et al. 2000) and Schreiber's groups (MacBeath and Schreiber 2000), printing arrays consisting of minute volumes of proteins, has paved the way for a variety of applications of the array format. Planar arrays are printed onto a solid support, traditionally a microscope slide, where the printed material is in pL-scale and can be either antibodies (antibody arrays), protein/peptides (antigen arrays) or the sample to be analyzed (reverse phase microarrays).

Antibody arrays are generally either dual-antibody sandwich arrays or single-capture, direct labeled arrays (Kingsmore 2006, Liu, Zhang et al. 2006, Schroder, Jacob et al. 2010). In sandwich assays, one capture antibody is printed and the bound proteins are detected using a second antibody targeting a different epitope of the protein analyte. Benefits of this approach include the inherent high specificity of using two antibodies and no need to label the sample. On the other hand, scaling up assays might prove difficult due to cross-reactivity that has been observed in arrays with more than 30 pairs, as well as the logistics of obtaining functional antibody sandwich pairs for all proteins of interest (Miller, Zhou et al. 2003). However, the sandwich array format is well-suited for low-plex assays, e.g. targeted cytokine arrays. The single-capture approach, where antibodies are printed and the proteins in the sample are labeled with e.g. a fluorescent tag, is particularly suited for large-scale studies and has been explored by our group and will be further discussed in chapter 4.

In antigen arrays, a wide range of proteins or peptides are printed, the array is probed with a sample and bound protein/antibodies are detected using a labeled affinity reagent. This approach has been utilized for detection of auto-antibody response to tumors or in auto-immune conditions, for instance by printing tumor associated antigens e.g. aberrant glycosylation patterns in different tumor associated proteins (Pedersen, Blixt et al. 2011). Other groups have studied IgE-response by large allergen arrays (Deinhofer, Sevcik et al. 2004). This format has also been explored by a number of commercial vendors

including ProtoArray® (Invitrogen.com) today printing >9000 protein per array and PEPperCHIP® (pepperprint.com) printing up to 8600 peptides per array.

Reverse-phase protein microarrays (RPPM) have evolved as a tool for pathway analysis (Pawlak, Schick et al. 2002, Spurrier, Ramalingam et al. 2008). Discrete volumes of tissue lysates or body-fluids are printed, and the arrays are then probed with detection antibodies, often targeting phosphorylation or other PTMs. Using RPPM denatured protein lysates can be analyzed, while using up to 10000 times less sample per analysis than western blots do. Another advantage of using the reverse approach is that the affinity reagents are kept in solution, and not subjected to harsh printing conditions. Comparing the throughput of antibody arrays versus reverse arrays, the antibody array format is more convenient for multiplexing (simultaneous analysis of many proteins), while the reverse format is more efficient for high sample throughput (Stoevesandt and Taussig 2012).

With the long-term goal of targeting the entire human proteome, the chosen analysis platform needs to be capable of substantial up-scaling towards untargeted, global proteome analysis, while still remaining sensitive, and capable of high-throughput analysis. Encompassing all these features, the single-capture, direct labeling antibody array platform has been the assay of choice in our group and will be further discussed in chapter 4.

4. Design and optimization of antibody microarrays

In the last decade, our group has developed a platform for affinity proteomics, based on recombinant scFvs (Ingvarsson, Larsson et al. 2007, Wingren, Ingvarsson et al. 2007). With the long-term goal of targeting the entire proteome, the assay format we have chosen is single-capture, direct labeling antibody microarrays. Briefly, scFvs are printed onto a solid support and allowed to dry out before the surface around the spots is blocked in order to prevent unspecific background binding. The clinical sample is labeled through biotinylation and then added to the array where labeled proteins are allowed to bind to their corresponding scFvs. After a second incubation with fluorescently labeled streptavidin, bound proteins are detected using a confocal scanner. Finally, by comparing protein binding patterns between different samples, differentially expressed protein profiles can be detected, and in the long run potentially be used as biomarkers signatures (Figure 2).

In this chapter, I will describe some of the key features we have addressed in the optimization process, including probe format (**paper I**), sample format (**paper II**), as well as more specific assay parameters, such as choice of substrate, printing, and detection.

4.1 Antibody fragments as affinity probes

The feasibility of using antibody fragments as affinity probes on microarrays has been demonstrated in several studies by our group (Borrebaeck and Wingren 2011) and others, (Pavlickova, Schneider et al. 2004, Seurynck-Servoss, Baird et al. 2008) where scFvs and Fabs have shown excellent on-chip performance, including functionality, sensitivity and specificity (Seurynck-Servoss, Baird et al. 2008, Borrebaeck and Wingren 2011). Large combinatorial

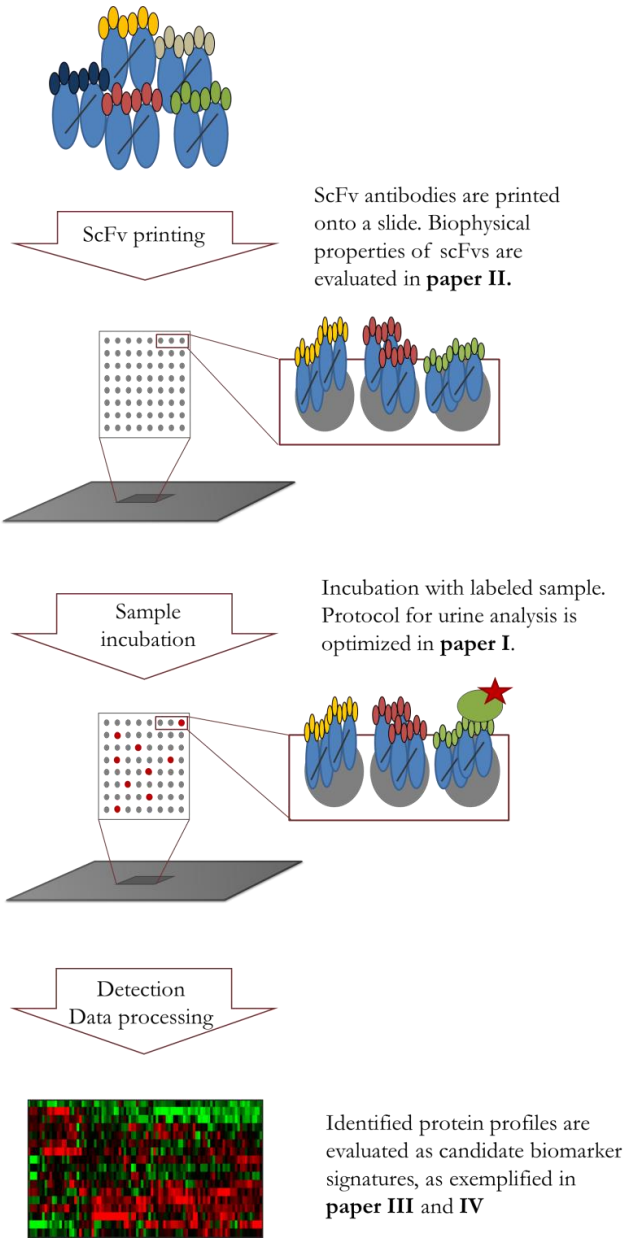

Figure 2. Schematic overview of a recombinant antibody microarray platform

libraries can provide binders of virtually any specificity (Barbas, Bain et al. 1992, Hoogenboom and Winter 1992), and once the binders have been selected, they are renewable and easily accessible (Borrebaeck and Wingren 2011).

Recombinant antibody fragments can be selected from libraries constructed around a single FW (Barbas, Bain et al. 1992, Soderlind, Strandberg et al. 2000, Lee, Liang et al. 2004) or multiple different FWs (Hanes, Schaffitzel et al. 2000, Knappik, Ge et al. 2000). Using libraries of multiple FWs allows for increased variability and potentially improved specificity and affinity among selected clones, since certain FW residues potentially participate in antigen binding (Carter, Presta et al. 1992, Lee, Liang et al. 2004). On the other hand, libraries constructed around a single FW instead offer more homogenous biophysical properties among the selected clones, and the possibility of engineering the common FW for the intended application (Lee, Liang et al. 2004, Borrebaeck and Wingren 2011). The antibody fragments predominantly used in our platform are scFvs selected from a phage-display library (n-CoDeR) constructed around a single, constant FW (VH3-23/VL1-47), where this master FW was chosen based on its excellent expression as soluble protein in bacteria and display in phage-based systems (Soderlind, Strandberg et al. 2000). The library is highly diverse, and the diversity was introduced by shuffling naturally occurring human complementary determining regions (CDRs) and grafting them to the constant FW, resulting in a library composed of $2x10^9$ members (Jirholt, Ohlin et al. 1998, Soderlind, Strandberg et al. 2000).

4.1.1 Stability of single-chain Fragment variables (scFvs)

The on-chip functionality of arrayed probes is essential for well-performing antibody arrays. The physical properties of antibody fragments have been evaluated in several studies, primarily addressing structural stability in solution (Kipriyanov, Moldenhauer et al. 1997, Worn and Pluckthun 2001, Ewert, Honegger et al. 2004). The structural stability has proven critical for improved shelf-life (in solution) and in-vivo applications (Willuda, Honegger et al. 1999), and is usually characterized in terms of half-life (time required for a 50% loss in protein activity) and melting temperature (the temperature at which a certain protein denatures, Tm). The functional on-chip stability of affinity probes do not always correlate with stability in solution and needs to be assessed separately (Steinhauer, Wingren et al. 2002). ScFvs selected from the n-CoDeR library have shown superior on-chip performance, as compared to competing FW

(Steinhauer, Wingren et al. 2002). For instance, arrayed n-CoDeR scFvs have been found to display an on-chip half-life of 4-6 months as compared to 42, 39 and 7 days for competing FWs. Still, additional improvements in stability could potentially reduce the observed scFv activity fluctuation over-time, as well as clone dependent differences, most likely conferred by differences in the CDRs. As for example, individual Vλ domains have shown low stability, but often form stable scFvs, accomplished through a strong interaction with VH, which in turn is dependent on the sequence of CDR-loop 3 (CDR-L3) (Ewert, Huber et al. 2003).

Design of even more stable and homogenous scFvs could also enable long-time storage on-chip, which would facilitate assay logistics. The on-chip stability can be targeted by i) addressing the surface chemistries and immobilizing of scFvs via e.g. affinity coupling (Seurynck-Servoss, Baird et al. 2008), ii) using surfaces as well as coating and blocking buffers with stabilizing properties (Kopf, Shnitzer et al. 2005, Kopf and Zharhary 2007), or iii) targeting the affinity molecules themselves, using protein-engineering, and screening for improved stability on-chip. In **paper II,** I have used the third approach, and I will focus the remaining discussion in this section on stability engineering of scFvs.

The stability of scFvs is a function of the intrinsic stability of each domain (VH and VL), and the stability conferred by the interactions (interface) between the two domains (Jager and Pluckthun 1999, Worn and Pluckthun 1999). Each individual domain has the characteristic immunoglobulin fold (Bork, Holm et al. 1994), with two tightly packed antiparallel β-sheets and 3 protruding loops forming the antigen-binding site together with 3 loops from the other domain (3 loops from VH and 3 loops from VL). The sheets are held together by hydrophobic side chains, closely packed in the core of each domain, and by a conserved disulfide bridge. Formation of rigid loops and hydrogen bonds also help in stabilizing the domain structure. The stability of the interface is influenced by the size of the surface area and favorable interactions between the two domains, again including hydrophobic side chains from each domain (Worn and Pluckthun 1999). The choice of FW domains and their compatibility is therefore crucial, and this has been investigated in detail by Pluckthun and co-workers, where different combinations of domains were evaluated in terms of stability (Worn and Pluckthun 1999) in solution. In their study, the domains were first evaluated individually, and then in different combinations. The results showed that an individual stable domain could rescue a less stable counterpart,

and also that two less stable domains could be rescued by a favorable interface. Notably, VH3-23/VL1-47 was found to be one of the most stable combinations of FW domains, and has also been the FW used in the on-chip applications described in **papers I, III and IV**.

Approaches for stability engineering of scFvs include both evolutionary and rational design experiments. Evolutionary design involves introducing random mutations to the FW and, by using a suitable selection pressure, more stable mutants can be selected using phage display or other panning systems. Selection pressures commonly used include elevated temperature and chemical denaturation, where temperature stress has yielded more stable mutants (Jung, Honegger et al. 1999). In a rational design approach, key residues are identified based on structural analysis or alignment studies, and then targeted using site-directed mutagenesis. Several key residues, crucial for high stability, have been identified through alignment of amino-acid sequences between scFv clones of different stability (Saul and Poljak 1993, Krauss, Arndt et al. 2004, Rodriguez-Rodriguez, Ledezma-Candanoza et al. 2012). Position 6 in the heavy domain (H6) of scFv has attracted much attention, indicating strong influence on the overall stability of the scFv (Kipriyanov, Moldenhauer et al. 1997, Honegger and Pluckthun 2001, Jung, Spinelli et al. 2001) (Figure 3). The H6 position in human scFvs can only accept Glutamatic acid (E) or Glutamine (Q) (Honegger and Pluckthun 2001). Q in H6 position confers a more stable scFv, and is the only tolerable amino acid for scFvs lacking the intrinsic di-sulfide bridge e.g. due to expression under reducing conditions. ScFvs carrying a di-sulfide bridge can tolerate E in H6, while resulting in a less stable scFv than with a Q (Langedijk, Honegger et al. 1998). ScFvs selected from n-CoDeR carry an E in H6, possibly leaving room for stability improvement.

Our group has adopted both an evolutionary and a rational design approach for stabilization of scFvs selected from n-CoDeR (Vallkil et al., unpublished observations and **paper II**). First, a randomized phage display library was constructed around a single n-CoDeR clone (α-FITC), through random mutations directed to the FW of the scFv (Vallkil et al.). The library was panned with heat (45-55°C) as selection pressure, and one dominant mutant clone was identified as substantially more stable, on phage-level, than wild-type (WT). Sequencing analysis revealed a single mutation in the light chain FW between CDR-L2 and CDR-L3, where a serine in a loop position had been replaced by a more rigid proline (S96P) (Figure 3). The importance of prolines in loop

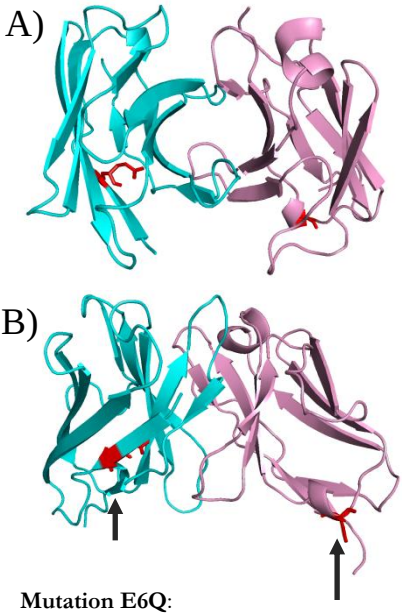

Figure 3. Structural homology model of a scFv clone (α-βGal) with VH in cyan, VL in magenta and mutation sites marked in red. A) Top-view B) Side-view

positions for stabilization of protein structure has previously been shown by others (Watanabe, Masuda et al. 1994, Tian, Wang et al. 2010). The mutant carrying the S96P mutation and WT α-FITC were then also produced as soluble proteins, and the stabilizing effect could be verified in solution using circular dichroism, as described in **paper II**.

Next, in order to investigate if the stabilizing effect of the S96P mutation was a clone-dependent phenomenon or generally applicable to other n-CoDeR clones, the mutation was introduced into three other clones directed against antigens of varying size (α-CT, α-βGal and α-C1q) (**paper II**). The type and size of the antigen determines the composition and shape of the antigen binding site (Webster, Henry et al. 1994) made up of the CDR-regions, which in turn might influence the biophysical properties of the scFv, such as their overall stability. Also, in a parallel rational design experiment, the above described H6 position was targeted and mutation E6Q was introduced into two n-CoDeR clones (α-CT and α-C1q) through site-directed mutagenesis. In addition, double mutants, carrying both S96P and E6Q were constructed. Circular dichroism measurements displayed a 1-3°C increase in T_m for the single mutations and additive effects for double mutants. The results showed that the two mutations, S96P and E6Q, indeed conferred increased stability in solution for all scFv clones included in the study, indicating that the stabilizing effects were not clone-dependent, but instead general for scFvs selected from the n-CoDeR library.

The on-chip performance of WT and mutated clones was assessed by printing un-stressed clones for a standard array-based analysis. The resulted showed that all mutants were active on-chip and that the activity was equal (or improved) to

their corresponding WT. Further, in order to assess the functional on-chip stability, arrayed mutants and WT were screened using elevated temperature (70 °C), and incubation in an denaturing agent (guanidine hydrochloride, GdmCl) (**paper II**). The use of elevated temperature as screening pressure has provided similar results as long-time storage in room temperature, and enables the conducting of stability studies in a reasonable time frame. Briefly, mutants and WT of all four clones were printed onto slides and then incubated dry in a 70°C incubator for 6-38 days, or in a serial dilution of GdmCl in room temperature. Incubated slides were analyzed according to standard protocol with pure antigens, and obtained array signals of mutants and WT were compared. The results displayed similar, slightly improved, or even slightly impaired on-chip stability for mutants as compared to WT, indicating that the E6Q and S96P mutants were functional on-chip, and that their stabilizing/destabilizing effects rather appeared to be clone dependent.

In more detail, the clone with lowest initial stability (α-CT: T_m=59°C), was found to be the one that benefited the most from the stabilizing mutations (on-chip stability). This indicated that the mutations could potentially reduce clone-dependent differences in stability, by making the clones more similar. Also, the two clones (α-βGal and α-C1q) which did not benefit or appeared to be slightly impaired by mutations in the on-chip temperature stress experiments (70°C), were the ones that had the highest initial T_m (75°C (α-C1q) and 69°C (α-βGal)). Therefore, the on-chip temperature stress probably did not affect the α-C1q and α-βGal molecules as much as it affected scFvs of lower initial T_m. In order to identify binders with pronounced increased on-chip activity, the initial selection of mutants should preferably be performed on-chip. An appealing approach is panning of a library of proteins produced by large-scale compatible approaches, such as on-chip protein production through self-assembly (e.g. NAPPA or PISA (He and Taussig 2001, Ramachandran, Hainsworth et al. 2004)), see section 4.3.2.

In silico homology modeling of mutants and WT revealed a number of structural alterations, which might explain stabilizing behavior of the mutations. The E6Q mutation conferred a more densely packed hydrophobic core by introducing a longer hydrophilic side-chain pointing towards the center of the domain, participating in a hydrogen binding network. The effect of the S96P mutation was most pronounced in the α-FITC clone, where the mutant structure displayed a larger interface area as well as more hydrogen bonds and van der Waals interactions.

4.2 Sample formats

Virtually any solubilized sample format, constituting of proteins with exposed and accessible epitopes, can be analyzed on antibody microarrays, but each format will need an individually optimized protocol. We have, in a step-by-step procedure, optimized protocols for analyzing most of the available clinical sample formats, including serum, plasma, tissue extract (freshly frozen as well as FFPE), cell lysates and intact cells (Ellmark, Ingvarsson et al. 2006, Ingvarsson, Larsson et al. 2007, Wingren, Ingvarsson et al. 2007, Dexlin, Ingvarsson et al. 2008, Dexlin-Mellby, Sandstrom et al. 2011). Recently, urine has been added to the list, and optimization of the urine protocol is described in **paper I**. In general, we have aimed at analyzing crude samples, with minimal pre-treatment, in order to minimize any factors influencing the reproducibility and sensitivity of the platform. Also, the consumed sample volume has been kept to a minimum through the miniaturized set-up and stringently optimized protocols.

4.2.1 Optimization of protocols for serum, plasma, tissue and cell culture profiling

Serum and plasma are well-established as sample formats in biomarker discovery (Anderson and Anderson 2002), and also the formats our antibody array platform was originally designed for (Ingvarsson, Larsson et al. 2007). Blood comprises a very complex mixture of proteins, with protein concentrations ranging over 10 orders of magnitude from low-abundant cytokines (pg/ml) to high-abundant complement proteins and albumin (30-50 mg/ml). This complexity and wide span of protein concentration poses high demands on any techniques used with respect to specificity, dynamic range, resolution and reproducibility of the assay. Also, complex samples can be associated with high unspecific background binding, reducing the signal-to-noise ratio (S/N) and thereby the sensitivity of the platform. In early applications of our platform, the analysis was either focused on high-abundant proteins (Ingvarsson, Larsson et al. 2007) or enabled by pre-fractionation based on size (Ingvarsson, Lindstedt et al. 2006). Since then, careful optimization of the protocol, with regard to choice of surface, sample handling, blocking and washing solutions, now allow for simultaneous detection of high as well as low abundant protein in a single analysis of a crude, non-fractionated sample, while still providing low non-specific background binding and high S/N (Wingren,

Ingvarsson et al. 2007, Carlsson, Wingren et al. 2008, Borrebaeck and Wingren 2009).

The optimized, sensitive assay now allow us to dilute serum and plasma samples 90-450 fold in PBS (45 fold dilution before labeling) and sample buffer (2-10 fold dilution after labeling), and thereby only using minute sample volumes, less than 1 μL for each analysis. Serum and plasma are analyzed with identical protocols, with one exception: in order to prevent coagulation of plasma samples, the anti-coagulant used in the original sampling tubes (e.g. EDTA) is added to all buffers throughout the protocol. These optimized protocols for serum and plasma analysis have been applied in **paper III and IV**, where they also are described in detail.

Protein extracts from tissue specimen have successfully been analyzed on antibody microarrays (Bartling, Hofmann et al. 2005, Ellmark, Ingvarsson et al. 2006). Analysis of water-soluble cytosolic proteins can be achieved through careful extraction protocols, partially denaturing the cell membrane while sparing proteins and the nuclear membrane, and has been pursued by our group (Ellmark, Ingvarsson et al. 2006, Dexlin-Mellby, Sandstrom et al. 2011) and others (Hudelist, Pacher-Zavisin et al. 2004, Bartling, Hofmann et al. 2005). A greater challenge lies in also targeting the hydrophobic and usually low-abundant membrane proteins, playing a vital role as cell-surface receptors and commonly targeted by therapeutics. Using a two-step fractionation protocol, Dexlin-Mellby et al. (Dexlin-Mellby, Sandstrom et al. 2011) managed to extract both soluble and membrane proteins from a single sample, and then to analyze them on a single microarray. Membrane proteins can also be targeted while still buried in the cell-membrane of the intact cell. This approach has been applied to both native blood cells purified from buffy coats and to various suspension cell cultures (Belov, Mulligan et al. 2006, Dexlin, Ingvarsson et al. 2008).

4.2.2 Optimization of protocol for urine profiling

Urine displays a number of inherent differences as compared to serum and plasma that could influence the microarray analysis. The total protein concentration of urine samples is in general 10-1000 times lower than serum and plasma (Decramer, Gonzalez de Peredo et al. 2008). This fact, together with the more pronounced inter-sample variation in pH and salt content (osmolality) of urine, could influence the labeling reaction, where all proteins in

the samples are tagged with a fluorescent dye, or other tags. The total protein concentration of a sample and its relation to the amount of labeling-tag added, require careful optimization in order to avoid epitope masking, which would inhibit antigen-antibody interaction and result in false-negative signals (Wingren, Ingvarsson et al. 2007). Further, the salt content and pH will also influence the efficiency of the labeling reaction, and fluctuation in pH could potentially introduce bias in the degree of labeling, why an initial standardization of pH is required.

In addition, the concentration of individual protein analytes is generally lower in urine than in plasma, placing higher demands on sensitivity of the analysis platform. This has forced other groups to substantial pretreatment of the urine samples before proteomic analysis (Thongboonkerd 2007, Voss, Goo et al. 2011). For urinary profiling using MS-based techniques, sample preparation methods for concentrating or isolating of proteins include precipitation with organic solvents, centrifugal filtration and lyophilization (Thongboonkerd 2007). These efforts will indeed increase the analyte concentration, but have also been associated with biased loss of certain proteins and reduced reproducibility of the analysis platform. In addition, removal of high-abundant proteins, predominantly albumin, has been applied in order to target low-abundant proteins (Thongboonkerd and Malasit 2005). However, concerns have been raised regarding loss of important biological information, including oxidation status of plasma albumin, as well as simultaneous removal of other proteins complex bound with albumin. Urinary proteomics have so far only been addressed in a few studies using antibody microarrays (Liu, Zhang et al. 2006, Schroder, Jacob et al. 2010). Low-plex sandwich assays have been applied for cytokine profiling (Liu, Zhang et al. 2006), only using a centrifugation for removal of cell and debris as sample preparation. Exemplifying a more extensive array set-up, Hoheisel and co-worker printed 810 polyclonal antibodies for analysis of proteins in directly labeled urine samples. This effort, however, included substantial sample pretreatment including desalting and lyophilization of urine samples before analysis (Schroder, Jacob et al. 2010). For large-scale urinary proteomics, the analysis platform will have to be i) capable of substantial up-scaling ii) based on renewable probe format and iii) sensitive enough to analyze crude urine without harsh pre-treatment. Meeting all these criteria, our single-capture antibody array platform, based on recombinant scFvs, can preferably be used for urinary protein profiling.

In **paper I**, I describe the optimization process of adjusting our antibody microarray platform for urine analysis. First, the pH of all included samples was standardized, in order to avoid pH bias in the labeling reaction. To this end, all urine samples were dialyzed against PBS (pH = 7.4) a gentle standardization method, keeping the proteins under physiological conditions. Of note, samples were dialyzed against an excessive amount of PBS, and in a manner keeping the sample volume constant. Next, due to the significantly lower total protein concentration than serum/plasma, the dialyzed samples were labeled un-diluted (45-fold dilution in serum/plasma protocol), and the degree of labeling was optimized in order to achieve strong signals while still avoiding epitope masking. For urine samples with a total protein concentration of 2 mg/ml, a NHS-biotin (labeling tag) concentration of 0.6 mM provided highest S/N. We initially adopted the serum protocol and labeled all samples with one fixed amount of biotin, however, during the course of the evaluation process we have concluded that an adjusted amount of biotin should be used. In order to achieve representative labeling of different urine samples of varying protein concentration, the amount of NHS-biotin should preferably be adjusted to the total protein concentration of each sample. Further, in accordance with the serum/plasma protocol, the choice of assay buffer showed crucial for obtaining minimal background while high S/N. A combination of milk and tween in PBS (0.5% (w/v) milk and 0.5% (v/v) Tween in PBS) was the preferred choice, similar to the serum protocol (1% (w/v) milk and 1% (v/v) Tween in PBS). In contrast to the serum protocol, the amount of added buffers in the optimized urine protocol was kept to a minimum in order to maintain high specific signals, why labeled samples only were diluted 1.3 times in sample buffer, instead of 10 times in serum protocol.

Urinary proteins are a mixture of intact proteins and fragments thereof, as a result of degradation during renal passage (Osicka, Panagiotopoulos et al. 1997). The amount of fragments has previously been underestimated, due to inability of various methods to detect degraded epitopes (Greive, Balazs et al. 2001). In addition, the molecular weight cut-off of the kidney increases substantially in renal disease, secreting larger proteins than the healthy kidney (Decramer, Gonzalez de Peredo et al. 2008). These two factors result in detection of urinary proteins far larger than the cut-off of a healthy kidney (30-40 kDa), both in our assay and others. For instance, we have detected high levels of complement protein C1q in urine from foremost patients with renal damage. This large four-domain protein (460 kDa) was not expected to be found in urine. We have not

elucidated to what extent this is due to degradation or increased filtration, but we can at least propose that our α-C1q scFv target an epitope that is not degraded in the renal passage.

Taken together, we have now added urine to the list of samples that can be analyzed on our microarray platform, and we can now target all of the most commonly used clinical sample formats. This will not only allow us to perform profiling of urine for identification of renal and systemic biomarkers, but also to perform studies where several sample formats are targeted, and where we can compare the impact and utility of different sample formats. This is exemplified in **paper III** where I have analyzed serum and urine samples from SLE patients for identification of novel biomarkers. The result in this study showed complementary results from serum and urine, thus motivating the analysis of both sample formats that combined will enable us to gain a deeper understanding of the disease status.

4.3 Assay

The design of a high-performing antibody microarray assay involves several key features ranging from choice of surface (substrate), method for dispensing reagents to the array, and detection system, to optimized protocols for blocking and washing of the array. Further, the array lay-out is dependent on availability of probes, printing logistics, and foremost the intended application of the array: condensed low-plex arrays can be used in e.g. targeted pathway analysis point-of-care application, while large high-density arrays are used in biomarker discovery studies.

Planar arrays are predominantly printed on microscope slides, while several efforts using 96 well plates or other well/vial formats have been pursued (Urbanowska, Mangialaio et al. 2006). Choice of format (well-based or slides) is usually based on demands on i) array size, ranging from low-plex arrays easily fit in well-based format to multiplex arrays with thousands of features, demanding a substantial portion of a microscope slide, ii) sample/reagent consumption, where valuable sample of expensive reagent motivates down-scaling of array size, or iii) practical limitations such as compatibility with scanners and other detection systems.

In the following sections I will discuss three key features in design of a protein microarray assay: substrate, printing, and detection.

4.3.1 Substrate

An ideal surface for antibody array analysis provides low background, high S/N, and homogeneous spot morphology (Kopf and Zharhary 2007, Sanchez-Carbayo 2011, Sauer 2011). This requires a surface with high binding capacity and bio-compatibility with arrayed probes, while low auto-fluorescence and non-specific background binding.

A plethora of slides of different material (glass, polymer and nitrocellulose), surface structure (2D or 3D for entrapment) and immobilization strategies (adsorption, covalent or affinity binding) are available and have been evaluated (Angenendt, Glokler et al. 2003, Pavlickova, Schneider et al. 2004, Seurynck-Servoss, Baird et al. 2008), (Sandström et al. unpublished observations). For analysis of complex biological samples (e.g. serum, urine, tissue) two slides of quite different properties have, in investigations by our group, shown to be superior with regard to S/N and background binding (Wingren, Ingvarsson et al. 2007), (Sandström et al. unpublished observations): First, Nexterion H slides with a 3D hydrogel surface and covalent coupling of printed protein and second, black polymer Maxisorp slides with a planar black polymer (coating proprietary) where proteins are adsorbed to the surface. Out of these, the Maxisorp slides have demonstrated a wider dynamic range, which is vital for analysis of plasma and other biological samples, and has therefore been the choice of surface in the majority of clinical studies in our group (Carlsson, Persson et al. 2010, Carlsson, Wuttge et al. 2011, Sandstrom, Andersson et al. 2012). I have applied the Maxisorp slide in the serum and plasma protocols used in **papers III and IV**, and in the optimization of urine analysis (**paper I**). The Maxisorp slide again demonstrated low background, high S/N and good spot morphology also for urinary proteins.

4.3.2 Printing

A key enabling factor for the production of high-density microarrays is high-precision printing robotics (Austin and Holway 2011, McWilliam, Chong Kwan et al. 2011). Depositing minute amounts of affinity reagents onto a solid support, with extreme precision in drop volume and spatial position, poses high

demands on the instrumentation used. High spot-to-spot reproducibility is required for quantification and comparisons between samples, and even minimal divergence in spatial position would prevent identification of individual spots. Printing robotics for microarrays employ contact or non-contact techniques (Austin and Holway 2011, McWilliam, Chong Kwan et al. 2011). Contact printers transfer minute volumes of affinity reagent to the substrate using solid steel pins, while non-contact printers dispense droplets of affinity reagent from glass capillaries. Contact printers are found to be faster and usually of low maintenance, while non-contact printers have fewer problems with carryover, and damage to vulnerable surfaces and a more controlled dispensing procedure, why non-contact printing has been the method of choice in our set-up (**papers I-IV**).

Alternative printing approaches include self-addressing (Wacker, Schroder et al. 2004) and self-assembly arrays (He and Taussig 2001, Ramachandran, Hainsworth et al. 2004, He, Stoevesandt et al. 2008). Both these approaches involve printing DNA instead of proteins, simplifying both printing and storage of arrays. In self-addressing arrays, each probe is tagged with an oligonucleotide and a complementary strand is printed on the array. The affinity probes can be added in bulk to the array, reducing logistics of purification, and will find and bind to its designated spot through DNA hybridization. In self-assembly arrays, DNA strands coding for the affinity probe is printed and the protein is produced on-chip using a cell-free expression system.

4.3.3 Detection

In single-capture antibody arrays, the bound proteins are identified by their position in the array, and the amount of bound protein is assessed by quantifying the signal from the applied labeling tag.

A majority of antibody arrays use fluorescence scanning as a read-out system (Angenendt 2005). Analyte proteins can be either directly labeled i.e. tagged with a fluorophore, or indirectly labeled e.g. via the biotin-avidin system. Fluorophores are small organic molecules, minimally affecting the antibody-protein interaction and common coupling chemistries include NHS (sulfo-groups targeting primary amines, i.e. side chain of lysine and N-terminal of protein) and ULS (platinum targeting sulfur and nitrogen containing side chains of methionine, cysteine and histidine). Commonly used fluorophores are Cy-

dyes and AlexaFlour dyes, where the Cy-dyes are the brighter of the two, but the AlexaFlours have shown to more resistant to quenching after multiple scans (Ballard, Peeva et al. 2007).

By using biotinylation of protein analytes and a second incubation with fluorescently labelled streptavidin, we have detected higher S/N than with direct labeling of proteins (Wingren, Ingvarsson et al. 2007). The unspecific background signals were substantially decreased in the biotin set-up, and at the same time, the specific spot signals where increased. This amplification of signal is probably explained by the fact that each biotin-group on the labeled protein allow for binding of a streptavidin-molecule that can carry multiple fluorophores. Biotinylation of serum and plasma proteins has been applied in **papers III** and **IV**, and in the optimized urine protocol (**paper I**) biotinylation was again demonstrated as a convenient labeling system resulting in high S/N arrays.

Further amplification of protein microarrays signals can be achieved using rolling circle amplification (RCA) (Lizardi, Huang et al. 1998, Schweitzer, Roberts et al. 2002) or tyramide signal amplification (TSA) (Chao, DeBiasio et al. 1996, Meany, Hackler et al. 2011). In RCA the signal is amplified through elongation of a primer conjugated to the detection reagent (e.g. antibodies or streptavidin), using a circular complementary DNA molecule. TSA is a horseradish peroxidase (HRP)-mediated signal amplification, in which tyramide molecules form radicals and bind tyrosines in the absolute proximity of the HRP conjugate. In an effort to design ultra-high sensitive arrays, TSA has been applied to serum and urine protocols using our antibody microarrays. The results did indeed show amplified specific spot signals, but also increased unspecific background signals. The largest benefit of using TSA on our arrays was a 10-fold reduction in required sample volume, paving the way for analysis of scarce and valuable samples. (Nordstrom et al. unpublished observations)

4.4 Data processing

The processing of data from scanned microarray images to candidate biomarker signature involves key steps of quantification of array signals, normalization of array data and statistical analysis. There are no clear guide lines for management

of protein array data, in contrast to DNA arrays (Perlee, Christiansen et al. 2004). The data processing strategy applied in clinical studies in **paper III and IV** has been developed within our group (Carlsson, Wingren et al. 2008, Carlsson, Wingren et al. 2011).

First, signal intensities from each spot are quantified, using a fixed spot diameter for all spots, and in this process it becomes evident that homogenous spot morphology is crucial for reliable quantification of data. Also, local background effects can substantially affect specific spot signals. To circumvent this, we have included 8 replicates of each antibody spot. After quantification of signals, the two spots with highest and the two spots with lowest signal intensities have automatically been identified and removed. A mean from the remaining four spots has been regarded as representative and used in further data analysis.

Next, array raw data is normalized in order to compensate for variation in sample handling e.g. labeling, or day-to-day differences. A semi-global normalization approach has been applied in the clinical applications in this thesis (**paper III and IV**). In this normalization approach, the fifteen percent of the analytes in a data set that display the lowest coefficient of variation are identified, and their signal intensities are used to calculate a normalization factor that is applied to all samples is the data set. The normalized data is then applied in different statistical analysis.

In order to evaluate the potential of identified protein profiles as biomarkers, we perform classification analysis. To this end, unsupervised or supervised learning methods can be applied. In unsupervised clustering, the learning method will not know *á priori* which sample group each sample belongs to. The samples are clustered into sub-groups, based on all available information from the data-set, and the obtained clusters can be compared to clinical information.

In contrast, in a supervised learning method, the samples are divided into sample groups, and then the learning method evaluates how well the array data can classify the samples into the correct sample group. Support vector machine (SVM) is a supervised learning method that creates a hyperplane between two pre-defined groups of data. The SVM require a training data set for creating the hyperplane, and then a test data set where the hyperplane can be evaluated. If the data-set is too small for subdivision into training and test sets, a leave-one-

out cross validation can be performed instead. In a leave-one-out cross-validation, one sample is left out while creating the hyperplane, after which the classifier tries to correctly classify the left-out sample. After each iteration, a decision value is calculated based on the distance between the sample and the hyperplane. Based on the decision values, a receiver operating characteristics (ROC) curve is constructed and an area under the curve (AUC) value can be calculated. The AUC-value can consequently be used as a measure of have well the data can classify the samples into the correct sample groups, where AUC = 1 represents a perfect classification, and AUC = 0.5 tells us that the data does not provide any information that can be used for correct classification of samples.

5. Clinical Applications

The quest for protein biomarkers for e.g. early detection and monitoring of complex diseases is an inherently challenging task (Hanash 2003). The effort involves detection of *á priori* unknown proteins, residing in tissue or body fluids at low concentration and in a mixture of thousands of other irrelevant proteins. Also, the proteins are often present at a dynamic range of several orders of magnitude, why low-abundant markers might be "masked" by more high-abundant proteins. Consequently, the study design and techniques used for the task at hand thus have a formidable challenge to live up to.

A very appealing approach is to take advantage of the body's own defense system, evolved over millions of years to discriminate non-self from self, and now highly specified at detecting even subtle changes throughout the body (Paul 2013). A non-invasive strategy for surveying the immune response to complex diseases is to study proteins released by immune cells into the blood stream or fluids proximal to affected organs (Ramachandran, Srivastava et al. 2008). Also, studying proteins in body fluids instead of tissue can help avoiding invasive sampling procedures. The proteins released from immune cells include a large variety of antibodies and immunoregulatory proteins. This complexity poses high demands on techniques used to survey them in order to identify disease specific patterns. The protein microarray format allows simultaneous analysis of thousands of proteins in a high-throughput manner, and can be designed to either target auto-antibodies (antigen arrays) or proteins (antibody arrays)(Kingsmore 2006). These highly sensitive platforms will enable researchers to target low-abundant proteins, such as cytokines, even in the presence of high-abundant proteins, such as albumin and complement factors.

In cancer, tumor associated antigens will evoke an immune response that can be detected in body fluids (Anderson and LaBaer 2005). Our current understanding of this immune surveillance of cancers is limited, and further insight might provide both novel markers of disease and targets for

immunotherapy. Similarly, in autoimmune disorders the immune system reacts to self-antigen, due to loss of immunological tolerance (Paul 2013). Auto-antibodies and immunoregulatory proteins play important roles in the pathogenesis of autoimmunity, but they also constitute a valuable source of candidate markers of disease.

The heterogeneity of both cancers and auto-immune disorders has resulted in low sensitivity of single markers and motivates the quest for multiplex signatures (Wallstrom, Anderson et al. 2013). Discovery- and validation studies in heterogeneous diseases also demand larger sample cohorts compared to more homogeneous disease (Wallstrom, Anderson et al. 2013).

Our recombinant antibody microarrays are based on a wide range of scFvs, primarily targeting immunoregulatory proteins and we have applied our optimized platform in several clinical studies, with a focus on cancer and autoimmunity (Ellmark, Ingvarsson et al. 2006, Carlsson, Persson et al. 2010, Carlsson, Wingren et al. 2011, Carlsson, Wuttge et al. 2011, Wingren, Sandstrom et al. 2012).

5.1 Prostate cancer

Since the PSA test was approved by FDA in 1994, the number of early diagnosed cases of prostate cancer has increased and mortality rates (proportion of all cases) have declined (Welch and Albertsen 2009). If this is a result of us now including more indolent, harmless cancer cases into the statistics, or if we really are curing more cases of dangerous cancers, remains to be elucidated. The debates on the feasibility of using PSA as a screening marker was intensified in 2009, when two large randomized screening studies where published in NEJM, one conducted in the US and one conduced in Europe. The results of the two studies were not consistent: While the US study (Andriole, Crawford et al. 2009) could not detect improved survival among PSA tested men after 7-10 years follow-up, the European study (Schroder, Hugosson et al. 2009) did show a reduction in prostate-cancer related mortality associated with PSA screening. It is, however, evident that tPSA has low specificity for malignant disease and that thousands of healthy men are subjected to biopsy testing causing them harm and risk of infections. The FDA approval of %fPSA in 1998 for men with mid-

range tPSA (4-10 ng/ml) led to a 20% reduction of unnecessary biopsies. Still, the specificity of the test needs further improvement (Figure 4).

Aiming for reduction of unnecessary biopsy-testing in prostate cancer, a large-scale longitudinal study was conducted by Catalona and co-workers during the last decade (Catalona, Partin et al. 2011, Loeb, Sokoll et al. 2012). From 2003 to 2009, 892 men with midrange tPSA (2-10 ng/ml) and benign prostate biopsies were enrolled in a prospective study. The success of this study was evident in June 2012 when FDA approved a new test for risk classification of potential prostate cancer patients, called the Phi-index. This new screening test combines tPSA, %fPSA with a third marker denoted -2 pro PSA. -2 pro PSA is a precursor of PSA more highly elevated in prostate cancer than in benign tissue (Mikolajczyk, Millar et al. 2000). The Phi-test was developed by Beckman Coulter Inc. and has been evaluated in several studies (Jansen, van Schaik et al. 2010) before the large validation study referred to above.

The newly approved Phi-test outperforms %fPSA or tPSA, with a ROC AUC of 0.7 as compared to 0.65 for %fPSA and 0.53 for tPSA (Loeb, Sokoll et al. 2012). There is still a long way to a perfect separation of patient groups, and due to the complexity of the disease and the large patient cohort addressed it is unlikely to find a better separator using a single or a few markers (Wallstrom, Anderson et al. 2013). The stratification of breast cancer patients shown by

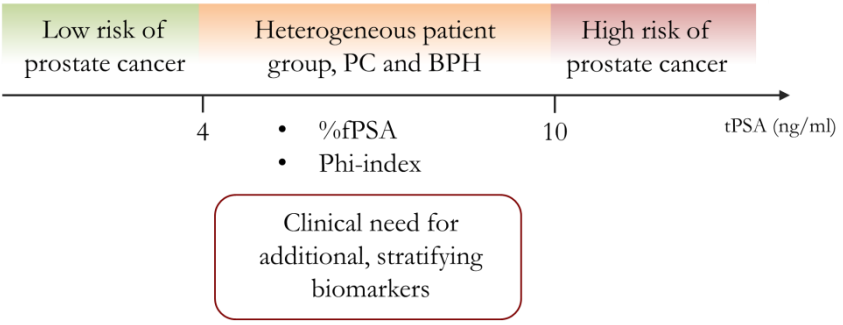

Figure 4. Men with total serum PSA between 4 and 10 ng/ml constitute a heterogeneous patient group including both men with prostate cancer (PC) and benign enlargement of the prostate (BPH). Markers for stratification of this patient group could reduce patient harm and reduce cost of unnecessary biopsies

Sorlie et al. (discussed in 2.1.1), encourages the use of complex gene and protein patterns for identification of disease sub-groups (Sorlie, Perou et al. 2001, Rees, Laversin et al. 2012). An attractive approach would be to use a similar strategy for risk stratification of prostate cancer patients, but instead identifying protein patterns in an easily accessible blood sample.

We have addressed this issue in **paper IV**, where we have analyzed plasma samples from 80 men *á priori* divided into four risk groups of prostate cancer based on tPSA and %fPSA. The groups reflect highly different categories of risk of PC diagnosis, or outcome, with group A having very low long-term risk of significant PC (tPSA ≤0.70 ng/ml), group B having modestly increased risk of prostate disease but low likelihood of clinically significant PC (tPSA: 2.1-8.0 ng/ml and %fPSA ≥27.9%), group C with considerably increased risk of PC (tPSA: 5.0-10-3 ng/ml and %fPSA ≤12.6) and group D having very high risk of clinically significant or advanced stages of PC (tPSA: 24.6-724 ng/ml). All 80 samples were analyzed on our antibody microarray platform optimized for plasma protein analysis, and protein profiles of individual samples were compared to each-other using classification analysis.

The classification analysis tells us to what extent the obtained protein profiles could be used to distinguish the four groups (A-D) from each-other, and is an indication of how well a biomarker signature based on the results would perform. The results showed that the high risk group D could be distinguished from low risk groups A and B with a reasonable accuracy (AUC = 0.68 and 0.72). Also, the two low risk groups, A and B, could be well separated from each-other (AUC = 0.82).

In contrast, the classification analysis also showed that risk group C (midrange tPSA and low %fPSA) could not be distinguished from any of the three other risk groups (A, B and D), with an AUC-value of 0.5 in all three cases. This means that the protein profiles of the C group to some/large extent overlapped with the profiles of the three other groups (A, B and D), and it also indicated that the C group could be a very heterogeneous sample group. These results are in accordance with the notion that the C group represents men who are all sent for biopsy testing, but only around half of them do have cancer, consequently constituting a highly diverse group of patients. Therefore, we set out to investigate if we could stratify this heterogeneous sample group further and performed an unsupervised clustering of the C group data. The results showed

that we indeed could stratify risk group C into two distinct subgroups, denoted C1 and C2. Of note, further classification analysis showed that group C1 appeared to have protein profiles similar to the profiles of low risk groups A and B while group C2 appeared to be more similar to high risk group D. These results indicate that protein microarray data can be used to stratify heterogeneous patient groups into sub-groups of higher of lower risk of a certain disease. As discussed in chapter 2, discovery studies like ours need to be validated in larger sample cohorts, where the samples are fully documented with regard to disease status and follow-up.

Targeting prostate cancer using antibody arrays was first performed by Haab and co-workers representing early clinical applications based on the protein array format (Miller, Zhou et al. 2003, Shafer, Mangold et al. 2007). In 2003, Miller et al. presented a five protein marker panel that had significantly different levels in serum from patients with prostate cancer as compared to controls. Further, in 2007 Shafer presented antibody array data showing that thrombospondin-1 levels were elevated in benign prostate enlargement, while not in patient with malignant disease. However, further pre- and validation studies of these discovery studies have not yet been reported.

On a genetic level, loss of tumor suppressor PTEN has been described in variety of human cancers and has in prostate cancer been associated with tumor progression and poor prognosis (McMenamin, Soung et al. 1999). With that as a starting-point, Cima and coworkers identified 775 N-linked glycoproteins from PTEN negative mice and used targeted proteomics (MRM) in order to identify protein profiles for diagnosis and grading of prostate cancer (Cima, Schiess et al. 2011). This study demonstrates the feasibility of integrating of genetics, proteomics as well as experimental mouse models.

The proximity of the prostate to the renal system, together with the ease of urine sampling gives urine a status as an attractive sample format for prostate cancer diagnosis and prognosis (Downes, Byrne et al. 2007, Jamaspishvili, Kral et al. 2010). Markers that have been verified in independent sample cohorts of prostate cancer versus controls include VEGF and matrix metalloproteinases (Chan, Moses et al. 2004). Still, the study of multiplex panels of urinary proteins in association with prostate cancer is so far limited, but indicated great opportunities (Jamaspishvili, Kral et al. 2010). With the recently optimized

protocol for urine analysis (**paper I**), exploring the potential of our platform for this purpose might prove rewarding.

5.2 Systemic lupus erythematosus (SLE)

The complexity of SLE pathophysiology calls for systematic and multiplex analysis in order to identify biomarkers and drug targets for improved prognosis of SLE patients. SLE has risen as a model autoimmune disease, and has in several efforts been addressed using different multiplexed platforms, including genetic and proteomic approaches (Balboni, Chan et al. 2006, Maecker, Lindstrom et al. 2012).

The heterogeneity of the disease has drawn attention to multiplexed analysis of protein and auto-antibodies using array-based approaches, aiming for diagnosis as well as prognosis (Balboni, Chan et al. 2006, Maecker, Lindstrom et al. 2012). Several efforts for identification of autoantibody repertoires revealed by antigen arrays have been investigated in SLE (Fattal, Shental et al. 2010, Maecker, Lindstrom et al. 2012, Papp, Vegh et al. 2012). Fattal et al. used arrays comprising 694 different antigens, mainly self-antigens, to study the humoral response of 40 SLE patients and compared them to 16 matched healthy controls. The results showed that the detected auto-antibody profiles, associated with active disease, persisted even after long term clinical remission. Further, Papp et al. used antigen arrays of 58 features and sera from 61 SLE patients of active and inactive disease, and identified both auto-antibodies as well as complement factors C3 and C4, complex bound to the array. Hence, antigen arrays can be utilized for identification of potential markers as well as mapping of SLE pathogenesis.

Comprehensive genetic studies of SLE pathogenesis have revealed a central role of the type 1 Interferon (IFN) pathway (Pascual, Farkas et al. 2006). Up-regulation of the type 1 IFN pathway has been associated with SLE both on gene- (Baechler, Batliwalla et al. 2003, Bennett, Palucka et al. 2003, Kirou, Lee et al. 2005) and protein level (Bengtsson, Sturfelt et al. 2000), and key players in the pathway could potentially be utilized both as markers and drug targets. To further test the role of type I IFN-regulated proteins in SLE, Bauer et al. conducted a survey of 160 serologic cytokines (Bauer, Baechler et al. 2006). Sera

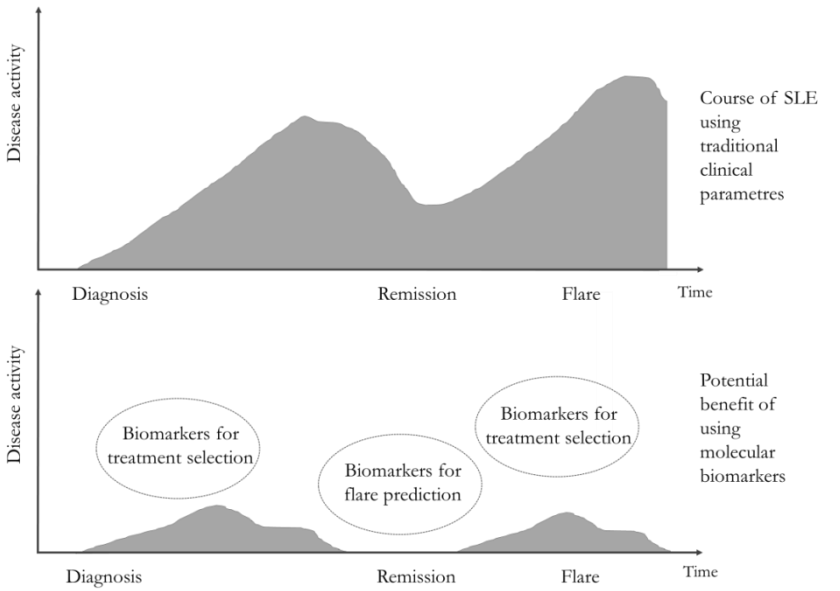

Figure 5. The use of molecular biomarkers, detected in serum and urine, could potentially improve quality of life for SLE patients. By using biomarkers for treatment selection and flare prediction, the flares could more effectively be suppressed and thereby reducing the symptoms for the patients, and further organ damage could also be prevented.

from 30 SLE patients (15 patients with high gene expression of 82 type I INF-related genes and 15 patients with lower levels of the same gene panel) and 15 matched controls were analyzed on a 160 cytokine dual antibody array using RCA as detection system. The results showed that 30 differentially expressed cytokines could be delineated between the group of high IFN gene expression versus the control group. Out of these 30 cytokines, a panel of three chemokines (IP-10, MCP-1 and MIP-3B) displaying the strongest correlations with disease activity, was chosen for validation in a second, independent follow-up study. In the validation study, 267 SLE patients were followed for one year in a longitudinal study (Bauer, Petri et al. 2009). Results from the validation study showed that this three-plex chemokine panel correlated with disease activity, supporting the use of multiplex protein panels for monitoring of SLE disease activity. Antibody arrays for profiling of cytokines have also been employed in other autoimmune conditions, including multiple sclerosis and

rheumatoid arthritis, where therapies utilizing either increase or decrease of cytokine levels have been effective (Balboni, Chan et al. 2006).

In this context, our group has applied our antibody microarray platform, targeting mainly cytokines and complement proteins, in a recent SLE biomarker discovery study (Carlsson, Wuttge et al. 2011). The results showed that multiplexed candidate serum protein panels for diagnosis, prognosis and classification could be identified. Of note, the patients included in this study had a wide variety of different manifestations, ranging from low to high severity, and the identified protein signatures could, with high accuracy, classify the patients into the correct subgroup.

The most severe manifestation of SLE, SLE nephritis, is characterized by pronounced renal involvement, and associated with high morbidity and mortality if left untreated (Rovin, Birmingham et al. 2007). The characteristics of nephritis, including both systemic and renal symptoms, motivate the use of both serum and urine as potential sources of markers for e.g. diagnosis, monitoring of disease and treatment selection (Figure 5). Proteins detected in serum are more likely to reflect the systemic characteristics of SLE pathogenesis, while urinary proteins to great extent (approx. 70%) originate from the renal system and more accurately reflect nephritis activity (Decramer, Gonzalez de Peredo et al. 2008). By studying both sample formats in a single study, routes of clearance and degradation of proteins can potentially be investigated, e.g. by identifying proteins decreasing in serum, while simultaneously increasing in urine.

In **paper III**, we have adopted the above described approach of harnessing the immune system as a sensitive sensor for disease activity in SLE nephritis. In order to identify protein signatures reflecting disease activity, we have analyzed serum (n = 59) and urine (n = 58) samples from patients with SLE nephritis and candidate protein biomarker panels associated with high versus low disease activity were delineated. The majority of analysis was performed as a case-control study, comparing patients with active disease to patients with inactive disease. In addition, a smaller portion of the sample cohort had been followed over time and a longitudinal analysis could be performed. The identified protein profiles included both known and novel markers of disease activity. Previously reported markers included complement proteins (C1q, C3 and C4 in serum) as

well as a number of cytokines (e.g. MCP-1, IL-6, IL- 10, IFN-γ and TNF-α) (Li, Tucci et al. 2006, Gigante, Gasperini et al. 2011, Brugos, Vincze et al. 2012)

Disease activity was defined both based on systemic symptoms (defined by SLEDAI-2K) (Gladman, Ibanez et al. 2002) as well on specific renal symptoms (defined by renal SLEDAI). We could identify unique protein panels for systemic versus renal symptoms, indicating that we could pin-point biomarkers for monitoring of renal flare activity. In more detail, serum protein panels reflecting systemic symptoms included down-regulated C1q, and up-regulated IL-6, IL-8, IL-10, IL-12, IL-16 and MCP-1, while the serum protein panels reflecting renal symptoms included up-regulated IL-6, IL-16, MCP-1 and CD40, and again down-regulated C1q. The identified urinary profiles included mainly up-regulated proteins and were of substantially different length and composition, as compared to the serum profiles, and are more extensively described in **paper III**.

Comparing the results from serum or urinary analysis revealed two key points. First, the two sample formats provided protein profiles of substantially differences with regard to both length and composition, indicating that complementary information was retrieved by analyzing the two sample formats. Second, serum profiles performed best in the region of higher disease activity, i.e. conferred most pronounced separation of the groups in term in AUC-values. In contrast, the urinary profiles performed better in the region of low disease activity, distinguishing no/low disease activity from medium or high disease activity. Taken together, the two sample formats could preferable be used in combination in order to obtain a more extensive view of the disease status. To the best of our knowledge, this is the first study deciphering multiplex panels in both serum and urine reflecting disease activity and renal involvement.

By analyzing a small subset of samples, we also carried out a longitudinal study, performing pair-wise comparisons of two samples from each patient, collected both during and between flares. The obtained protein profiles overlapped to large extent with the case-control results, while some additional proteins were identified in the pair-wise analysis, including GM-CSF, IL-1ra and IL-1β. These results demonstrated that a refined view of SLE disease pathogenesis might be obtained through pair-wise comparisons, due to elimination of biological (patient-to-patient) variation. Future studies, targeting larger sample cohorts,

will be required for validation of identified protein panels. Also, pathway analysis of identified proteins would bring further insight into disease pathogenesis.

6. Concluding remarks

The process of identifying novel markers for diagnosis, prognosis, and treatment decision has proven a difficult task. A successful biomarker discovery requires clinical sample cohorts of high quality and quantity, as well as technologies that can identify disease related differences between clinical samples that then can be used to differentiate the targeted patient groups with high specificity and sensitivity.

The aim of this thesis, based on four original papers, has been to further optimize and apply an affinity proteomics platform for biomarker discovery, antibody microarrays. In **papers I and II**, I have addressed two key assay features, probe format and sample format, and in **papers III and IV**, I have applied the optimized platform in two clinical studies, targeting prostate cancer and SLE.

The on-chip functionality of arrayed probes is essential for well-performing antibody arrays. In **paper II**, I addressed the on-chip performance of scFv selected from the n-CoDeR library by introducing stabilizing mutations into the common FW of the scFvs. The results showed that the point-mutations E6Q (VH) and S96P (VL) conferred improved stability on soluble proteins and clone-dependent effects on the on-chip stability of the scFvs. For future application of the E6Q and S96P mutations in microarray experiments, the mutations should be introduced into the master FW of a combinatorial scFv library, and all selected clones would then carry the novel mutations. The mutants have so far not been applied in large-scale microarray analysis, simply due to logistic issues. The stabilizing mutations E6Q and S96P could be especially beneficial for scFvs with intended use under de-stabilizing conditions, such as i) long-time storage, ii) recombinant tagging of scFv conferring decreased stability or iii) cytosolic applications, where conserved disulfide-bridges not are formed.

In **paper I**, I have optimized a protocol for urinary proteomics using antibody microarrays, now allowing us to target all of the most commonly used clinical sample formats. This will enable us to perform studies including several different sample formats, as was done in **paper III** where both urine and serum was analyzed for identification of novel marker associated with SLE nephritis activity. This combined approach proved to be very beneficial as we could not only detect candidate biomarkers from both urine and serum, but we could also show that the two sample formats provided complementary information. The biomarker panels identified in serum and urine were of substantially different length and composition. In addition, the classification analysis indicated that the urine markers worked best in distinguishing high from medium disease activity, while serum performed better when distinguishing medium from low disease activity. Taken together, urine and serum biomarkers could be used in parallel in clinical settings, providing complementary information, in the end giving a more comprehensive view of the disease status.

The application studies (**paper III and IV**) demonstrate the feasibility of using antibody microarrays for identification of candidate protein biomarker signatures for risk classification and monitoring of disease. In **paper III**, protein signatures associated with SLE nephritis disease activity and renal involvement were identified, including both previously reported and novel markers. In **paper IV**, identified protein profiles could be utilized for stratification of a heterogeneous sample group into two groups of high or low risk of having prostate cancer. These two clinical applications (**paper III and IV**) together show that we now have at hand an analysis platform that can be used for identification of protein profiles with potential use as biomarker signatures. We can print up to 2000 antibodies per cm^2 and analyze hundreds of samples per day (per workstation). Consequently, the current bottleneck is the availability of a sufficient number of well-characterized affinity reagents and in particular large, well-documented sample cohorts.

Taken together, key determinant factors for bringing biomarker candidates from discovery studies into the clinics are the issues of development and production of affinity reagents, and sample collection and banking. Thousands of affinity reagents will be required for un-targeted analysis of the human proteome. Also, the discovery and validation studies will need to enroll numerous patients and controls, or utilize banked material from large biobanks. Extensive collaborations between academia, public sector, and industry will be

required for both the development of affinity reagents as well as for performing large-scale studies. However, if the above described logistics can be resolved, there are great hopes of affinity proteomics delivering biomarker signatures to the clinics in the near future. With high-performing biomarkers for early diagnosis, prognosis, and stratification of patients for treatment selection at hand, personalized medicine could truly impact the survival and quality of life of for thousands of patients.

Populärvetenskaplig sammanfattning

Cancer är inte bara en sjukdom, utan väldigt många olika. För det första kan cancertumörer befinna sig på många olika ställen i kroppen, till exempel bröstcancer, prostatacancer och tjocktarmscancer. För det andra finns det många olika cancervarianter som är olika farliga för patienten. Vissa cancervarianter är aggressiva, sprider sig snabbt och kräver tufft behandling, medan andra gånger växer tumören så långsamt att den inte behöver behandlas alls. Det finns också olika cancervarianter som svarar olika bra på olika behandlingsmetoder. Problemet är att det ofta är svårt att avgöra vilka tumörer som är ofarliga och vilka som behöver behandlas, utan alla får samma tuffa behandling vilket ofta ger svåra biverkningar för patienten. För att snabbt kunna ta riktiga beslut om vilka patienter som bör få en viss behandling behöver läkarna nya hjälpmedel. Ett sådant hjälpmedel skulle kunna vara så kallade biomarkörer.

En biomarkör kan vara en gen eller ett protein, som kan mätas i ett vävnadsprov (biopsi) eller i ett blodprov. I vissa fall har forskare lyckats identifiera biomarkörer som kan ge läkaren besked om en viss patient kan förväntas ha nytta av en viss behandling eller inte. Forskningen för att identifiera nya biomarkörer har tagit ordentlig fart senaste decennierna. Detta är mycket tack vare kartläggningen av människans DNA (arvsmassa) i början av 2000-talet och utveckling av ny tekniker som möjliggör att vi nu kan mäta upp till tusentals olika proteiner på en gång. Sådan storskalig undersökning av proteiner kallas proteomik.

Den proteomikteknik vi har utvecklat i vår forskargrupp kallas antikroppsmikromatriser (antibody microarrays), där vi använder oss av antikroppar för att analysera proteiner i blodet eller andra biologiska prover. Antikroppar är specialiserade på att binda till olika proteiner för att därmed särskilja mellan friskt och sjukt, och den förmågan använder vi oss av i vår teknik. Vi tar hjälp av en robot för att placera små (ca 0.0000003 ml) droppar av olika antikroppar på en plastyta så att de bildar ett ordnat mönster, en matris. Vi

använder sedan antikropparna som metspön för att fiska ut specifika proteiner ur provet, och dessa proteiner kan sedan i förlängningen användas som biomarkörer. Fördelen med att placera ut så små droppar med antikroppar är att vi då från ett och samma prov kan analysera hundratals till tusentals proteiner på en gång. Genom att utföra analysen t.ex. på prover från patienter med en viss sjukdom och jämföra med prover från friska personer kan vi identifiera proteinmönster som i förlängningen kan användas till biomarkörer för diagnostik av cancer eller andra sjukdomar.

På senare år har många forskargrupper presenterat potentiella biomarkörer som i deras studier visat att de kan särskilja mellan prover från sjuka och friska personer. Det har däremot visat sig vara svårt att upprepa resultaten i nya, större provsamlingar, vilket är ett krav för att kunna börja använda biomarkören på patienter i sjukvården. Jag har i min avhandling diskuterat några anledningar till varför detta är så svårt. För det första är upplägget av studien viktig, d.v.s. hur patienterna och proverna som ska analyseras väljs ut. För det andra spelar det stor roll vilket provformat (blod, urin etc.) som används och hur proverna behandlas vid provtagning och därefter. Slutligen är kraven stora på analysmetoderna som används för att identifiera biomarkör. Metoden måste t.ex. kunna mäta väldigt låga proteinkoncentrationer, i ett komplext prov där det finns tusentals andra proteiner.

Min avhandling är baserad på fyra artiklar. Två av dessa handlar om teknikutveckling av vår proteomikplattform och två av dem handlar om hur jag har använt vår plattform för att analysera blod- och urinprover med målet att finna nya biomarkör.

I den första artikeln har jag optimerat plattformen så att vi nu även kan analysera proteiner i urinprover. Proteiner som finns i människans urin härrör till största del från njurarna, men det är även en stor del som har filtrerats ut från blodet och som kan spegla sjukdomstillstånd i hela kroppen. Detta medför att vi kan använda vår analysplattform i kliniska studier för att leta efter biomarkörer både för njursjukdomar, men även för sjukdomar som berör resten av kroppen.

I min andra artikel har jag fokuserat på de antikroppar som vi använder på våra mikromatriser. För att producera våra matriser tar roboten en lösning innehållande antikroppar och skjuter ut små droppar som sedan torkar ut på plastytan där matriserna bildas. Detta är en väldigt hård påfrestning för antikropparna och de måste ha en stabil struktur för att inte denaturera, d.v.s.

förlora sin struktur och därmed sin funktion. För att försöka förbättra antikropparnas stabilitet har jag infört några mutationer d.v.s. jag har gjort några förändringar i antikropparnas DNA. Jag har sedan jämfört antikroppar med och utan mutationer för att se vad mutationerna hade för effekt. Det visade sig att när antikropparna befann sig i lösning hade de nytta av mutationerna och uppvisade mer stabil struktur. Däremot var det svårare att uttala sig om antikropparna som hade torkat ut på plastyta hade lika mycket nytta av mutationerna.

I artikel tre och fyra har jag använt mikromatriserna för att identifiera biomarkörer i den autoimmuna sjukdomen SLE (lupus) och prostatacancer. SLE är en sjukdom som kan påverka större delen av kroppen, allt från leder till njurar. Sjukdomen går i skov, vilket betyder att sjukdomssymptom kommer och går och det är svårt att förutspå när nästa skov kommer. Om patienten får behandling när skovet börjar, eller helst innan, kan symptomen dämpas kraftigt vilket både betyder att patienten mår bättre för stunden och dessutom att kroppen inte bryts ned så mycket av skovet. I artikel tre har jag analyserat både blod- och urinprover från SLE-patienter med hjälp av våra mikromatriser. Proverna var tagna både under och mellan skov, vilket gjorde det möjligt att identifiera proteinmönster som speglade skovet. Dessa proteinmönster skulle i förlängningen kunna användas som biomarkörer som kan avgöra om ett skov är på gång så att patienten snabbt kan påbörja behandling.

Prostatacancer är den vanligaste diagnostiserade cancerformen bland män i industrialiserade länder. I Sverige får nästan 9000 män diagnosen varje år. För att bedöma om någon har förhöjd risk för prostatacancer kan läkaren med hjälp av ett blodprov mäta nivån på biomarkören PSA (prostata specifikt antigen). Utifrån PSA-nivån bedöms sedan om man bör ta ett vävnadsprov och med hjälp av detta vävnadsprov ställer sedan läkaren diagnos. Problemet med PSA-testet är det medför att väldigt många män får beskedet att de har förhöjd risk för cancer och bör lämna vävnadsprov, men av dessa män är det bara cirka en fjärdedel som visar sig ha prostatacancer. Det betyder att tusentals män varje år tvingas genomgå provtagning i onödan, och att de felaktigt behöver oroa sig för att de kanske har cancer. För att undvika detta behövs nya, bättre biomarkörer.

I min fjärde artikel har vi analyserat blodprover från män med olika hög risk för prostatacancer. Vi kunde identifiera proteinmönster som skulle kunna användas som ett komplement till PSA för att förhindra att så många män tvingas lämna

vävnadsprov i onödan. Detta skulle leda till bättre livskvalitet för männen och lägre kostnader för sjukvården.

Sammanfattningsvis har jag med mina artiklar bidragit till utveckling av vår proteomikplattform och sedan visat på dess användning för identifiering av nya biomarkörer. Detta kan i förlängningen leda till att sjukvården kommer ha tillgång till fler och bättre biomarkörer som kan vara viktiga hjälpmedel både vid diagnostik av olika sjukdomar, och för att välja behandlingsmetoder som passar den enskilde patienten.

Acknowledgement

Tack till min handledare Christer för engagemang, entusiasm och peppning genom åren och för att man alltid kan komma förbi men en fråga eller flera. Tack för det stora lasset du drog sista veckorna, nu hoppas jag att du får mer tid för cykelträning.

Tack till min biträdande handledare Carl för dina visioner och inspirerande diskussioner. Tack för att jag fått möjligheten att doktorera på en avdelning där du har samlat både spännande och viktiga projekt och dedikerade forskare!

Tack Per för mentorsmöten som gett mig nya perspektiv och en spark i baken. Jag har alltid lämnat våra möten med ny kraft och energi.

Tack till alla samarbetspartner för möjliggörande av spännande projekt samt värdefull input och diskussioner. Tack främst till Anders B, Hans L, Charlotte B i prostatastudien samt Gunnar S och Anders B i SLE-studien.

Tack till studenter jag fått handleda i projektarbeten (Karolina, Oskar och Linn) och exjobb (Jonas, Elin och Staffan), för fint arbete och fina resultat, samt att jag får möjligheten att lära mig att lära ut.

Tack Elin, Niclas, Frida och mamma för er feedback på avhandlingen. Tack Laura för smarta tips angående bildredigering.

Tack till mina rumskamrater. Först Linda som lotsade in mig på labbet, alltid med ett leende på läpparna och en uppmuntrande ringsignal på mobilen. Senaste åren har Linn har fått stå ut med mina suckar och när jag pratar för mig själv. Det har skönt att alltid kunna kasta några ord över axeln om jobb eller övriga livet, tack!

Tack till övriga protein-chip gruppen, för hjälp, samarbete och vänskap. Tack Anna för långa stunder i spotterrummet och flertaligt konferenssällskap. Tack Frida för att jag får låna lite av din skärpa ibland. Tack Niclas för att du alltid tar dig tid när jag står i dörren med en fundering. Tack Anna, Payam och Petter för

hejarop sista tiden. Tack Anders för allt från skaka-islådan-knep till statistikskript. Tack Jacob som handledde mig första tiden på avdelningen.

Tack alla på IT för utflykter, gåsamiddagar, torsdagsbullar (I miss you), onsdagsfikor och alla fantastiska disputationsfester genom åren. Sådana kollegor växer inte på träd. Tack Marianne och Tommie för att ni håller reda på oss.

Tack till vänner i Skåne och i övriga landet för engagemang, stöttning och källor till glädje!

Tack till mamma och pappa för alla spännande samtal vid middagsbordet som födde min nyfikenhet, och för att ni alltid trott på mig. Tack till mina syskon för att ni alltid sporrar mig, för alla roliga stunder och för ert stöd

Tack Johan för din kärlek. När jag förtjänat den som minst har jag behövt den som mest. Tack för fantastisk markservice senaste månaderna, och för din (irriterande) förmåga att se ljuspunkter och möjligheter i alla situationer.

Tack Stina för att du alltid lockar fram ett leende hos mig och för att du gett mig nya perspektiv på livet. Nu ska vi ha en fantastisk sommar tillsammans!

References

Adachi, J., C. Kumar, Y. Zhang, J. V. Olsen and M. Mann (2006). "The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins." Genome Biol **7**(9): R80.

Agmon-Levin, N., M. Mosca, M. Petri and Y. Shoenfeld (2012). "Systemic lupus erythematosus one disease or many?" Autoimmun Rev **11**(8): 593-595.

Anderson, K. S. and J. LaBaer (2005). "The sentinel within: exploiting the immune system for cancer biomarkers." J Proteome Res **4**(4): 1123-1133.

Anderson, N. L. and N. G. Anderson (2002). "The human plasma proteome: history, character, and diagnostic prospects." Mol Cell Proteomics **1**(11): 845-867.

Anderson, N. L., N. G. Anderson, L. R. Haines, D. B. Hardie, R. W. Olafson and T. W. Pearson (2004). "Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA)." J Proteome Res **3**(2): 235-244.

Anderson, N. L., A. S. Ptolemy and N. Rifai (2013). "The riddle of protein diagnostics: future bleak or bright?" Clin Chem **59**(1): 194-197.

Andriole, G. L., E. D. Crawford, R. L. Grubb, 3rd, S. S. Buys, D. Chia, T. R. Church, M. N. Fouad, E. P. Gelmann, P. A. Kvale, D. J. Reding, J. L. Weissfeld, L. A. Yokochi, B. O'Brien, J. D. Clapp, J. M. Rathmell, T. L. Riley, R. B. Hayes, B. S. Kramer, G. Izmirlian, A. B. Miller, P. F. Pinsky, P. C. Prorok, J. K. Gohagan, C. D. Berg and P. P. Team (2009). "Mortality results from a randomized prostate-cancer screening trial." N Engl J Med **360**(13): 1310-1319.

Angenendt, P. (2005). "Progress in protein and antibody microarray technology." Drug Discov Today **10**(7): 503-511.

Angenendt, P., J. Glokler, J. Sobek, H. Lehrach and D. J. Cahill (2003). "Next generation of protein microarray support materials: evaluation for protein and antibody microarray applications." J Chromatogr A **1009**(1-2): 97-104.

Austin, J. and A. H. Holway (2011). "Contact printing of protein microarrays." Methods Mol Biol **785**: 379-394.

Baechler, E. C., F. M. Batliwalla, G. Karypis, P. M. Gaffney, W. A. Ortmann, K. J. Espe, K. B. Shark, W. J. Grande, K. M. Hughes, V. Kapur, P. K. Gregersen and T. W. Behrens (2003). "Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus." Proc Natl Acad Sci U S A **100**(5): 2610-2615.

Balboni, I., S. M. Chan, M. Kattah, J. D. Tenenbaum, A. J. Butte and P. J. Utz (2006). "Multiplexed protein array platforms for analysis of autoimmune diseases." Annu Rev Immunol **24**: 391-418.

Ballard, J. L., V. K. Peeva, C. J. deSilva, J. L. Lynch and N. R. Swanson (2007). "Comparison of Alexa Fluor and CyDye for practical DNA microarray use." Mol Biotechnol **36**(3): 175-183.

Bansal, A. and M. Sullivan Pepe (2013). "When does combining markers improve classification performance and what are implications for practice?" Stat Med.

Barbas, C. F., 3rd, J. D. Bain, D. M. Hoekstra and R. A. Lerner (1992). "Semisynthetic combinatorial antibody libraries: a chemical solution to the diversity problem." Proc Natl Acad Sci U S A **89**(10): 4457-4461.

Bartling, B., H. S. Hofmann, T. Boettger, G. Hansen, S. Burdach, R. E. Silber and A. Simm (2005). "Comparative application of antibody and gene array for expression profiling in human squamous cell lung carcinoma." Lung Cancer **49**(2): 145-154.

Bates, S. (2010). "Progress towards personalized medicine." Drug Discov Today **15**(3-4): 115-120.

Bauer, J. W., E. C. Baechler, M. Petri, F. M. Batliwalla, D. Crawford, W. A. Ortmann, K. J. Espe, W. Li, D. D. Patel, P. K. Gregersen and T. W. Behrens (2006). "Elevated serum levels of interferon-regulated chemokines are biomarkers for active human systemic lupus erythematosus." PLoS Med **3**(12): e491.

Bauer, J. W., M. Petri, F. M. Batliwalla, T. Koeuth, J. Wilson, C. Slattery, A. Panoskaltsis-Mortari, P. K. Gregersen, T. W. Behrens and E. C. Baechler (2009). "Interferon-regulated chemokines as biomarkers of systemic lupus erythematosus disease activity: a validation study." Arthritis Rheum **60**(10): 3098-3107.

Belov, L., S. P. Mulligan, N. Barber, A. Woolfson, M. Scott, K. Stoner, J. S. Chrisp, W. A. Sewell, K. F. Bradstock, L. Bendall, D. S. Pascovici, M. Thomas, W. Erber, P. Huang, M. Sartor, G. A. Young, J. S. Wiley, S. Juneja, W. G. Wierda, A. R. Green, M. J. Keating and R. I. Christopherson (2006). "Analysis of human leukaemias and lymphomas using extensive immunophenotypes from an antibody microarray." Br J Haematol **135**(2): 184-197.

Bengtsson, A. A., G. Sturfelt, L. Truedsson, J. Blomberg, G. Alm, H. Vallin and L. Ronnblom (2000). "Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies." Lupus **9**(9): 664-671.

Bennett, L., A. K. Palucka, E. Arce, V. Cantrell, J. Borvak, J. Banchereau and V. Pascual (2003). "Interferon and granulopoiesis signatures in systemic lupus erythematosus blood." J Exp Med **197**(6): 711-723.

Binz, H. K., M. T. Stumpp, P. Forrer, P. Amstutz and A. Pluckthun (2003). "Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins." J Mol Biol **332**(2): 489-503.

Bork, P., L. Holm and C. Sander (1994). "The immunoglobulin fold. Structural classification, sequence patterns and common core." J Mol Biol **242**(4): 309-320.

Borrebaeck, C. A. K. and M. Ohlin (2002). "Antibody evolution beyond Nature." Nat Biotechnol **20**(12): 1189-1190.

Borrebaeck, C. A. K. and C. Wingren (2009). "Design of high-density antibody microarrays for disease proteomics: key technological issues." J Proteomics **72**(6): 928-935.

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

Boschetti, E., M. C. Chung and P. G. Righetti (2012). ""The quest for biomarkers": are we on the right technical track?" Proteomics Clin Appl **6**(1-2): 22-41.
Brennan, D. J., D. P. O'Connor, E. Rexhepaj, F. Ponten and W. M. Gallagher (2010). "Antibody-based proteomics: fast-tracking molecular diagnostics in oncology." Nat Rev Cancer **10**(9): 605-617.
Brugos, B., Z. Vincze, S. Sipka, G. Szegedi and M. Zeher (2012). "Serum and urinary cytokine levels of SLE patients." Pharmazie **67**(5): 411-413.
Canney, P. A., M. Moore, P. M. Wilkinson and R. D. James (1984). "Ovarian cancer antigen CA125: a prospective clinical assessment of its role as a tumour marker." Br J Cancer **50**(6): 765-769.
Cardoso, F., L. Van't Veer, E. Rutgers, S. Loi, S. Mook and M. J. Piccart-Gebhart (2008). "Clinical application of the 70-gene profile: the MINDACT trial." J Clin Oncol **26**(5): 729-735.
Carlsson, A., O. Persson, J. Ingvarsson, B. Widegren, L. Salford, C. A. K. Borrebaeck and C. Wingren (2010). "Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients." Proteomics Clin Appl **4**(6-7): 591-602.
Carlsson, A., C. Wingren, J. Ingvarsson, P. Ellmark, B. Baldertorp, M. Ferno, H. Olsson and C. A. K. Borrebaeck (2008). "Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays." Eur J Cancer **44**(3): 472-480.
Carlsson, A., C. Wingren, M. Kristensson, C. Rose, M. Ferno, H. Olsson, H. Jernstrom, S. Ek, E. Gustavsson, C. Ingvar, M. Ohlsson, C. Peterson and C. A. K. Borrebaeck (2011). "Molecular serum portraits in patients with primary breast cancer predict the development of distant metastases." Proc Natl Acad Sci U S A **108**(34): 14252-14257.
Carlsson, A., D. M. Wuttge, J. Ingvarsson, A. A. Bengtsson, G. Sturfelt, C. A. K. Borrebaeck and C. Wingren (2011). "Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays." Mol Cell Proteomics **10**(5): M110 005033.
Carter, P., L. Presta, C. M. Gorman, J. B. Ridgway, D. Henner, W. L. Wong, A. M. Rowland, C. Kotts, M. E. Carver and H. M. Shepard (1992). "Humanization of an anti-p185HER2 antibody for human cancer therapy." Proc Natl Acad Sci U S A **89**(10): 4285-4289.
Catalona, W. J., A. W. Partin, M. G. Sanda, J. T. Wei, G. G. Klee, C. H. Bangma, K. M. Slawin, L. S. Marks, S. Loeb, D. L. Broyles, S. S. Shin, A. B. Cruz, D. W. Chan, L. J. Sokoll, W. L. Roberts, R. H. van Schaik and I. A. Mizrahi (2011). "A multicenter study of [-2]pro-prostate specific antigen combined with prostate specific antigen and free prostate specific antigen for prostate cancer detection in the 2.0 to 10.0 ng/ml prostate specific antigen range." J Urol **185**(5): 1650-1655.
Catalona, W. J., A. W. Partin, K. M. Slawin, M. K. Brawer, R. C. Flanigan, A. Patel, J. P. Richie, J. B. deKernion, P. C. Walsh, P. T. Scardino, P. H. Lange, E. N. Subong, R. E. Parson, G. H. Gasior, K. G. Loveland and P. C. Southwick (1998). "Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: a prospective multicenter clinical trial." JAMA **279**(19): 1542-1547.
Chan, L. W., M. A. Moses, E. Goley, M. Sproull, T. Muanza, C. N. Coleman, W. D. Figg, P. S. Albert, C. Menard and K. Camphausen (2004). "Urinary VEGF and MMP

levels as predictive markers of 1-year progression-free survival in cancer patients treated with radiation therapy: a longitudinal study of protein kinetics throughout tumor progression and therapy." J Clin Oncol **22**(3): 499-506.
Chao, J., R. DeBiasio, Z. Zhu, K. A. Giuliano and B. F. Schmidt (1996). "Immunofluorescence signal amplification by the enzyme-catalyzed deposition of a fluorescent reporter substrate (CARD)." Cytometry **23**(1): 48-53.
Chatterjee, S. K. and B. R. Zetter (2005). "Cancer biomarkers: knowing the present and predicting the future." Future Oncol **1**(1): 37-50.
Cheng, L., R. Montironi, D. G. Bostwick, A. Lopez-Beltran and D. M. Berney (2012). "Staging of prostate cancer." Histopathology **60**(1): 87-117.
Chugh, P. K. and B. S. Kalra (2013). "Belimumab: targeted therapy for lupus." Int J Rheum Dis **16**(1): 4-13.
Cima, I., R. Schiess, P. Wild, M. Kaelin, P. Schuffler, V. Lange, P. Picotti, R. Ossola, A. Templeton, O. Schubert, T. Fuchs, T. Leippold, S. Wyler, J. Zehetner, W. Jochum, J. Buhmann, T. Cerny, H. Moch, S. Gillessen, R. Aebersold and W. Krek (2011). "Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer." Proc Natl Acad Sci U S A **108**(8): 3342-3347.
Clamp, M., B. Fry, M. Kamal, X. Xie, J. Cuff, M. F. Lin, M. Kellis, K. Lindblad-Toh and E. S. Lander (2007). "Distinguishing protein-coding and noncoding genes in the human genome." Proc Natl Acad Sci U S A **104**(49): 19428-19433.
Coon, J. J., P. Zurbig, M. Dakna, A. F. Dominiczak, S. Decramer, D. Fliser, M. Frommberger, I. Golovko, D. M. Good, S. Herget-Rosenthal, J. Jankowski, B. A. Julian, M. Kellmann, W. Kolch, Z. Massy, J. Novak, K. Rossing, J. P. Schanstra, E. Schiffer, D. Theodorescu, R. Vanholder, E. M. Weissinger, H. Mischak and P. Schmitt-Kopplin (2008). "CE-MS analysis of the human urinary proteome for biomarker discovery and disease diagnostics." Proteomics Clin Appl **2**(7-8): 964.
Cordero, O. J., L. De Chiara, Y. Lemos-Gonzalez, M. Paez de la Cadena and F. J. Rodriguez-Berrocal (2008). "How the measurements of a few serum markers can be combined to enhance their clinical values in the management of cancer." Anticancer Res **28**(4C): 2333-2341.
D'Amico, A. V., A. Desjardin, A. Chung, M. H. Chen, D. Schultz, R. Whittington, S. B. Malkowicz, A. Wein, J. E. Tomaszewski, A. A. Renshaw, K. Loughlin and J. P. Richie (1998). "Assessment of outcome prediction models for patients with localized prostate carcinoma managed with radical prostatectomy or external beam radiation therapy." Cancer **82**(10): 1887-1896.
Decramer, S., A. Gonzalez de Peredo, B. Breuil, H. Mischak, B. Monsarrat, J. L. Bascands and J. P. Schanstra (2008). "Urine in clinical proteomics." Mol Cell Proteomics **7**(10): 1850-1862.
Deinhofer, K., H. Sevcik, N. Balic, C. Harwanegg, R. Hiller, H. Rumpold, M. W. Mueller and S. Spitzauer (2004). "Microarrayed allergens for IgE profiling." Methods **32**(3): 249-254.
Dexlin-Mellby, L., A. Sandstrom, L. Antberg, J. Gunnarsson, S. R. Hansson, C. A. K. Borrebaeck and C. Wingren (2011). "Design of recombinant antibody microarrays for membrane protein profiling of cell lysates and tissue extracts." Proteomics **11**(8): 1550-1554.

Dexlin, L., J. Ingvarsson, B. Frendeus, C. A. K. Borrebaeck and C. Wingren (2008). "Design of recombinant antibody microarrays for cell surface membrane proteomics." J Proteome Res **7**(1): 319-327.
Downes, M. R., J. C. Byrne, S. R. Pennington, M. J. Dunn, J. M. Fitzpatrick and R. W. Watson (2007). "Urinary markers for prostate cancer." BJU Int **99**(2): 263-268.
Ekins, R. P. (1998). "Ligand assays: from electrophoresis to miniaturized microarrays." Clin Chem **44**(9): 2015-2030.
Ellington, A. D. and J. W. Szostak (1990). "In vitro selection of RNA molecules that bind specific ligands." Nature **346**(6287): 818-822.
Ellmark, P., J. Ingvarsson, A. Carlsson, B. S. Lundin, C. Wingren and C. A. K. Borrebaeck (2006). "Identification of protein expression signatures associated with Helicobacter pylori infection and gastric adenocarcinoma using recombinant antibody microarrays." Mol Cell Proteomics **5**(9): 1638-1646.
Euser, A. M., C. Zoccali, K. J. Jager and F. W. Dekker (2009). "Cohort studies: prospective versus retrospective." Nephron Clin Pract **113**(3): c214-217.
Ewert, S., A. Honegger and A. Pluckthun (2004). "Stability improvement of antibodies for extracellular and intracellular applications: CDR grafting to stable frameworks and structure-based framework engineering." Methods **34**(2): 184-199.
Ewert, S., T. Huber, A. Honegger and A. Pluckthun (2003). "Biophysical properties of human antibody variable domains." J Mol Biol **325**(3): 531-553.
Fattal, I., N. Shental, D. Mevorach, J. M. Anaya, A. Livneh, P. Langevitz, G. Zandman-Goddard, R. Pauzner, M. Lerner, M. Blank, M. E. Hincapie, U. Gafter, Y. Naparstek, Y. Shoenfeld, E. Domany and I. R. Cohen (2010). "An antibody profile of systemic lupus erythematosus detected by antigen microarray." Immunology **130**(3): 337-343.
Ferlay, J., H. R. Shin, F. Bray, D. Forman, C. Mathers and D. M. Parkin (2010). "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008." Int J Cancer **127**(12): 2893-2917.
Gibson, D. S., J. Banha, D. Penque, L. Costa, T. P. Conrads, D. J. Cahill, J. K. O'Brien and M. E. Rooney (2010). "Diagnostic and prognostic biomarker discovery strategies for autoimmune disorders." J Proteomics **73**(6): 1045-1060.
Gigante, A., M. L. Gasperini, A. Afeltra, B. Barbano, D. Margiotta, R. Cianci, I. De Francesco and A. Amoroso (2011). "Cytokines expression in SLE nephritis." Eur Rev Med Pharmacol Sci **15**(1): 15-24.
Gladman, D. D., D. Ibanez and M. B. Urowitz (2002). "Systemic lupus erythematosus disease activity index 2000." J Rheumatol **29**(2): 288-291.
Gleason, D. F. and G. T. Mellinger (1974). "Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging." J Urol **111**(1): 58-64.
Goonetilleke, K. S. and A. K. Siriwardena (2007). "Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer." Eur J Surg Oncol **33**(3): 266-270.
Grantzdorffer, I., S. Yumlu, Z. Gioeva, R. von Wasielewski, M. P. Ebert and C. Rocken (2010). "Comparison of different tissue sampling methods for protein extraction from formalin-fixed and paraffin-embedded tissue specimens." Exp Mol Pathol **88**(1): 190-196.

Greive, K. A., N. D. Balazs and W. D. Comper (2001). "Protein fragments in urine have been considerably underestimated by various protein assays." Clin Chem **47**(9): 1717-1719.
Griffiths, A. D., M. Malmqvist, J. D. Marks, J. M. Bye, M. J. Embleton, J. McCafferty, M. Baier, K. P. Holliger, B. D. Gorick, N. C. Hughes-Jones and et al. (1993). "Human anti-self antibodies with high specificity from phage display libraries." EMBO J **12**(2): 725-734.
Guh, J. Y. (2010). "Proteinuria versus albuminuria in chronic kidney disease." Nephrology (Carlton) **15 Suppl 2**: 53-56.
Gygi, S. P., Y. Rochon, B. R. Franza and R. Aebersold (1999). "Correlation between protein and mRNA abundance in yeast." Mol Cell Biol **19**(3): 1720-1730.
Haab, B. B., M. J. Dunham and P. O. Brown (2001). "Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions." Genome Biol **2**(2): RESEARCH0004.
Haab, B. B., B. H. Geierstanger, G. Michailidis, F. Vitzthum, S. Forrester, R. Okon, P. Saviranta, A. Brinker, M. Sorette, L. Perlee, S. Suresh, G. Drwal, J. N. Adkins and G. S. Omenn (2005). "Immunoassay and antibody microarray analysis of the HUPO Plasma Proteome Project reference specimens: systematic variation between sample types and calibration of mass spectrometry data." Proteomics **5**(13): 3278-3291.
Hanash, S. (2003). "Disease proteomics." Nature **422**(6928): 226-232.
Hanes, J., C. Schaffitzel, A. Knappik and A. Pluckthun (2000). "Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display." Nat Biotechnol **18**(12): 1287-1292.
He, M., O. Stoevesandt, E. A. Palmer, F. Khan, O. Ericsson and M. J. Taussig (2008). "Printing protein arrays from DNA arrays." Nat Methods **5**(2): 175-177.
He, M. and M. J. Taussig (2001). "Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method)." Nucleic Acids Res **29**(15): E73-73.
Hewitt, R. E. (2011). "Biobanking: the foundation of personalized medicine." Curr Opin Oncol **23**(1): 112-119.
Hoeppe, S., T. D. Schreiber, H. Planatscher, A. Zell, M. F. Templin, D. Stoll, T. O. Joos and O. Poetz (2011). "Targeting peptide termini, a novel immunoaffinity approach to reduce complexity in mass spectrometric protein identification." Mol Cell Proteomics **10**(2): M110 002857.
Honegger, A. and A. Pluckthun (2001). "The influence of the buried glutamine or glutamate residue in position 6 on the structure of immunoglobulin variable domains." J Mol Biol **309**(3): 687-699.
Hoogenboom, H. R. and G. Winter (1992). "By-passing immunisation. Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro." J Mol Biol **227**(2): 381-388.
Hu, S., J. A. Loo and D. T. Wong (2006). "Human body fluid proteome analysis." Proteomics **6**(23): 6326-6353.
Hudelist, G., M. Pacher-Zavisin, C. F. Singer, T. Holper, E. Kubista, M. Schreiber, M. Manavi, M. Bilban and K. Czerwenka (2004). "Use of high-throughput protein array for profiling of differentially expressed proteins in normal and malignant breast tissue." Breast Cancer Res Treat **86**(3): 281-291.

Ingvarsson, J., A. Larsson, A. G. Sjoholm, L. Truedsson, B. Jansson, C. A. K. Borrebaeck and C. Wingren (2007). "Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins." J Proteome Res **6**(9): 3527-3536.

Ingvarsson, J., M. Lindstedt, C. A. K. Borrebaeck and C. Wingren (2006). "One-step fractionation of complex proteomes enables detection of low abundant analytes using antibody-based microarrays." J Proteome Res **5**(1): 170-176.

Jager, M. and A. Pluckthun (1999). "Domain interactions in antibody Fv and scFv fragments: effects on unfolding kinetics and equilibria." FEBS Lett **462**(3): 307-312.

Jamaspishvili, T., M. Kral, I. Khomeriki, V. Student, Z. Kolar and J. Bouchal (2010). "Urine markers in monitoring for prostate cancer." Prostate Cancer Prostatic Dis **13**(1): 12-19.

Jansen, F. H., R. H. van Schaik, J. Kurstjens, W. Horninger, H. Klocker, J. Bektic, M. F. Wildhagen, M. J. Roobol, C. H. Bangma and G. Bartsch (2010). "Prostate-specific antigen (PSA) isoform p2PSA in combination with total PSA and free PSA improves diagnostic accuracy in prostate cancer detection." Eur Urol **57**(6): 921-927.

Jirholt, P., M. Ohlin, C. A. K. Borrebaeck and E. Soderlind (1998). "Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework." Gene **215**(2): 471-476.

Joske, D. J. (2008). "Chronic myeloid leukaemia: the evolution of gene-targeted therapy." Med J Aust **189**(5): 277-282.

Jung, S., A. Honegger and A. Pluckthun (1999). "Selection for improved protein stability by phage display." J Mol Biol **294**(1): 163-180.

Jung, S., S. Spinelli, B. Schimmele, A. Honegger, L. Pugliese, C. Cambillau and A. Pluckthun (2001). "The importance of framework residues H6, H7 and H10 in antibody heavy chains: experimental evidence for a new structural subclassification of antibody V(H) domains." J Mol Biol **309**(3): 701-716.

Kasman, L. M., A. A. Lukowiak, S. F. Garczynski, R. J. McNall, P. Youngman and M. J. Adang (1998). "Phage display of a biologically active Bacillus thuringiensis toxin." Appl Environ Microbiol **64**(8): 2995-3003.

Katz, M. S., J. A. Efstathiou, A. V. D'Amico, M. W. Kattan, M. G. Sanda, P. L. Nguyen, M. R. Smith, P. R. Carroll and A. L. Zietman (2010). "The 'CaP Calculator': an online decision support tool for clinically localized prostate cancer." BJU Int **105**(10): 1417-1422.

Kingsmore, S. F. (2006). "Multiplexed protein measurement: technologies and applications of protein and antibody arrays." Nat Rev Drug Discov **5**(4): 310-320.

Kipriyanov, S. M., G. Moldenhauer, A. C. Martin, O. A. Kupriyanova and M. Little (1997). "Two amino acid mutations in an anti-human CD3 single chain Fv antibody fragment that affect the yield on bacterial secretion but not the affinity." Protein Eng **10**(4): 445-453.

Kirou, K. A., C. Lee, S. George, K. Louca, M. G. Peterson and M. K. Crow (2005). "Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease." Arthritis Rheum **52**(5): 1491-1503.

Knappik, A., L. Ge, A. Honegger, P. Pack, M. Fischer, G. Wellnhofer, A. Hoess, J. Wolle, A. Pluckthun and B. Virnekas (2000). "Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides." J Mol Biol **296**(1): 57-86.

Kopf, E., D. Shnitzer and D. Zharhary (2005). "Panorama Ab Microarray Cell Signaling kit: a unique tool for protein expression analysis." Proteomics **5**(9): 2412-2416.

Kopf, E. and D. Zharhary (2007). "Antibody arrays--an emerging tool in cancer proteomics." Int J Biochem Cell Biol **39**(7-8): 1305-1317.

Koprowski, H., Z. Steplewski, K. Mitchell, M. Herlyn, D. Herlyn and P. Fuhrer (1979). "Colorectal carcinoma antigens detected by hybridoma antibodies." Somatic Cell Genet **5**(6): 957-971.

Krauss, J., M. A. Arndt, Z. Zhu, D. L. Newton, B. K. Vu, V. Choudhry, R. Darbha, X. Ji, N. S. Courtenay-Luck, M. P. Deonarain, J. Richards and S. M. Rybak (2004). "Impact of antibody framework residue VH-71 on the stability of a humanised anti-MUC1 scFv and derived immunoenzyme." Br J Cancer **90**(9): 1863-1870.

LaBaer, J. (2012). "Improving international research with clinical specimens: 5 achievable objectives." J Proteome Res **11**(12): 5592-5601.

Langedijk, A. C., A. Honegger, J. Maat, R. J. Planta, R. C. van Schaik and A. Pluckthun (1998). "The nature of antibody heavy chain residue H6 strongly influences the stability of a VH domain lacking the disulfide bridge." J Mol Biol **283**(1): 95-110.

Lee, C. V., W. C. Liang, M. S. Dennis, C. Eigenbrot, S. S. Sidhu and G. Fuh (2004). "High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold." J Mol Biol **340**(5): 1073-1093.

Lescuyer, P., D. Hochstrasser and T. Rabilloud (2007). "How shall we use the proteomics toolbox for biomarker discovery?" J Proteome Res **6**(9): 3371-3376.

Li, Y., M. Tucci, S. Narain, E. V. Barnes, E. S. Sobel, M. S. Segal and H. B. Richards (2006). "Urinary biomarkers in lupus nephritis." Autoimmun Rev **5**(6): 383-388.

Liang, S. L. and D. W. Chan (2007). "Enzymes and related proteins as cancer biomarkers: a proteomic approach." Clin Chim Acta **381**(1): 93-97.

Lilja, H., A. Christensson, U. Dahlen, M. T. Matikainen, O. Nilsson, K. Pettersson and T. Lovgren (1991). "Prostate-specific antigen in serum occurs predominantly in complex with alpha 1-antichymotrypsin." Clin Chem **37**(9): 1618-1625.

Liu, B. C., L. Zhang, L. L. Lv, Y. L. Wang, D. G. Liu and X. L. Zhang (2006). "Application of antibody array technology in the analysis of urinary cytokine profiles in patients with chronic kidney disease." Am J Nephrol **26**(5): 483-490.

Lizardi, P. M., X. Huang, Z. Zhu, P. Bray-Ward, D. C. Thomas and D. C. Ward (1998). "Mutation detection and single-molecule counting using isothermal rolling-circle amplification." Nat Genet **19**(3): 225-232.

Loeb, S., L. J. Sokoll, D. L. Broyles, C. H. Bangma, R. H. van Schaik, G. G. Klee, J. T. Wei, M. G. Sanda, A. W. Partin, K. M. Slawin, L. S. Marks, I. A. Mizrahi, S. S. Shin, A. B. Cruz, D. W. Chan, W. L. Roberts and W. J. Catalona (2012). "Prospective Multicenter Evaluation of the Beckman Coulter Prostate Health Index Using WHO Calibration." J Urol.

Lueking, A., M. Horn, H. Eickhoff, K. Bussow, H. Lehrach and G. Walter (1999). "Protein microarrays for gene expression and antibody screening." Anal Biochem **270**(1): 103-111.

MacBeath, G. (2002). "Protein microarrays and proteomics." Nat Genet **32 Suppl**: 526-532.
MacBeath, G. and S. L. Schreiber (2000). "Printing proteins as microarrays for high-throughput function determination." Science **289**(5485): 1760-1763.
Maecker, H. T., T. M. Lindstrom, W. H. Robinson, P. J. Utz, M. Hale, S. D. Boyd, S. S. Shen-Orr and C. G. Fathman (2012). "New tools for classification and monitoring of autoimmune diseases." Nat Rev Rheumatol **8**(6): 317-328.
Maidhof, W. and O. Hilas (2012). "Lupus: an overview of the disease and management options." P T **37**(4): 240-249.
Mair, J., E. Artner-Dworzak, P. Lechleitner, J. Smidt, I. Wagner, F. Dienstl and B. Puschendorf (1991). "Cardiac troponin T in diagnosis of acute myocardial infarction." Clin Chem **37**(6): 845-852.
Mann, C. J. (2003). "Observational research methods. Research design II: cohort, cross sectional, and case-control studies." Emerg Med J **20**(1): 54-60.
Marko-Varga, G., A. Vegvari, C. Welinder, H. Lindberg, M. Rezeli, G. Edula, K. J. Svensson, M. Belting, T. Laurell and T. E. Fehniger (2012). "Standardization and utilization of biobank resources in clinical protein science with examples of emerging applications." J Proteome Res **11**(11): 5124-5134.
McMenamin, M. E., P. Soung, S. Perera, I. Kaplan, M. Loda and W. R. Sellers (1999). "Loss of PTEN expression in paraffin-embedded primary prostate cancer correlates with high Gleason score and advanced stage." Cancer Res **59**(17): 4291-4296.
McWilliam, I., M. Chong Kwan and D. Hall (2011). "Inkjet printing for the production of protein microarrays." Methods Mol Biol **785**: 345-361.
Meany, D. L., L. Hackler, Jr., H. Zhang and D. W. Chan (2011). "Tyramide signal amplification for antibody-overlay lectin microarray: a strategy to improve the sensitivity of targeted glycan profiling." J Proteome Res **10**(3): 1425-1431.
Mehta, R., R. K. Jain and S. Badve (2011). "Personalized medicine: the road ahead." Clin Breast Cancer **11**(1): 20-26.
Mikolajczyk, S. D., L. S. Millar, T. J. Wang, H. G. Rittenhouse, L. S. Marks, W. Song, T. M. Wheeler and K. M. Slawin (2000). "A precursor form of prostate-specific antigen is more highly elevated in prostate cancer compared with benign transition zone prostate tissue." Cancer Res **60**(3): 756-759.
Miller, J. C., H. Zhou, J. Kwekel, R. Cavallo, J. Burke, E. B. Butler, B. S. Teh and B. B. Haab (2003). "Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers." Proteomics **3**(1): 56-63.
Mischak, H., J. P. Ioannidis, A. Argiles, T. K. Attwood, E. Bongcam-Rudloff, M. Broenstrup, A. Charonis, G. P. Chrousos, C. Delles, A. Dominiczak, T. Dylag, J. Ehrich, J. Egido, P. Findeisen, J. Jankowski, R. W. Johnson, B. A. Julien, T. Lankisch, H. Y. Leung, D. Maahs, F. Magni, M. P. Manns, E. Manolis, G. Mayer, G. Navis, J. Novak, A. Ortiz, F. Persson, K. Peter, H. H. Riese, P. Rossing, N. Sattar, G. Spasovski, V. Thongboonkerd, R. Vanholder, J. P. Schanstra and A. Vlahou (2012). "Implementation of proteomic biomarkers: making it work." Eur J Clin Invest **42**(9): 1027-1036.
Mitchell, P. (2002). "A perspective on protein microarrays." Nat Biotechnol **20**(3): 225-229.

Nord, K., E. Gunneriusson, J. Ringdahl, S. Stahl, M. Uhlen and P. A. Nygren (1997). "Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain." Nat Biotechnol **15**(8): 772-777.
Olsson, N., C. Wingren, M. Mattsson, P. James, D. O'Connell, F. Nilsson, D. J. Cahill and C. A. K. Borrebaeck (2011). "Proteomic analysis and discovery using affinity proteomics and mass spectrometry." Mol Cell Proteomics **10**(10): M110 003962.
Osicka, T. M., S. Panagiotopoulos, G. Jerums and W. D. Comper (1997). "Fractional clearance of albumin is influenced by its degradation during renal passage." Clin Sci (Lond) **93**(6): 557-564.
Paez, J. G., P. A. Janne, J. C. Lee, S. Tracy, H. Greulich, S. Gabriel, P. Herman, F. J. Kaye, N. Lindeman, T. J. Boggon, K. Naoki, H. Sasaki, Y. Fujii, M. J. Eck, W. R. Sellers, B. E. Johnson and M. Meyerson (2004). "EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy." Science **304**(5676): 1497-1500.
Paik, S., S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F. L. Baehner, M. G. Walker, D. Watson, T. Park, W. Hiller, E. R. Fisher, D. L. Wickerham, J. Bryant and N. Wolmark (2004). "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer." N Engl J Med **351**(27): 2817-2826.
Papp, K., P. Vegh, R. Hobor, Z. Szittner, Z. Voko, J. Podani, L. Czirjak and J. Prechl (2012). "Immune complex signatures of patients with active and inactive SLE revealed by multiplex protein binding analysis on antigen microarrays." PLoS One **7**(9): e44824.
Parekh, D. J., D. P. Ankerst, D. Troyer, S. Srivastava and I. M. Thompson (2007). "Biomarkers for prostate cancer detection." J Urol **178**(6): 2252-2259.
Pascual, V., L. Farkas and J. Banchereau (2006). "Systemic lupus erythematosus: all roads lead to type I interferons." Curr Opin Immunol **18**(6): 676-682.
Paul, W. E. (2013). Fundamental immunology. Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.
Pawlak, M., E. Schick, M. A. Bopp, M. J. Schneider, P. Oroszlan and M. Ehrat (2002). "Zeptosens' protein microarrays: a novel high performance microarray platform for low abundance protein analysis." Proteomics **2**(4): 383-393.
Pavlickova, P., E. M. Schneider and H. Hug (2004). "Advances in recombinant antibody microarrays." Clin Chim Acta **343**(1-2): 17-35.
Pedersen, J. W., O. Blixt, E. P. Bennett, M. A. Tarp, I. Dar, U. Mandel, S. S. Poulsen, A. E. Pedersen, S. Rasmussen, P. Jess, H. Clausen and H. H. Wandall (2011). "Seromic profiling of colorectal cancer patients with novel glycopeptide microarray." Int J Cancer **128**(8): 1860-1871.
Perlee, L., J. Christiansen, R. Dondero, B. Grimwade, S. Lejnine, M. Mullenix, W. Shao, M. Sorette, V. Tchernev, D. Patel and S. Kingsmore (2004). "Development and standardization of multiplexed antibody microarrays for use in quantitative proteomics." Proteome Sci **2**(1): 9.
Phizicky, E., P. I. Bastiaens, H. Zhu, M. Snyder and S. Fields (2003). "Protein analysis on a proteomic scale." Nature **422**(6928): 208-215.
Poetz, O., R. Ostendorp, B. Brocks, J. M. Schwenk, D. Stoll, T. O. Joos and M. F. Templin (2005). "Protein microarrays for antibody profiling: specificity and affinity determination on a chip." Proteomics **5**(9): 2402-2411.

Puntmann, V. O. (2009). "How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease." Postgrad Med J **85**(1008): 538-545.
Ramachandran, N., E. Hainsworth, B. Bhullar, S. Eisenstein, B. Rosen, A. Y. Lau, J. C. Walter and J. LaBaer (2004). "Self-assembling protein microarrays." Science **305**(5680): 86-90.
Ramachandran, N., S. Srivastava and J. Labaer (2008). "Applications of protein microarrays for biomarker discovery." Proteomics Clin Appl **2**(10-11): 1444-1459.
Rees, R., S. Laversin, C. Murray and G. Ball (2012). Current Approaches to Identify and Evaluate Cancer Biomarkers for Patient Stratification. Vaccinology, Wiley-Blackwell**:** 452-463.
Ridker, P. M. (2009). "C-reactive protein: eighty years from discovery to emergence as a major risk marker for cardiovascular disease." Clin Chem **55**(2): 209-215.
Rifai, N., M. A. Gillette and S. A. Carr (2006). "Protein biomarker discovery and validation: the long and uncertain path to clinical utility." Nat Biotechnol **24**(8): 971-983.
Rodriguez-Rodriguez, E. R., L. M. Ledezma-Candanoza, L. G. Contreras-Ferrat, T. Olamendi-Portugal, L. D. Possani, B. Becerril and L. Riano-Umbarila (2012). "A single mutation in framework 2 of the heavy variable domain improves the properties of a diabody and a related single-chain antibody." J Mol Biol **423**(3): 337-350.
Ross, J. S. (2011). "Cancer biomarkers, companion diagnostics and personalized oncology." Biomark Med **5**(3): 277-279.
Ross, J. S., E. A. Slodkowska, W. F. Symmans, L. Pusztai, P. M. Ravdin and G. N. Hortobagyi (2009). "The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine." Oncologist **14**(4): 320-368.
Rovin, B. H., D. J. Birmingham, H. N. Nagaraja, C. Y. Yu and L. A. Hebert (2007). "Biomarker discovery in human SLE nephritis." Bull NYU Hosp Jt Dis **65**(3): 187-193.
Rovin, B. H., A. M. McKinley and D. J. Birmingham (2009). "Can we personalize treatment for kidney diseases?" Clin J Am Soc Nephrol **4**(10): 1670-1676.
Rovin, B. H. and X. Zhang (2009). "Biomarkers for lupus nephritis: the quest continues." Clin J Am Soc Nephrol **4**(11): 1858-1865.
Rowley, J. D. (1973). "Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining." Nature **243**(5405): 290-293.
Sakr, W. A., G. P. Haas, B. F. Cassin, J. E. Pontes and J. D. Crissman (1993). "The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients." J Urol **150**(2 Pt 1): 379-385.
Sanchez-Carbayo, M. (2011). "Antibody microarrays as tools for biomarker discovery." Methods Mol Biol **785**: 159-182.
Sandstrom, A., R. Andersson, R. Segersvard, M. Lohr, C. A. K. Borrebaeck and C. Wingren (2012). "Serum proteome profiling of pancreatitis using recombinant antibody microarrays reveals disease-associated biomarker signatures." Proteomics Clin Appl **6**(9-10): 486-496.

Sauer, U. (2011). "Impact of substrates for probe immobilization." Methods Mol Biol **785**: 363-378.
Saul, F. A. and R. J. Poljak (1993). "Structural patterns at residue positions 9, 18, 67 and 82 in the VH framework regions of human and murine immunoglobulins." J Mol Biol **230**(1): 15-20.
Schechter, A. L., D. F. Stern, L. Vaidyanathan, S. J. Decker, J. A. Drebin, M. I. Greene and R. A. Weinberg (1984). "The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen." Nature **312**(5994): 513-516.
Schroder, C., A. Jacob, S. Tonack, T. P. Radon, M. Sill, M. Zucknick, S. Ruffer, E. Costello, J. P. Neoptolemos, T. Crnogorac-Jurcevic, A. Bauer, K. Fellenberg and J. D. Hoheisel (2010). "Dual-color proteomic profiling of complex samples with a microarray of 810 cancer-related antibodies." Mol Cell Proteomics **9**(6): 1271-1280.
Schroder, F. H., J. Hugosson, M. J. Roobol, T. L. Tammela, S. Ciatto, V. Nelen, M. Kwiatkowski, M. Lujan, H. Lilja, M. Zappa, L. J. Denis, F. Recker, A. Berenguer, L. Maattanen, C. H. Bangma, G. Aus, A. Villers, X. Rebillard, T. van der Kwast, B. G. Blijenberg, S. M. Moss, H. J. de Koning, A. Auvinen and E. Investigators (2009). "Screening and prostate-cancer mortality in a randomized European study." N Engl J Med **360**(13): 1320-1328.
Schrohl, A. S., S. Wurtz, E. Kohn, R. E. Banks, H. J. Nielsen, F. C. Sweep and N. Brunner (2008). "Banking of biological fluids for studies of disease-associated protein biomarkers." Mol Cell Proteomics **7**(10): 2061-2066.
Schweitzer, B., S. Roberts, B. Grimwade, W. Shao, M. Wang, Q. Fu, Q. Shu, I. Laroche, Z. Zhou, V. T. Tchernev, J. Christiansen, M. Velleca and S. F. Kingsmore (2002). "Multiplexed protein profiling on microarrays by rolling-circle amplification." Nat Biotechnol **20**(4): 359-365.
Schwenk, J. M. and P. Nilsson (2011). "Antibody suspension bead arrays." Methods Mol Biol **723**: 29-36.
Seurynck-Servoss, S. L., C. L. Baird, K. D. Miller, N. B. Pefaur, R. M. Gonzalez, D. O. Apiyo, H. E. Engelmann, S. Srivastava, J. Kagan, K. D. Rodland and R. C. Zangar (2008). "Immobilization strategies for single-chain antibody microarrays." Proteomics **8**(11): 2199-2210.
Shafer, M. W., L. Mangold, A. W. Partin and B. B. Haab (2007). "Antibody array profiling reveals serum TSP-1 as a marker to distinguish benign from malignant prostatic disease." Prostate **67**(3): 255-267.
Shariat, S. F., A. Semjonow, H. Lilja, C. Savage, A. J. Vickers and A. Bjartell (2011). "Tumor markers in prostate cancer I: blood-based markers." Acta Oncol **50 Suppl 1**: 61-75.
Soderlind, E., L. Strandberg, P. Jirholt, N. Kobayashi, V. Alexeiva, A. M. Aberg, A. Nilsson, B. Jansson, M. Ohlin, C. Wingren, L. Danielsson, R. Carlsson and C. A. K. Borrebaeck (2000). "Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries." Nat Biotechnol **18**(8): 852-856.
Sorlie, T., C. M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, T. Thorsen, H. Quist, J. C. Matese, P. O. Brown, D. Botstein, P. E. Lonning and A. L. Borresen-Dale (2001). "Gene

expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications." Proc Natl Acad Sci U S A **98**(19): 10869-10874.
Spurrier, B., S. Ramalingam and S. Nishizuka (2008). "Reverse-phase protein lysate microarrays for cell signaling analysis." Nat Protoc **3**(11): 1796-1808.
Stein, C. J. and G. A. Colditz (2004). "Modifiable risk factors for cancer." Br J Cancer **90**(2): 299-303.
Steinhauer, C., C. Wingren, A. C. Hager and C. A. K. Borrebaeck (2002). "Single framework recombinant antibody fragments designed for protein chip applications." Biotechniques **Suppl**: 38-45.
Stoevesandt, O. and M. J. Taussig (2007). "Affinity reagent resources for human proteome detection: initiatives and perspectives." Proteomics **7**(16): 2738-2750.
Stoevesandt, O. and M. J. Taussig (2012). "Affinity proteomics: the role of specific binding reagents in human proteome analysis." Expert Rev Proteomics **9**(4): 401-414.
Stoevesandt, O. and M. J. Taussig (2012). "European and international collaboration in affinity proteomics." N Biotechnol **29**(5): 511-514.
Telli, M. L., S. A. Hunt, R. W. Carlson and A. E. Guardino (2007). "Trastuzumab-related cardiotoxicity: calling into question the concept of reversibility." J Clin Oncol **25**(23): 3525-3533.
Thadikkaran, L., M. A. Siegenthaler, D. Crettaz, P. A. Queloz, P. Schneider and J. D. Tissot (2005). "Recent advances in blood-related proteomics." Proteomics **5**(12): 3019-3034.
Thavasu, P. W., S. Longhurst, S. P. Joel, M. L. Slevin and F. R. Balkwill (1992). "Measuring cytokine levels in blood. Importance of anticoagulants, processing, and storage conditions." J Immunol Methods **153**(1-2): 115-124.
Thongboonkerd, V. (2007). "Practical points in urinary proteomics." J Proteome Res **6**(10): 3881-3890.
Thongboonkerd, V. and P. Malasit (2005). "Renal and urinary proteomics: current applications and challenges." Proteomics **5**(4): 1033-1042.
Tian, J., P. Wang, S. Gao, X. Chu, N. Wu and Y. Fan (2010). "Enhanced thermostability of methyl parathion hydrolase from Ochrobactrum sp. M231 by rational engineering of a glycine to proline mutation." FEBS J **277**(23): 4901-4908.
Tillett, W. S. and T. Francis (1930). "Serological Reactions in Pneumonia with a Non-Protein Somatic Fraction of Pneumococcus." J Exp Med **52**(4): 561-571.
Tsokos, G. C. (2011). "Systemic lupus erythematosus." N Engl J Med **365**(22): 2110-2121.
Tuck, M. K., D. W. Chan, D. Chia, A. K. Godwin, W. E. Grizzle, K. E. Krueger, W. Rom, M. Sanda, L. Sorbara, S. Stass, W. Wang and D. E. Brenner (2009). "Standard operating procedures for serum and plasma collection: early detection research network consensus statement standard operating procedure integration working group." J Proteome Res **8**(1): 113-117.
Tuerk, C. and L. Gold (1990). "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science **249**(4968): 505-510.
Uhlen, M., P. Oksvold, L. Fagerberg, E. Lundberg, K. Jonasson, M. Forsberg, M. Zwahlen, C. Kampf, K. Wester, S. Hober, H. Wernerus, L. Bjorling and F. Ponten

(2010). "Towards a knowledge-based Human Protein Atlas." Nat Biotechnol **28**(12): 1248-1250.
Urbanowska, T., S. Mangialaio, C. Zickler, S. Cheevapruk, P. Hasler, S. Regenass and F. Legay (2006). "Protein microarray platform for the multiplex analysis of biomarkers in human sera." J Immunol Methods **316**(1-2): 1-7.
Wacker, R., H. Schroder and C. M. Niemeyer (2004). "Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavidin-biotin attachment: a comparative study." Anal Biochem **330**(2): 281-287.
Wallstrom, G., K. S. Anderson and J. Labaer (2013). "Biomarker Discovery for Heterogeneous Diseases." Cancer Epidemiol Biomarkers Prev.
van 't Veer, L. J., H. Dai, M. J. van de Vijver, Y. D. He, A. A. Hart, M. Mao, H. L. Peterse, K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards and S. H. Friend (2002). "Gene expression profiling predicts clinical outcome of breast cancer." Nature **415**(6871): 530-536.
Watanabe, K., T. Masuda, H. Ohashi, H. Mihara and Y. Suzuki (1994). "Multiple proline substitutions cumulatively thermostabilize Bacillus cereus ATCC7064 oligo-1,6-glucosidase. Irrefragable proof supporting the proline rule." Eur J Biochem **226**(2): 277-283.
Webster, D. M., A. H. Henry and A. R. Rees (1994). "Antibody-antigen interactions." Current opinion of structural biology **4**: 123-129.
Welch, H. G. and P. C. Albertsen (2009). "Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005." J Natl Cancer Inst **101**(19): 1325-1329.
Willuda, J., A. Honegger, R. Waibel, P. A. Schubiger, R. Stahel, U. Zangemeister-Wittke and A. Pluckthun (1999). "High thermal stability is essential for tumor targeting of antibody fragments: engineering of a humanized anti-epithelial glycoprotein-2 (epithelial cell adhesion molecule) single-chain Fv fragment." Cancer Res **59**(22): 5758-5767.
Wingren, C., J. Ingvarsson, L. Dexlin, D. Szul and C. A. K. Borrebaeck (2007). "Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support." Proteomics **7**(17): 3055-3065.
Wingren, C., A. Sandstrom, R. Segersvard, A. Carlsson, R. Andersson, M. Lohr and C. A. K. Borrebaeck (2012). "Identification of serum biomarker signatures associated with pancreatic cancer." Cancer Res **72**(10): 2481-2490.
von Rosen, A., S. Linder, U. Harmenberg and S. Pegert (1993). "Serum levels of CA 19-9 and CA 50 in relation to Lewis blood cell status in patients with malignant and benign pancreatic disease." Pancreas **8**(2): 160-165.
von Schantz, L., F. Gullfot, S. Scheer, L. Filonova, L. Cicortas Gunnarsson, J. E. Flint, G. Daniel, E. Nordberg-Karlsson, H. Brumer and M. Ohlin (2009). "Affinity maturation generates greatly improved xyloglucan-specific carbohydrate binding modules." BMC Biotechnol **9**: 92.
Worn, A. and A. Pluckthun (1999). "Different equilibrium stability behavior of ScFv fragments: identification, classification, and improvement by protein engineering." Biochemistry **38**(27): 8739-8750.

Worn, A. and A. Pluckthun (2001). "Stability engineering of antibody single-chain Fv fragments." J Mol Biol **305**(5): 989-1010.
Voss, J., Y. A. Goo, K. Cain, N. Woods, M. Jarrett, L. Smith, R. Shulman and M. Heitkemper (2011). "Searching for the noninvasive biomarker holy grail: are urine proteomics the answer?" Biol Res Nurs **13**(3): 235-242.
Zhu, H., J. F. Klemic, S. Chang, P. Bertone, A. Casamayor, K. G. Klemic, D. Smith, M. Gerstein, M. A. Reed and M. Snyder (2000). "Analysis of yeast protein kinases using protein chips." Nat Genet **26**(3): 283-289.