



**ROYAL INSTITUTE
OF TECHNOLOGY**

EPITOPE MAPPING OF ANTIBODIES TOWARDS HUMAN PROTEIN TARGETS

Barbara Hjelm

Royal Institute of Technology
School of Biotechnology
Stockholm 2011

© Barbara Hjelm
Stockholm 2011
Royal Institute of Technology
School of Biotechnology
AlbaNova University Center
SE-106 91 Stockholm
Sweden

Printed by Universitetsservice US-AB
Drottning Kristinas väg 53B
SE- 100 44 Stockholm
Sweden

ISBN 978-91-7501-222-3
TRITA BIO-Report 2012:1
ISSN 1654-2312

Cover illustration by Helena Malm: «Antikroppar ända in i kaklet»

«One day you're in, and the next day, you're out»

Heidi Klum

Abstract

This thesis, based on five research papers, presents results from development and evaluation of methods for identifying the interaction site of antibodies on their antigens and the functional investigation of these in different assays. As antibodies have proven to be invaluable tools in diagnostics, therapy and basic research, the demand of characterizing these binding molecules has increased. Techniques for epitope mapping in a streamlined manner are therefore needed, particularly in high throughput projects as the Human Protein Atlas that aims to systematically generate two antibodies with separate epitopes towards all human proteins.

In paper I we describe an approach to map the epitopes of polyclonal and monoclonal antibodies for the first time using staphylococcal display. This method was combined with peptide scanning and alanine scanning using suspension bead arrays, to create a streamlined approach of high-resolution characterization of epitopes recognized by antibodies as demonstrated in paper II. Single epitopes were identified for the monoclonal antibodies and several (one to five) separate epitopes scattered throughout the antigen sequence were determined for each polyclonal antibody. Further, antibodies of different species origin showed overlapping binding epitopes. In paper III we studied the epitope patterns of polyclonal antibodies generated with the same antigen in different animals. Although common epitope regions could be identified the exact epitope pattern was not repeated, as some epitopes did not reoccur in the repeated immunizations. In paper IV, a potential biomarker for colon cancer, RBM3, was investigated using validated antibodies by epitope mapping and siRNA analysis. Finally, in paper V, a method for generating epitope-specific antibodies based on affinity purification of a polyclonal antibody is described. The generated antibodies were used in several immunoassays and showed a great difference in functionality. Paired antibodies with separate epitopes were successfully generated and could be used in a sandwich assay or validate each other in immunohistochemistry.

Taken together, in these studies we have demonstrated valuable concepts for the characterization of antibody epitopes.

Keywords: antibody, antibody validation, biomarker, epitope mapping, peptide array, proteomics, RBM3, staphylococcal surface display

List of publications

This thesis is based upon the following five papers, which are referred to in the text by the corresponding Roman numerals (I-V). The papers are included in the Appendix.

- I Rockberg J., Löfblom J., **Hjelm B.**, Ståhl S. and Uhlén M. (2008) Epitope mapping of antibodies using bacterial display. *Nat Methods* 5: 1039-1045

- II **Hjelm B.**, Fernández C.D., Löfblom J., Ståhl S., Johannesson H., Rockberg J. and Uhlén M. (2010) Exploring epitopes of antibodies towards the human tryptophanyl-tRNA synthetase. *N Biotechnol* 27: 129-137

- III **Hjelm B.**, Forsström B., Löfblom J., Rockberg J. and Uhlén M. (2011) Immunizations of inbred rabbits using the same antigen yield antibodies with similar, but not identical, epitopes. *Manuscript*

- IV **Hjelm B.**, Brennan D.J., Zendehrokh N., Eberhard J., Nodin B., Gaber A., Pontén F., Johannesson H., Smaragdi K., Frantz C., Hober S., Johnson L.B., Pählman S., Jirstrom K. and Uhlén M. (2011) High nuclear RBM3 expression is associated with an improved prognosis in colorectal cancer. *Proteomics Clinical Applications In Press*

- V **Hjelm B.**, Forsström B., Igel U., Johannesson H., Stadler C., Lundberg E., Pontén F., Sjöberg A., Rockberg J., Schwenk J.M., Nilsson P., Johansson C. and Uhlén M. (2011) Generation of monospecific antibodies based on affinity capture of polyclonal antibodies. *Protein Science* 20: 1824-1835

All papers are reproduced with permission from the copyright holders

Abbreviations

3D	Three-dimensional
ANLN	Actin-binding protein anillin
ABP	Albumin binding protein
C	Constant
CDR	Complementarity-determining region
CNDPI	Carnosine dipeptidase 1
DNA	Deoxyribonucleic acid
ELISA	Enzyme-linked immunosorbent analysis
Fab	Fragment antigen binding
FACS	Fluorescence activated cell sorting
Fc	Fragment crystallizable
FCA	Freund's complete adjuvant
H	Heavy
HER2	Human epidermal growth factor receptor 2
His ₆ -tag	Hexahistidine-tag
HPA	Human Protein Atlas
HSA	Human serum albumin
IF	Immunofluorescence
IgG	Immunoglobulins
IHC	Immunohistochemistry
IP	Immunoprecipitation
L	Light
mAb	Monoclonal antibody
MS	Mass spectrometry
NMR	Nuclear Magnetic Resonance
pAb	Polyclonal antibody
PCR	Polymerase chain reaction
PrEST	Protein epitope signature tags
PSA	Prostate specific antigen
RBM3	RNA-binding protein 3
RIA	Radioimmunoassay
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
SATB2	Special AT-rich binding protein 2
scFv	Single chain fragment variable
TMA	Tissue microarray
TYMP	Thymidine phosphorylase
V	Variable
WARS	Tryptophanyl-tRNA synthetase

Contents

<i>Abstract</i>	v
<i>List of publications</i>	vii
<i>Abbreviations</i>	ix

INTRODUCTION

<i>Proteins</i>	2
<i>Antibodies</i>	6
Natural polyclonal and monoclonal antibodies.....	7
Recombinant antibodies and antibody fragments.....	8
Leaving nature behind.....	9
<i>Epitopes</i>	12
Epitope mapping methods	13
<i>Proteomics</i>	16
Affinity-based proteomics	17
Biomarker discovery.....	17
Human Protein Atlas	18

PRESENT INVESTIGATION

<i>Aims of the thesis</i>	24
Epitope mapping of antibodies using bacterial display (Paper I)	25
Exploring epitopes of antibodies toward the human tryptophanyl-tRNA synthetase (Paper II)	26
Immunizations of inbred rabbits using the same antigen yields antibodies with similar, but not identical, epitopes (Paper III)	28
High nuclear RBM3 expression is associated with an improved prognosis in colorectal cancer (Paper IV) .	29
Generation of monospecific antibodies based on affinity capture of polyclonal antibodies (Paper V)	29
<i>Concluding remarks and future perspectives</i>	32
<i>Acknowledgements</i>	34
<i>References</i>	38

INTRODUCTION

Proteins

In the beginning was the word, and the word was with Berzelius and the word was protein [1]. Well, kind of. The Swedish godfather of chemistry proposed this name - "protein" - for the core substance that his fellow Dutch chemist Gerardus Johannes Mulder observed and hypothesized as common for animal substances [2]. This perception (clearly somewhat revised) is today a proven fact, as we know that the human body consists of a plethora of different proteins that together form the basis for almost all of its biological functions. Regulation, signal transduction, transportation, protection, structural support, you name it. These macromolecules build up the cellular machinery in all living organism and should make it work every day.

The constituents of most proteins are 20 amino acids. These building blocks exhibit the same structural elements, with a central carbon to which an amino-, carboxyl-, hydrogen and a variable group are bonded. The distinctive group, also called side chain, differs between the amino acids and appoints the chemical and physical properties of the resulting proteins. Two amino acids can form a peptide bond by a condensation reaction [3,4], and consecutive reactions adding one amino acid after another creates a polypeptide chain. It has been agreed upon to describe a polypeptide chain starting from the N-terminal end (containing an amine group) and end at the C-terminus (with a carboxyl group). The order of the amino acids in a protein is encoded in genes by deoxyribonucleic acid (DNA), which is transcribed into messenger ribonucleic acid (mRNA) before it is translated into the polypeptide chain forming the protein.

These polypeptide chains can be folded into the complex structures of proteins on three additional levels (Figure 1). The primary structure consists of the linear sequence of amino acids in the peptide chain. Linus Pauling and colleagues concluded that considering the planar nature of the peptide bond [5,6] the peptide chain could be arranged by two major secondary structures: alpha helices and beta sheets. Hydrogen bonds are the stabilizing forces involved in these "local" structural features. Zooming out further, these secondary structures can fold into a tertiary structure, the three dimensional (3D) structure of the protein, which is important for the function of the protein. The tertiary structure is stabilized by, amongst other interactions: the hydrophobic effect, hydrogen bonds between side chains and occasionally by disulfide bridges.

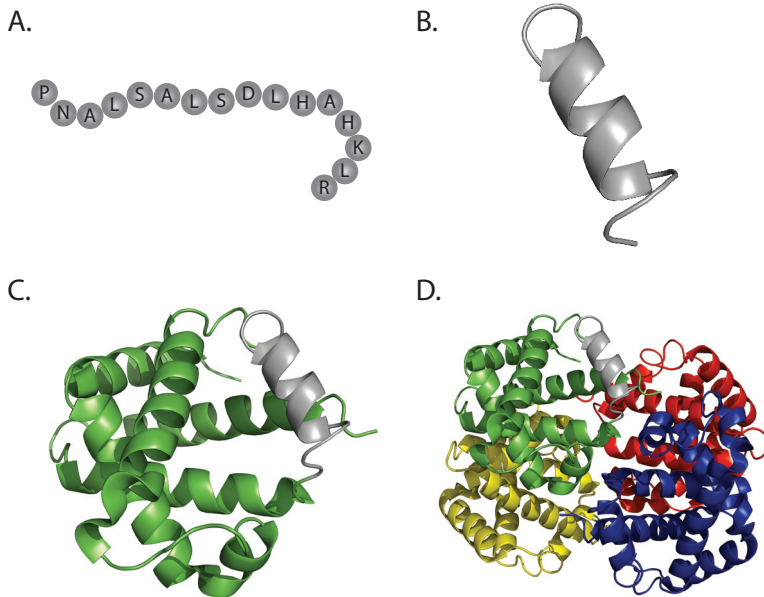


Figure 1. Levels of protein structure, here exemplified by the structure of human hemoglobin (1HHB.pdb). (A) Primary structure is the sequence of a chain of amino acids. (B) Secondary structure occurs when regions of the polypeptide chain are organized by hydrogen bonds into regular structures, here exemplified by an alpha helix. (C) Tertiary structure refers to the three dimensional organization of a folded polypeptide chain. (D) Quarternary structure describes the arrangement of several folded polypeptide chains in a single larger protein.

The hydrophobic effect drives the non-polar side chains to interact inside the protein molecule while folding and the polar or charged side chains are therefore often found on the surface of many proteins, interacting with the aqueous surrounding in cells. Finally, the quaternary structure refers to protein complexes composed by several polypeptide chains. Anfinsen and co workers showed that the primary structure dictates the tertiary structure of the protein [7], on account of the different properties of the amino acids side chains. But even though we can decode the amino acid sequence from our genes, it is extremely difficult to predict the tertiary structure of the protein. Since the first 3D structure of a protein was solved about 50 years ago [8] tens of thousands of proteins have been structurally determined using techniques as i.e. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Today, the Protein Data Bank [9] collects all structural data that is accessible to download for viewing and analysis using computer software programs.

The biological processes carried out by proteins rely on interactions between these through specific molecular recognition. Protein-protein interactions can take place in protein complexes (quaternary structure, see above) to stabilize the structure of the monomeric polypeptides but also between non-obligated protein complexes such as hormone-receptor or antibody-antigen [10]. Detailed knowledge of these interactions is desirable in the area of life science to understand biology and disease, as well as to develop new tools in the field of biotechnology industry. Recombinant DNA technologies emancipated the field when methods like production of recombinant proteins [11] together with site directed mutagenesis [12] for example, enabled such studies by simply modifying proteins and exploring the effects of such changes. The constant development of the protein science toolbox but more importantly the expanding knowledge of proteins as the acting components in our cells, has led to that most modern drug targets are proteins [13] and that protein therapeutics, although many are already in clinical use, have so far only showed a glimpse of the possibilities of future medicine [14].

Antibodies

Immunoglobulins (Ig), more commonly known as antibodies, are proteins that have specialized to function through molecular recognition. They are a part of our immune system where they are produced by B-cells to protect our body from pathogens by recognizing and triggering elimination of the foreign substances called antigens. The ability of antibodies to bind to other molecules has been exploited in diagnostics, therapeutics and basic research.

The quaternary structure of an antibody typically forms a Y-shaped molecule (Figure 2) where the two “arms” and the “stem” display two distinct functions: antigen binding and mediation of effector functions respectively. Each arm is commonly termed as fragment antigen binding (Fab) and as the name indicates, binds to the antigen. The stem is named fragment crystallizable (Fc) and shows no antigen binding properties when isolated [15], however other studies have shown the significance of this part in biological functions as it is bound by Fc receptors on different cells and the complement system. Four polypeptide chains, two identical heavy (H) chains and two identical light (L) chains are the basic units of the antibody held together by disulfide bonds and non-covalent interactions [16,17]. Amino acid sequencing of these immunoglobulin chains revealed that the polypeptide chains consist of a variable (v) region at the N-terminal end and a constant (c) region at the C-terminal end [18,19]. Human antibodies can be divided into five major subclasses or isotypes based on the sequence of the c region of the H chain; IgA, IgD, IgE, IgG, IgM. The most abundant isotype in serum is IgG and henceforth the word antibody in this thesis mainly refers to this subclass.

The molecular recognition of the antigen is performed by the heterogeneous v region at the tip of each arm of the antibody. This bivalency increases the functional affinity through avidity effects [20]. The complementarity-determining regions (CDRs) constitute the binding surface formed by six highly variable loops, three on the H chain and three on the L chain, and define the specificity of the antibody. The great diversity of antibody specificities is created by imprecise combinatorial joining of multiple antibody gene segments and somatic hypermutation during B-cell maturation and activation [21].

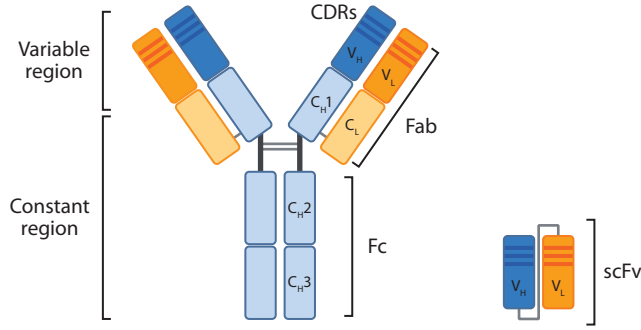


Figure 2. A schematical representation of an IgG-molecule (antibody) and a single chain fragment variable (scFv). Heavy chain illustrated in shades of blue and light chain illustrated in shades of orange. The IgG consists of variable domains (V) and constant domains (C). Several disulphide bonds stabilize the molecule. The scFv antibody derivative consists of the variable domains of the heavy and light chain (V_H , V_L) joined by a peptide linker.

Natural polyclonal and monoclonal antibodies

After appearing inside the body, antigens are recognized by membrane bound immunoglobulins on a B-cell. Following a complicated series of activation signals and cell proliferation, antibodies are secreted into the blood where the B-cells circulate. By collecting the blood, the antibodies can be used in different ways. This phenomenon of immunization was used to produce a serum containing antibodies towards a particular antigen, which was used as a therapeutic agent to treat diphtheria and tetanus as described by von Behring and Kitasato in 1890 [22]. von Behring received the first Nobel Prize in physiology or medicine in 1901 for this work.

The obtained serum from an immunization is polyclonal, as each B-cell recognizes one antigenic determinant, commonly called epitope, on a protein and hence the antibody mixture will cover multiple epitope specificities. When producing polyclonal antibodies the immunizing antigen is often administered together with an adjuvant to enhance the immune response and achieve high titers of antibodies. These substances are used to prolong the exposure of presented antigen and to stimulate cells of the immune system for activation. The most widely used adjuvant for antibody production has been Freund's complete adjuvant (FCA) [23], a mixture of light mineral oil, surfactant agent and heat-killed and dried mycobacterial cells. The choice of adjuvant should be considered as it could affect the antigen protein and whether antibodies towards native or denatured antigens are desired. Some of the first examples

of molecular recognition using polyclonal antibodies are blood typing [24], radioimmunoassay (RIA) for measuring insulin levels [25] and antigen localization in tissue using fluorescein [26].

On the contrary, monoclonal antibodies, all derived from a single B-cell clone, exhibit specificity for a single epitope. Monoclonal antibodies were very important in the characterization of antibodies, multiple myeloma cells producing antibodies of a single specificity were originally used for sequencing and crystallography experiments. The reactivity of these antibodies was however unknown and a major breakthrough was achieved by Köhler and Milstein in 1975 [27] when they succeeded to fuse a B-cell from an immunized mouse with a mouse myeloma cell. The continuous cultures of so called hybridomas give a renewable source of monoclonal antibodies with a desired specificity.

Generating polyclonal antibodies for a selected target is currently more rapid, less expensive and less laborious than generating monoclonal antibodies [28]. Nevertheless, the limited amount of a polyclonal antibody is a pronounced disadvantage. The ability to detect several epitopes of an antigen in the final application makes polyclonal antibodies less sensitive to assay conditions that could change the epitope accessibility. On the other hand, in those applications where you want to minimize the risk of cross-reactivity, as for example in immunotherapy or diagnostics [29], the choice of a highly homogenous product of consistent quality as a monoclonal antibody is nearer at hand. When choosing your weapon one should therefore consider in which application the antibody will be used in, in addition to the matter of production time and cost.

Recombinant antibodies and antibody fragments

The pursuit of a “magic bullet” that would selectively target disease [30] was immensely affected by the hybridoma technology in combination with recombinant DNA technologies. Unfortunately, when administered to humans, monoclonal antibodies of murine origin can evoke immune responses and also show inefficient activation of effector functions because of the foreign origin of the protein. Techniques such as chimerization [31,32] and humanization [33] have been developed with the aim to retain the binding properties of the murine monoclonal antibody, but minimizing immunogenicity and promoting effector functions with a human Fc part. A chimeric antibody consists of murine v regions and a human c regions while in humanized antibodies

sequences of murine origin are only found in the CDRs. Yet another way of producing “fully” human antibodies was the creation of transgenic mice with human antibody gene loci introduced [34,35,36].

With the expansion of antibody applications, the antibodies’ Fc part has shown to be unwanted or unnecessary in certain applications [20]. Fab fragments can be produced by papain cleavage [37] or by recombinant expression in for example *E. coli* [38]. Another, even smaller monovalent antibody fragment is the single chain fragment variable (scFv), which is composed of the variable domains of the H chain and L chain joined by a flexible linker [39]. For increased avidity or multispecificity for different targets, scFv can be modified into multivalent homomers or heteromers like diabodies [40], triabodies [41] or tetrabodies [42,43]. The smaller recombinant antibody fragments are possible to produce in prokaryotic hosts such as *E. coli*, which is usually less cumbersome and less expensive than mammalian cell production needed for full-length antibodies, although these processes are continuously improved.

Leaving nature behind

The cloning of antibody genes from immunized animals [44] in combination with bidning analysis of secreted antibody fragments was a first step of imitating the natural immune system [45,46] in efforts to bypass hybridoma production. A technique called phage display [47] (described below) broke ground for the *in vitro* selection of antibodies. The display of scFv on filamentous phage was described in 1990 [48] as an efficient way to select antibody clones through antigen binding. Antibody selections today employ immunized, naïve or synthetic libraries for biased or unbiased specificities towards antigens [49]. These selection systems also allow for affinity maturation in order to generate high affinity binders, if desired, which is not possible in hybridoma technology.

Selection systems of combinatorial protein libraries utilize a physical link between the genotype, the gene, and the phenotype, the protein, of its different members. This allows for manipulation of the gene using recombinant DNA technologies to diversify the displayed protein. The different library members can then be subjected to screening for a desired trait such as binding to a target protein. George Smith pioneered this in 1985 when he described a genetic fusion of a peptide to a coat protein expressed on the surface of filamentous phages. Since then the phage display selection of a binder towards a target is

most often performed by a procedure called “biopanning” [50]. The phages with displayed proteins are incubated with the target of interest and the target-bound phages are captured onto a solid phase. After washing for removal of background interactions the captured phages are eluted for infection into *E. coli*. The latter enables amplification since phages are non-replicating and after recovering the phages a new round of selection is performed. After several rounds of panning, the phage-displayed peptides are screened for binding in enzyme-linked immunosorbent analysis (ELISA) [51] and positive clones are sequence analyzed. Phage display is a well-established and standardized technique with no generally need of expensive lab resources, which makes it the most extensively used selection system today.

Cell display systems using yeast and bacteria have also been developed as they offer some advantages over phage display. The larger cell size and the multivalent display of the recombinant peptide allow the binding analysis to be performed in a flow cytometer where fine affinity discrimination of binders can be monitored in real-time during screening and sorting of target-bound cells. The yeast strain *S. cerevisiae* [52], gram-negative *E. coli* [53] and gram-positive *S. carnosus* [54] among others have been employed for selection of affinity proteins. An advantage of the eukaryotic yeast display system is the ability to express more complex proteins but a drawback exists in that library sizes are somewhat limited compared to *E. coli* and phage display. *S. carnosus* display is also hampered with smaller library sizes, however these libraries have shown to be suitable in affinity maturation of binders [55].

To avoid the limitations implied by transformation in order to create larger libraries, one can turn to cell free display systems like ribosome display [56,57] or mRNA display [58]. All screening and diversification is done *in vitro* and when performing evolutionary selection experiments diversification is easily introduced in the regular step of polymerase chain reaction (PCR) amplification after each selection round. Ribosome display uses the ribosome as a linker between the translated peptide (phenotype) and the mRNA (genotype) while in mRNA display a covalent linkage is formed by puromycin between the peptide chain and the mRNA.

The advances in library display technologies have led to that affinity ligands other than antibodies have emerged. These different molecules originate from human or other species but all possess the intrinsic feature of molecular recognition. Examples of protein based affinity reagents are Affibody molecules (derived from the staphylococcal protein A) [59], Adnectins (based on a human fibronectin domain) [60] and DARPins (based on ankyrin-repeat proteins)

[61] among others. These “protein scaffolds” are smaller than antibodies, demonstrate high stability and could be produced at high yield in bacterial expression systems [62,63,64]. Nucleic-acid based Aptamers are yet another kind of alternative affinity reagents [65].

Epitopes

Antibodies recognize antigenic determinants, also called epitopes [66] on their antigen. The corresponding interface of the antibody involved in the binding is called the paratope, which is combined of different amino acids in the CDRs. Throughout the years, investigations of interactions between an antibody and an antigen has led to some general conclusions. In the second part of the 1900's, epitopes were postulated to locate on the surface of a protein [67] and were divided into sequential (Figure 3), i.e. depending on the amino acid sequence in a random coil form of a peptide from the protein, and conformational (Figure 3), i.e. dependent on the native conformation of the antigen protein by residues brought together by folding [68]. Another nomenclature uses continuous and discontinuous [69], where not only the primary sequence but also a distinct conformation is important in the continuous epitopes. However, the distinction between these epitope classes is not so clear cut as a discontinuous epitope can consist of several continuous and a sequential/linear epitope can be autonomous without some residues, which would classify it as discontinuous [70]. The latter example is described in Paper II, where a discontinuous epitope on an alpha-helical structure was detected using synthetic peptides that are designed to characterize linear epitopes. Epitope mapping is important in understanding the mode of action of therapeutic antibodies and for guiding antigen design used in vaccine development and antibody generation for example. The interest in determining the interaction site between antibodies and antigen has lead to the emergence of many methods to map epitopes, all with different advantages and disadvantages. Often, a combination of different methods is preferred to characterize the epitope.

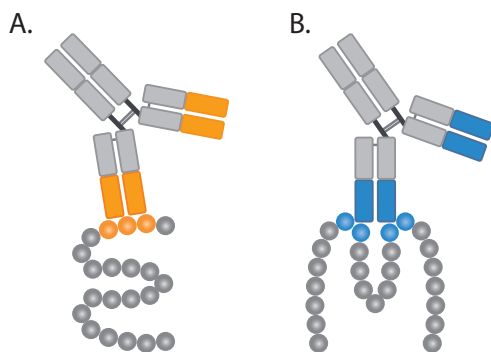


Figure 3. Simplified illustrations of (A) a continuous epitope and (B) a discontinuous antibody bound by their respective antibodies.

Epitope mapping methods

X-ray crystallography of an antibody-antigen complex is the gold standard in epitope determination [71]. The method provides a high-resolution structure of the complex where the molecular details of the interaction are defined, for both epitope and paratope. The technique is however dependent on production of highly purified proteins and the co-crystallization of these. Similarly detailed data is achieved with NMR spectroscopy, where the interaction is measured in solution. The method is though limited in molecular size, and only small antigens are well suited for NMR analysis. Both methods resolve both linear and conformational epitopes, and often render the highest resolution when studying antibodies binding to discontinuous epitopes. The defined epitope will include structural data of the interaction, but only indirect evidence of which residues contribute most to the binding energy of the interaction. Other methods suitable for detection of discontinuous epitopes include hydrogen/deuterium exchange of protein followed by pepsin digestion and mass spectrometry (MS) analysis [72].

Less cumbersome approaches to discover epitopes involve manipulation of the antigen in different ways followed by binding analysis to the antibody. Partitioning of the antigen by various means identifies the minimal stretch of amino acids needed for binding. The approach is considered limited for the detection of continuous epitopes since the partitioning of the protein antigen can disrupt the spatial organization of the epitope. Mutagenesis of the antigen on the other hand pinpoints single residues involved in the binding interaction, these amino acids form the functional epitope [73]. There are many techniques that use these two ways of antigen perturbation.

Peptide scanning

Peptide scanning involves most often the chemical synthesis of overlapping peptides covering the antigen sequence targeted by the investigated antibodies. In 1984, Geysen et al [74] described parallel synthesis of overlapping oligopeptides on a solid surface followed by probing of antibody binding by ELISA directly on the same surface. This technique (pin synthesis) allowed a far more throughput and straightforward way of epitope mapping antibodies towards proteins of known sequence than earlier for example by Kazim and Atassi [75] where peptides were synthesized on one surface and then assayed on another. Other techniques followed in this path, such as the SPOT synthesis technique [76]. This technique utilizes cellulose filter membranes as solid support where the peptides are synthesized manually in a standard microtiter plate format

or by automated manner to produce more dense arrays. The development of peptide synthesis by photolithography [77] has helped to overcome the limitations of miniaturization of above-mentioned systems.

The length of the synthesized peptides have been suggested to range between 6 and 18 amino acids [78] and the larger overlap between peptides, the higher resolution of the determined epitopes. Truncations of peptides are used to further narrow down the epitope sequence and mutagenesis of each amino acid of a peptide can also indicate the energetically favorable residues in the epitope. This can be achieved by either alanine substitution of one residue at a time or by a replacement set analysis [79] where each residue is substituted systematically by another of the remaining 19 amino acids to investigate changes in binding affinity. The vast number of peptides needed for characterization of an antigen at a high resolution still leaves these methods relatively expensive.

Phage display and other display technologies

Combinatorial libraries are widely used in selection and screening strategies of proteins for desired traits as enzyme activity or binding ability as for example in antibody selections mentioned in the previous chapter. These display systems are also employed in epitope mapping, where different peptide or protein libraries can be expressed on the surface of different hosts circumventing the costs of peptide synthesis described above. The sequences of selected binders or non-binders (derived by standard DNA sequencing) are aligned and an epitope or relevant residues for binding can be identified. The peptide libraries used in display systems are either random peptide libraries or antigen-derived libraries, fragmented or randomized by mutation [80]. The random peptide libraries are convenient for their universal nature as theoretically one library of greater size could be used to epitope map any antigen. The binding peptides from these libraries are mimics of the true epitope, “mimotopes” that can reveal residues of importance in epitopes rather than sequential stretches. Together with known structure of proteins this can aid in identifying discontinuous epitopes. In a comparison between random peptide libraries and antigen fragment libraries, the latter showed a higher success rate of identifying epitopes [81]. However, the outcome of a mapping depends on the nature of the epitope why random libraries would be more effective for discontinuous epitopes [80]. Combinatorial alanine scanning using phage display [82] is a recent version of the rather tedious, but precise alanine scanning mutagenesis [83]. The more rapid way of depicting the energy contributions of residues in protein interaction is done by generating a combinatorial library with multiple alanine substitutions and statistical analysis of DNA sequences from binding clones calculates the ener-

getic contributions of each individual side chain. However, mutations outside of the epitope could be detrimental for the conformation of the protein and this could render false positive results.

Selections of binding peptides on the surface of yeast and bacteria are performed for retained binding (positive selection) or loss of binding (negative selection). The ability to use flow cytometric analysis for sorting cells allow good discrimination between different affinities which enable a relatively sensitive negative sorting. Using *E. coli* display, linear epitopes have been defined by gene-derived random peptide libraries [84] by positive selection. Yeast display libraries have been able to map conformational epitopes on different proteins using domain level mapping and random mutagenesis of full-length protein [85,86,87,88] by negative selection.

Random peptide libraries have been used in ribosome display where direct protein synthesis of binding clones is used to verify binding clones in Western blot [89]. Using mRNA display Baggio *et al* [90] showed that epitope-like consensus motifs as well as the wild type ligand sequence could be identified in a random peptide library. In this thesis the *S. carnosus* display system has been used for epitope mapping.

In silico epitope predictions

Accurate computational tools for prediction of epitopes have been highly sought for in vaccine design to overcome the experimental determination of epitopes. The first report of prediction of antigenic features from a linear sequence was based on a scale of hydrophilic features for each amino acid [91]. More algorithms followed based on the physicochemical properties of amino acids e.g. mobility, surface exposure, turn, propensity and flexibility [92,93,94,95]. Combinations of these methods are also used [96,97]. An extensive study in 2004 by Blythe and Flower [98] showed that these propensity scales are futile in predicting epitopes and the need of more sophisticated approaches are needed to succeed in determining linear epitopes computationally. The increase in structural data of proteins has been incorporated in prediction methods to compute discontinuous epitopes [99,100]. The need of common datasets and standardized formats for epitope predictions to finally break through [101], emphasizes that still many experimental epitope mapping efforts are necessary in the nearest future. There is also an open question if epitopes are an intrinsic feature of a protein. If not, this would complicate the matter for prediction methods to succeed [102].

Proteomics

The proteome defines the entire protein content of a given cell, tissue or whole organism [103] and proteomics encompasses the study of these (in different ways) at a large scale. The number of protein-coding genes is currently predicted to be around 20,500 [104] however; each gene can lead to several different isoform of proteins due to alternative splice variants and post-translational modifications, thus resulting in a factorial multitude of functionally distinct proteins. Unraveling localization, structure, interaction partners, modifications and abundance of proteins, would advance our knowledge of biological processes as it would allow us to understand the different physiological conditions of life. To fill in all the existing blanks for molecules as heterogeneous as proteins, of different abundance in complex samples, clearly necessitates highly sensitive and selective analytical techniques. This is one of the driving forces pushing forward the development of new methods for protein analysis that can complement each other and support different strategies of proteome studies.

Mass spectrometry analysis preceded by separation of the complex sample material, using either one- or two-dimensional gel electrophoresis (1DE and 2DE, respectively) has been the more traditional approach for proteomic studies [105]. Shifting to a higher sample and target throughput, shotgun proteomics approaches have been used for protein identification studies where the protein sample is digested enzymatically prior to chromatographic separation and analysis by tandem MS [106]. Strategies using for example tagging or isotope labeling [107,108] of samples have moved MS proteomics from rather being an identification strategy into a quantitative strategy [109] as they compare the relative abundance of proteins between two samples. The field of MS proteomics is technically intense where hybrid instruments combined with spike in and depletion strategies (each named with an acronym more inventive or tongue twisting than the other) have emerged. Still they are faced with limitations of dynamic range, reproducibility and sampling rates of analysis [105,110,111,112].

Affinity-based proteomics

In addition to MS, proteomics can also be performed in a more directed manner where affinity reagents (described in chapter 2) are used for protein profiling. To be able to analyze the protein products of the entire human genome, large collections of well-validated binders have to be created. The lack of these has recently been emphasized [113] but several initiatives have been started to fill in this void i.e. the Human Protein Atlas (HPA) [114], the Proteome Binders initiative [51], the NCI Clinical Proteomics initiative [115] and the German Antibody Factory project [116]. One of the advantages of using affinity tools is their versatility in techniques such as ELISA, Western Blotting, immunoprecipitation (IP), immunohistochemistry (IHC), which allow to answer questions regarding protein expression levels, localization and function [117]. For high throughput protein analysis using IHC, tissue microarrays (TMAs) [118] for example enable a simultaneous detection of a protein of interest in multiple tissue sections. Miniaturization and multiplexing of immunoassays is performed on different types of protein arrays, which are mainly built using planar or bead surfaces [119,120]. An essential issue in affinity proteomics is to confirm the specificity of the reagent towards its antigen, with regards to minimize off-target interactions, as cross-reactivity will impede the results. Incorporating two independent recognition events, as in sandwich assays, increases specificity [117]. Due to diverse sample processing protocols of different assays, one should also consider how a protein is presented to its binder. Denaturation or tethering of proteins could affect epitope appearance and accessibility, which are important for interaction. Therefore validation of the affinity reagent for the intended application is of great importance, especially when they are generated in different lots, by different manufacturers, or from different polyclonal antisera [121].

The massive amount of data generated in both affinity and MS based proteomics relies on fast and unbiased methods of data assessment. Tools for bioinformatics and statistical analysis are evolving alongside the technical developments and should be carefully evaluated and standardized for implementation [120].

Biomarker discovery

A result of proteomic profiling is that potential candidates for new protein biomarker of diseases can be identified. A biomarker can be defined as a characteristic that is objectively measured and evaluated as an indicator of nor-

mal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [122]. The identification of diagnostic, prognostic and predictive biomarkers would help to stratify patients and their ability to respond to a certain treatment [123] and is therefore an important and needed component in patient care. A prognostic and predictive biomarker is for example the HER2 protein, which overexpressed in tumor cells in breast cancer indicates poor prognosis but also susceptibility for treatment with trastuzumab (Herceptin) [124]. Another routinely used biomarker is the prostate specific antigen (PSA), which is used to diagnose and predict prostate cancer by simple blood sampling [125]. Although there is debate about PSA's specificity as a biomarker, it exemplifies an ideal diagnostic assay as it is performed non-invasively in a blood test. Plasma represents the most comprehensive human proteome as it contains proteins leaked from tissue of a body in addition to the proteins acting in plasma such as immunoglobulins. However, the dynamic range of different protein concentrations span over more than twelve orders of magnitude [126], which makes the detection of leaked proteins a grand quest. The challenge has however been accepted by many scientists and examples of aspiring techniques are shown in the use of antibody arrays [127,128], aptamer arrays [129] and the "ProteinSeq" approach [130].

The path from discovering a candidate biomarker to its approved clinical use is though long and seldom successful. To address this, a comprehensive pipeline [110] has been proposed for a more effective and successful biomarker development. Phases included herein are verification of candidates certified by reproducibility of results in the technical assay, on related platforms and across different patient cohorts. This may then allow to perform follow up studies aiming to translate the marker into clinical assays (most often immunoassays) and subsequently validating it in a larger collections of samples that reflect the variation in the targeted population.

Human Protein Atlas

The Human Protein Atlas project [131] was launched in 2003 with the aim to map the expression and localization of protein targets using antibody-based proteomics. By generating protein-specific polyclonal antibodies to one representative protein from every human gene as defined by Ensembl [132], proteins can be studied in a wide range of normal tissue, cancer tissue and cell lines by several techniques. The goal is to provide a draft of the human pro-

teome by 2015 but the generated data is continuously published on the website www.proteinatlas.org.

The high-throughput generation of antibodies [114] is outlined in Figure 4. The workflow starts with *in silico* selection of antigens called protein epitope signature tags (PrEST) based on low sequence similarity to other human proteins [133]. Designed oligonucleotide primers are used for amplification of the fragment using reverse transcriptase polymerase chain reaction (RT-PCR) from human RNA pools and the fragment is subsequently cloned into an expression vector [134]. The PrESTs are expressed in *E. coli* as fusion proteins with a hexahistidine-tag (His₆-tag) and an albumin binding protein (ABP), purified [135] and immunized into rabbits with Freund's adjuvant for generation of polyclonal antibodies. The obtained sera are purified in a two-step protocol where tag-specific antibodies are removed before affinity purification on the PrEST used for immunization [136]. The specificity of these antigen-purified polyclonal antibodies is validated by protein microarrays composed of 384 purified PrEST fusions and Western blotting in a routine sample setup of IgG/HSA-depleted serum and protein lysates from two human cell lines and two human tissues. Validated antibodies are used to screen for protein profiling in 46 normal tissue types, 20 different cancer types [114] and 47 cell lines [137]. Three of the cell lines are used for a subcellular profiling of proteins [138] using immunofluorescence (IF) and confocal microscopy. These three cell lines have also been analyzed by next generation RNA-sequencing to calculate the RNA levels of transcribed genes [139].

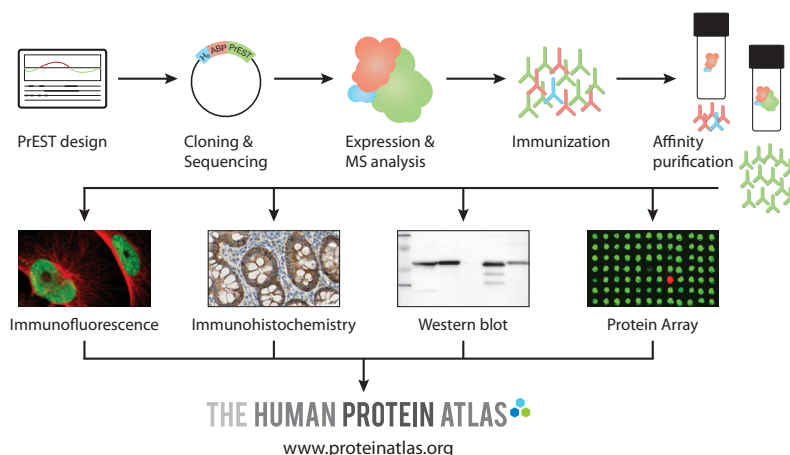


Figure 4. Overview of the pipeline of the Human Protein Atlas for high-throughput generation and validation of antibodies.

All annotation from the different protein profiling assays is publicly available on the website and scored according to different criteria. A concept of annotated protein expression was launched in version 7 of the atlas [140] where a reliability score estimates the degree of certainty of the reported expression file. This is done by validating data of a protein target with two independent antibodies, why the long-term goal of the Human Protein Atlas project is to generate at least two antibodies with separate and non-overlapping epitopes for all human protein coding genes. The growing database of annotated protein profiles provides an important source of information that can be used in biomarker discovery efforts [141].

PRESENT
INVESTIGATION

Aims of the thesis

The general aim of this thesis was to develop methods for epitope mapping and to implement these for the validation and characterization of antibodies used in proteomic studies and biomarker discovery.

Paper I

The aim of this study was to develop a new method for epitope mapping monoclonal and polyclonal antibodies. The method is based on presentation of a library based on antigen fragments on the surface of Gram-positive *S. carnosus*.

Paper II

The aim of this study was to develop a pipeline for streamlined epitope mapping at high resolution combining the method developed in Paper I with peptide and alanine scanning using a suspension bead array.

Paper III

The aim of this study was to investigate the epitope patterns of polyclonal antibodies from parallel immunizations using the same antigen.

Paper IV

The aim of this study was to develop “paired” antibodies towards a potential biomarker RBM3 of colorectal cancer. Different epitope mapped antibodies were used to stratify the expression of this protein in tumor tissue from colon cancer patients as means to validate each other.

Paper V

The aim of this study was to “break down” polyclonal antibodies into single epitope-specific entities to evaluate their functionality in different assays.

Epitope mapping of antibodies using bacterial display (Paper I)

The generation of antibodies in both proteomic endeavors as well as therapeutic or diagnostic applications is dependent on methods for exact characterization and validation. The identification of the antigen-binding site of the antibody is essential in this and therefore means of systematic epitope mapping is needed. We used the gram-positive bacterium *Staphylococcus carnosus* [142] to display a peptide library assayed for antibody binding in fluorescence activated cell sorting (FACS) followed by sequencing identification as a strategy for efficient epitope mapping. The workflow of the method is outlined in Figure 5. The antigen DNA is amplified by PCR and fragmented by sonication to sizes of 50-150 nucleotides, thereafter fragments are blunt-end cloned into a staphylococcal display vector and transformed by electroporation into *S. carnosus*. The fragments are fused C-terminally to ABP, which works as a spacer between the cell wall and the displayed library member, but also as a normalization tag in the FACS analysis [143]. The surface-displayed library is then incubated with labeled antibodies and labeled human serum albumin (HSA) followed by flow-cytometric sorting and subsequent sequencing. The normalization signal from the fluorescently labeled HSA enables discrimination between binders of different apparent affinity. This allows for positive selection of different epitope populations (as in polyclonal antibodies), as well as negative selection where non-binding peptides can be identified and used to confirm the epitope definition.

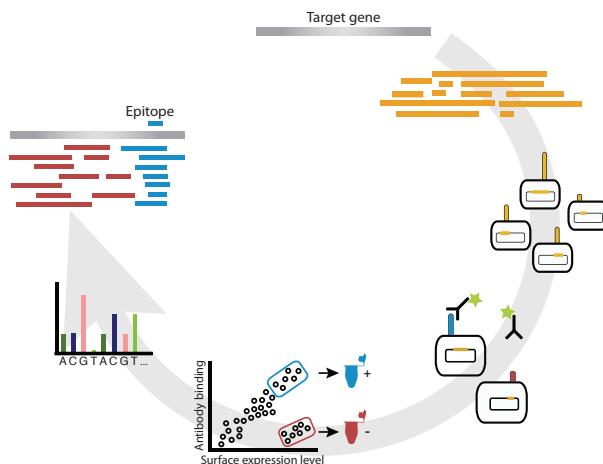


Figure 5. Epitope mapping of antibodies using bacterial display. Random fragments of the antigen are displayed on the surface of *S. carnosus*. After incubation of cells with the antibody of interest, bound and non-bound cells are sorted for sequence analysis. Derived sequences are aligned back to the antigen sequence for the identification of epitopes.

The method was validated by epitope mapping antibodies towards three different proteins. Three separate libraries were constructed covering the extracellular part of HER2, a recombinant fragment of Ephrin-B3 and a recombinant fragment of SATB2. Using the first library of HER2, a proof of principle was conducted by epitope mapping one commercial mAb and one commercial pAb against HER2. The identified epitopes were within the 17 amino acid peptide sequence used for immunization and a high resolution of only five amino acids was achieved for the mAb. Four different anti-HER2 domain polyclonal antibodies were also mapped revealing no cross-reactivity between domains and a pattern of a few distinct epitopes was identified in each case.

The Ephrin-B3 library was used to epitope map three polyclonal antibodies from the Human Protein Atlas generated with the same antigen in separate animals. The mapping revealed a common C-terminal epitope for all three immunizations and an additional epitope for two of the immunizations. To further validate the epitope mapping, one staphylococcal epitope clone was used as a whole-cell affinity ligand to generate an epitope-specific polyclonal antibody. A competition assay of this antibody with a synthesized peptide corresponding to the same epitope showed complete blocking of the antibody binding, thus confirming the epitope specificity.

The SATB2 library was used to compare the epitopes of one polyclonal antigen purified rabbit antibody and four mouse monoclonal antibodies. The mapping identified single epitopes for the monoclonal antibodies and four epitopes for the polyclonal antibody. Furthermore, the epitopes of the monoclonal antibody overlapped in pairs with two different epitopes of the polyclonal antibody.

Exploring epitopes of antibodies toward the human tryptophanyl-tRNA synthetase (Paper II)

Even though synthetic peptide mapping [79] can define epitopes at a high resolution of a single amino acid, the cost of synthesis to cover the entire antigen sequence can end up relatively high. Similarly, when performing alanine scanning throughout the antigen sequence [83] to identify residues important for the interaction, the time expense for producing many mutants for evaluation of binding is not appealing. Therefore we combined the bacterial display method described in the previous study, with peptide scanning and alanine

scanning using synthetic peptides as a pipeline for high-resolution epitope mapping of antibodies (Figure 6). By localizing the epitope regions for the antibody by bacterial display a more narrow peptide array can be used for the fine mapping, which is finally used as a guide in the design of the alanine scan analysis.

This streamlined analysis was performed on a polyclonal and a monoclonal antibody generated towards a recombinant fragment of the human tryptophanyl-tRNA synthetase. The bacterial display identified three epitope regions for the polyclonal antibody and one region, as expected, for the monoclonal antibody. The epitope of the monoclonal antibody was found to overlap with the epitope demonstrating the highest apparent binding of the polyclonal antibody. A common peptide array could therefore be used for fine mapping of both antibodies. Two peptide arrays of different length were designed, one 10-mer and one 15-mer, with a sliding window of two amino acids. Both antibodies demonstrated no binding to the 10-mer peptide array, however using the 15-mer peptide array epitopes could be determined and showed to differ by only a few amino acids.

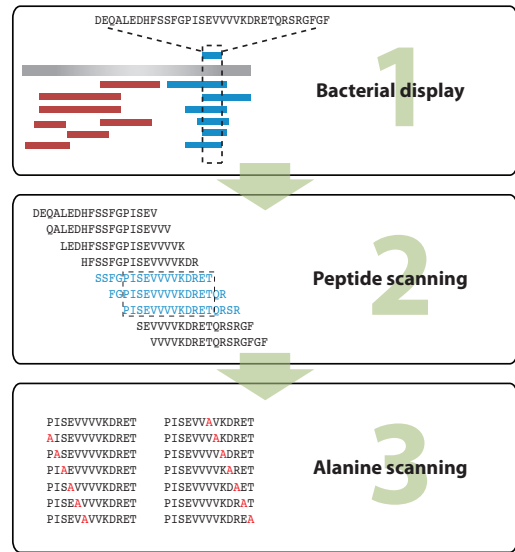


Figure 6. A streamlined approach for high-resolution epitope mapping of antibodies. Bacterial display is used for locating epitopes throughout the whole antigen sequence. A fine mapping of the discovered epitopes is performed with overlapping synthetic peptides covering the sequence of the epitope region. In the final step, an alanine scan with synthetic peptides based on the fine-mapped epitope in the previous step is performed.

The monoclonal antibody was finally investigated by an alanine scan array based on a seventeen amino acid long peptide fragment corresponding to the fine mapped epitope. Binding loss was observed for four of the alanine substituted peptides, concluding that those substituted residues are important for binding. Structural analysis of the protein showed that these residues were oriented outwards from the surface of an alpha-helical structure.

Immunizations of inbred rabbits using the same antigen yields antibodies with similar, but not identical, epitopes (Paper III)

Polyclonal antibody generation is cost-effective as compared to the generation of monoclonal antibodies [28]. However, the limited amount of a polyclonal antibody from one animal immunization is a major drawback, as new immunization will not yield the same product [144,145]. To investigate the reproducibility of polyclonal antibodies, we mapped the epitopes of antibodies targeting ten different proteins. For each protein, a recombinant PrEST fragment was used as antigen for immunization in three separate in-bred rabbits. The antibodies were validated for target binding in Western blot and in three cases IHC. All antibodies were indeed binding to their target protein, however several antibodies exhibited cross-reactivity in Western blot where additional bands could be observed. Ten different staphylococcal libraries were used for epitope mapping their respective antibodies. Also, for six targets, peptide mapping using suspension bead arrays was performed. The mapping showed that in between immunizations of the same antigen some overlapping epitope regions could be identified, but that the epitope pattern was not identical. This was especially noted in the bacterial display mapping as the peptide array was of quite low resolution. Some epitopes were only noted for one antibody

On the basis of the peptide mapping, the three antibodies targeting the protein TYMP were separated to generate single epitope-specific fractions, as described in paper V, for in depth analysis of the different peptide epitopes. This showed that the relative abundance of antibodies from different immunizations, but specific for the same epitope, differed. One common epitope-specific fraction for all three immunizations showed good performance in Western blot. Altogether, the polyclonal antibodies from repeated immunization using the same antigen exhibited similar, but not identical patterns.

High nuclear RBM3 expression is associated with an improved prognosis in colorectal cancer (Paper IV)

By using the Human Protein Atlas [131,141], a differential expression of the RNA binding protein RBM3 was discovered in breast cancer tissue. This was evaluated in two breast cancer cohorts, and a high nuclear expression of RBM3 was associated with a prolonged survival [146]. Following studies have shown that RBM3 is a potential biomarker in ovarian cancer [147], malignant melanoma [148] and prostate cancer [149]. We generated one polyclonal and four monoclonal antibodies with the purpose to investigate the prognostic value of RBM3 in colorectal cancer. Preferably, two antibodies with separate and non-overlapping epitopes should be used in for example IHC to validate the staining pattern of each other [150]. Again, the bacterial display method and a synthetic peptide array spanning the whole antigen was used for epitope mapping the RBM3 targeting antibodies. Interestingly, the bacterial display could identify epitopes for all antibodies, while the protein array could only detect binding for the polyclonal antibody and one of the monoclonal antibodies. Furthermore, the epitopes of the mouse monoclonal antibodies overlapped with different epitopes of the polyclonal rabbit antibody. The polyclonal antibody and one monoclonal antibody were further validated for specificity using siRNA mediated knock-down of RBM3 in cells prior immunohistochemical staining of TMA's with specimens collected from two independent patient cohorts. The staining was evaluated and denoted with a nuclear score, with high correlation between both antibodies. Statistical analysis with different methods showed that, in both cohorts, tumors with high nuclear staining had significantly prolonged the overall survival.

Generation of monospecific antibodies based on affinity capture of polyclonal antibodies (Paper V)

Although the multi-specific nature of polyclonal antibodies to several epitopes of a target protein is of advantage in many detection assays, this also elevates the risk of cross-reactivity to other proteins. Therefore single epitope-specific antibodies as monoclonal antibodies or recombinant affinity reagents are needed. Antibodies specific for one epitope could also be generated from partitioning of the multiple epitope clones of one polyclonal antibody, offering a possibility to achieve paired antibodies against one target protein. We have put up a strategy where mapping of polyclonal sera using peptide bead

arrays followed by affinity purification using synthetic peptides yields single epitope-specific antibody fractions (Figure 7). The synthetic peptides covering the epitopes are coupled to separate columns, a tag-coupled column and an antigen-coupled column are also prepared. The polyclonal serum is passed through these columns in a serial manner and elution of these is performed separately in a standardized manner using an automated chromatography system. The resulting epitope-specific fractions as well as the last fraction containing antibodies to remaining epitopes on the antigen are mapped using the same peptide array as previously to confirm the epitope identity.

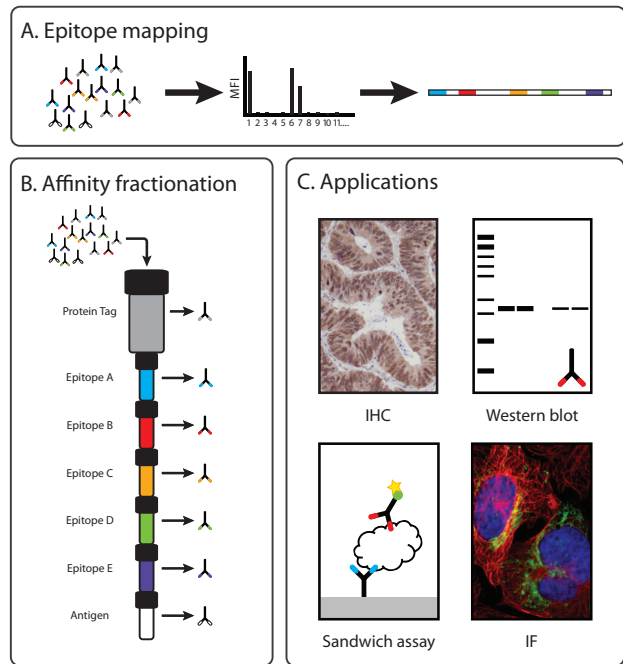


Figure 7. Principle for generation of polyclonal single epitope-specific antibodies. (A) Epitopes are determined using a synthetic peptide bead array. (B) Peptides corresponding to the epitopes as well as the recombinant antigen are coupled to separate columns used for fractionation of polyclonal sera. (C) The different single epitope-specific antibodies are evaluated in different paired-antibody applications.

We generated single epitope specific antibodies towards four protein targets and evaluated them in different assays and compared them to the originating polyclonal antibody results. The different epitope-specific antibodies toward their respective target showed diverse results. Some epitope-directed antibodies were functional in all assays, while antibodies directed to another binding region were not functional in any of the assays. Cross-reactivity was identified for one polyclonal antibody targeting the human enzyme ANLN, and could be assigned to one of the epitope specific fractions. This observation was confirmed by siRNA mediated knockdown analysis.

For the two protein targets RBM3 and SATB2, paired antibodies towards separate, non-overlapping epitopes were generated and an annotated protein expression profile [140] was assigned. One antibody pair towards the human enzyme CNDP1 was used in a sandwich assay for capture and detection. Lastly, the results from the epitope-specific antibodies were used as a guide for generation of monoclonal antibodies towards RBM3. Peptides corresponding to two epitopes were used in immunization of mice and subsequent generation of mouse hybridoma cells. The supernatants of the hybridoma were tested in Western blot and immunohistochemistry and showed the same binding pattern as the corresponding epitope-specific antibodies.

Concluding remarks and future perspectives

Antibodies have become indispensable tools in the research area of life sciences. The quality of these is dependent on their selectivity and affinity for the target of interest and therefore methods for the validation of this are needed. The work behind this thesis has focused on the development and evaluation of methods to identify the interaction site of antibodies on their antigen.

In paper I we have demonstrated the applicability of staphylococcal surface display in epitope mapping of antibodies. The method was implemented into a strategy for high-resolution definition of epitopes together with peptide scanning and alanine scanning on suspension bead arrays in paper II. In this thesis, polyclonal and monoclonal antibodies of rabbit and murine origin were mapped, but one should also be able to profile the epitope patterns of antibodies in sera from individuals with autoimmune disorders with both bacterial display and peptide suspension bead arrays. The fact that shorter peptides were not able to detect any antibody binding in paper II, and that only one of four monoclonal antibodies was successfully mapped using the peptide array in paper III, points out the versatility of the staphylococcal display method compared to the peptide arrays. The displayed peptides on the bacterial surface are of random lengths and were able to detect epitopes for all assayed antibodies.

In paper III, we performed a systematic study of epitope patterns obtained when an antigen is separately immunized into several inbred animals. Staphylococcal display of ten different antigen libraries was used to map the epitopes of 30 antibodies and six peptide bead arrays were used to map the epitope of 18 of these antibodies. The batch-to-batch variability of polyclonal antibodies was confirmed as some minor epitopes differed between the immunizations and this could be an explanation to some cross-reactivity observed in a functional assay. However, common dominating epitopes were also identified suggesting that the immune system of the different animals are capable to direct the response to the same regions of the antigen. This study emphasizes the need of thorough validation of polyclonal antibodies to ascertain that the functionality in particular immune assays is the same as for the original antibody.

The overlap of epitopes targeted by antibodies of different species origin noted in paper I, II and IV, awakes the notion to attempt translating validated epitopes of rabbit polyclonal antibodies into renewable mouse monoclonal antibodies. As shown in paper V, the information from polyclonal single epitope-specific antibodies of rabbit origin could be used to generate mouse monoclonal antibodies to selected epitopes. Peptides corresponding to the epitope of interest were used directly in the immunization, but one could think of alternative routes; to use the original recombinant antigen for immunization and screen the obtained B-cells or generated hybridoma for binding to selected peptide-epitopes and isolate these clones.

The staphylococcal display system holds the potential of being used in whole-proteome studies. A library could be constructed to cover all genes of a proteome and used for epitope mapping of antibodies at a sufficient resolution, while at the same time analyzing cross-reactivity to other proteins. In the near future technical advances should present the potential to produce high-density microarray chips (Rockberg, Schaffer and Uhlén, unpublished) covering the whole proteome of an organism with a couple of million overlapping synthetic peptides which would also allow for characterization of antibodies in a systematic manner.

The possibility to capture epitope-specific antibodies from one polyclonal antibody has been illustrated both in paper I and paper V. This could fast track the generation of paired antibodies towards separate and non-overlapping epitopes for use in protein expression/localization analysis and sandwich assays. The method is also suited to study the relative abundance of antibodies toward continuous respectively discontinuous epitopes.

Altogether, the epitope mapping strategies presented in this thesis provide a promising contribution to already existing methods. The expanding knowledge of antibody epitopes on a molecular level will help us to understand mechanisms in molecular recognition and to optimize the use of antibodies in research, diagnostics and therapy.

Acknowledgements

Ropen skalla, tacktal till alla! För det här skulle lilla jag aldrig klarat utan hjälp och stöd från många personer. Därför vill jag tacka bland andra:

Först och främst min handledare **Mathias**. Vilken resa! Tack för att du erbjöd mig att stanna kvar efter exjobbet. Det har varit minst sagt lärorikt och utvecklande att jobba med dig. Din smittande entusiasm och avslappnade inställning gör arbetet lättare och roligare. **Stefan**, min problemlösande bihandledare. Tack för att du inkluderade mig i din grupps olika aktiviteter och alltid haft tid att lyssna eller diskutera något man haft på hjärtat.

Det tog nog två år innan jag vågade gå in på ditt kontor **Per-Åke** men det var tur att jag tog mod till mig. Du svarar på mina dumma frågor som ”Vad gör jag här?” utan att dra på munnen och så har du en skön humor också. **Amelie** vill jag tacka för värdefull granskning av den här avhandlingen. Även alla andra **professorer och PI's** som peppar och pushar fram avdelningen.

ProNova VINN Excellence Centre for Protein Technology för finansiering .

Mina exjobbshandledare för ett roligt projekt, **Johan** som drog in mig i detta äventyr och tålmodige **John** som svarat på miljoner frågor och även stått ut med mina pipande öron och snarkningar i San Diego. **Carmen** som dök upp från ingenstans och fick leva med en lagom förvirrad handledare. **Björn** för att jag fick hänga och lära mig massa hos Immunotech och för att du sedan hoppade på tåget och körde vidare på epitop-spåret. **Henrik** på Atlas, det har varit väldigt roligt att samarbeta med dig och jag fastnar alltid uppe hos dig för långa diskussioner om våra projekt. Mycket uppskattat! **Jochen** som inflikat med värdefulla tips och feedback lite här och där förutom i våra samarbetsprojekt. **Karin** som styrde upp RBM3-peket nerifrån Skåne, vad kul att vi fick ihop något till slut! Alla som hjälpt mig med avhandlingen på olika sätt. **Alfred, Helena W, Helena M, Jochen, Johanna, John och Kalle**.

Mitt första kontor med **Cristina, Johan, Linn, Lisa och Åsa** för alla promenader till kaffe-, jag menar temaskinen. Sedan **Snygggrummet** och alla dess självklart snygga medlemmar för vår härliga stämning. Jag anser att vi vidhåller en utmärkt balans mellan forskning och andra, minst lika viktiga, saker. Oavsett vad andra säger. **Anna P, Camilla, Helena, Johanna, Magda, Ullis och**

andra fantastiska tjejer på jobbet som förgyllt min vardag, konferenser och fester. **Anna K** för att du är lika knäpp som jag och så himla grym. **Bahram** för att du ser till att man går hem, levererat utomordentliga sekvenser och alltid bjudit på ett leende i korridoren. Mina **labbkamrater i Mellangården** för att ni fått långa inkubationstider att plötsligt kännas korta. **Inger, Kristina och Mona** för hjälp med blanketter, ekonomi och frågor om administration. Alla **andra på plan 3 och i HPA** för all hjälp och det fantastiska arbete ni genomför. Men också för att ni är personerna som gör vår arbetsplats till ett roligt ställe att jobba på. Ett särskilt tack till alla er som tar sig lite extra tid och arrangerar roliga tillställningar och event för oss andra!

Familjen Neiman, Maja och Märta som kommit förbi och fikat under pekskrivande och avhandlingsarbete, **Mårten** för att du vågar gå mot strömmen samt för att du låtit mig demoralisera dig och gladeligen provat alla polska delikatesser. Vidare har vi **Daniel**, min "arbetskamrat" som förklarat saker för mig sedan den första dagen på KTH. När ska du förklara mig hur man slutar känna sig dum? **Martina, Josefine och Cajsas**, vi möttes tio år sen under mottagningen till KTH, vi hade kul då och det har vi fortfarande. Let's keep it up!

Tack **Patrik och Erik** för alla luncher. Nu är det äntligen min tur! Det har varit väldigt skönt att prata av sig med er.

Haben och Kina. Våra spelkvällar med de stående vinnarna (you know it Kina!) och en legendarisk resa i Provence (hoppas på uppföljare!).

Cicci, Ebba, Ellinor, Karin och Lena för vinmiddagar, seglingar, Milano, London. Era pojkar är inte så pjåkiga heller och vi har spenderat otaliga middagar och nyårsaftnar ihop. Vi har alltid så sjukt roligt ihop! Särskilt **Alfred** vill jag tacka för blixtnabb hjälp med InDesign.

Brommapojkarna Anders, Daniel, Gustav, Jocke, Jonas, Markus, Micke och norske gutten Kristian med tillhörande flickor. Ni är outtröttliga debattörer vilket jag fått vänja mig vid, men nu kan jag inte tänka mig en middag eller efterfest som inte avslutas med extremt viktiga (eller var de meningslösa?) diskussioner. Tack även till dem av er som spenderade en trevlig vecka i Italien med mig a.k.a. min disputationresa.

Anna, Mia, Petra Tove, Åsa och tillhörande pojkar. Fina rosa flickor som klappar och kramar när det behövs. Många peppsamtal har vi haft och dessutom är ni de som har mest förståelse för min något barnsliga humor.

Och sen kommer AF:s fyror 1992. **Helena, Helena, Lisa, Leon, Jocke, Hanna och Jenny**. Jag är så glad att vi håller kontakten och ses med mer eller mindre jämna mellanrum. Ett särskilt tack till Helena M för den fina bilden till framsidan!

Min nya extrafamilj, **Agneta** för att du är så härlig och Kalles största fan. **Göran** för underbara Koster-vistelser, **Bosse och Elisabeth** för avkoppling och familjemys i Bingsjö, och givetvis **resten av hela släkten** för att ni välkomnat mig så hjärtligt i era liv. **Johan och Nina** för trevliga middagar hos **Anders**, samt för lammfester och helger ute på Sannerby tillsammans med resten av **Axnäs-klanen**.

Ania i Jacek för att ert hem alltid är öppet för mig. Jag mår alltid så himla bra av att få komma över, umgås med er och leka med era underbara flickor.

Rodzicom za to że jesteście, dziękuje **Mamusie** za między innymi podwójne obiady i **Tatusia** który mnie wspiera w swój cichy sposób. **Babcia Hexa**, niestety daleko od nas w Polsce ale zawsze mi bliska w myślach, zawsze byłam i jestem zaimponowana tobą. Mina bröder **Micke och Alexander** som orkar med sin bossiga syster i cała rodzina **Szmilichowska** za przyjemne spotkania oraz radosne wspomnienia. Jesteście wszyscy wspaniali!

Tack även till alla som jag missat få med på pränt men finns i mina tankar ändå.

Sist av alla, men viktigast, **min underbare Kalle**. Jag älskar dig, Du och Jag.

References

1. Berzelius J, Söderbaum H, et al. (1916) Bref. 2:2, Brefväxling mellan Berzelius och G. J. Mulder (1834-1847). Uppsala: Almqvist & Wiksell.
2. Vickery HB (1950) The origin of the word protein. *Yale J Biol Med* 22: 387-393.
3. Fischer E (1902) Ueber die hydrolyse der proteinstoffe. *Chem Ztg* 26: 939-940.
4. Hofmeister F (1902) Über Bau und Gruppierung der Eiweisskörper. *ErgebPhysiol* 1: 759-802.
5. Pauling L, Corey RB, et al. (1951) The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* 37: 205-211.
6. Pauling L, Corey RB (1951) The Pleated Sheet, A New Layer Configuration of Polypeptide Chains. *Proceedings of the National Academy of Sciences* 37: 251-256.
7. Anfinsen CB (1973) Principles that Govern the Folding of Protein Chains. *Science* 181: 223-230.
8. Kendrew JC, Bodo G, et al. (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* 181: 662-666.
9. Berman HM, Westbrook J, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235-242.
10. Jones S, Thornton JM (1996) Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences* 93: 13-20.
11. Cohen SN, Chang ACY, et al. (1973) Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proceedings of the National Academy of Sciences* 70: 3240-3244.
12. Hutchison CA, Phillips S, et al. (1978) Mutagenesis at a specific position in a DNA sequence. *Journal of Biological Chemistry* 253: 6551-6560.
13. Drews Jr (2000) Drug Discovery: A Historical Perspective. *Science* 287: 1960-1964.
14. Leader B, Baca QJ, et al. (2008) Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* 7: 21-39.
15. Porter RR (1958) Separation and isolation of fractions of rabbit gamma-globulin containing the antibody and antigenic combining sites. *Nature* 182: 670-671.
16. Edelman GM (1959) Dissociation of gamma-globulin. *Journal of the American Chemical Society* 81: 3155-3156.
17. Edelman GM, Poulik MD (1961) Studies on structural units of the gamma-globulins. *The Journal of Experimental Medicine* 113: 861-884.
18. Hilschmann N, Craig LC (1965) Amino Acid Sequence Studies with Bence-Jones Proteins. *Proceedings of the National Academy of Sciences of the United States of America* 53: 1403-1409.
19. Putnam FW, Easley CW (1965) Structural Studies of the Immunoglobulins. *Journal of Biological Chemistry* 240: 1626-1638.

20. Holliger P, Hudson PJ (2005) Engineered antibody fragments and the rise of single domains. *Nat Biotech* 23: 1126-1136.
21. Goldsby RA, Kindt TJ, et al. (2006) *Immunology*: W H Freeman & Co.
22. von Behring E, Kitasato S (1991) [The mechanism of diphtheria immunity and tetanus immunity in animals. 1890]. *Molecular immunology* 28: 1317, 1319-1320.
23. Freund J, Casals J, et al. (1937) Sensitization and Antibody Formation after Injection of Tubercle Bacilli and Paraffin Oil. *Proceedings of the Society for Experimental Biology and Medicine Society for Experimental Biology and Medicine* (New York, NY) 37: 509-513.
24. Landsteiner K (1901) Über Agglutinationserscheinungen normalen menschlichen Blutes. *Wien Klin Wschr* 14: 1132-1134.
25. Yalow RS, Berson SA (1960) Immunoassay of endogenous plasma insulin in man. *The Journal of clinical investigation* 39: 1157-1175.
26. Coons A.H., Creech H.J., et al. (1942) The Demonstration of Pneumococcal Antigen in Tissues by the Use of Fluorescent Antibody. *The Journal of Immunology* 45: 159-170.
27. Kohler G, Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256: 495-497.
28. Lipman NS, Jackson LR, et al. (2005) Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources. *ILAR journal / National Research Council, Institute of Laboratory Animal Resources* 46: 258-268.
29. Borrebaeck CAK (2000) Antibodies in diagnostics, Åi from immunoassays to protein chips. *Immunology Today* 21: 379-382.
30. Streibhardt K, Ullrich A (2008) Paul Ehrlich's magic bullet concept: 100 years of progress. *Nat Rev Cancer* 8: 473-480.
31. Boulianne GL, Hozumi N, et al. (1984) Production of functional chimaeric mouse/human antibody. *Nature* 312: 643-646.
32. Morrison SL, Johnson MJ, et al. (1984) Chimeric human antibody molecules: mouse antigen-binding domains with human constant region domains. *Proceedings of the National Academy of Sciences* 81: 6851-6855.
33. Jones PT, Dear PH, et al. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature* 321: 522-525.
34. Green LL, Hardy MC, et al. (1994) Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. *Nature genetics* 7: 13-21.
35. Lonberg N, Taylor LD, et al. (1994) Antigen-specific human antibodies from mice comprising four distinct genetic modifications. *Nature* 368: 856-859.
36. Fishwild DM, O'Donnell SL, et al. (1996) High-avidity human IgG kappa monoclonal antibodies from a novel strain of minilocus transgenic mice. *Nature biotechnology* 14: 845-851.
37. Porter RR (1959) The hydrolysis of rabbit γ -globulin and antibodies with crystalline papain. *The Biochemical journal* 73: 119-126.

38. Better M, Chang CP, et al. (1988) Escherichia coli secretion of an active chimeric antibody fragment. *Science* 240: 1041-1043.
39. Bird RE, Hardman KD, et al. (1988) Single-chain antigen-binding proteins. *Science* 242: 423-426.
40. Holliger P, Prospero T, et al. (1993) «Diabodies»: small bivalent and bispecific antibody fragments. *Proceedings of the National Academy of Sciences of the United States of America* 90: 6444-6448.
41. Iliades P, Kortt AA, et al. (1997) Triabodies: single chain Fv fragments without a linker form trivalent trimers. *FEBS Letters* 409: 437-441.
42. Todorovska A, Roovers RC, et al. (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *Journal of Immunological Methods* 248: 47-66.
43. Le Gall F, Kipriyanov SM, et al. (1999) Di-, tri- and tetrameric single chain Fv antibody fragments against human CD19: effect of valency on cell binding. *FEBS Letters* 453: 164-168.
44. Orlandi R, Gussow DH, et al. (1989) Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* 86: 3833-3837.
45. Ward ES, Gussow D, et al. (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. *Nature* 341: 544-546.
46. Huse WD, Sastry L, et al. (1989) Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* 246: 1275-1281.
47. Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 228: 1315-1317.
48. McCafferty J, Griffiths AD, et al. (1990) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* 348: 552-554.
49. Hoogenboom HR (2005) Selecting and screening recombinant antibody libraries. *Nat Biotech* 23: 1105-1116.
50. Parmley SF, Smith GP (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* 73: 305-318.
51. Taussig MJ, Stoevesandt O, et al. (2007) ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome. *Nat Meth* 4: 13-17.
52. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotech* 15: 553-557.
53. Francisco JA, Earhart CF, et al. (1992) Transport and anchoring of beta-lactamase to the external surface of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* 89: 2713-2717.
54. Samuelson P, Hansson M, et al. (1995) Cell surface display of recombinant proteins on Staphylococcus carnosus. *Journal of Bacteriology* 177: 1470-1476.
55. Kronqvist N, Malm M, et al. (2011) Combining phage and staphylococcal surface display for generation of ErbB3-specific Affibody molecules. *Protein engineering, design & selection : PEDS* 24: 385-396.

56. Mattheakis LC, Bhatt RR, et al. (1994) An in vitro Polysome Display System for Identifying Ligands from very Large Peptide Libraries. *Proceedings of the National Academy of Sciences of the United States of America* 91: 9022-9026.
57. Hanes J, Plückthun A (1997) In vitro selection and evolution of functional proteins by using ribosome display. *Proceedings of the National Academy of Sciences* 94: 4937-4942.
58. Roberts RW, Szostak JW (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences* 94: 12297-12302.
59. Nord K, Gunneriusson E, et al. (1996) Binding proteins selected from combinatorial libraries of an [alpha]-helical bacterial receptor domain. *Nat Biotech* 15: 772-777.
60. Koide A, Bailey CW, et al. (1998) The fibronectin type III domain as a scaffold for novel binding proteins. *Journal of Molecular Biology* 284: 1141-1151.
61. Binz HK, Amstutz P, et al. (2004) High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotech* 22: 575-582.
62. Nygren P-Å, Skerra A (2004) Binding proteins from alternative scaffolds. *Journal of Immunological Methods* 290: 3-28.
63. Binz HK, Amstutz P, et al. (2005) Engineering novel binding proteins from nonimmunoglobulin domains. *Nat Biotech* 23: 1257-1268.
64. Löfblom J, Frejd FY, et al. (2011) Non-immunoglobulin based protein scaffolds. *Current Opinion in Biotechnology* 22: 843-848.
65. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249: 505-510.
66. Jerne NK (1960) Immunological speculations. *Annual review of microbiology* 14: 341-358.
67. Benjamin DC, Berzofsky JA, et al. (1984) The antigenic structure of proteins: a reappraisal. *Annual review of immunology* 2: 67-101.
68. Sela M, Schechter B, et al. (1967) Antibodies to Sequential and Conformational Determinants. *Cold Spring Harbor Symposia on Quantitative Biology* 32: 537-545.
69. Atassi MZ, Smith JA (1978) A proposal for the nomenclature of antigenic sites in peptides and proteins. *Immunochemistry* 15: 609-610.
70. Van Regenmortel MHV, de Marcillac GD (1988) An assessment of prediction methods for locating continuous epitopes in proteins. *Immunology Letters* 17: 95-107.
71. Gershoni JM, Roitburd-Berman A, et al. (2007) Epitope Mapping: The First Step in Developing Epitope-Based Vaccines. *BioDrugs* 21: 145-156.
72. Baerga-Ortiz A, Hughes CA, et al. (2002) Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein. *Protein Science* 11: 1300-1308.
73. Cunningham BC, Wells JA (1993) Comparison of a Structural and a Functional Epitope. *Journal of Molecular Biology* 234: 554-563.

74. Geysen HM, Meloen RH, et al. (1984) Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc Natl Acad Sci U S A* 81: 3998-4002.
75. Kazim AL, Atassi MZ (1980) A novel and comprehensive synthetic approach for the elucidation of protein antigenic structures. Determination of the full antigenic profile of the alpha-chain of human haemoglobin. *The Biochemical journal* 191: 261-264.
76. Frank R (1992) Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. *Tetrahedron* 48: 9217-9232.
77. Fodor SP, Read JL, et al. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science* 251: 767-773.
78. Hilpert K, Winkler DFH, et al. (2007) Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion. *Nat Protocols* 2: 1333-1349.
79. Geysen HM, Rodda SJ, et al. (1987) Strategies for epitope analysis using peptide synthesis. *Journal of Immunological Methods* 102: 259-274.
80. Irving MB, Pan O, et al. (2001) Random-peptide libraries and antigen-fragment libraries for epitope mapping and the development of vaccines and diagnostics. *Current Opinion in Chemical Biology* 5: 314-324.
81. Fack F, Hügler-Dörr B, et al. (1997) Epitope mapping by phage display: Random versus gene-fragment libraries. *Journal of Immunological Methods* 206: 43-52.
82. Weiss GA, Watanabe CK, et al. (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proceedings of the National Academy of Sciences* 97: 8950-8954.
83. Cunningham BC, Wells JA (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244: 1081-1085.
84. Christmann A, Wentzel A, et al. (2001) Epitope mapping and affinity purification of monospecific antibodies by *Escherichia coli* cell surface display of gene-derived random peptide libraries. *Journal of Immunological Methods* 257: 163-173.
85. Chao G, Cochran JR, et al. (2004) Fine Epitope Mapping of anti-Epidermal Growth Factor Receptor Antibodies Through Random Mutagenesis and Yeast Surface Display. *Journal of Molecular Biology* 342: 539-550.
86. Cochran JR, Kim Y-S, et al. (2004) Domain-level antibody epitope mapping through yeast surface display of epidermal growth factor receptor fragments. *Journal of Immunological Methods* 287: 147-158.
87. Oliphant T, Engle M, et al. (2005) Development of a humanized monoclonal antibody with therapeutic potential against West Nile virus. *Nat Med* 11: 522-530.
88. Levy R, Forsyth CM, et al. (2007) Fine and Domain-level Epitope Mapping of Botulinum Neurotoxin Type A Neutralizing Antibodies by Yeast Surface Display. *Journal of Molecular Biology* 365: 196-210.
89. Osada E, Shimizu Y, et al. (2009) Epitope Mapping Using Ribosome Display in a Reconstituted Cell-Free Protein Synthesis System. *Journal of Biochemistry* 145: 693-700.

90. Baggio R, Burgstaller P, et al. (2002) Identification of epitope-like consensus motifs using mRNA display. *Journal of Molecular Recognition* 15: 126-134.
91. Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Sciences* 78: 3824-3828.
92. Westhof E, Altschuh D, et al. (1984) Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature* 311: 123-126.
93. Novotny J, Handschumacher M, et al. (1986) Antigenic determinants in proteins coincide with surface regions accessible to large probes (antibody domains). *Proceedings of the National Academy of Sciences* 83: 226-230.
94. Pellequer J-L, Westhof E, et al. (1993) Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunology Letters* 36: 83-99.
95. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins. *Naturwissenschaften* 72: 212-213.
96. Alain J.P A (1999) Predictive estimation of protein linear epitopes by using the program PEOPLE. *Vaccine* 18: 311-314.
97. Odorico M, Pellequer J-L (2003) BEPITOPE: predicting the location of continuous epitopes and patterns in proteins. *Journal of Molecular Recognition* 16: 20-22.
98. Blythe MJ, Flower DR (2005) Benchmarking B cell epitope prediction: Underperformance of existing methods. *Protein Science* 14: 246-248.
99. Haste Andersen P, Nielsen M, et al. (2006) Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Science* 15: 2558-2567.
100. Kulkarni-Kale U, Bhosle S, et al. CEP: a conformational epitope prediction server. *Nucleic Acids Research* 33: W168-W171.
101. Greenbaum JA, Andersen PH, et al. (2007) Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *Journal of Molecular Recognition* 20: 75-82.
102. Ponomarenko JV, Bourne PE (2007) Antibody-protein interactions: benchmark datasets and prediction tools evaluation. *BMC structural biology* 7: 64.
103. Wilkins MR, Pasquali C, et al. (1996) From Proteins to Proteomes: Large Scale Protein Identification by Two-Dimensional Electrophoresis and Amino Acid Analysis. *Nat Biotech* 14: 61-65.
104. Clamp M, Fry B, et al. (2007) Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19428-19433.
105. Aebersold R, Mann M (2003) Mass spectrometry-based proteomics. *Nature* 422: 198-207.
106. Domon B, Aebersold R (2006) Mass Spectrometry and Protein Analysis. *Science* 312: 212-217.
107. Gygi SP, Rist B, et al. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotech* 17: 994-999.
108. Zhao Y, Lee WN, et al. (2009) Quantitative proteomics and biomarker discovery in human cancer. *Expert review of proteomics* 6: 115-118.

109. Schulze WX, Usadel B (2010) Quantitation in mass-spectrometry-based proteomics. *Annual review of plant biology* 61: 491-516.
110. Rifai N, Gillette MA, et al. (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotech* 24: 971-983.
111. Yates JR, Ruse CI, et al. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering* 11: 49-79.
112. Walsh GM, Rogalski JC, et al. (2010) Mass spectrometry-based proteomics in biomedical research: emerging technologies and future strategies. *Expert Reviews in Molecular Medicine* 12: null-null.
113. Edwards AM, Isserlin R, et al. (2011) Too many roads not taken. *Nature* 470: 163-165.
114. Nilsson P, Paavilainen L, et al. (2005) Towards a human proteome atlas: high-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* 5: 4327-4337.
115. Haab BB, Paulovich AG, et al. (2006) A Reagent Resource to Identify Proteins and Peptides of Interest for the Cancer Community. *Molecular & Cellular Proteomics* 5: 1996-2007.
116. Mersmann M Fau - Meier D, Meier D Fau - Mersmann J, et al. Towards proteome scale antibody selections using phage display.
117. Stoevesandt O, Taussig MJ (2007) Affinity reagent resources for human proteome detection: initiatives and perspectives. *Proteomics* 7: 2738-2750.
118. Kononen J, Bubendorf L, et al. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 4: 844-847.
119. Kingsmore SF (2006) Multiplexed protein measurement: technologies and applications of protein and antibody arrays. *Nat Rev Drug Discov* 5: 310-321.
120. Ayoglu B, Hysggmark A, et al. (2011) Systematic antibody and antigen-based proteomic profiling with microarrays. *Expert Review of Molecular Diagnostics* 11: 219-234.
121. Bordeaux J, Welsh A, et al. (2010) Antibody validation. *Biotechniques* 48: 197-209.
122. Mishra A, Verma M (2010) Cancer Biomarkers: Are We Ready for the Prime Time? *Cancers* 2: 190-208.
123. Brennan DJ, O'Connor DP, et al. (2010) Antibody-based proteomics: fast-tracking molecular diagnostics in oncology. *Nat Rev Cancer* 10: 605-617.
124. Ludwig JA, Weinstein JN (2005) Biomarkers in Cancer Staging, Prognosis and Treatment Selection. *Nat Rev Cancer* 5: 845-856.
125. Lilja H, Ulmert D, et al. (2008) Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. *Nat Rev Cancer* 8: 268-278.
126. Anderson NL, Anderson NG (2002) The Human Plasma Proteome. *Molecular & Cellular Proteomics* 1: 845-867.
127. Carlsson A, Wingren C, et al. (2008) Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays. *European Journal of Cancer* 44: 472-480.

128. Schwenk JM, Igel U, et al. (2010) Toward Next Generation Plasma Profiling via Heat-induced Epitope Retrieval and Array-based Assays. *Molecular & Cellular Proteomics* 9: 2497-2507.
129. Gold L, Ayers D, et al. (2010) Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS ONE* 5: e15004.
130. Darmanis S, Nong RY, et al. (2011) ProteinSeq: High-Performance Proteomic Analyses by Proximity Ligation and Next Generation Sequencing. *PLoS ONE* 6: e25583.
131. Uhlén M, Pontén F (2005) Antibody-based Proteomics for Human Tissue Profiling. *Molecular & Cellular Proteomics* 4: 384-393.
132. Berglund L, Björling E, et al. (2008) A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. *Molecular & Cellular Proteomics* 7: 2019-2027.
133. Berglund L, Björling E, et al. (2008) A whole-genome bioinformatics approach to selection of antigens for systematic antibody generation. *Proteomics* 8: 2832-2839.
134. Larsson M, Gräslund S, et al. (2000) High-throughput protein expression of cDNA products as a tool in functional genomics. *Journal of Biotechnology* 80: 143-157.
135. Steen J, Uhlén M, et al. (2006) High-throughput protein purification using an automated set-up for high-yield affinity chromatography. *Protein Expression and Purification* 46: 173-178.
136. Agaton C, Falk R, et al. (2004) Selective enrichment of monospecific polyclonal antibodies for antibody-based proteomics efforts. *J Chromatogr A* 1043: 33-40.
137. Andersson A-C, Strömberg S, et al. (2006) Analysis of Protein Expression in Cell Microarrays: A Tool for Antibody-based Proteomics. *Journal of Histochemistry & Cytochemistry* 54: 1413-1423.
138. Barbe L, Lundberg E, et al. (2008) Toward a Confocal Subcellular Atlas of the Human Proteome. *Molecular & Cellular Proteomics* 7: 499-508.
139. Lundberg E, Fagerberg L, et al. (2010) Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol* 6.
140. Uhlén M, Oksvold P, et al. (2010) Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 28: 1248-1250.
141. Pontén F, Jirstrom K, et al. (2008) The Human Protein Atlas--a tool for pathology. *J Pathol* 216: 387-393.
142. Götz F (1990) *Staphylococcus carnosus*: a new host organism for gene cloning and protein production. *Society for Applied Bacteriology symposium series* 19: 49S-53S.
143. Löfblom J, Wernerus H, et al. (2005) Fine affinity discrimination by normalized fluorescence activated cell sorting in staphylococcal surface display. *FEMS microbiology letters* 248: 189-198.
144. Colwill K, Graslund S (2011) A roadmap to generate renewable protein binders to the human proteome. *Nature methods* 8: 551-558.
145. Bradbury A, Velappan N, et al. (2003) Antibodies in proteomics I: generating antibodies. *Trends in Biotechnology* 21: 275-281.

- 146. Jogi A, Brennan DJ, et al. (2009) Nuclear expression of the RNA-binding protein RBM3 is associated with an improved clinical outcome in breast cancer. *Mod Pathol* 22: 1564-1574.
- 147. Ehlén A, Brennan DJ, et al. (2010) Expression of the RNA-binding protein RBM3 is associated with a favourable prognosis and cisplatin sensitivity in epithelial ovarian cancer. *J Transl Med* 8: 78.
- 148. Jonsson L, Bergman J, et al. (2011) Low RBM3 protein expression correlates with tumour progression and poor prognosis in malignant melanoma: an analysis of 215 cases from the Malmo Diet and Cancer Study. *Journal of translational medicine* 9: 114.
- 149. Jonsson L, Gaber A, et al. (2011) High RBM3 expression in prostate cancer independently predicts a reduced risk of biochemical recurrence and disease progression. *Diagnostic pathology* 6: 91.
- 150. Uhlén M (2007) Mapping the Human Proteome Using Antibodies. *Molecular & Cellular Proteomics* 6: 1455-1456.