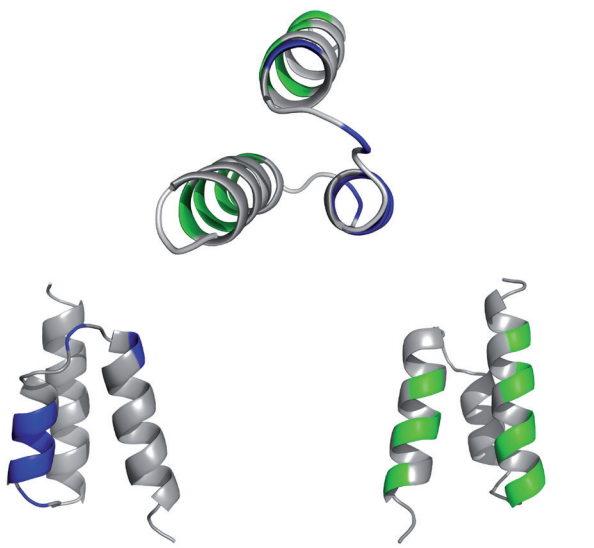
An albumin-binding domain as a scaffold for bispecific affinity proteins

JOHAN NILVEBRANT



KTH VETENSKAP OCH KONST

KTH Biotechnology

Doctoral Thesis in Biotechnology
Stockholm, Sweden 2012

ROYAL INSTITUTE
OF TECHNOLOGY

An albumin-binding domain as a scaffold for bispecific affinity proteins

JOHAN NILVEBRANT

Royal Institute of Technology
School of Biotechnology
Stockholm 2012

Stockholm 2012

Royal Institute of Technology
School of Biotechnology
AlbaNova University Center
SE-106 91 Stockholm
Sweden

Printed by E-Print
Oxtorgsgatan 9
SE-111 57 Stockholm
Sweden

ISBN 978-91-7501-569-9
TRITA-BIO Report 2012:21
ISSN 1654-2312

Johan Nilvebrant (2012): **An albumin-binding domain as a scaffold for bispecific affinity proteins.**
Division of Proteomics, School of Biotechnology, Royal Institute of Technology (KTH), Stockholm, Sweden.

Abstract

Protein engineering and *in vitro* selection systems are powerful methods to generate binding proteins. In nature, antibodies are the primary affinity proteins and their usefulness has led to a widespread use both in basic and applied research. By means of combinatorial protein engineering and protein library technology, smaller antibody fragments or alternative non-immunoglobulin protein scaffolds can be engineered for various functions based on molecular recognition. In this thesis, a 46 amino acid small albumin-binding domain derived from streptococcal protein G was evaluated as a scaffold for the generation of affinity proteins. Using protein engineering, the albumin binding has been complemented with a new binding interface localized to the opposite surface of this three-helical bundle domain. By using *in vitro* selection from a combinatorial library, bispecific protein domains with ability to recognize several different target proteins were generated. In paper I, a bispecific albumin-binding domain was selected by phage display and utilized as a purification tag for highly efficient affinity purification of fusion proteins. The results in paper II show how protein engineering, *in vitro* display and multi-parameter fluorescence-activated cell sorting can be used to accomplish the challenging task of incorporating two high affinity binding-sites, for albumin and tumor necrosis factor-alpha, into this new bispecific protein scaffold. Moreover, the native ability of this domain to bind serum albumin provides a useful characteristic that can be used to extend the plasma half-lives of proteins fused to it or potentially of the domain itself. When combined with a second targeting ability, a new molecular format with potential use in therapeutic applications is provided. The engineered binding proteins generated against the epidermal growth factor receptors 2 and 3 in papers III and IV are aimed in this direction. Over-expression of these receptors is associated with the development and progression of various cancers, and both are well-validated targets for therapy. Small bispecific binding proteins based on the albumin-binding domain could potentially contribute to this field. The new alternative protein scaffold described in this thesis is one of the smallest structured affinity proteins reported. The bispecific nature, with an inherent ability of the same domain to bind to serum albumin, is unique for this scaffold. These non-immunoglobulin binding proteins may provide several advantages as compared to antibodies in several applications, particularly when a small size and an extended half-life are of key importance.

Keywords: albumin-binding domain, bispecific, albumin, affinity protein, phage display, staphylococcal display, orthogonal affinity purification, TNF-α, ErbB2, ErbB3

Papers in this thesis

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

I Alm, T., Yderland, L., **Nilvebrant, J.**, Halldin, A., Hober, S. (2010). A small bispecific protein selected for orthogonal affinity purification, *Biotechnol. J.*, 5: 605-617

II **Nilvebrant, J.**, Alm, T., Hober, S., Löfblom, J. (2011). Engineering bispecificity into a single albumin-binding domain, *PLoS ONE*, 6(10): e25791

III **Nilvebrant, J.**, Åstrand, M., Georgieva, M., Björnmalm, M., Löfblom, J., Hober, S. Engineering of bispecific affinity proteins with nanomolar affinity for both ErbB2 and albumin, *manuscript*

IV **Nilvebrant*, J.**, Åstrand*, M., Löfblom, J., Hober, S. Development and characterization of small bispecific three-helical ErbB3/albumin-binding domains aimed at therapeutic applications, *manuscript*

* Authors contributed equally

Papers not included in the thesis

Boström*, T., **Nilvebrant*** **J.**, Hober, S. (2012). Purification systems based on bacterial surface proteins. *Protein Purification*, R. Ahmad (Ed.), ISBN: 978-953-307-831-1, InTech

Nilvebrant, J., Alm, T., Hober, S. (2012). Orthogonal protein purification facilitated by a small bispecific affinity tag. *J. Vis. Exp.* (59), e3370

Nilvebrant, J., Dunlop, C., Wurch, T., Falkowska, E. Helguera, G., Piccione, E., Reichert, J. (2012). Meeting report of the 2011 IBC's 22nd Annual International Conference on Antibody Engineering and Antibody Therapeutics, *mAbs*, 4(2): 153-181

Nilvebrant, J., Kuku, G., Björkelund, H., Nestor, M. (2012). Selection and *in vitro* characterization of human CD44v6-binding antibody-fragments, *Biotech. Appl. Biochem.*, 59(5): 367-380

* Authors contributed equally

“A classic is something that everyone wants to have read and nobody wants to read”

– Mark Twain

CONTENTS

Common abbreviations

ABD	Albumin-binding domain derived from streptococcal protein G (G148-ABD3)
ALB8-GA	Albumin-binding domain derived from protein albumin binding of *F. magna*
CDR	Complementarity determining region
ErbB2	Epidermal growth factor receptor 2 (HER2)
ErbB3	Epidermal growth factor receptor 3 (HER3)
Fab	Fragment antigen binding (of an antibody molecule)
FACS	Fluorescence-activated cell sorting
Fc	Fragment crystallizable (of an antibody molecule)
FcRn	Neonatal Fc-receptor (Brambell receptor)
HSA	Human serum albumin
IgG	Immunoglobulin G
K_D	Dissociation equilibrium constant
mAb	Monoclonal antibody
scFv	Single-chain fragment variable
SPA	Staphylococcal protein A
SPG	Streptococcal protein G
SPR	Surface plasmon resonance
TNF-α	Tumor necrosis factor-alpha

Denna avhandling är tillägnad mormor Ulwa

INTRODUCTION

The aim of this thesis was to apply protein engineering approaches to develop bispecific binding proteins based on a small albumin-binding domain employed as a template. These efforts represent a field denoted alternative scaffold engineering, in which non-immunoglobulin proteins are developed which may complement, or potentially outperform, antibodies in various applications based on molecular recognition. Correspondingly, the main focus of the introductory part of the thesis will be on protein engineering for molecular recognition and generation of affinity proteins. To understand how the overall objective in this work was approached in principle and experimentally, a background is first given to protein engineering in general followed by an in-depth discussion of so-called *in vitro* display systems.

This thesis is focused on bacterial albumin-binding domains, which generally are small, chemically and physically stable proteins with favorable properties for protein engineering, such as a high tolerance for amino acid substitutions in surface exposed positions. To become familiar with these domains and the large number of alternative protein scaffolds that is available and used for affinity protein development, one introductory section is devoted to scaffold proteins. This section also contains a comparison of these small scaffolds to conventional antibodies and their fragments. The albumin-binding domain and its homologs are discussed in detail in a separate section, which includes a description of the specific characteristic of this domain in its ability to associate with albumin, an abundant serum protein. This property can be used to extend the plasma half-lives of proteins fused to this domain or potentially of the domain itself. The albumin-binding property is crucial for the scope of this thesis in which efforts have been devoted to develop variants of the domain that contain, in addition to the native albumin-binding site, a second binding surface to make them bispecific. Small protein domains can be used as affinity tags for protein purification and, when combined with albumin binding in a bispecific format, the second high affinity targeting ability can provide a new molecular format with potential use in therapeutic applications.

1. PROTEIN ENGINEERING

Protein engineering includes methods to modify proteins, particularly by introducing mutations into a parent sequence, followed by functional screening of the resulting variants (Figure 1A). This field is commonly divided into rational- and combinatorial protein engineering, respectively. Rational protein engineering principles generally rely on hypotheses predicting how certain modifications may affect a trait of the protein, often its affinity for a specific target protein, but also for example folding, structural stability or, for enzymes, catalytic properties. In contrast, combinatorial strategies rely on screening of larger collections of mutated proteins to select variants with the desired properties. *In vitro* selection methods are used to screen so-called libraries generated by mutagenesis of a parent protein sequence. They utilize a physical linkage between the genotype and the phenotype of each member in the sequence library (Figure 1B) to select desired variants from such large collections, rather than to screen a more limited number of variants generated by rational means.

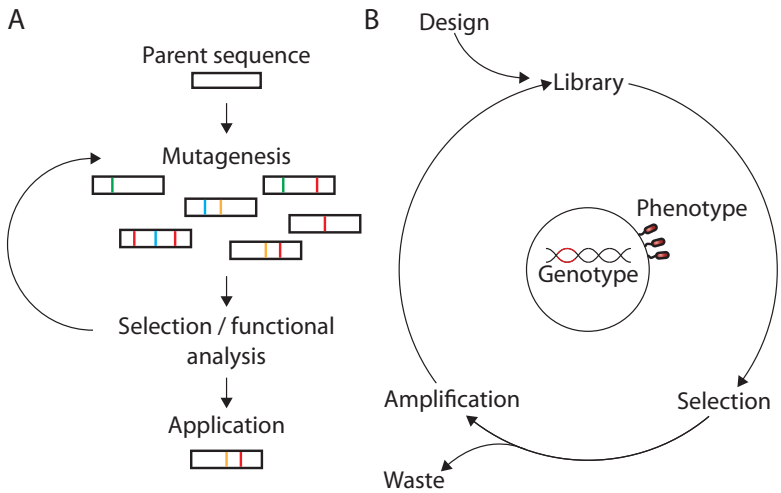


Figure 1. Strategies for protein engineering. A. Mutagenesis of a parental sequence and subsequent functional analysis is performed, in iterative cycles if desired, to modify a protein for an intended application. This is either done on individual, isolated variants one-by-one or in mixture (i.e. in a library format), which requires a methodology to provide a physical link between the genotype (gene encoding the protein) and phenotype (protein) during the functional analysis, as illustrated in **B**. Many *in vitro* selection systems rely on such a genotype-phenotype link to display variants present in so-called sequence libraries. This facilitates selection and amplification of members with desired properties from a background on non-functional variants.

Natural proteins consist of one or more polypeptides, which are polymers of amino acids of different length and combinations that fold into functional entities. There are 20 amino acids commonly found in proteins and the proteins are generally encoded by nucleic acids carried in the genomes of their host organisms. Billions of years of evolution have produced an abundance of proteins in living systems with highly specialized functions such as catalysis (e.g. enzymes), structural support (e.g. collagen), transport (e.g. hemoglobin), motility (e.g. actin), electron transport (e.g. cytochromes) or defense against infection (e.g. antibodies). The width in functional capacity displayed by proteins makes them attractive for many applications in industry, as tools for research and diagnostics or as therapeutics. During the last decades, progress in recombinant deoxyribonucleic acid (DNA) technology such as the development of the polymerase chain reaction (PCR) (Mullis 1986; Saiki 1985), methods for cutting and pasting genetic elements (Arber 1969; Weiss 1967), methods for DNA sequencing like the Sanger chain terminating method (Sanger 1977) and the use of various host cells as specialized production units, has facilitated the expression and purification of exogenous proteins for detailed structural and functional characterization. Inspired by the functional diversity and ingenuity of natural proteins to engage in a wide range of tasks, the field of protein engineering emerged (Ulmer 1983). Protein engineers aim to modify or design proteins in order to improve their properties, to generate novel features or to better understand basic principles of protein folding and function (Brannigan 2002).

Many traits of a protein can be addressed by sequence modification or by chemical manipulation. Some examples include stability, solubility, expression level, catalytic activity, immunogenicity, pharmacokinetics or molecular recognition. The common feature of most of these strategies is to introduce mutations (substitutions, insertions or deletions) in the amino acid sequence and analyze the influence on a specific property. Functions may also be added for example through genetic fusion of two or more proteins. The correlation between sequence and function is often illustrated as a protein fitness landscape, where all possible protein sequences are represented as points on a two-dimensional grid where neighboring sequences only differ by one substitution (Smith 1970). A simplified view of how single- or combinations of substitutions may affect a trait is illustrated in figures 2A and B. Each protein can be assigned a fitness, which for natural evolution involving living organisms would be the individual's ability to survive and reproduce, but in protein engineering rather could be defined by the objective of the experiment (the desired molecular trait). A simplified fitness landscape is illustrated in figure 2C. Many protein engineering efforts can be viewed as optimization processes or "adaptive walks" to reach the highest possible fitness in a given sequence space (Romero 2009). Excitingly, since these processes can be performed under controlled conditions in a laboratory, new and "non-natural" definitions of fitness can

also be explored, such as binding to a defined region of a protein or stability to non-natural physical or chemical conditions.

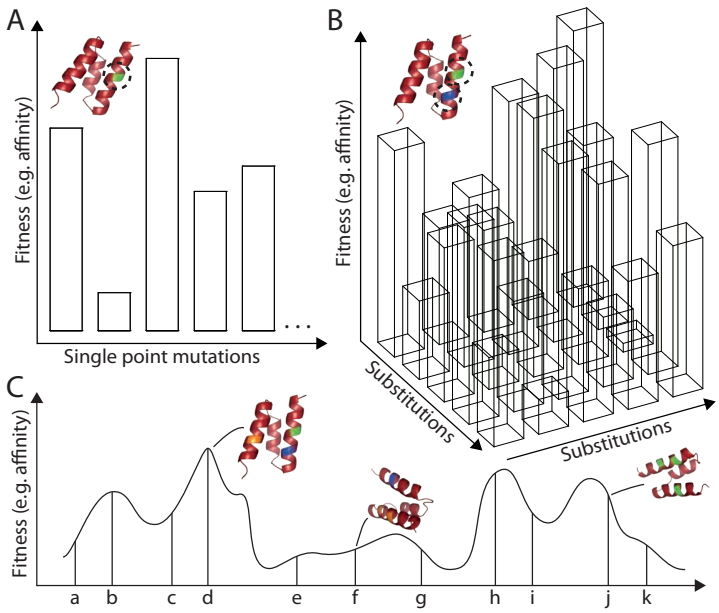


Figure 2. Substitutions, combinations of substitutions and protein fitness. A. Mutation, here represented by substitution of a single amino acid to other residues, may affect fitness to various extents (exemplified here by affinity for a target protein). A high signal on the y-axis represents a beneficial sequence change. **B.** Simultaneous substitutions at several positions result in more complex patterns where some combinations may be beneficial and others cannot be well tolerated. **C.** A schematic fitness landscape drawn in two dimensions to provide a simplified view. Candidate d, located at the highest possible fitness, contains an optimal combination of substitutions. Clones corresponding to this maximum or even to local maxima may not be represented in a library, as illustrated for example by the library members h and j. The challenge for protein engineering is to identify fitness maxima and, ideally, find the global fitness maximum.

As mentioned above, many experimental strategies for protein engineering have been described and they are often grouped into either rational or combinatorial methods. A clear distinction does not exist as elements from both groups are often combined in library-based methods. Several methods have a random mutagenesis component and those techniques are commonly grouped together with the combinatorial systems. Methods that expand the protein complexity beyond the naturally occurring set of 20 amino acids are also available, for example the recombinant incorporation of more than 40 different non-natural

amino acids by the use of designed orthogonal aminoacyl-tRNA-synthetase/tRNA pairs in several host organisms (Liu 2010), or methods that rely on chemical incorporation of non-natural characteristics by solid phase peptide synthesis (SPPS) (Albericio 2004; Dawson 2000). Modifications at the sequence level of a protein rather than chemical or post-translational modifications will be discussed below.

1.1 Rational protein engineering

Rational protein engineering can be regarded as hypothesis-driven and specified modification of proteins, i.e. defined positions are generally targeted to address a certain function or trait. It generally requires detailed knowledge, preferably structural data, to provide assumptions about functional relationships. Many of the rational strategies for protein modification were first applied to enzymes (Bloom 2005; Quin 2011), but the increasing number of solved structures of proteins and protein-protein complexes, especially during the last decade (www.rcsb.org, (Berman 2003; Berman 2007)), as well as advances in DNA sequencing and characterization of newly identified genes (Mardis 2008) will facilitate exploitation of proteins where such information has been missing. Still, small changes in sequence or structure can have large effects on protein function. Thus, prediction of beneficial sequence alterations is a major challenge even with structural information available (Romero 2009).

Changes in the amino acid sequence of a protein are normally introduced on the genetic level by altering a codon to affect the corresponding amino acid after translation. The genetic code consists of 64 (4^3) possible triplets corresponding to combinations of four different nucleotides: adenine (A), cytosine (C), guanine (G) and thymine (T). In the genetic code, some amino acids are represented by several codons and two are represented only once. This redundancy influences the relative ability to modify different amino acids and may also cause mutational biases in some protein engineering experiments that rely on random mutagenesis (Wong 2007), as exemplified later in this section.

A straightforward strategy to elucidate the functional contributions of specific residues in a protein, for example in binding to another protein, is by site-directed mutagenesis (Hutchison 1978), a method where a gene is mutated to specifically alter a single or several amino acids in the encoded protein. All mutagenesis methods require a suitable assay to screen a desired function or property of the variants that are generated. A common procedure to identify regions in the protein associated with the trait to be engineered is to substitute a chosen set of amino acids to alanine one at a time, a technique known as alanine-scanning (Cunningham 1989; Matouschek 1990; Schreiber 1993). By determining the function of

the substituted proteins, compared to the wild type, the influence of each residue may be deduced. However, in the interpretation, detrimental functional effects due to an overall loss of structure rather than alterations at, for example, a binding interface, need to be analyzed by complementary methods such as circular dichroism spectroscopy to rule out such effects. Alanine is considered a rather neutral residue at a functional interface due to the small size and non-polar nature of its methyl side chain. In contrast to the smaller residue glycine, alanine is both chiral and rigid and therefore does not influence the backbone structure to the same extent. A similar, but much more laborious method for a rational approach, is site-saturation mutagenesis, where a targeted residue is substituted to all 19 alternative amino acids (Seeburg 1984; Wells 1985). To increase the throughput and make analysis of variants with multiple substitutions possible, alanine scanning has been explored in a combinatorial fashion, similar to the methods described in the following section (Gregoret 1993; Kouadio 2005; Morrison 2001; Weiss 2000). One of the most comprehensive studies of a protein-protein interaction has been collected from mutant forms of human growth hormone where 20 substitutions at each of 35 sites have been analyzed (Pal 2006). Interestingly, this example indicates that sequence conservation across species may be a poor predictor of the actual importance of a residue at an interface and that many chemically conservative substitutions may not be tolerated.

Structural stability, for example to heat, represents a popular trait to address by rational engineering approaches (O'Fagain 2011). Incorporation of disulfide bonds into a protein structure (Matsumura 1989), modification of the hydrophobic core (Munson 1996), substitutions at the protein surface (Predki 1996) or exploration of consensus motifs derived from homologous sequences (Cole 2011) can improve stability. For very stable proteins, an initial destabilization may be required to identify potentially stabilizing modifications (Linhult 2004). Moreover, an interesting relationship between stability and evolvability has been described. Introduction of stabilizing mutations can increase the robustness of the protein to accommodate additional functionally beneficial, yet destabilizing alterations, and thereby more changes can be tolerated (Bloom 2006; Tokuriki 2009). Many random mutations are destabilizing (Guo 2004; Steipe 1999) and it is often assumed that a proper folding and a sufficient stability are required for function, especially for affinity proteins. As a consequence, the use of a stabilized protein as the starting point for engineering may open up new adaptive routes for improvement (Lappe 2009; Romero 2009). Structural stability has also been related to better resistance to aggregation and tolerance to extended periods of use or storage, which is preferred in many industrial and biomedical applications (Buchanan 2012; Cole 2011; Forrer 1999; O'Fagain 2011).

Additional approaches have been described to engineer protein function. Some examples from engineering of domains of staphylococcal protein A (SPA) and streptococcal protein G (SPG), two extensively studied proteins that will have a central part later in this thesis, are given below. Transplantation of a function from one protein to another has been demonstrated by the grafting of an elastin-derived loop (Reiersen 2000) or a constrained protease inhibiting peptide (Bratkovic 2006) onto an immunoglobulin-binding domain of SPA. Genetic fusion is another widespread way to modify a protein, which is commonly performed to facilitate protein expression, purification, and to improve or add further functions. Since the early examples of gene fusion for protein purification were reported (for example (Uhlen 1983)), several such systems have been described for domains of SPA and SPG (Boström 2012), to complement other common gene fusion strategies for protein purification (Hedhammar 2005).

Some striking examples of how a few substitutions can have a profound impact on structure and function come from studies of domains of SPG as models for protein folding. It has for example been demonstrated that single substitutions can switch a protein domain from being an immunoglobulin-binding protein of a 4ß+α structure to an albumin-binding protein of a 3α structure (Figure 3, (Alexander 2009; He 2012)). Here, separate albumin binding- and immunoglobulin-binding domains of SPG with no significant similarity were first engineered to show a high degree of sequence identity with retained original functionalities (Alexander 2007). The mutational pathway between the two folds could be studied thanks to the reduced complexity of the sequence space that separated them. To move one substitution at a time requires that there are a continuous number of variant proteins that retain function (Smith 1970).

Despite the difficulty of rational protein engineering, several engineered proteins have gained a wide use, as exemplified by subtilisin and insulin. Subtilisin has been extensively modified to tolerate bleach in laundry detergents, for improved stability and catalysis (Bryan 2000). Modified analogs to human insulin with tailored pharmacokinetic profiles, as a result of engineered homo-oligomerization (Zn^{2+}-coordination) and precipitation tendencies, have also been developed (Vajo 2000).

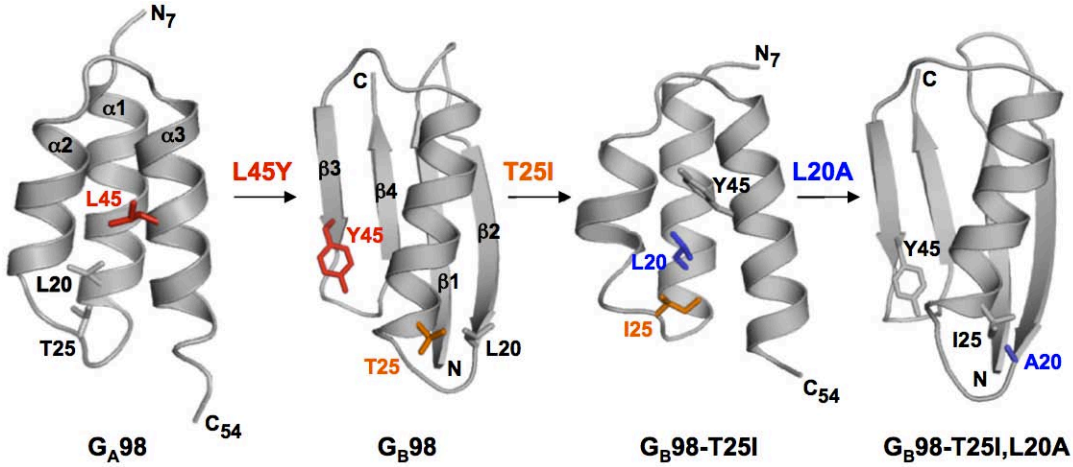


Figure 3. Examples of substitutions that switch protein folds. Substituted residues are highlighted in this very specific, and probably uncommon, example of fold switching. Reproduced from (He 2012) with permission from Cell Press.

The ultimate challenge for rational protein engineering is *de novo* design of proteins. However, even for the design of a relatively small protein of 100 amino acids, there are 20^{100} possible sequences, which is more than there are atoms in the universe (Romero 2009) and the combined mass of one copy of each variant would exceed the mass of the sun a trillion-fold (Magliery 2004). It has been suggested that evolution only has explored an infinitesimal fraction of this massive sequence space (Mandecki 1998) and, even though the accuracy of the above statement has been argued (Dryden 2008), for an experimental approach to design a protein to be conceivable, even by computational means, the complexity needs to be reduced. Nevertheless, designed proteins that experimentally correspond to their designs have been reported, for example a zinc finger protein that was inspired by a natural fold (Dahiyat 1997). Several achievements have followed (Floudas 2006; Pantazes 2011), including the popular designed four-helix bundle proteins (Hecht 2004; Kamtekar 1993; Regan 1988). Despite these accomplishments, protein design is still in its infancy and protein structure prediction constitutes a major part of the *de novo* design problem (Floudas 2006). To form a protein for a practical application by *de novo* design still requires much more work as compared to more conventional methods that start from a functional protein as a template. However, a few astonishing examples of designed functionality into new, designed protein frameworks have been reported (Fleishman 2011; Koder 2009). Yet, due to the inherent limitations in rational protein engineering and design, combinatorial methods have emerged as powerful alternatives.

1.2 Combinatorial protein engineering

In contrast to the rational approach, combinatorial (library-based) methods rely on sequence libraries and selection- or screening methods to find desired protein variants (Figure 4). These methods can mimic the natural evolution process at increased speed, especially if mutations are gradually introduced during the procedure, and are often referred to as directed evolution. Iterative rounds of selection and screening can be applied to generate new proteins and select rare variants with increased fitness or a new desired property. The selection methods are based on a physical link between the phenotype of the protein variant and the genetic information that encodes it during selection, which facilitates the identification of a selected protein with a preferred characteristic. Experimental platforms handling such collections, denoted selection systems, will be covered in more detail in a later section. Common methods to generate diversity in sequence libraries and considerations related to this are discussed below.

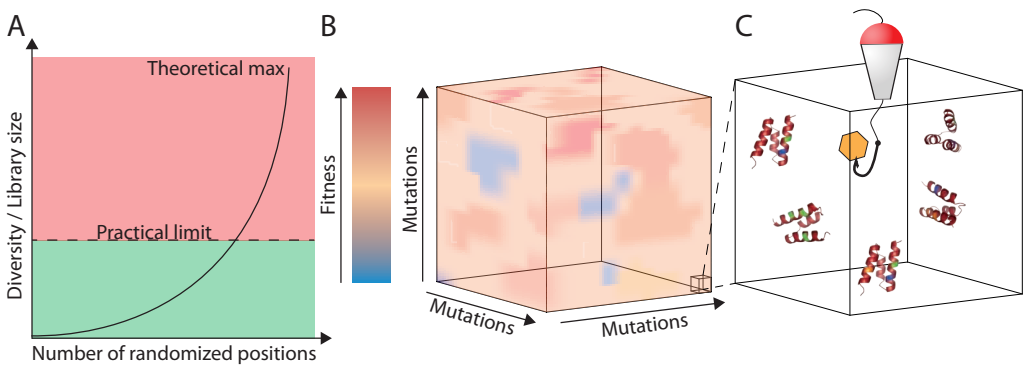


Figure 4. Library complexity and selection systems. **A.** Library complexity quickly increases as more positions are diversified simultaneously. Each selection system has a practical limit for the diversity it may be able to screen (see next section for examples). **B.** A library can be viewed as a volume of sequence variants of different fitness, but usually all theoretical variants are not present during the experiment. **C.** A selection system, illustrated as a float hooked up to a target protein, probes this sequence volume for variants with sought properties. Each protein variant in the library has a link to its corresponding genotype, the appearance of which depends on the choice of selection system (not shown in the picture).

To construct gene pools that encode more than one amino acid in the same position, DNA synthesis can be applied to produce mixtures of oligonucleotides by means of so called degenerate codons. For example the codon NNN (where N represents an equal blend of A, T, C and G) includes all 64 codons whereas NNK (K = G or T) limits the diversity to 32 codons, which still represent all possible 20 amino acids at

least once. Combinatorial protein engineering allows exploration of structure-function relationships that are not easily approached by hypotheses or modeling, such as cooperative contributions of several residues to a trait (Brannigan 2002; Wells 1990). It can also facilitate a disconnection of a protein from its *in vivo* function to be engineered for a new purpose (Shivange 2009). The principles behind the use of sequence libraries used in combinatorial protein engineering are inspired by the natural diversity found for example in the molecules of the adaptive immune systems of vertebrates, and in similar systems among invertebrates and certain types of peptide neurotoxins (Boehm 2011; Cole 2011; Danilova 2012; Olivera 1990; Pancer 2004; Sollod 2005; Vita 1995). A library is a collection of sequences and can be seen as the volume surrounding a progenitor sequence in the protein fitness landscape (Steipe 1999). Random changes spread out over the entire sequence can be introduced by different mutational strategies. Mutagenesis may also be targeted to specific regions or even particular positions in the protein. Libraries are generally divided into natural or synthetic depending on their origin. In terms of antibody libraries derived from natural sources, they can either be naïve, i.e. not intentionally enriched for any specific binding property, or derived from hosts immunized with or exposed to a protein of interest. Synthetic libraries generally allow better control of the composition and diversity (Koide 2009b).

Investigation involving multiple simultaneous mutations increases the library size exponentially and the likelihood of finding a desired variant increases with the library size since a larger range of adaptive scenarios can be explored (Romero 2009). However, the conceivable library size is limited by the selection or screening system, which is referred to as the “numbers problem” (Reetz 2008). For synthetic libraries, a careful library design, which often incorporates rational considerations, is important to generate a rich diversity that is enriched in potentially active variants to maximize the functionality in the limited sequence space that can be sampled experimentally (Shivange 2009; Wong 2007). A common way to limit library size is by careful selection of degenerate codons at the sequence positions targeted for randomization in the library using oligonucleotide directed mutagenesis (Dalbadie-McFarland 1982). As mentioned earlier, NNK instead of NNN reduces the complexity on the nucleotide level by 50 % and removes two of three possible, and most often undesired, stop codons, but is still redundant because 32 possibilities encode 20 amino acids. Therefore, NNK is more widely used than NNN. Another popular codon to limit redundancy is NDT (D = A, G or T), which encodes 12 different amino acids with a balanced mix of properties in 12 possible codons (Reetz 2008). Unfortunately the KTH codon (H= A, C or T) is not very useful in this context since it only encodes three amino acids by its six possible codons. More useful alternatives biased towards hydrophilic, hydrophobic, charged etc amino acids are available. Based on amino acid frequencies observed in the interfaces of protein-protein complexes, limited

diversity strategies that include residues common at binding interfaces have been proposed, such as a 4-letter code (Tyr, Ser, Ala, Asp) (Fellouse 2004) and a binary code (Tyr and Ser) (Fellouse 2005), generally with an emphasis on tyrosine (Koide 2009b). Even though a limited diversity often is a practical advantage and despite the fact that functional binding proteins have been selected from limited diversity libraries, some reports imply that a richer amino acid diversity is superior and that, generally, a sparse search of a large library may be better than a thorough exploration of a smaller one (Birtalan 2008; Gilbreth 2008; Hackel 2010; Munoz 2008; Nov 2011). Several computational tools have also been developed to aid library design, reduce redundancy and to optimize choices of degenerate codons (see (Craig 2010; Firth 2008; Wong 2007) for examples). Methods based on mixtures of degenerate primers have been proposed to limit the redundancy (Tang 2012). However, a general consensus on a best design practice has not yet been reached. Full diversity is considered to enable better shape and chemical complementarity and may increase tolerance to alterations (Koide 2009b), but it may also include many non-functional sequences. Well-designed libraries with reduced complexity can increase the frequency of binders, but may be unable to provide equally high affinity binders as a full diversity library (Hackel 2010). A hybrid design biased towards some key amino acids is another interesting approach (Fellouse 2007; Koide 2009b) and methods also exist to fully control the diversity using specified trinucleotide building blocks in the DNA-synthesis (Neylon 2004; Van der Brulle 2008; Virnekas 1994). However, these methods have not yet become the general option due to the high cost associated with them. A similar method to encode all amino acids at a specific position without redundancy bias is the so-called theoretical maximum efficiency (MAX) randomization (Hughes 2003). It is more complex than regular degenerate codons and relies on assembly and ligation of a mixture of primers onto a template oligonucleotide (Hughes 2003; Tang 2012). Several alternative randomization methods have been reported (Neylon 2004; Shivange 2009), one of the more straightforward options is split and mix synthesis of oligonucleotides (Glaser 1992; Lahr 1999).

Another group of library construction strategies rely on mutagenesis of random sequence positions. The most common method to create diversity based on a single gene is error-prone PCR (Cadwell 1994; Leung 1989), because of its experimental simplicity (Kazlauskas 2009; McCullum 2010). Error-prone PCR is typically performed using conditions that reduce the fidelity of a DNA polymerase during DNA synthesis or by using low fidelity enzymes. The error rate is thereby enhanced and mutagenesis can be restricted to a region of interest in a template using specific primer pairs. It is generally considered best to keep the mutation rate low to maintain function in mutated variants of a template, which is intended to be improved, and rather incorporate more mutations in successive cycles (Steipe 1999). The error-rate may

be controlled by several factors such as the introduction of small amounts of Mn^{2+} instead of the natural cofactor Mg^{2+}, by adjustment of the concentration of free nucleotides in the reaction and the number of amplification cycles (Cirino 2003; McCullum 2010). PCR-based mutagenesis is often limited, and sometimes biased, by the redundancy and organization of the genetic code and properties of the polymerases (Neylon 2004). However, recent strategies using a mixture of enzymes have limited this bias (Cadwell 1994; Vanhercke 2005). The pool of mutated genes can be used as a library and undergo selection, or the mutagenesis strategy may be used in consecutive cycles of mutation and selection.

Error-prone rolling-circle amplification has also been described as a simple mutagenesis technique that circumvents the need for cloning steps (Fujii 2004, 2006). A related approach is to use mutator strains with deficiencies in the DNA proofreading or editing machinery, which gives rise to unusually high rates of spontaneous DNA mutagenesis (Blagodatski 2012; Greener 1996; Low 1996; Nguyen 2012). A recent development, which has not yet been used for directed evolution applications, is induced expression of antisense RNAs to silence genes involved in DNA repair (Nakashima 2009). A drawback with such *in vivo* methods is that mutations are introduced throughout the entire plasmid, which may cause unexpected effects that are not directly linked to mutations in the targeted gene, such as altered expression rate or unrelated advantages gained through some mutations. DNA shuffling (Stemmer 1994) is a popular technique to combine existing (or generated) diversity in new ways by randomly fragmenting and recombining homologous sequences. This mix and match of successful variants, which may already have been selected during evolution, may yield a higher ratio of functional sequences as compared to random mutagenesis of a single template (Harayama 1998; Joern 2003). Other examples based on recombination (Voigt 2002), circular permutation (Neylon 2004; Qian 2005) or random insertion/deletion of sequences of a defined length (Murakami 2002; Murakami 2012) are available. Several systems can be combined or alternated to incorporate additional diversity.

Mutations and selections are powerful tools to explore the vast sequence space and complex nature of protein function in order to generate proteins with new functions or modify existing proteins (Romero 2009). Although fully random mutagenesis methods face enormous library sizes and do not rely on a starting structure or function (Golynskiy 2010; Kazlauskas 2009), a few amazing examples of new functions identified from such libraries have been reported. For example, the continuous evolution or selection of RNA-ligases (Seelig 2007; Wright 1997) and ATP-binding proteins (Keefe 2001). However, a combination of a well designed combinatorial library and a suitable selection system has turned out to be a more reliable means to generate affinity proteins or peptides.

2. *IN VITRO* SELECTION SYSTEMS

A number of *in vitro* selection systems have been developed to select peptides or proteins with desired properties from large libraries of variants. The function of the selection system is to enable isolation of desired library members based on a specific trait, and thereby co-select the encoding DNA to facilitate amplification and identification. Therefore, the most important property of all selection systems is a link between the phenotype (displayed peptide/protein) and the corresponding genotype (DNA or RNA) and this link can be either physical or spatial. The emphasis in this introductory section is on selection of affinity proteins from combinatorial libraries for binding to a target protein. The development of selection systems was epitomized by the invention of phage display almost 30 years ago (Smith 1985), which since then has inspired the development of several clever linking- or encapsulation strategies to connect genotype and phenotype (see Figure 5 for examples). The systems are commonly divided into cell-dependent (mainly phage- and cell-surface display), cell-independent (essentially mRNA and ribosomal display) and non-display systems (including protein fragment complementation assays), examples of which are discussed below. This section discusses general advantages of selection systems as well as details of some selection systems for the interested reader. Phage display has been used in all papers of this thesis. Staphylococcal display was applied in papers II and III.

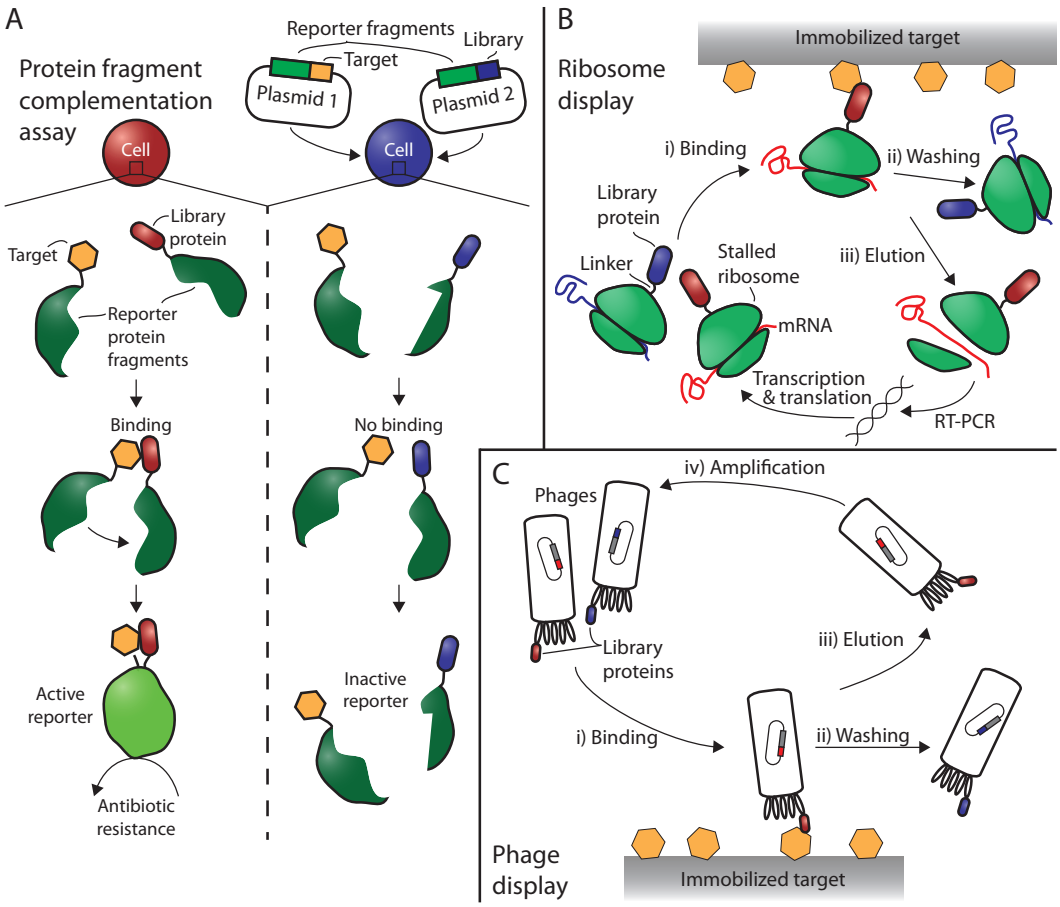


Figure 5. Examples of *in vitro* selection systems. A. A protein complementation assay (PCA), which is an example of a non-display system. **B.** Ribosome display, a cell-independent system. **C.** Phage display, the most common platform that depends on the use of cells. All systems rely on a physical link between the genotype and the phenotype of each protein variant. Functional library proteins, i.e. variants binding to a target protein (yellow), are indicated in red and non-functional variants in blue.

There is a wide range of applications of display technologies (Hoess 2001; Li 2000; Li 2002; Pelletier 2001; Sergeeva 2006) and for a long time the selection of antibodies or alternative protein binders from combinatorial libraries, displayed on phages in particular, has successfully generated affinity proteins for research, diagnostics and therapeutics (Bradbury 2011; Lofblom 2011b; Rader 1997; Wurch 2012). All methods have common steps based on diversification, functional expression and selection followed by amplification of positive clones (Figure 5). The selection of library candidates for binding to a target protein is often performed in successive rounds in a process called biopanning. Different systems have

certain advantages and limitations, the most important of which are mentioned below, which may direct interest to a specific system depending on the task at hand. Direct access to the gene encoding the selected affinity protein, provided by the genotype-phenotype linkage, is a major advantage of all *in vitro* display platforms. It does not only provide a renewable source of the affinity protein, but also means for dimerization, tag-fusion or other desired manipulations (Bradbury 2011). Preferably, a selection system should not suffer from biases such as variation in amplification rate or expression level and have potential both for monitoring enrichment and for automation (Hoogenboom 2005). However, it is also an advantage if selections can be performed without need for special resources, instrumentation or require complicated procedures (Geyer 2012). A large library is considered to give a higher chance of finding a high affinity clone (Bradbury 2004; Gold 2001; Lancet 1993; Perelson 1979), and the ability to screen large repertoires, or conveniently introduce mutations between selection rounds, is advantageous. On the other hand, systems with ability to manage the largest libraries, i.e. the cell-free methods discussed below, have not become dominating since their introduction roughly 15 years ago. The challenge of handling large libraries and more demanding protocols often make other, complementary, methods more attractive. Wide accessibility, low requirements for specific instrumentation and relative ease of use makes phage display an established platform that often fulfill all requirements.

The focus in this section is on development of novel affinity proteins. Even though libraries of short peptides are frequently used, this is not discussed here. There are also several examples where the affinity of a natural ligand for its receptor has been improved or analyzed using different selection systems (for example phage display to study human growth hormone (Bass 1990; Lowman 1991), yeast display to select high affinity variants of an epidermal growth factor (Cochran 2006) or ribosome display to map functional residues of the Nogo receptor (Schimmele 2005)). Initial validations of selection systems have commonly been performed using peptides, various antibody fragments or libraries thereof. More recently, applications with alternative binding proteins have become more common. Scaffolds for molecular recognition, some of which are mentioned in examples in this section, are covered in more detail in section 3. The current part will discuss selection systems with a primary focus on the cell-dependent systems, which are the most extensively used systems and include the methods used in the work in this thesis. In general, the phenotype is represented by a peptide or protein and the genotype by a nucleic acid. Nevertheless, systems exist where the genetic material itself constitutes the phenotype, such as the aptamers (Ellington 1990; Tuerk 1990), or where pure chemical diversity is bar-coded by DNA-tags (Mannocci 2011).

A selection for affinity is generally based on an incubation of the library with the target molecule, where high affinity clones will eventually outperform weaker binders and non-binders. Since affinity is a central concept in this thesis, it is explained in some detail here. Many biological functions of proteins and applications of engineered proteins depend on their binding to other molecules, and in particular to other proteins (Sidhu 2012; Uhlén 2008). Affinity is a measure of the strength of an interaction between two molecules, for example the proteins A and B, and is expressed as the association equilibrium constant K_A (M^{-1}) (Equation 1). [A], [B] and [AB] are the concentrations of the free proteins and the protein complex, respectively, at equilibrium under standard conditions and a 1:1 stoichiometry. Affinity between proteins is often expressed as the dissociation equilibrium constant K_D (M), which is the inverse of K_A (Equation 1). Furthermore, the equilibrium constants can be related to the dissociation rate constant (k_d; s^{-1}) and the association rate constant (k_a; $M^{-1}s^{-1}$), which describe the kinetics of complex formation between A and B (Equation 1). Affinities (K_D) in natural systems range from the order of micromolar (major histocompatibility complex-T-cell receptor complex), nanomolar (typical receptor-ligand interaction) to femtomolar (biotin-streptavidin) (Wang 2004; Wilchek 2006). The sufficient or optimal affinity depends on the intended application of the affinity protein. The Gibbs free energy of reaction (ΔG, $Jmol^{-1}$) between the bound and free states is the driving force for binding of two proteins. It is related to K_A according to equation 2, where R is the general gas constant ($JK^{-1}mol^{-1}$) and T is the absolute temperature (K). Equation 3 describes the relation between ΔG, enthalpy (ΔH), entropy (ΔS) and temperature and may sometimes help in understanding why certain interactions have a high affinity.

$$K_A = \frac{[AB]}{[A][B]} = \frac{1}{K_D} = \frac{k_a}{k_d} \quad (1)$$

$$\Delta G = -RT\ln(K_A) \quad (2)$$

$$\Delta G = \Delta H - T\Delta S \quad (3)$$

To obtain binders with higher affinities, for example in selections aimed at affinity maturation, the target concentration is lowered and the incubation time may need to be extended in order to reach equilibrium (Hulme 2010; Levin 2006). The increased stringency achieved by lowered target concentration may lead to an increased proportion of non-functional background binders relative to the desired clones as the number of selected high-affinity clones becomes small. A less stringent selection round after a more stringent one may resolve such problems when the number of copies of rare high affinity clones has increased (Pluckthun 2012). Another advantageous strategy to select high affinity clones is so-called off-rate selection (Hawkins 1992; Zahnd 2010). Here, the library is incubated with labeled target, washed and

incubated for a suitable time period in an excess of unlabeled target (Boder 1998). Faster dissociating binders will to a greater extent rebind soluble competitors as compared to binders with slower off-rates that will remain bound to the labeled target and, thus, become selected. Another motivation for off-rate selection is that association rate constants for protein-protein complexes generally fall in a very narrow window between 10^5 and 10^6 $M^{-1}s^{-1}$ (Northrup 1992; Schreiber 2002), which leaves more room for improvement of the dissociation rate to improve the K_D. Rational engineering for improved association rate has been reported, for example by Selzer et al. (Selzer 2000) and Marvin et al. (Marvin 2003), although not nearly as frequently as off-rate improvements. Generally, the first selection round should emphasize yield - i.e. all favorable clones should be recovered at the expense of collecting neutral clones. After re-growth, when several copies of selected clones are expected to be present, smaller fractions of clones can be collected under more stringent selection pressure. A monovalent display format is generally preferred for identification of the clones with the highest affinities (O'Connell 2002; Qi 2012) and to avoid avidity effects, which cause an apparent increase in functional affinity observed when multiple monovalent interactions combine synergistically.

2.1 Cell-dependent systems

Cell-dependent systems mainly include phage display and cell-surface display. Phage display is the most extensively used system and it requires a bacterial host cell for replication of phage particles displaying a protein library. The cell-surface display systems use the host cell as a compartment to encapsulate the genotype of a protein displayed on the cell surface. Yeast display is a widely used cell-surface display system for library applications and similar platforms have also been developed for gram-negative *Escherichia coli* and gram-positive *Staphylococcus carnosus*. The main disadvantage with cell-dependent systems is the requirement for transformation of the library to cells, which sometimes limits the library size by the transformation frequency of the host cells. Issues of practical handling of larger libraries may also limit the manageable library sizes of these systems. However, the technical simplicity of phage display and possibility to use fluorescence-activated cell sorting (FACS) to monitor cell-displayed libraries have led to a widespread use of these system and made them the dominating techniques for *in vitro* selection. Systems utilizing for instance retrovirus (Urban 2005) or baculovirus (Crawford 2006) have been conceived, as well as protein display on other cells than bacteria and yeast, for example on insect cells (Ernst 1998) or mammalian cells (Ho 2006; Wolkowicz 2005; Zhou 2010). However, this section will be focused on better-known systems.

2.1.1 Phage display

Filamentous phages have been used for genetic experiments since the 1970s and in the mid 1980s George Smith demonstrated how phages that displayed a fragment of the bacterial enzyme EcoRI could be enriched from a background of wild-type phages (Smith 1985; Smothers 2002). Soon thereafter, the first random peptide libraries fused to phage coat protein III (pIII) were reported (for example (Cwirla 1990) and (Scott 1990)). Those initial achievements were quickly followed by display of growth hormone variants (Bass 1990). Display of functional recombinant antibody fragments, for example single chain fragment variable (scFv) (McCafferty 1990) and fragment antigen binding (Fab) (Barbas 1992; Hoogenboom 1991) was enabled by progress in cloning of antibody gene repertoires (Orlandi 1989) and expression of functional immunoglobulin fragments in *E. coli* (Better 1988; Skerra 1988). More recently bivalent Fab ($F(ab')_2$) (Lee 2004) and even full-length antibodies, for example (Mazor 2010), have been displayed on phages, as well as several other forms of antibody derivatives. This progress has paved the way to generate synthetic antibodies without immunization, which allows rigorous control over the selection process and the design of the binding proteins. *In vitro* systems circumvent several drawbacks of immunization and hybridoma strategies such as difficulty to generate binders to lethal toxins, some pathogens or highly conserved antigens. Furthermore, non-physiological conditions may be applied and, importantly, the gene encoding the binder is immediately available. Synthetic sources of diversity also avoid inherent limitations of the natural immune repertoire, such as lack of control over sequence diversity of selected antibodies, and these synthetic sources have generated antibodies with higher affinities than those obtained by immunization (Bradbury 2011; Geyer 2012).

Non-lytic filamentous phages that do not kill their host cell upon infection, and thereby allow continued propagation and growth, are the most commonly used for *in vitro* selection (M13 and fd being the most widespread (Hoogenboom 2005)). However, other phages (such as λ, T4 and T7) have been used as well (Benhar 2001; Miersch 2012). Filamentous phages (Figure 6) are rod-shaped bacterial viruses about 900 nm long and 6-7 nm in diameter with a viral coat built by five coat proteins (pIII, pVI, pVII, pVIII and pIX) (see (Smith 1997) or (Miersch 2012) for more comprehensive references). A ~6400 bases single stranded DNA genome enclosed in the phage coat contains eleven genes. The coat is built from ~2700 copies of the major coat protein along the length of the particle, which is capped at one end with 3-4 copies each of pVII and pIX and at the other end by 3-5 copies of pIII and pVI. All molecular details of the infection mechanisms have not been resolved, but pIII is known to mediate the critical first interaction by binding to the F-pilus on a gram-negative host cell. Retraction of the pilus triggers subsequent

interactions that eventually lead to injection of the viral genome into the host cell. ssDNA is converted to dsDNA inside the host cell to initiate viral replication, assembly and release of new phage particles. Since the initial display on pIII (Smith 1985), all the other coat proteins have been used for display of heterologous proteins (Miersch 2012). However, pIII remains the most popular for display of larger proteins and pVIII is often used for smaller fusions (Bradbury 2004; Miersch 2012; Qi 2012). Thanks to the stability of the phage particle, selections can be performed under harsh conditions with maintained phage infectivity (Qi 2012).

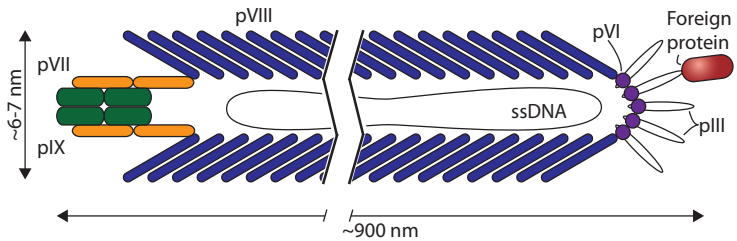


Figure 6. Schematic illustration of a filamentous phage. A heterologous protein fusion to phage coat protein pIII is shown.

At first, gene fusions were incorporated directly into the phage genome and this could generate functional fusions with efficient multivalent display and little perturbation of infectivity and propagation (Miersch 2012). However, it was soon recognized that such phage vectors might indeed interfere with phage infectivity, production and stability, which inspired the now common use of phagemid systems. Here, the protein to be displayed is fused to an additional copy of pIII on an additional plasmid and phage production is accomplished by the use of helper phages containing wild-type pIII (Qi 2012). The phagemid is designed to lack all critical elements for the formation of functional phages except an engineered form of pIII and therefore depend on co-infection with helper phages (using a second antibiotic selection marker) that contributes with components required for viral assembly and replication. The phagemid-encoded pIII-fusion and wild-type pIII will compete for incorporation into the virions, which typically results in oligovalent display (0-5 copies), but most phages that display a heterologous protein do it monovalently in this system (Miersch 2012; Qi 2012). In fact, only a small fraction (ca 1 %) of all produced phagemid or phage particles actually display a recombinant pIII fusion. Phagemids are generally more stable, as compared to phage vectors, and they are easier to transform and manipulate, which has made them popular (Bradbury 2004; Qi 2012). In general, N-terminal fusions of either full-length or N-terminally truncated pIII variants are used for library insertions. Several systems exist with various alterations to improve display level, transformation frequency, infectivity etc to manage the

complex interplay between phage, host and phagemid (see (Bradbury 2004; Miersch 2012; Qi 2012) for more details). Moreover, systems to eliminate the need for helper phages altogether have been devised (Chasteen 2006). Engineered packaging bacteria carrying helper plasmids are utilized to supply the packaging components and allow formation of various types of functional phages upon infection with phagemid.

Phages are usually non-specifically eluted from their bound target by incubation in an acidic solution, but competitive strategies as well as affinity independent methods, such as proteolytic cleavage (for example (Thomas 2010)) or reduction of a disulfide bond (for example (Prassler 2011)), have been employed. Other elegant approaches link infectivity or replication to target binding (Duenas 1994; Gramatikoff 1994; Spada 1997). Since its discovery, several remarkable phage display experiments and selection strategies have been reported including intravenous administration for *in vivo* panning of phage display libraries in human cancer patients (Krag 2006), phage display with libraries containing unnatural amino acids (Liu 2008; Tian 2004), selections against specific protein conformations (Gao 2009; Nizak 2003) or against subtle chemical differences such as specific post-translational modifications independent of the sequence context (Hoffhines 2006; Kehoe 2006), development of bispecific antibodies (Bostrom 2009; Fagete 2009) or phage-assisted continuous evolution (PACE) by connecting a selectable phenotype to the expression of pIII (Esvelt 2011). Many of the affinity proteins identified in these examples would be impossible to generate by traditional immunization methods.

2.1.2 Cell-display

Currently, cell-display based systems are rapidly expanding *in vitro* selection systems (Lofblom 2011a). Eukaryotic yeast still provides the main platform, but prokaryotic systems based on *E. coli* and, more recently, *S. carnosus* have moved forward. Given the relatively large size of yeast cells, larger libraries could potentially be more conveniently handled when using prokaryotic cells. The diameter of a yeast cell is roughly twice that of a staphylococcal cell, which in turn has a size comparable to a rod-shaped *E. coli* (Feldhaus 2004; Kronqvist 2010). Smaller sample volumes may provide an experimental advantage for prokaryotic systems. In spite of the widespread use of *E. coli*, high transformation frequencies and availability of molecular biology tools, the capability for display of heterologous proteins on their cell surface has not been fully exploited (Binder 2010). A main reason is that an inner membrane, a periplasmic space and an outer membrane need to be traversed for surface-display on those gram-negative cells. Gram-positive staphylococci provide a promising alternative; they only have a single membrane and

are highly tolerant to mechanical stress because of their thick cell wall, which results in a high viability after sorting (Wernerus 2004).

The large size of cells and multivalent expression of library proteins on the cell surface enables analysis and sorting by FACS, which cannot be done using smaller phages. This strategy also circumvents the need for elution of bound library proteins. The antigen is generally labeled with a fluorophore and incubated in solution with the cell-displayed library (Figure 7). In principle, any property of interest that can be linked to laser light scattering or fluorescence can be measured quantitatively in real time by flow cytometry (Daugherty 2007). It provides detailed information and, in addition to the general biopanning strategies using lowered target concentration and increased washing, the positioning of sort gates and utilization of both negative and positive screens can control selection stringency. Fine affinity discrimination is achieved by simultaneously assaying a reporter tag to monitor expression level (Lofblom 2005; VanAntwerp 2000). Impressing sorting rates can be achieved (ca 70000 cells per second or in the order of 10^8 events per hour; (Ackerman 2009; Lofblom 2011a)), but to ensure proper oversampling of larger libraries sorting becomes time-consuming. Magnetic-activated cell sorting (MACS) is often employed for pre-sorting or libraries exceeding around 10^8 variants (Chao 2006; Yeung 2002). Cell-surface display is also well suited for affinity maturation using focused libraries based on identified first-generation binders or immune libraries, which often have smaller sizes. When combined with appropriate pre-enrichment strategies, library construction and transformation are limiting rather than the selection method itself. Relative affinities ranked on cell-surfaces generally correlate with ranking obtained by other methods (Gai 2007; Lofblom 2007), even though absolute values may be different if avidity effects are involved when analyzing multivalent targets. Because selections may identify large numbers of positive clones, examples of more than 1000 confirmed binders from a selection exist (Edwards 2003), the demand for fast and easy characterization or ranking to identify the best clones can be met by flow-cytometric analysis.

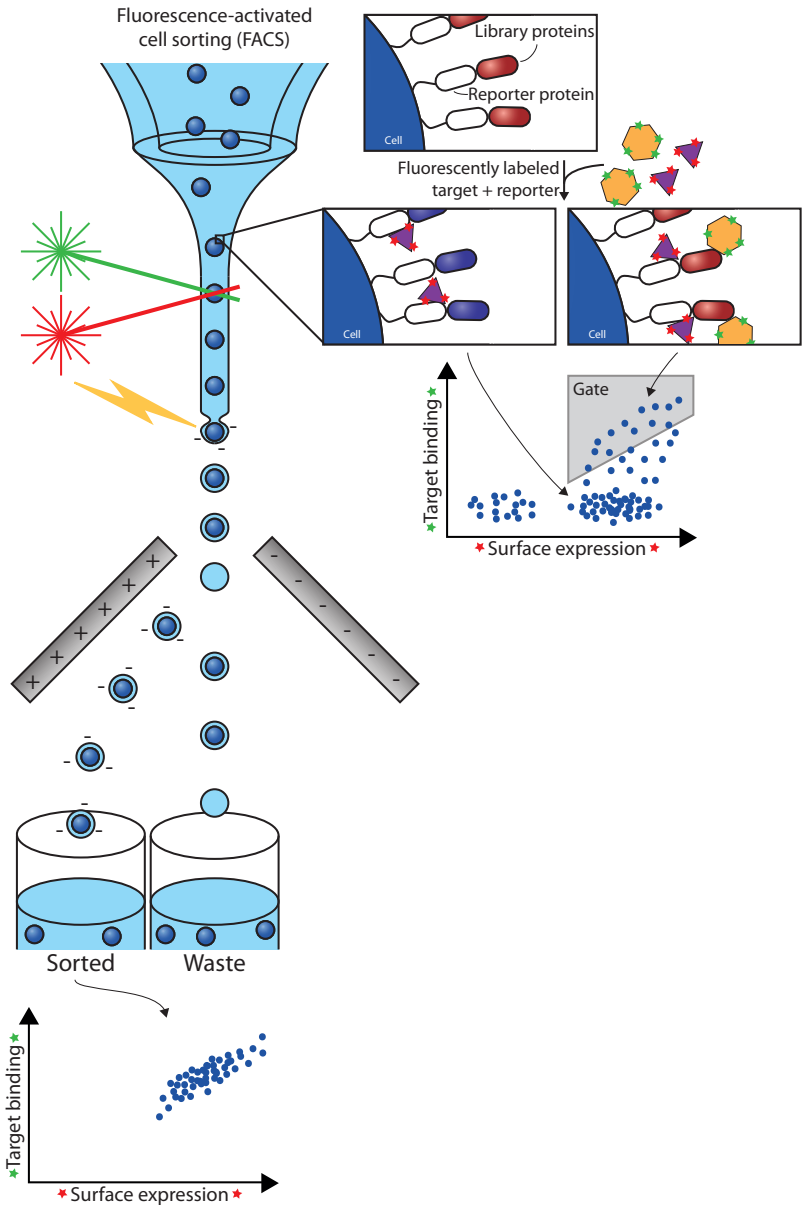


Figure 7. Fluorescence-activated cell sorting (FACS) and cell display. FACS is used to isolate cells that express a target-binding library protein (red) on their surface from cells displaying non-binding variants (blue). Simultaneous monitoring of expression level, using a labeled reporter protein (red fluorophore), and target-binding, detected by an additional (green) fluorophore, allows affinity-based selection that is not biased by varying expression levels. Sort gates are used to select cells to be enriched, which are charged and electrostatically deflected. Empty droplets and droplets containing non-binding cells end up in the waste.

The first reports of bacterial display of heterologous proteins fused to carrier proteins appeared around the same time that phage display was discovered (Charbit 1986; Freudl 1986) and a variety of prokaryotic hosts and anchoring strategies have been evaluated for surface expression of passenger proteins (Daugherty 2007; Jostock 2005; Samuelson 2002; Wernerus 2004). However, only a few systems have been used for combinatorial protein engineering and selection of affinity proteins (Daugherty 2007; Georgiou 1997; Lofblom 2011a; Wittrup 2001). Yeast display (Boder 1997) has become a popular and successful platform for this purpose but, due to the ease of manipulating bacteria and their rapid growth rate, those systems are expected to develop further as versatile platforms for protein engineering and may provide competitive alternatives to the well-established yeast display system (Daugherty 2007; Lofblom 2011a).

2.1.2.1 Yeast display

As mentioned, yeast display (Boder 1997) is a widely used cell surface display technique (reviewed in (Gai 2007; Gera 2012; Pepper 2008)). It has been suggested that the eukaryotic host may have a protein processing machinery that is better suited for proteins from higher organisms since it promotes efficient oxidative protein folding and N-linked glycosylation (Gai 2007). It has also been anticipated that engineered yeast strains possessing a human glycosylation pathway could be used as expression- and display hosts (Gai 2007; Gera 2012). A protein of interest is generally fused N- or C-terminally to a subunit (Aga2p) of the agglutinin receptor, which attaches to the cell wall through disulfide bonds to Aga1p. N- and C-terminal fusion tags facilitate monitoring of expression level and degree of full-length clones, respectively (Chao 2006; Pepper 2008). The largest libraries reported for yeast display included in the order of 10^9 clones (Feldhaus 2003; Gera 2012) and each cell typically displays 10^4-10^5 protein copies. This platform has been extensively used for affinity maturation (for example (Boder 2000) and (Jin 2006)) and epitope mapping (for example (Han 2011)). A correlation between display level and stability of the displayed protein has also been demonstrated and used for engineering of protein stability (Shusta 1999). Mating of haploid yeast that carries smaller libraries to form diploid cells can be used to combine the diversity of two libraries. This approach has been used to generate large Fab-libraries from mating of yeast transformed with a heavy- and light chain library, respectively (Weaver-Feldhaus 2004). Several kinds of affinity protein libraries have been expressed on yeast, for example scFvs from both immune and synthetic libraries (Pepper 2008), Fab-fragments (Weaver-Feldhaus 2004), and libraries based on several other protein frameworks (Gera 2012).

2.1.2.2 *E. coli* display

Early library applications of *E. coli* display for a form of affinity selection included panning of a peptide library against iron oxide particles (Brown 1992), followed by sorting of cells displaying scFvs from a background of cells displaying non-binding scFvs (Francisco 1993) and eventually sorting of an scFv library (Daugherty 1998). Most systems have focused on display of peptides (Daugherty 2007; van Bloois 2011) or small peptide-like scaffolds, for example a cysteine constrained knottin (Getz 2012), but this section is focused on library selection applications of larger, structured affinity proteins. A step towards current systems was periplasmic expression with cytometric screening (PECS), which relies on diffusion of small, labeled targets into the periplasmic space where they interact with secreted library proteins (Chen 2001). Although soluble expression may be advantageous for some proteins, this system preferentially selects for well-expressed variants. A refinement of PECS led to a more generally applicable method utilizing anchored periplasmic expression (APEx). Here, proteins from a combinatorial library are anchored to the inner membrane of *E. coli* by fusion either to a short peptide derived from native lipoprotein NlpA or to the phage minor coat protein III of M13 phage (Harvey 2004). To enable flow-cytometric analysis of bound fluorescently labeled antigen, disruption of the outer membrane is required (at least for larger antigens). The latter fusion system facilitates a hybrid strategy in combination with phage display without sub-cloning, which may aid screening of repertoires pre-enriched by phage display. Since the outer membrane is permeabilized or removed prior to sorting, clones need rescuing by PCR due to decreased viability. On the other hand, further mutagenesis may be performed during this step if desired (Harvey 2004; Harvey 2006). Selections on cells have also been demonstrated using the APEx system (Qiu 2010). In a further development of this system, APEx 2-hybrid, the same host expresses both antigen and affinity protein and the interaction between them in the periplasm can be monitored by flow cytometry to allow affinity-based selection (Jeong 2007). Methods for display and selection of libraries of full-length antibodies through periplasmic capture on an engineered immunoglobulin G (IgG) binding domain (Z-domain, (Nilsson 1987)) have been developed recently (Mazor 2008; Mazor 2007b). Moreover, five autotransporter systems were recently evaluated for display of Anticalins (Binder 2010), which are single-chain binding proteins based on the lipocalin ß-barrel fold with variegated loops (Skerra 2008). This study identified an engineered autotransporter domain of EspP as a promising vehicle for display on the outer membrane of *E. coli*. It has lately been demonstrated that disulfide-containing proteins, known to fold in the periplasm and considered difficult to translocate to the cell surface, can be efficiently and functionally displayed on *E. coli* through the autotransporter antigen 43 (Ag43) (Ramesh

2012). These reports indicate that *E. coli* display will continue to advance as a platform for combinatorial protein engineering.

2.1.2.3 Staphylococcal display

Many applications of staphylococcal display have been described (Wernerus 2002; Wernerus 2004) since the first display and flow-cytometric analysis of heterologous proteins expressed on the surface of these bacteria (Samuelson 1995). *S. carnosus* exists almost exclusively as single cells with a diameter of 0.5-1.5 μm. The surface-displayed proteins are not membrane spanning but covalently attached to the cell wall through the activity of a sortase. Proteins are attached via the C-termini to have free protruding N-termini that can tolerate large protein fusions (Kronqvist 2010; Lofblom 2011a). Apart from the single membrane that needs to be traversed, another advantage of this system is the lack of extracellular proteolytic activity and general robustness of the cells, which may be advantageous for tough selection conditions such as low pH or presence of proteases (Daugherty 2007; Kronqvist 2010).

An average surface expression level of between 10^4 and 10^5 copies per cell (Andreoni 1997; Kronqvist 2008b) and fine affinity discrimination through normalization against expression level (Lofblom 2005) make this system suitable for combinatorial protein engineering. The thick cell wall complicates transformation, but a recent improvement of the protocol now allows effective library transformations (Lofblom 2011a). A heat treatment of cells, increased amount of DNA per electroporation and addition of an osmotic stabilizer in the electroporation buffer improved the transformation rates markedly (Lofblom 2007). The display vectors can be propagated both in staphylococci and in *E. coli*, by two separate antibiotic selection markers and origins of replication, to facilitate cloning work. Following initial proofs of concept such as the display of an scFv (Gunneriusson 1996) and sorting of a model library based on Affibody molecules (Wernerus 2003), which are alternative binding proteins based on the Z-domain of SPA (Lofblom 2010; Nygren 2008), the first real library application of the staphylococcal system was reported (Kronqvist 2008a). An Affibody library, pre-enriched for binders to tumor necrosis factor alpha (TNF-α) by one round of phage display, was sub-cloned to a staphylococcal display vector for three subsequent rounds of FACS-sorting. Interestingly, this selection identified a similar but different set of clones as compared to a selection from the same library performed with phage display only (Jonsson 2009). This indicates that the systems are complementary. A comparison between phage- and yeast display, using another library and target, gave a similar result where the yeast display selection identified a larger set of clones (Bowley 2007). A method for epitope mapping using the staphylococcal display

system has been demonstrated (Rockberg 2008) and affinity maturation of Affibody molecules binding the human epidermal growth factor receptor 3 (ErbB3) (Kronqvist 2011) and bispecific binders to TNF-α based on an albumin-binding domain (Paper II in this thesis). Recent developments include the incorporation of a new affinity matured albumin-binding domain (Jonsson 2008) for normalization of expression level, the construction and transformation of a large, naïve Affibody library based on trinucleotide gene synthesis that exceeds 10^9 variants (Lofblom 2011a) and selection of binders from a nanobody library (Fleetwood 2012).

2.2 Cell-independent systems

The two most important cell-free systems are ribosome display and mRNA display. Their main advantages are the cell-free application, which enables screening of very large libraries without the need for laborious transformation of cells, and *in vitro* amplification to provide opportunity to integrate random PCR based mutagenesis between selection rounds (Lipovsek 2004). Such iteration between diversification and selection can closely mimic natural evolution and it may even be more attractive to introduce diversity between rounds, to tailor evolving binders as the selection progresses, than to construct large initial libraries. However, this strategy does not easily allow user-defined, site-specific positioning of mutations. The most significant difference between the two methods is the physical link between the protein and the mRNA. The methods can only be applied to affinity proteins encoded by a single, continuous gene. Libraries with full randomization of as few as 7-8 positions, which corresponds to approximately 10^9-10^{10} variants, approach the current practical limits of phage display. If a full coverage of the theoretical diversity is desired, the unmatched library sizes that may be achieved by the cell-free, fully *in vitro*, methods enable screening of a bigger sequence space in several orders of magnitudes larger libraries ($>10^{12}$) (Seelig 2011), at least in theory.

Following the first experimental demonstration of ribosome display selection of short peptides (Mattheakis 1994), the first validations for whole proteins were reported (Hanes 1997; He 1997). Soon thereafter, the potential to accumulate beneficial mutations was demonstrated in immune (Hanes 1998) and synthetic (Hanes 2000) libraries. In ribosome display, the *in vitro* translated protein remains connected to the ribosome and its encoding mRNA and this so-called ternary complex (non-covalent) is used for selection. This is achieved by the absence of a stop codon in the mRNA that causes the ribosome to stall and remain physically connected to the displayed protein, usually via a spacer sequence that fills the ribosomal tunnel and provides flexibility for folding (Pluckthun 2012). mRNA is recovered, converted

to cDNA and amplified by PCR for the next round of selection. The stability of the ternary complexes and enzymatic degradation of mRNA are concerns, but selection can be performed successfully for example at ambient temperature. In practice, the functional diversity of the library is limited by the amount of ribosomes in the *in vitro* translation extract rather than by the number of DNA templates (Pluckthun 2012).

mRNA-display is similar to ribosome display and was independently invented by two groups (Nemoto 1997; Roberts 1997). The mRNA-molecule that encodes a library protein is covalently linked to the encoded protein after translation, generally by using puromycin as an adaptor molecule. If the mRNA stability is a concern, it can be converted to its corresponding cDNA in a variation of the method (Kurz 2001; Seelig 2011; Ueno 2012). A commonly cited advantage with mRNA- over ribosome display is the overall smaller size of the complex when a ribosome is not required (see for example (Gold 2001)). However, experimental evidence for this assumption is still missing (Pluckthun 2012).

Both ribosome and mRNA display have been used in a number of affinity selections and both methods have become established in several laboratories. They have even been used for incorporation of unnatural amino acids to further expand the functional diversity (for example (Li 2002) and (Watts 2012)). Fewer reports are available for mRNA display than ribosome display and fewer binder-target systems have been explored (Lipovsek 2004), but both systems should in principle be applicable to the same types of problems. Similar techniques have also been reported, including CIS-display (Odegrip 2004), microbead display (Nord 2003; Sepp 2002), various forms of DNA-display (Bertschinger 2007; Reiersen 2005; Tabuchi 2001) and plasmid display (Cull 1992; Schatz 1993) among others. However, many of these methods may need additional validation to demonstrate advantages over existing and more established methods in combinatorial library applications to achieve a more widespread use.

2.3 Non-display systems

Many non-display systems are inspired by the yeast two-hybrid system (Fields 1989; Parrish 2006), which is a genetic method where binding of two proteins activates the transcription of a reporter gene. This system has been applied on antibodies (Visintin 1999), but techniques based on bacteria are generally more attractive for combinatorial protein engineering. Similar to the yeast two-hybrid, protein fragment complementation assays (PCA) are based on the interaction between two protein partners, which are genetically fused to the two halves of a dissected marker protein being expressed in the same cell. Binding

of the two partners reassembles the marker and reconstitutes its activity as a phenotype for selection (Secco 2009; Shekhawat 2011). Common marker proteins are split ß-lactamase (Wehrman 2002) or split dihydrofolate reductase (Pelletier 1998), which allow selection based on growth, and various forms of split green fluorescent protein (GFP; for example (Cabantous 2005)) that can be linked to whole cell fluorescence. PCA-systems have the advantages of experimental simplicity and that an antigen does not need to be expressed and purified. However, some human proteins may be difficult to express in bacteria and the control over target and binder concentrations is very limited. Several systems have been described for combinatorial protein engineering of antibody fragments (Koch 2006; Mossner 2001; Secco 2009), alternative scaffold proteins (exemplified by Affibody molecules (Lofdahl 2009; Lofdahl 2010)) as well as a library versus library selection of leucine zippers (Pelletier 1999). An interesting example using a domain of SPG utilized a split GFP fused to a small library of a split form of the C1 domain to study the stability of various C1-variants (Lindman 2010).

Comparison between Affibody molecules selected by PCA against TNF-α (Lofdahl 2009; Lofdahl 2010) with binders selected by phage or cell-surface display (Jonsson 2009; Kronqvist 2008a) demonstrates that different systems can identify distinct sets of clones, in this example various sequences of binders targeted to the same epitope. This finding confirms the functionality of the selection systems and, considering the large sequence libraries used, it is not surprising that non-identical sequences were obtained. Although direct comparisons between selection systems and scaffolds are difficult due to the broad range of parameters involved, combinations of systems and complementary advantages of the techniques may be useful for different applications.

Several related systems based on compartmentalization in water droplets combined with micro-fluidics have been reported (for instance (Brouzes 2009; Doi 1999; Kaltenbach 2012; Miller 2006; Paegel 2010; Tawfik 1998)), which provides an interesting new direction for protein engineering experiments. Such systems can be either cell-dependent or cell-independent.

3. AFFINITY PROTEINS AND PROTEIN SCAFFOLDS FOR MOLECULAR RECOGNITION

Affinity-based tools are of incredible importance in many areas of life science, including basic research, diagnostics and drug development. Many proteins possess a remarkable capacity for molecular recognition, and immunization or *in vitro* selection methods have the power to produce binding proteins to virtually any biomolecule (Dubel 2010; Uhlen 2010).

Antibodies have been the archetype of binding proteins with desired specificities for more than a century and still remain at the forefront of versatility and applicability (Gebauer 2009; Miersch 2012). They have had a tremendous impact on biomedical research and are increasingly used as diagnostic and therapeutic agents (An 2010; Imai 2006). Several thousands of antibodies, both mono- and polyclonal, have been developed for research and diagnostics ((Gronwall 2009; Michaud 2003), www.antibodypedia.com), some of which however have a questionable quality (Couchman 2009). Several antibodies and antibody derivatives have been approved for clinical use or are under clinical evaluation for several indications. Around 30 affinity protein biopharmaceuticals (excluding other biopharmaceutical proteins such as hormones and blood factors) have been approved for human use by the food and drug administration (FDA) (Beck 2011a; Chames 2009), of which the anti-TNF-α antibody adalimumab (or Humira®) was the first to be generated by *in vitro* selection technology (Nelson 2010). Adalimumab was recently followed by belimumab (Benlysta®), also identified through phage display, for the treatment of systemic lupus erythematosus (SLE) (Stohl 2012). As of January 2012, 165 protein-based drug candidates, including mainly standard bivalent, monospecific full-length IgGs, but also antibody-drug conjugates, bispecific or otherwise engineered forms, were under evaluation for several types of cancer (Reichert 2012). These candidates were aimed at 92 distinct targets, 65 of which were addressed by a single mAb. However, well-validated targets such as epidermal growth factor receptors are still the most common targets. Today, clinical safety and efficacy of antibodies have been well established, which has generated a wide acceptance of these affinity proteins among regulatory agencies, physicians and patients (Gebauer 2009; Nelson 2010).

Humanized and/or recombinant antibodies have become more common over time and several forms of antibody fragments as well as fusions of non-immunoglobulin domains to antibody fragments (especially the fragment crystallizable (Fc) region) have gained attention for therapeutic applications (Beck 2011b).

These antibody formats, especially scFv, are suitable for expression in microorganisms and for the selection systems mentioned earlier (Geyer 2012). In parallel to this development, major progress in the use of other proteins of various origins as scaffolds to modify an existing- or implement a new binding site has taken place. Synthetic libraries and *in vitro* evolution technologies can in principle make immunization unnecessary and could potentially also make the antibody molecule itself dispensable in the future (Boersma 2011). Natural proteins are regularly involved in molecular recognition events and can be used as starting points to develop new affinity proteins, but a natural interaction is not a required departure point. This field, in which so-called scaffolds or engineered protein scaffolds of both immunoglobulin- and, as discussed further below, non-immunoglobulin origin are utilized, has expanded quickly and more than 50 scaffolds have been described in the literature as summarized in a large number of reviews (Binz 2005a; Binz 2005b; Gill 2006; Hosse 2006; Miao 2010; Nygren 2004; Skerra 2007; Ståhl 2012; Wurch 2012). The scaffolds range from small, constrained peptides to large protein domains, representing different folds and architectures of the, not necessarily pre-existing, binding sites or surfaces. They are usually derived from a small, stable and soluble monomeric protein or domain of an extracellular cell surface receptor. Even though monoclonal antibodies are excellent affinity proteins, alternative scaffolds may be able to address several limitations or disadvantages of antibodies and antibody fragments and thereby provide alternatives in certain applications. For example, the relatively large size and complex composition of antibodies and fragments thereof as well as the common requirement for eukaryotic hosts for production may be avoided. Current efforts aim to arm alternative scaffold proteins or smaller antibody fragments with effector functions, to adapt them to new routes of administration, to improve their pharmacokinetics and plasma half-life and to better understand the important issue of immunogenicity (Gebauer 2009). The following section will describe the structure and advantages of antibodies and how many of these properties can be utilized or improved by using either smaller antibody derivatives or other engineered non-immunoglobulin protein scaffolds. Carefully designed, systematic studies will be needed to define the capabilities and limitations of each display scaffold (Daugherty 2007), but a trend is a consolidation around a few scaffolds that show promise to be used for many targets and have beneficial properties in different applications (Gebauer 2009).

3.1 Antibodies and antibody derivatives

Polyclonal antibodies, i.e. mixtures of immunoglobulin molecules secreted against an antigen that can generally recognize several epitopes, are obtained from sera of immunized animals. Polyclonal antibodies often provide useful tools, but their supply and composition cannot easily be reproduced since they

usually come from a non-renewable source. A method to fuse antibody-producing B-lymphocytes from an immunized mouse with immortal myeloma cells to generate monoclonal antibodies (mAbs), i.e. renewable and identical immunoglobulin molecules that recognize a specific epitope of an antigen, from *in vitro* cultured hybridomas (Kohler 1975) provided a foundation for the central role of antibodies in basic and applied sciences (Miersch 2012). Since then, several methods have been developed to modify non-human monoclonal antibodies to become more human-like, for example chimerization and humanization (Almagro 2008; Jones 1986; Osbourn 2005), either by recombinant means or by using transgenic animals expressing repertoires of human antibody gene sequences (Fishwild 1996; Jakobovits 2007; Lonberg 2005).

Antibodies of the most common class (IgG) are composed of two heavy (H) and two light (L) chains held together by 12 intra-domain and four inter-domain disulfide bonds (Figure 8). The heavy chains consist of three constant domains and one variable (V) domain. The light chains contain one constant and one variable domain. The antigen-binding site of an antibody is composed of six hypervariable loops called complementarity-determining regions (CDRs), three in each VH and VL domain, respectively. A large collection of binding sites is created by the immune system through rearrangements of germline-encoded gene segments. This results in a naïve repertoire of B-cells in which each B-cell expresses and displays a unique antibody. After a complex selection to remove antibodies reactive to self-proteins, exposure to an antigen selects cells able to produce antigen-reactive antibodies and triggers incorporation of additional mutations by somatic hypermutation to fine-tune the complementarity of the antibody for the antigen. IgG is the most common antibody isotype in serum (85 % of all immunoglobulins in human serum) and is generally the preferred format for research and clinical applications (Lobo 2004). This section will therefore focus on antibodies of the IgG isotype. As mentioned in the previous section, antibodies with tailored specificity can nowadays be generated totally *in vitro*, without immunization, by using recombinant and combinatorial gene techniques (Bradbury 2011; Geyer 2012; Sidhu 2007, 2012). Several design approaches are available for synthetic antibodies and they vary in the number of framework regions they use, the design of CDR diversity and the library construction methods (Geyer 2012; Miersch 2012). In 2002, adalimumab became the first human mAb of this generation to be approved for clinical use for rheumatoid arthritis and also the first such product generated through *in vitro* selection technology (An 2010).

Full-length antibodies have several potentially beneficial properties such as bivalency and long half-life mediated by the constant Fc region. Fc is also responsible for complement activation, cell-dependent

cytotoxicity (CDC) and antibody-dependent cell-mediated cytotoxicity (ADCC) and, production-wise it may increase expression levels and facilitate purification (Beck 2011a). However, the Fc-mediated effector mechanisms are not necessary for many applications of affinity proteins (Binz 2005a). This, in combination with some difficulties in cloning and, commonly, the requirement for eukaryotic production hosts for full-length immunoglobulins, has gradually directed interest to smaller antibody fragments with retained antigen binding (de Marco 2011; Fernandez 2004; Holliger 2005). Antibody fragments were originally generated by enzymatic treatment of IgG, but such methods have now been replaced by recombinant techniques. The most common antibody fragments are the scFv (Bird 1988; Huston 1988) and Fab-fragments, which usually have the same affinity and specificity for their antigens as full size antibodies. These miniaturized formats may also be converted into full-length IgG, which often is a format of choice for clinical purposes where a long half-life or other Fc-mediated functions may be desired, although reformatting may lead to loss of activity (Mazor 2010). The Fab-fragment format (with or without an inter-chain linker) is popular during the discovery/selection step because of a more reliable conversion to bivalent IgG molecules (Miersch 2012), but it has other disadvantages. Fab-fragments are generally more difficult to assemble, more likely to be degraded, have lower yields in the form of soluble fragments and are more likely to cause problems with DNA instability in phages (but not in phagemids) due to the larger size of the encoding DNA compared to smaller singe-chain antibody derivatives (Bradbury 2004). All of these problems arise from the fact that Fab-fragments contain two protein chains, which need to assemble. Each chain is roughly the same size as a single scFv chain. However, Fab-fragments usually do not appear to suffer from the problem of dimerization and aggregation, which sometimes affect scFvs (Bradbury 2004; Holliger 2005). Trimers and higher multimers of antibody fragments can also be formed (for example (Holliger 2005; Hudson 1999)). One example is Diabodies, which are dimers of a VH and VL domain connected by a peptide linker too short to allow pairing between the two domains when expressed in fusion to each other. This drives pairing with domains on a different chain to recreate two antigen-binding sites, which also enables a bispecific format, in a molecule the size of a Fab-fragment (Holliger 1993; Holliger 1997).

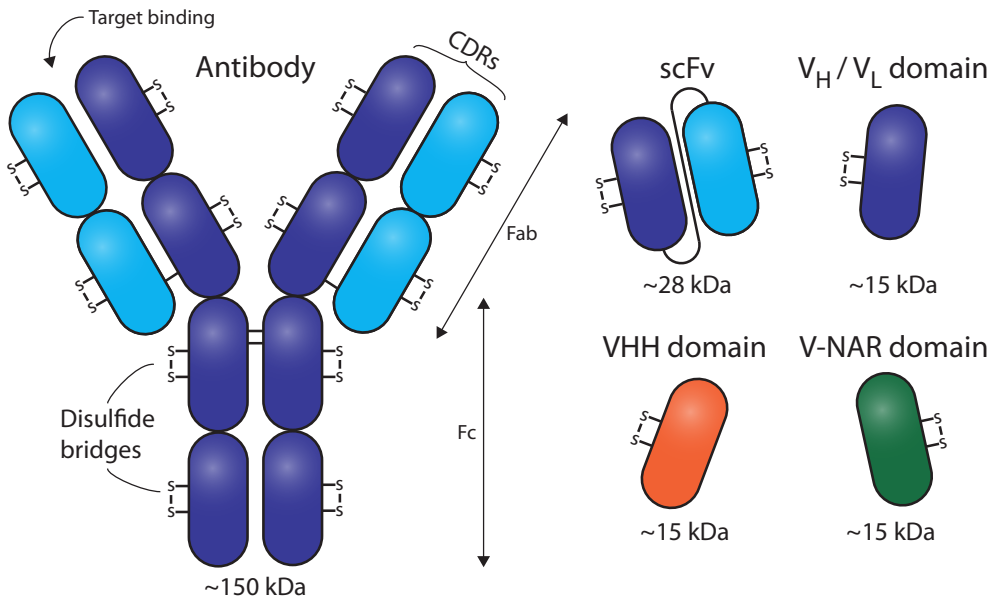


Figure 8. Schematic structure of an antibody and various antibody fragments. IgG is a multi-domain protein built up from two heavy- (dark blue) and two light chains (light blue) held together by several disulfide bonds (S-S). Complementarity determining regions (CDRs) in variable domains at the tips of the antibody arms confer binding specificity. The fragment crystallizable (Fc) and fragment antigen binding (Fab) are indicated and approximate molecular masses of the illustrated proteins (in kDa) are shown. Smaller antibody fragments, such as the single chain fragment variable (scFv) and a single domain antibody (derived from the variable heavy VH or variable light VL domain), are also available. The variable domain of a camelid heavy chain antibody is called a VHH or a nanobody. The variable-new antigen receptor (V-NAR) domain represents the smallest binding domain of heavy chain antibodies derived from cartilaginous fish.

Single domain antibodies (domain antibodies, dAbs) represent the smallest binding domains of a conventional monoclonal antibody and consist of either the VH or the VL domain (Gill 2006; Holliger 2005). These 11-15 kDa proteins carry a set of three CDRs and, in the late 1980s, it was demonstrated that a repertoire of VH domains, isolated by PCR from immunized mice, could be screened for binding to the antigen used for immunization (Saerens 2008; Ward 1989). Initial problems of insolubility based on exposure of the normally buried domain interface between VH and VL have been addressed, for example by targeted mutagenesis and selection of more soluble frameworks (Enever 2009; Ewert 2003; Jespers 2004; Nuttall 2008). Overall, the most common immunoglobulin-based scaffolds for synthetic binding proteins have been Fab-fragments, scFv and dAbs (Koide 2009b). Further attempts to generate smaller

binding proteins inspired by antibodies have also been described, for example the fusion of two CDR-loops to a portion of a heavy chain framework to produce a 3 kDa binding domain (Qiu 2007).

Interestingly, single binding domains naturally occur in so-called heavy chain immunoglobulins from camelids (Hamers-Casterman 1993) and cartilaginous fish (Greenberg 1996), which illustrates that evolution also arrived at a reductionist solution using only three CDR-regions on a single binding domain in these cases. These molecules are thought to have lost their light chains during evolution (Holt 2003). The single antigen-binding domain of a camelid antibody is called a VHH or nanobody (Saerens 2008) and the corresponding domain derived from cartilaginous fish, from the immunoglobulin new antigen receptor (IgNAR), is called a V-NAR (Barelle 2009; Saerens 2008). These single-domain antibody derivatives can be well expressed and form monomeric species in solution and both have been utilized for combinatorial protein engineering (Dooley 2003; Harmsen 2007; Nuttall 2001). In contrast to regular antibodies and their fragments, which often aggregate irreversibly on thermal denaturation, these domains often have the ability to regain activity after heat treatment. The high solubility and lack of aggregation is generally attributed to a long, disulfide stabilized CDRH3 loop, which covers the hydrophobic interface that would have been formed with a second variable domain, as well as a to few characteristic framework mutations (Holt 2003; Vincke 2009).

3.2 Alternative, non-immunoglobulin, scaffold proteins

In the early 1990s, affinity maturation and changes in specificity of protease inhibitors, using rational protein engineering and phage display, provided initial examples of the use of scaffolds other than antibody-like proteins for selection of specific binders (Binz 2005b; Nygren 1997). Over 100 different protein domain folds exist in nature, several represented in proteins with different biological functions (Kaneko 2011). Examples of this kind of functions include peptide recognition mediated by loop conformations in SH2 domains or variable peptide-binding clefts of PDZ domains (Ferrer 2005; Kaneko 2011). In these cases, the fold provides a framework for the optimal spatial disposition of binding residues that together define the specificity of the protein (Kaneko 2011). Alternative scaffolds are sometimes referred to as the missing link between small molecule drugs and antibodies, which aim both to fill the gap between them and to address inherent limitations of antibody-derived binding molecules and small molecules, respectively (Binz 2005a; Hey 2005). They may replace large, complex multi-chain immunoglobulin molecules with more convenient protein formats. Interest in this field has grown during recent years as well as the associated commercial attention (Sheridan 2007). Ideally, the target-binding

affinities and selectivity in binding displayed by antibodies should be present in binders of an alternative format. Instead of relying on Fc-mediated effector functions, a small binding protein based on a simple and robust scaffold may provide distinct mechanisms of action that are difficult to achieve using natural antibodies (Gebauer 2009), for example more efficient delivery of a payload (Wu 2005) carried by a small protein deep into a solid tumor. Affinity proteins based on alternative scaffolds have proven useful or promising in a growing number of applications including diagnostic imaging, biotherapy, affinity chromatography, for half-life extension, viral retargeting to tumors, in structural biology as crystallization chaperones and for intracellular applications.

As mentioned above, the general idea is to create novel binding sites on proteins that are recruited to serve as molecular scaffolds. This is typically achieved through combinatorial protein engineering and selection from designed protein libraries. A suitable scaffold should be tolerant to multiple amino acid substitutions or insertions directed to the molecular surface, to integrate a new affinity function (Zoller 2011). Substitutions or insertions may be performed in single loops, multiple loops or in more rigid elements of secondary structure. Loops have been diversified in most of the scaffolds evaluated and, more rarely secondary structural elements such as α-helices or ß-sheets (Binz 2005a). The number of scaffolds based on α-helix architecture is limited, as compared to the numerous examples of more antibody-like, ß-sheet frameworks (Hosse 2006). Residues to be randomized should be chosen carefully, since modifications may lead to instability or misfolding (Hosse 2006), an issue that has also been addressed computationally for several scaffolds (Wiederstein 2005). A sufficient number of residues should be randomized to allow diverse binding surfaces to be formed, but mutagenesis should preferably not disrupt the underlying structure. In cases where the native template protein has a binding function, residues involved in that interaction could be the first choice for the randomization (Nygren 1997). Obviously, an available structure of the starting scaffold is almost essential and this section will mainly focus on more validated systems where a structure of the starting scaffold, as well as several structures of generated affinity proteins in complex with their targets, can shed light on their specific modes of molecular recognition. Structures of four examples of alternative scaffold proteins are shown in figure 9. Examples of structural complexes with targets are also given later (see figure 11). Ideally, structural characterization of engineered binding proteins can reveal how the molecules achieve their function and may give insights to improve library designs in a positive feedback loop (Gilbreth 2012).

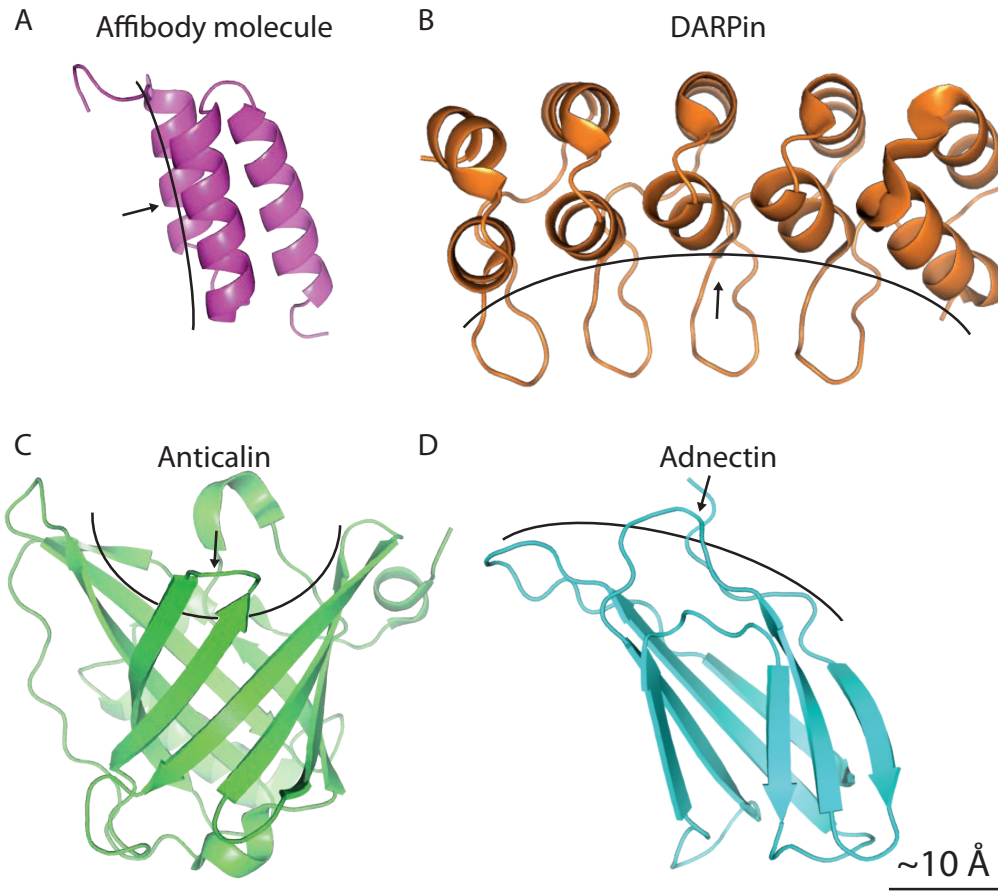


Figure 9. Examples of alternative scaffold proteins. A. Affibody molecule (here represented by a structure of the Z-domain, 1Q2N.pdb, (Zheng 2004)). **B**. Designed ankyrin repeat protein (DARPin) (2JAB.pdb, (Zahnd 2007)). **C**. Anticalin (generated from 1N0S.pdb, (Korndorfer 2003)). **D**. Adnectin (1TTG.pdb, (Main 1992)). These four scaffold proteins represent well-validated alternative binding protein scaffolds, on which binders to multiple targets have been generated. In Affibody molecules and DARPins, residues distributed on elements of secondary structure are randomized to generate combinatorial libraries. Anticalins and Adnectins generally utilize randomized loops.

Ability to refold after thermal denaturation, high solubility, and absence of disulfide bonds or free cysteines are desired characteristics to generate large amounts of non-aggregating, homogeneous preparations of a scaffold protein. Proteins without disulfides may facilitate higher yields in bacterial expression and enable intracellular applications in the reducing cytoplasm (Binz 2005a; Boersma 2011; Gronwall 2009). It also allows introduction of a unique cysteine residue for site-specific thiol-based modification (Feldwisch 2012). Moreover, a chemically and thermally robust scaffold generally tolerates

harsh labeling conditions better (Wu 2005), provided it has the ability to refold after denaturation. However, many of the smallest scaffolds, for example cysteine-knot miniproteins (knottins) (Kolmar 2008; Moore 2012; Zoller 2011), are used to present randomized loops on a disulfide-constrained structure. The 23 amino acid Min-23 is one of the smallest scaffolds reported (Heitz 1999; Souriau 2005) and a similar strategy, based on chemical cross-linking of phage-displayed peptides on an organic core, has been developed by Heinis and Winter (Heinis 2009). Presentation of random sequences in a structurally restricted manner should decrease the loss of conformational entropy upon complex formation, as compared to for example linear peptides (Binz 2005b; Marshall 1992). In contrast to peptides or disulfide-constrained structures, the independent folding and binding activity of a cysteine-free, folded domain is more likely to be maintained during production, purification and when it is used as a binding unit in a protein fusion. This facilitates generation of multivalent or multispecific binding proteins (Binz 2005a; Boersma 2011; Nygren 1997). Furthermore, a higher proteolytic stability can be expected for a structured scaffold than for more unordered peptides or loops and may also be an advantage considering aggregation, denaturation and chemical degradation. Higher affinities can theoretically be expected for well-ordered proteins because of a potentially smaller loss of entropy upon binding as a consequence of a smaller structural rearrangement upon binding (Gebauer 2009; Ladner 1995; Wahlberg 2006). However, very stable proteins may in other cases not be able to undergo structural rearrangements that could be required for binding.

Sometimes, for example in certain medical applications, smaller scaffolds have advantages compared to larger alternatives (see figure 10 for a size comparison of the albumin-binding domain and an antibody). Smaller scaffold proteins may be produced by solid phase peptide synthesis, which may facilitate site-specific and homogeneous modification (Nygren 2008), and be easier to study by nuclear magnetic resonance (NMR) spectroscopy. Low molecular mass is also an advantage because the protein will have a higher specific activity, i.e. the large size of antibodies requires considerably larger doses (in gram) as compared to smaller proteins to administer an equal amount of moles (Feldwisch 2012; Gebauer 2009). Moreover, high antibody doses in small volumes usually demand extensive pharmaceutical formulation development to obtain the protein concentrations desired (Feldwisch 2012). Importantly, a small size allows a better access to certain tissues and antigens that may have a poor availability to targeting by large antibodies (Gill 2006; Holliger 2005). The size and pharmacokinetics of antibodies lead to slow biodistribution and elimination, which is often desirable in many therapeutic settings but limits their use in applications where rapid clearance of unbound affinity protein is desired – i.e. in molecular imaging (Miao 2010). In contrast, smaller protein scaffolds are more likely to quickly accumulate in solid tumors

due to their improved tissue penetration, but they are also quickly eliminated through the kidneys (Feldwisch 2012). Modification of the protein sequence and labeling strategies may be used to alter clearance rate and route of excretion (Miao 2010), but small proteins are generally characterized by quick elimination. Strategies to extend plasma half-life of proteins, for example through association with albumin, are often useful for therapeutic intentions and will be discussed in a later section of this thesis.

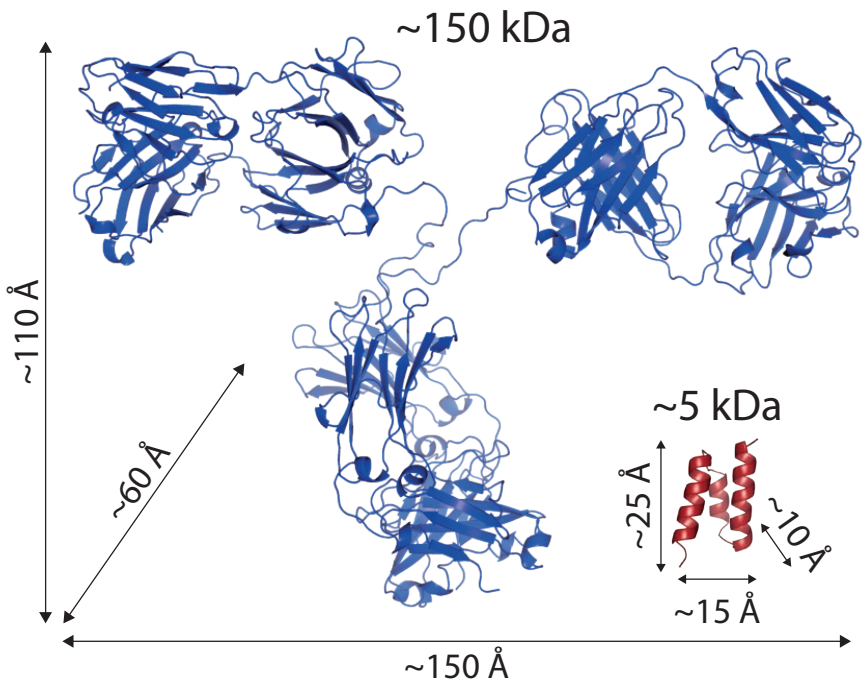


Figure 10. Size comparison of a small alternative scaffold protein and an antibody. Several alternative scaffold proteins are much smaller as compared to antibodies. This picture illustrates the relative sizes and molecular masses of the 46 amino acid albumin-binding domain utilized as a scaffold in the present investigation and an IgG antibody. Pictures were generated from the PDB entries 1IGT (antibody) (Harris 1997) and 1GJT (albumin-binding domain) (Johansson 2002a).

Four well-validated alternative scaffolds for molecular recognition are: Anticalins (Beste 1999; Gebauer 2012; Skerra 2008), designed ankyrin repeat proteins (DARPins) (Binz 2003; Boersma 2011; Tamaskovic 2012), Adnectins (Koide 1998; Koide 2012b; Lipovsek 2011) and Affibody molecules (Feldwisch 2012; Lofblom 2010; Nord 1997; Nygren 2008). Affibody molecules are described in greater detail, because they have been extensively evaluated and represent a scaffold very similar to the albumin-binding domain utilized throughout the work in this thesis. A growing number of crystal structures of alternative scaffold-based affinity proteins in complex with their antigens/targets illustrate mechanisms for molecular

recognition that are common or unique to a specific scaffold (Gilbreth 2012). These four scaffolds are all based on proteins with a known structure, they represent different architectures of the binding sites and several examples of binder-target complexes are available for all systems (see figure 9 and figure 11 for examples). Moreover, binding molecules based on these scaffolds have been selected by at least two different *in vitro* selection methods. Nearly 30 structures related to these four scaffolds are currently available (Gilbreth 2012). In contrast, many other alternative binding proteins have only been used in a single initial study. Even for many of the more extensively evaluated variants, little or no data exist on serum half-life, tissue penetration, tissue-to-blood ratio or immunogenicity. The success of several scaffold systems in producing high affinity binders to diverse targets suggests that several molecular architectures are feasible as long as diverse amino acid sequences can be presented without compromising overall structure (Gilbreth 2012; Koide 2009a). However, other properties will be equally or more important to move this class of affinity proteins into more demanding applications. The first candidate drugs derived from alternative protein scaffolds have already entered clinical trials (Feldwisch 2012; Wurch 2012) or even reached approval (for example a kallikrein inhibitor for treatment of acute hereditary angioedema (Ecallantide (Kalbitor), Dyax (Thompson 2010)). It has also been suggested that the robustness of some scaffolds may make them amendable to alternative routes of administration as compared to traditional antibody formats (Gebauer 2009; Kolmar 2008).

3.2.1 Anticalins

Lipocalins constitute a family of proteins with binding sites composed of four peptide loops on a stable, structurally conserved eight-stranded ß-barrel scaffold (Figure 9C, Figure 11B) (Skerra 2008). Artificial binding proteins based on the lipocalin fold are called Anticalins and were initially selected by phage display (Beste 1999), but can also be identified by bacterial display (Binder 2010; Gebauer 2012). Anticalins have been successfully selected to a range of target antigens and a distinguishing feature of this scaffold is the ability to bind small hapten-like molecules in the deep ligand-binding pocket. Anticalins have several advantages in common with other alternative scaffolds such as high stability, ease of manipulation and high expression yields in microbial hosts (Gebauer 2009, 2012; Gill 2006). They typically contain two disulfide bonds, but do not strictly require this (Gebauer 2012; Gronwall 2009). They are composed of a single polypeptide chain of 160-180 residues (~20 kDa) where 16-24 residues are randomized in or around the ligand-binding pocket at one end of the ß-barrel (Gebauer 2012). A new generation of Anticalins has been engineered on human lipocalins and structural data demonstrate that the architecture of the scaffold was preserved in several such Anticalin-target complexes (Gebauer 2012; Kim 2009; Schonfeld 2009). Bispeficic or bivalent dimers of Anticalins called Duocalins have also been described (Schlehuber 2001) as well as half-life extended variants (Skerra 2008). A current lead candidate binding vascular endothelial growth factor (VEGF) aimed for cancer therapy has recently passed a clinical phase I study (http://www.pieris-ag.com/).

3.2.2 Designed ankyrin repeat proteins

Designed ankyrin repeat proteins (DARPins) are modular binding proteins generally based on a consensus ankyrin repeat domain that supposedly implements favorable structural features that have been preserved during evolution (Binz 2004; Binz 2003). These binding proteins are composed of a number of centrally positioned 33 residue repeat domains consisting of a ß-turn followed by a helix-loop-helix motif (Figure 9B, Figure 11D), in which 6-7 amino acids usually are randomized. Binding proteins are commonly composed of 2-3 repeats capped at the termini with hydrophilic so-called capping repeats (133-166 residues in total, 14 kDa and larger) (Boersma 2011; Gronwall 2009; Stumpp 2008; Tamaskovic 2012). Similar strategies have been described for other repeat scaffolds as well, for example armadillo repeats (see (Boersma 2011; Varadamsetty 2012)). This scaffold is distinct from Anticalins and Adnectins because randomization is performed on rigid secondary structure elements on the repeats to form a

continuous, concave shape (Gilbreth 2012). Interestingly, natural binding proteins composed of repeat structures, in essence analogous to the DARPin strategy, have been identified in jawless vertebrates (Danilova 2012; Pancer 2004). DARPins have been selected to a large number of targets by ribosome display, an *in vitro* complementation assay (Amstutz 2006) and a modified form of phage display that allows display of fast-folding proteins (Steiner 2006). Recent design improvements of the scaffold have been implemented and evaluated by structural biology (Tamaskovic 2012). Different forms of DARPins have been evaluated for applications in for instance structural biology, tumor imaging and therapy (Boersma 2011; Stumpp 2008; Tamaskovic 2012). A current lead therapeutic candidate against VEGF has passed initial clinical testing for ophthalmology indications (http://www.molecularpartners.com/).

3.2.3 Adnectins

Adnectins are stable alternative binding proteins based on the 10th extracellular domain of human fibronectin III (Koide 1998), which adopts an immunoglobulin-like ß-sandwich fold (Figure 9D, Figure 11C). Generally, three loops on this 94 residue (10 kDa), cysteine-free scaffold are diversified to generate affinity proteins that mimic the antigen-binding site of antibodies. Binding proteins have been selected by phage display, yeast display, a yeast-2-hybrid screen and mRNA display (Koide 2012a; Lipovsek 2011). Adnectins have become one of the most widely used non-antibody binding proteins and are sometimes referred to as monobodies or Trinectins depending on their origin (Koide 2012b). Binders to more than 15 targets have been published (Lipovsek 2011) using several engineering approaches including loop-length diversity (Hackel 2008) and restricted or biased amino acid codes (Gilbreth 2008; Hackel 2010; Koide 2007; Wojcik 2010). Several fusion-protein constructs for various applications have been described as well (Koide 2012b; Lipovsek 2011; Lofblom 2011b). Adnectins often bind to clefts in the target through the protruding randomized loops and are thereby similar to camelid single-domain antibodies (Gilbreth 2012). Interestingly, a binder to lysozyme with an inter-loop disulfide bond reminiscent of single-domain shark- and camelid antibodies has been identified in one selection (Lipovsek 2007). Adnectins with concave binding sites, generated through randomization of an alternative binding area on the same scaffold, directed against a convex epitope on the target, have also been identified (Gilbreth 2012; Koide 2012a). This demonstrates that several surfaces on the same scaffold may be utilized to accommodate binding sites with different properties. An advanced anti-VEGF candidate for treatment of glioblastoma is under clinical evaluation (phase II) (http://www.adnexustx.com).

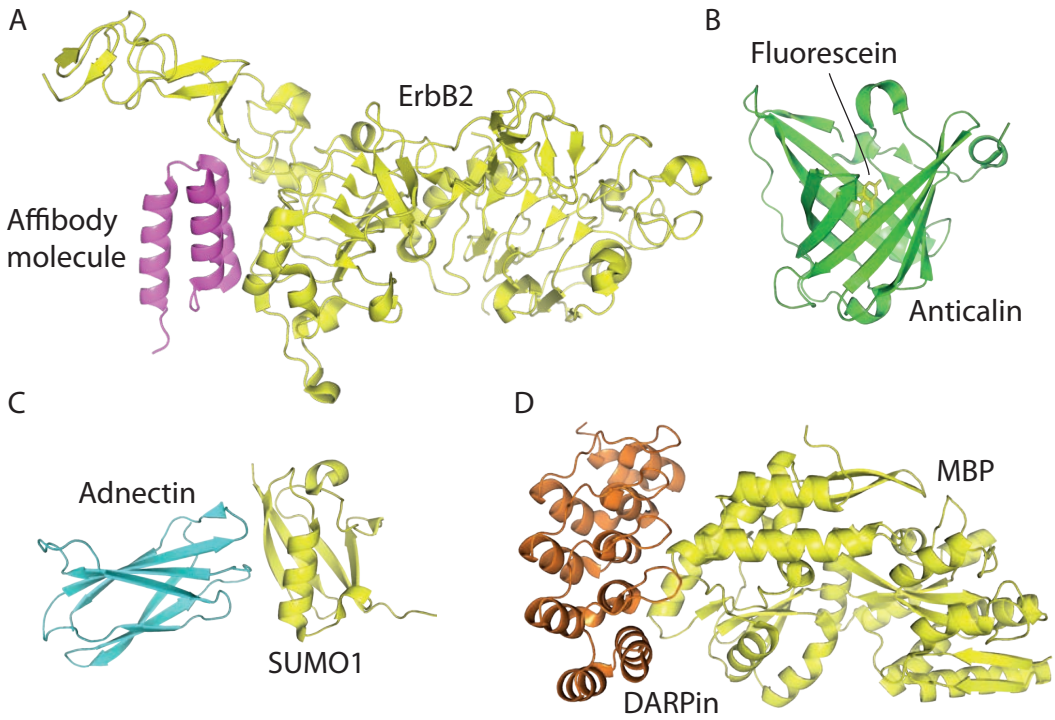


Figure 11. Examples of crystal complexes of alternative binding proteins and their targets. A. Affibody molecule in complex with the extracellular domain of human ErbB2 (3MZW.pdb, (Eigenbrot 2010)). **B.** Anticalin in complex with fluorescein, to represent hapten binding in this particular example (1NOS.pdb, (Korndorfer 2003)). **C.** Adnectin in complex with small ubiquitin-related modifier 1 (SUMO1) (3RZW.pdb, (Gilbreth 2008)). **D.** DARPin in complex with maltose-binding protein (MBP) (1SVX.pdb, (Binz 2004)).

3.2.4 Affibody molecules and other affinity proteins derived from the Z-domain

Affibody molecules (Figure 9A, Figure 11A) are affinity proteins based on a modified form of the B-domain of staphylococcal protein A (SPA), which essentially is a consensus domain of its five immunoglobulin-binding domains (Nilsson 1987; Nygren 2008). SPA is commonly used for purification of antibodies and Fc-fusion proteins. All domains of SPA have a high sequence similarity and all of them can individually bind Fc and, more weakly, Fab of immunoglobulins from several species and subclasses (Boström 2012; Jansson 1998). The modified consensus domain is called the Z-domain (Nilsson 1987), it consists of 58 amino acids and was generated by rational mutagenesis to increase the chemical stability and introduce an N-terminal (5´) restriction site, respectively. These modifications also resulted in almost

complete loss of the weak Fab-binding (Ljungberg 1993; Nilsson 1987), which is thought to be mainly due to one of the sequence modifications (namely the substitution G29A, (Graille 2000)), but the Fc-binding was retained (Jendeberg 1995). It is a highly soluble, chemically- and thermally stable domain composed of an anti-parallel three-helix bundle structure (see for example (Jendeberg 1996; Tashiro 1997)) similar to the structure of the B-domain (Deisenhofer 1981). Z folds very quickly (Arora 2004), lacks cysteines and has been used as an affinity ligand or affinity fusion tag for affinity chromatography (Boström 2012). Many of these properties are shared with the albumin-binding domain derived from streptococcal protein G, which is described in more detail in the next section, and make these molecules interesting as scaffolds for protein engineering. Helix bundle structures are generally expected to have a conformation relatively tolerant to changes in the side chains of residues not involved in helix-helix packing interfaces (Cedergren 1993; Nord 1995). Experimental evaluation of random sequences biased for certain secondary structures has also demonstrated promising characteristics for α-helices with soluble, monomeric proteins, whereas ß-strand proteins tended to form fibrils and have solubility problems (Matsuura 2002). In that study, the presence of ß-strands seemed to require precise topological arrangement to prevent aggregation.

Affibody molecules (Lofblom 2010; Nord 1997; Nygren 2008) are usually generated by randomization of 13 residues, including the amino acids involved in Fc-binding, as determined from a crystal complex and mutational analysis (Cedergren 1993; Deisenhofer 1981), and additional residues on the same molecular surface (Nord 1997; Nord 1995). The randomized residues are distributed on two of the α-helices and represent one of very few scaffolds where residues engaged in secondary structures (α-helices in particular) are diversified (Binz 2005a; Hosse 2006). In contrast to the structurally most closely related scaffolds, Affibody molecules have been selected against many targets (Gronwall 2009; Lofblom 2010) by several selection methods such as phage-, staphylococcal- and ribosome display as well as a ß-lactamase complementation assay (Lofblom 2010). Furthermore, numerous structural evaluations of Affibody molecules and Affibody-target complexes have been reported (Eigenbrot 2010; Gilbreth 2012; Hogbom 2003; Hoyer 2008; Lendel 2004; Lendel 2006; Wahlberg 2006; Wahlberg 2003). Most binding-surfaces are predominantly non-polar, as judged from available structural data, and comprise areas of 800-900 $Å^2$, which is typical for protein-protein interactions (Nygren 2008) and Affibody molecules tend to recognize flat surfaces on their targets (Gilbreth 2012). The Affibody scaffold was recently modified by 11 amino acid substitutions in non-binding regions, which increased its hydrophilicity, thermostability and amenability to peptide synthesis (Feldwisch 2010). With these positions added to the 13 that are generally randomized in Affibody libraries, more than 40 % of the domain sequence has been changed.

Affibody molecules have found a widespread use as tools in biotechnology (Boström 2012; Lofblom 2010), but a promising area of application is the use as molecular tracers for imaging of cancer (Lofblom 2010; Miao 2010; Nygren 2008). Common medical imaging methods, such as radiography and magnetic resonance imaging, do not provide information about the expression levels of specific target molecules. However, Affibody molecules produced recombinantly or by peptide synthesis have shown great promise as targeting agents in such applications (Ahlgren 2010). This success is mainly due to their small size (6.5 kDa), which leads to quick tumor accumulation and renal elimination of excess tracer to yield high imaging contrast a short time after administration. An Affibody molecule specific for the human epidermal growth factor receptor 2 (HER2, ErbB2) represents the most advanced candidate of the Affibody molecules for human cancer diagnostics and it may potentially be adapted for therapy as well (Andersen 2011; Baum 2010; Sandberg 2011). Several radio nuclides, labeling strategies and chelators have been evaluated for Affibody molecules (Lofblom 2010; Nygren 2008; Tolmachev 2008). Accumulation in kidneys and liver is, however, sometimes problematic for several small alternative binding proteins including the Affibody molecules (Hosseinimehr 2012; Miao 2010).

Interestingly, a systematic approach to miniaturize the Z-domain, using iterative structure-based design and phage display, has generated a functional 38-residue domain by stabilizing the first two helices to facilitate removal of the third helix (Braisted 1996). The structure of this minimized variant has been solved and used to construct a disulfide-bonded, cyclic 34-residue form with retained immunoglobulin-binding properties (Starovasnik 1997; Wells 1999). Loop-insertions in this miniaturized 2-helix variant have been performed (Reiersen 1999, 2000) and the miniaturization approach was recently modified through backbone cyclization (Jarver 2010). The 2-helix concept has also been applied to variants with specific binding properties to generate smaller and potentially improved molecular tracers (Fedorova 2011; Ren 2012; Ren 2009; Rosik 2011; Webster 2009). Such size reduction may simplify synthetic production and improve tracer characteristics, but it often results in destabilized forms. An alternative synthetic approach of full-sized Affibody molecules relies on native chemical ligation of two shorter, more readily produced, synthetic peptides (Lindgren 2012). In summary, α-helical scaffolds have many desirable properties for protein engineering. This is demonstrated by the massive protein engineering performed on the Z-domain to generate an array of affinity proteins and other modified variants for widely different applications.

Beside the Affibody molecules that represent the best-validated member of this structural subgroup, for instance cytochrome b_{562} (Ku 1995), the *E. coli* colichin E7 immunity protein (ImmE7) (Juraja 2006), the

immunity protein (Im9) (Bernath 2005; Levin 2009), a small zinc-finger protein (Bianchi 1995) and binding proteins based on the C1-domain of SPG (Mandal 2012) have been used for similar purposes. However, these examples often utilize randomized loop regions or random PCR-based mutagenesis and neither has yet been explored much beyond an initial study.

4. THE ALBUMIN-BINDING DOMAIN (ABD), HOMOLOGS AND ENGINEERED VARIANTS

Streptococcal protein G (SPG), which was identified in group C and G streptococci (Björck 1984; Reis 1984), is a multi-domain surface protein with independent binding sites for immunoglobulins and albumins (Akerstrom 1987; Fahnestock 1986; Guss 1986). It was reported in 1980 that various streptococci could bind human serum albumin (HSA) (Myhre 1980), and protein G was shown to be responsible for this interaction (Bjorck 1987). Various smaller fragments, and eventually isolated domains, of SPG have been identified and found to interact with albumin from several species (Nygren 1988; Nygren 1990; Sjobring 1988). SPG, here specifically from streptococcus G148 if not mentioned otherwise, contains three homologous immunoglobulin-binding domains and three homologous albumin-binding domains (Figure 12). The third albumin-binding domain, G148-ABD3, has been studied most extensively and it is simply referred to as ABD in this text. It has an affinity (K_D) for HSA of approximately 5 nM and also binds to albumin from mouse, rat and monkey (Johansson 2002a; Linhult 2002; Nygren 1990).

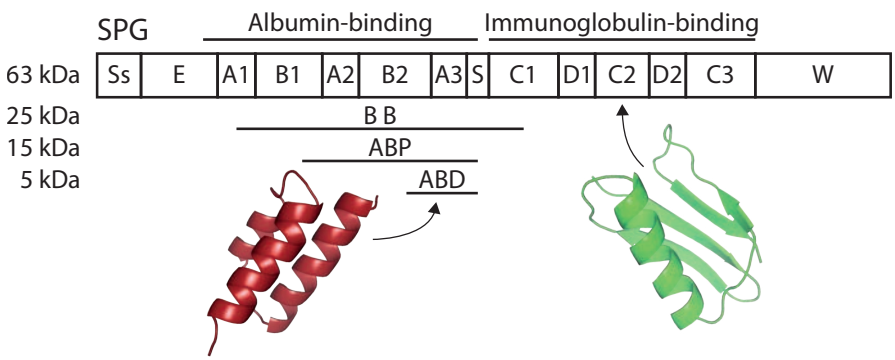


Figure 12. Schematic domain structure of streptococcal protein G and its derived albumin-binding regions. The albumin-binding regions (BB and ABP) and the third albumin-binding domain (ABD) are indicated. SPG also contains an immunoglobulin-binding region. C2 represents one of its immunoglobulin-binding domains. The figure was redrawn from (Linhult 2002) and domain structures were generated from PDB-entries 1GJT.pdb (Johansson 2002a) and 1FCC.pdb (Sauer-Eriksson 1995).

Several related albumin- and immunoglobulin-binding proteins have been identified (Boström 2012; Johansson 2002a), but SPG is one of the most investigated. It has found several applications in biotechnology, primarily for purification of antibodies, antibody fragments or Fc-fusion proteins (Grodzki 2010). The albumin-binding properties can be used for depletion of albumin from serum samples (Boström 2012). A related albumin-binding protein, peptostreptococcal albumin-binding protein (PAB), which contains another well-evaluated albumin-binding domain called ALB8-GA, was isolated from *Finegoldia magna* (formerly known as *Peptostreptococcus magnus*) (de Chateau 1994). These surface proteins have been suggested to help bacteria avoid detection by the immune system by covering themselves with host proteins (Figure 13) or to scavenge albumin-bound nutrients (Achari 1992; de Chateau 1996; Sauer-Eriksson 1995). Their difference in affinity for serum proteins from different species often reflect the host-specificities of the corresponding bacteria and may also be related to pathogenicity (Boström 2012; Johansson 2002a). Albumin-association is particularly interesting since albumin is the most abundant protein in plasma where it regulates the colloidal osmotic pressure. It also acts as a transporter for several compounds, which include fatty acids, amino acids, bile acids and steroids, and is known to carry several therapeutic substances (Curry 1998; Evans 2002). In contrast to antibodies, albumin is present at high concentration in the interstitial compartment and thereby has a larger volume of distribution (Anderson 2006).

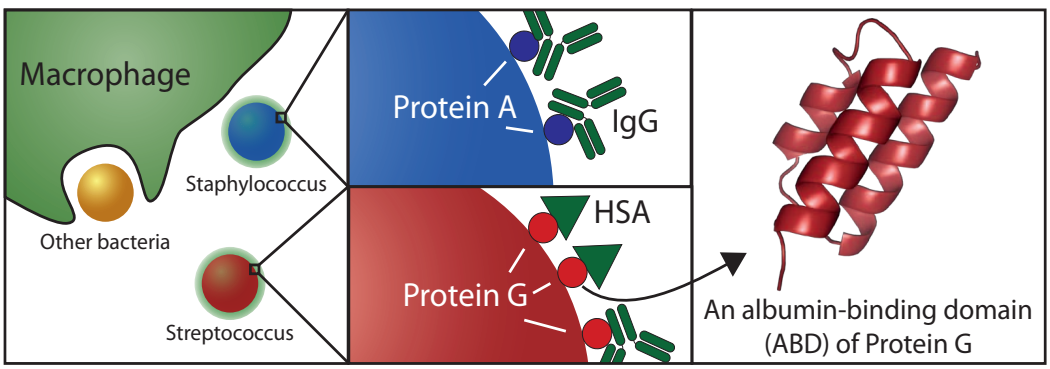


Figure 13. Hypothetical bacterial immune evasion strategy. Bacterial surface proteins such as the immunoglobulin-binding staphylococcal protein A (SPA, blue) and the bi-functional streptococcal protein G (SPG, red) have been suggested to help bacteria avoid detection by the immune system. This may be achieved by binding to abundant serum proteins. ABD is derived from SPG as indicated (1GJT.pdb, (Johansson 2002a)).

The albumin-binding domains of PAB, called ALB1-GA or ALB8-GA depending on the strain they originate from, are homologs to ABD. ALB8-GA, a name derived from protein G-related albumin binding, shares approximately 60 % sequence identity with ABD (de Chateau 1994). Both domains have stable hydrophobic cores and are remarkably tolerant to both pH and temperature (Johansson 1995; Johansson 1997; Kraulis 1996; Rozak 2005). Sequences of 16 albumin-binding domains derived from six proteins found in four bacterial species are shown in figure 14 to exemplify the sequence similarity between various albumin-binding domains. The figure also includes an artificial albumin-binding domain (PSD-1, (Rozak 2006)), a stabilized variant (ABDstable, (Gulich 2000)) and an affinity-matured ABD (ABD035, (Jonsson 2008)), which are described in more detail later.

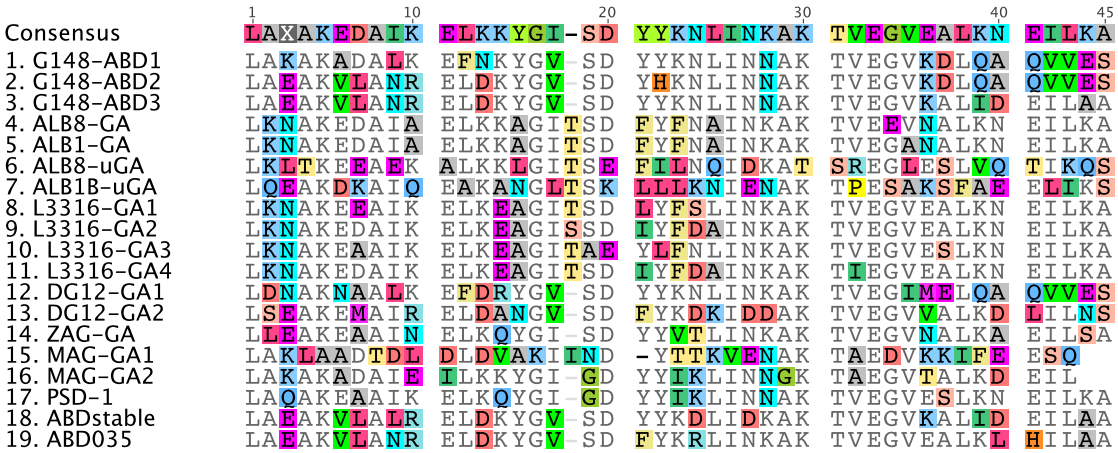


Figure 14. Sequence alignment of 16 homologous albumin-binding domains (1-16) and three engineered variants (17-19). Conserved amino acids are shown in gray and differences are highlighted in color. A consensus sequence is shown at the top of the alignment. Many homologs are not mentioned in the text and only included for comparison. Mainly G148-ABD3 (sequence number 3) and ALB8-GA (sequence 4) are discussed. The numbering above the alignment is most often used in the text; note the gap in some sequences at position 18 in the alignment. The picture was generated in Geneious Pro version 5.5.7 created by Biomatters and is based on similar pictures published in (Johansson 1997; Johansson 2002a).

ALB8-GA is commonly called a GA-binding module rather than domain because this motif represents the first example of a so-called module to shuttle between prokaryotic species (de Chateau 1994). Since its divergence from SPG, it is likely that the only important binding function of ALB8-GA in *F. magna*, which has only been isolated from humans, has been to recognize HSA. This may also explain the

stronger affinity of ALB8-GA for HSA as compared to ABD and its narrower species specificity. This subject has attracted much research attention, especially since the structures of these domains were solved. ABD has been thoroughly investigated using NMR (Johansson 2002a; Johansson 2002b; Kraulis 1996). Structures of free ALB8-GA (Cramer 2007) and ALB8-GA in complex with HSA exist (Lejon 2008; Lejon 2004) as well as several NMR-analyses of ALB8-GA (Johansson 1995; Johansson 1997; Johansson 2002a; Johansson 2002b). However, when NMR-derived structures are used to evaluate the binding interface they usually identify larger regions than the actual intermolecular contacts (Johansson 2002a).

Structural analyses and comparisons of ABD and ALB8-GA demonstrate that their structures can be overlain (Johansson 2002a) and that they share a common binding-site on albumin (Figure 15). The location of the binding-residues in ABD is indicated by conserved sequences in these regions among homologs ((Johansson 2002a), figure 14), which was confirmed by a mutational analysis (Linhult 2002). Mainly the second helix and the loops surrounding it are responsible for HSA-binding and the binding site on HSA is in domain II, as indicated by a structure of the homolog ALB8-GA in complex with HSA (Lejon 2004). At one point, residues in the first helix were thought to be of importance for binding (Johansson 1995), but this could not be verified in later studies (Johansson 1997) or in the structural complexes (Lejon 2008; Lejon 2004). The mutational analysis of ABD mainly emphasized the importance of residues located in the second helix (Linhult 2002), but, in addition to a hydrophobic interface, the binding site also consists of two hydrogen bond networks (Lejon 2004). A few residues in the domains seem to be of particular importance depending on their interaction with different albumin species (Cramer 2007; He 2006; Lejon 2004). For example a hydrophobic Phe21 in ALB8-GA interacts with a Met in HSA whereas Tyr20 in the corresponding position (position 21 in the alignment where the one residue shorter ABD has a gap) of ABD may interact more broadly with alternative residues on albumin from other species (Lejon 2004). The narrower specificity of ALB8-GA, which has a strong preference for primate serum albumins, has also been suggested to be a result of lowered flexibility as compared to ABD (Johansson 2002b). An inserted residue in the loop before helix 2 in ALB8-GA (Figure 14) may also be responsible for some of the observed differences (Lejon 2004). The more recently solved second crystal complex of HSA with ALB8-GA demonstrates that the GA-module is capable of binding a somewhat different conformation of HSA as well (Lejon 2008).

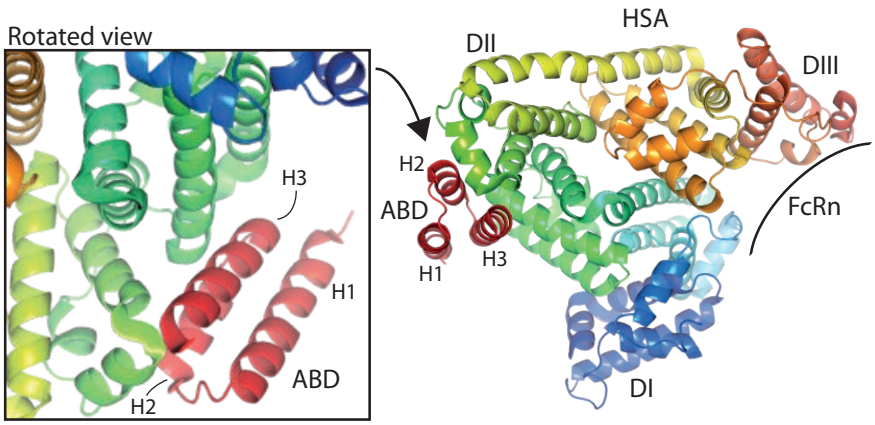


Figure 15. The binding site of an albumin-binding domain (ALB8-GA) on human serum albumin (HSA). The three helices of the albumin-binding domain are indicated (H1, H2 and H3) as well as the domains of HSA. Mainly helix 2 and the "backside" of helix 3 are in contact with albumin in the crystal complex. As indicated in the zoomed-in view, a novel binding site may potentially be accommodated on the surface formed by helices 1 and 3. The binding site of the neonatal Fc receptor (FcRn) on a different part of HSA is also indicated. The picture was generated from PDB-entry 1TFO.pdb (Lejon 2004).

The genotypic and phenotypic variations that exist among the group of albumin-binding domains have been used to explore the functional determinants within this limited sequence space (Rozak 2006). In this study, a library of albumin-binding domains was generated based on the existing diversity and this library was screened for variants capable of binding two albumins that represent opposite ends of the diverse phylogenetic range of molecules bound by ABD. These selections, carried out by phage display, identified a broadly reactive variant called phage-selected domain 1 (PSD-1, figure 14). NMR-analysis of PSD-1, which has less flexibility compared to ABD and includes Tyr20 (position 21 in figure 14) as opposed to Phe21 ALB8-GA, indicates that increased flexibility is not a requirement for broadened specificity in this case (He 2006). Instead, PSD-1 contains a core substitution that stabilizes its backbone in a conformation consistent with the albumin-bound ALB8-GA domain (He 2006; Rozak 2006). Together, these observations strengthen the previous suggestion that the broad specificity can mainly be attributed to Tyr20 in ABD (He 2006; Lejon 2004). Epitope-mapping of PSD-1 using NMR also indicated that the albumin-binding epitope includes several residues in the third helix, in addition to the most conserved and important amino acids in the second helix and the loop to the third helix (Figure 14, (He 2007)). Contributions to binding from residues in the third helix can also be seen in the crystal complexes of ALB8-GA and HSA (Lejon 2008; Lejon 2004). Interestingly, the suggested significance of residues in the

third helix of PSD-1, especially E40 (position 41 in figure 14), was not identified when an E40A-variant of ABD was analyzed in an earlier study (Linhult 2002).

Since asparagine residues are known to be sensitive to alkaline conditions, four such residues in ABD have been substituted to other amino acids represented among homologous sequences to generate an alkali stabilized form of ABD referred to as ABDstable ((Gulich 2000), Figure 14). This molecule has an increased melting temperature of at least 10°C, is more tolerant to chemical denaturation and has a much higher tolerance to alkaline treatment in chromatographic applications as compared to the original domain (Gulich 2000). It was hypothesized that one of the four substitutions (N9L) resulted in a stabilization of the hydrophobic core and, hence, would explain the improved thermostability. This stabilized ABD molecule has been used as a starting scaffold for the work presented in this thesis.

ABD has also been subjected to affinity maturation to yield a variant with very strong affinity for HSA (K_D 120 fM) (Frejd 2012; Jonsson 2008). Substitutions of up to 15 residues in helices 2-3 were investigated, using more conservative randomizations for residues thought to be directly involved in albumin binding. Two sub-libraries were used to also account for the extra residue in the loop between helices 1-2 found in some homologous domains (figure 14), but selections however only returned sequences with the shorter loop. Wild-type residues re-occurred in a majority of the selected clones in 9/15 positions, which implies that they are important for binding (S18, N23, L24, K29, T30, E32, G33, L37 and I41). Interestingly, the main candidate observed after affinity maturation (called ABD035, figure 14) contains phenylalanine at the position where a tyrosine is thought to be important for the broader specificity of ABD as compared to ALB8-GA and it seems like ABD035, still, has a similar specificity as the parental ABD. An Asn substituted with Arg (N23R) in ABD035 was hypothesized to be responsible for the improved solubility and lack of aggregation of this clone as compared to six closely related variants. All these related clones experienced such problems and shared a commonly observed and conserved Asn in this position (Jonsson 2008). A difference in the results from this study compared to observations from PSD-1 is that ABD035 had a lower thermal stability compared to its ancestor. A similar observation has been made for the affinity matured anti-ErbB2 Affibody molecule (Orlova 2006), where structural characterization demonstrated that improved affinity required recruitment (substitution) of residues that were originally involved in stabilizing intra-molecular hydrogen bonds (Eigenbrot 2010). Affinity maturation may also generate more- or as stable variants, and improved stability may increase affinity. For example, a rationally designed stabilizing intra-molecular disulfide bond has been shown to improve the affinity of an Affibody molecule by reducing the conformational entropy (Wahlberg 2006)

and a similar stabilization of the ErbB2-binding Affibody molecule has also been reported previously (Ekblad 2009), which indicates that such strategies may be useful if needed.

A library to select affinity proteins based on ABD was recently reported (Ahmad 2012). The approach taken in this study is similar to the strategy used to generate Affibody molecules (Nord 1997; Nord 1995) where an inherent binding surface (here for albumin) was replaced by a new interaction area generated by mutagenesis. This study is illustrated in some more detail in the present investigation since it is the only combinatorial library for affinity-protein selection based on ABD that has been reported, apart from results presented in papers included in this thesis.

4.1 *In vivo* half-life extension and its relation to ABD

Most proteins have a limited lifetime in the circulatory system and are either degraded or, if they are small enough, filtered out and eliminated by the kidneys. Protein half-life depends on several factors such as size, shape, hydrophilicity and sensitivity to degradation (Kontermann 2009). The openings in the glomerular filters in the kidneys are thought to be 50-100 nm fenestrations in the epithelial cell layer, which allow free diffusion of molecules (Haraldsson 2008; Kontermann 2011b). A porous structure inside the epithelium is filled with negatively charged proteoglycans that partially prevent passage of negatively charged macromolecules in the plasma. This filter has no sharp molecular mass cutoff, but hinders proteins with large hydrodynamic radii to pass to the urine and also introduces charge selectivity (Haraldsson 2008; Kontermann 2011b; Obeidat 2012; Plückthun 2009). Molecules up to the size of an scFv of roughly 25 kDa will largely pass through the filter, whereas molecules above 65 kDa will be almost completely retained. Albumin and IgG have extraordinarily long half-lives, 19 days for albumin and up to 23 days for IgG in humans, which distinguish them from other serum proteins of similar size. An important reason for these particularly long half-lives is an active rescue from endosomal degradation facilitated by the so-called neonatal Fc receptor (FcRn) (Andersen 2009; Anderson 2006; Roopenian 2007). Although referred to as the neonatal Fc receptor, FcRn influences IgG and albumin levels and tissue distribution at all stages of life (Christianson 2012; Roopenian 2007). FcRn is a heterodimer of the major histocompatibility complex class I-like H chain and the ß2-microglobulin L chain (Simister 1989) situated in acidic endosomes of virtually all nucleated cells (Anderson 2006; Kim 2007). It binds to endocytosed IgG and albumin at low pH, using two distinct binding sites, and releases them at physiological pH (Andersen 2009; Anderson 2006; Christianson 2012). FcRn thereby rescues bound proteins from degradation in the lysosomal compartment and transports them to the cell surface for release

at neutral extracellular pH. Structures of FcRn and FcRn in complex with Fc are available and demonstrate the importance of histidine residues for this pH-dependent interaction (Anderson 2006; Burmeister 1994a; Burmeister 1994b). Interestingly, the cycling between a neutral and an acidic environment of antibodies has been utilized to promote lysosomal degradation of an antigen with a pH-dependent dissociation (Igawa 2010). This strategy facilitates release of free antibody to encounter antigen in several cycles, in contrast to many other antibodies that remain bound to their targets throughout the FcRn-mediated endosomal cycle.

The role of FcRn in the regulation of IgG has been known for some time. That it also regulates the half-life of albumin was realized by serendipity when it was observed that albumin was co-purified with FcRn (Chaudhury 2003). The FcRn-binding site on albumin has not yet been determined by crystallography, but been suggested to reside in domain III of albumin (Andersen 2012; Anderson 2006; Chaudhury 2006). Thus, FcRn-mediated recycling and ABD-binding should not be competing processes since they occur at distinct sites on albumin (figure 15). This has also been experimentally validated for an ABD-fusion to an anti-ErbB2 Affibody molecule (Andersen 2011), and, therefore, a potential conformational change of albumin in the more acidic endosomal environment (Anderson 2006) does not prevent ABD-binding. FcRn binds Fc at the hinge-region to the same site that immunoglobulin-binding domains of SPA and SPG recognize (Roopenian 2007), which is discussed further below. The importance of FcRn in albumin homeostasis is illustrated by studies in FcRn deficient mice where albumin levels are 60 % lower as compared to levels in wild type mice (Andersen 2009). Studies of IgG have shown that its elimination rate in the knockout strain was increased 10-15 fold (Lobo 2004). A mouse FcRn knockout model that is transgenic for human FcRn is also available, which may facilitate translation of results between species (Andersen 2009).

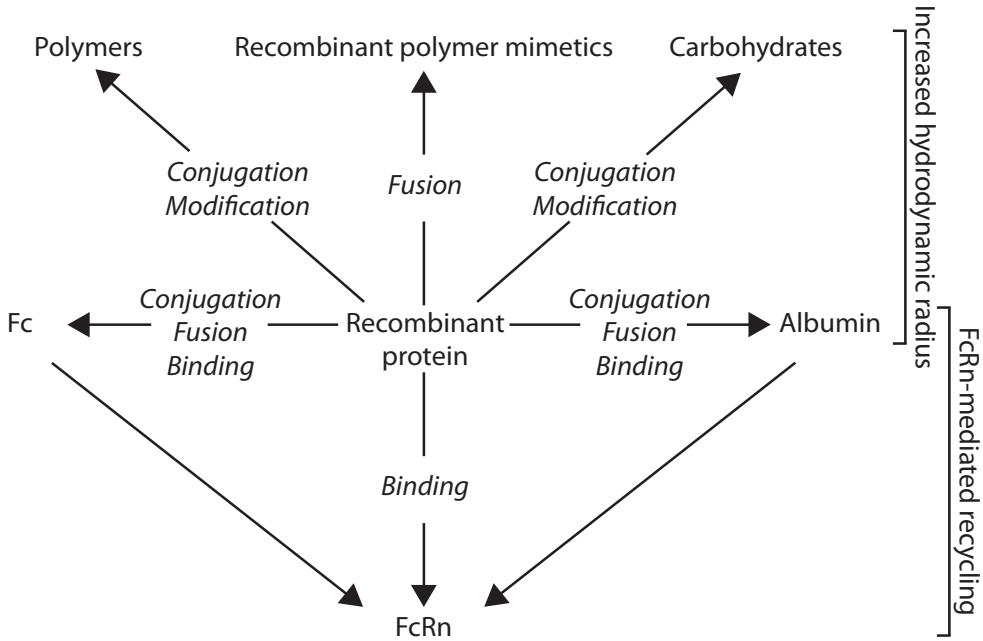


Figure 16. Summary of strategies to extend protein half-life and how they may be related to an increased hydrodynamic radius or FcRn-mediated mechanisms. Redrawn from (Kontermann 2009). Strategies that rely on non-covalent association with albumin are the focus of this section.

For the reasons mentioned before, protein therapeutics often suffer from short serum half-life. Therefore, several strategies to extend the half-lives of proteins, especially the ones with sizes below the renal filtration cutoff, have been developed (Figure 16). These efforts aim at increasing the total drug exposure, reduce dosing frequency and may, sometimes, also limit requirements for hospitalization (Kontermann 2009, 2011b). Glycosylation or conjugation to polyethylene glycol (PEG) has been performed to increase the hydrodynamic radius of proteins and thereby prevent or decrease renal elimination (Kontermann 2011b). Although site-specific conjugation strategies exist in approved drugs, high cost, intrinsic heterogeneity of PEG itself and potential problems with accumulation of metabolites in the kidneys present disadvantages with this strategy (Constantinou 2010; Gaberc-Porekar 2008). Similar approaches based on fusion to extended and unstructured polypeptides have also been employed (Kontermann 2009; Schellenberger 2009; Schlapschy 2007). Fusions to long-lived serum proteins, to the Fc-region of antibodies or albumin in particular, represent another means to achieve longer circulation times. Several albumin-fusion proteins have been described (for example (Flisiak 2010; Metzner 2009)) including conjugation of ErbB2-binding Affibody molecules to HSA (Hoppmann 2011), and such products have been approved as well (for example interferon-α linked to human serum albumin, Albuferon®).

Antibodies already contain an Fc-region to mediate FcRn-binding, but this interaction may be engineered to modulate their pharmacokinetics (Datta-Mannan 2007; Strohl 2009). Fc-fusion proteins can also make use of this mechanism, but it has been observed that some Fc-fused proteins have a weaker FcRn-binding, seemingly dependent on the fused domain, and thereby a shortened half-life (Suzuki 2010). Furthermore, pulmonary delivery of Fc-fusion proteins, relying on FcRn-mediated transport to the systemic circulation, has successfully made Fc-fusion proteins bioavailable, but it will need further improvement to find broader application (Bitonti 2004).

Non-covalent association to long-lived serum proteins has been explored, especially methods relying on albumin binding (Frejd 2012). Owing to its high concentration in serum (35-50 mg/ml or ~600 μM; (Kontermann 2009)) and long half-life, even weak interactions with albumin can play an important role (Nguyen 2006). Initial verification of this half-life extension strategy was performed using an albumin-binding region of SPG as fusion partner to soluble CD4 (Nygren 1991) or complement factor receptor type 1 (Makrides 1996). ABD-fusion has more recently been applied to recombinant antibodies (here a scDb) (Stork 2007), a bispecific diabody (Stork 2009) and Affibody molecules (Andersen 2011; Lee 2008; Tolmachev 2007; Tolmachev 2009). Other albumin-binding moieties have been successfully utilized as well, for instance albumin-binding linear peptides (Dennis 2007; Dennis 2002; Nguyen 2006), a cyclic peptide (Angelini 2012), several antibody fragments or domains (Holt 2008; Schlapschy 2007), a V-NAR domain (Müller 2012) or nanobodies (Coppieters 2006; Tijink 2008).

The studies mentioned above revealed that great improvements in pharmacokinetics can be achieved with half-lives approaching that of albumin itself and that this may also translate into improved accumulation in tumors for fused tumor-targeted affinity proteins (Kontermann 2011b; Stork 2007). Interestingly, direct comparisons to PEGylation (Stork 2009) and albumin-fusion (Walker 2010) have shown advantages of non-covalent association to albumin through ABD and the subsequent FcRn-mediated rescue. Surprisingly, a comparative study (Hopp 2010) using fusion constructs with a high affinity ABD (Jonsson 2008), a low affinity ABD (Linhult 2002) or a dual fusion of wild-type ABD, indicated that the half-life extending property of ABD was only weakly influenced by affinity for albumin or by the use of bivalent constructs. This result was unexpected considering the direct relationship between albumin-binding affinity and half-life previously observed for peptides (Nguyen 2006). One plausible interpretation is that a weak binding to albumin is essential to achieve half-life extension and that this property is related to albumin-binding affinity up to a certain limit, above which further improvements in affinity only have a marginal effect. It has also been argued that a higher affinity, which translates into a lower free fraction

that is not associated to albumin, will decrease the kidney uptake, especially of smaller fusion proteins (Frejd 2012). The idea that stronger affinity reduces kidney uptake has been verified using Affibody molecules fused to ABD035 ((Frejd 2012), Affibody AB, unpublished data). The existence of an affinity limit for half-life extension was demonstrated by using the very weak (with a K_D of approximately 0.5 mM) S18Y20K22/A ABD variant (Linhult 2002) in fusion to an Affibody molecule ((Frejd 2012), Affibody AB, unpublished data).

However, circulation times as long as for IgG molecules are generally not achieved for ABD-fusion proteins. This difference may arise from the higher affinity of IgG for FcRn at pH 6 compared to HSA or from differences in the serum concentration of these proteins (Andersen 2011). Consequently, the interaction between albumin and FcRn, rather than the binding affinity of the fusion protein for albumin, limits the recycling efficiency and, thus, the half-life (Hopp 2010). As an example, antibodies engineered for improved binding to FcRn at acidic pH have achieved prolonged half-lives (Zalevsky 2010). Therefore, several strategies to investigate immunoglobulin-binding domains in a manner analogous to work on ABD have been undertaken. An early report on such approach included fusions to either a single C2-domain or the BB-region from SPG (Figure 12) and resulted in a longer persistence for the albumin-binding fusion (Nygren 1991), which may have been due to the bivalency of the BB-domain (figure 12). The B-domain of SPA has also been used to improve the bioavailability of a bispecific scDb (Unverdorben 2012) and similar results were obtained in a different study using a dimer of the Z-domain (Mazor 2007a). Furthermore, a comparative study of immunoglobulin-binding domains from SPA, SPG or *F. magna* protein L (all composed of 50-60 amino acids) fused to an scFv or scDb demonstrated promising results, especially for the C3-domain of SPG. The observed differences are hypothesized to be due to differences in pH-dependency of the Fc-binding domains (Hutt 2011). A comparative study of half-lives of albumin and IgG in wild type and ß2-microglobulin knockout mice, lacking a part of the FcRn, demonstrated that the half-life of albumin was reduced from 35 h to 25 h whereas that of IgG was reduced from 95 h to 19 h (Chaudhury 2003). Consequently, IgG is more dependent on FcRn-recycling than albumin. As mentioned before, the main binding site of SPA-domains on IgG is located at the CH2-CH3 domain interface of the Fc-region, which is overlapping with the binding site for FcRn (Burmeister 1994b; Deisenhofer 1981; Kim 2007) and inhibition of binding of rat FcRn to human IgG using the B-domain of SPA has been demonstrated (Raghavan 1994). Results obtained using fusions to the B-domain (Unverdorben 2012) were similar to what was observed in FcRn knockout mice (Chaudhury 2003). Therefore, strategies relying on non-covalent association to this site on Fc may have an inherent limitation as a result of competition with FcRn-binding. An alternative approach relies on fusion of FcRn-binding

peptides, with a pH-dependent binding profile, to proteins (Sockolosky 2012). This strategy, however, does not address the renal elimination problem of smaller proteins. An important issue for all strategies described above is the potential influence on the activity of the fused therapeutic protein. Such effects need to be evaluated for each new combination and may be dependent on target, epitope etc (Tolmachev 2012). If half-life improvement comes at the expense of impaired function, higher doses may be required (Dennis 2002).

Since ABD is derived from a bacterial protein, potential immunogenicity needs to be considered. Yet, protein G probably evolved to help limiting immune detection. Previous experiments using BB as a carrier for vaccine candidates have shown that it may be immunogenic under such circumstances (Libon 1999; Sjolander 1997; Sjölander 1993). A T-cell epitope has been identified and localized to a sequence within the ABD (Goetsch 2003) as well as a, non-overlapping, antigenic region in helix 1 and the loop to helix 2 (based on studies of antibody epitopes in ABP, Johanna Steen et al. unpublished data). However, these studies generally intended to provoke an immune response via addition of adjuvants and, in contrast, an ABD gene fusion strategy has in fact been shown to reduce the immunogenicity of the partner moiety and to be well tolerated in repeated administrations (Frejd 2012; Kontermann 2009). Even using entirely human scaffolds does not guarantee that an immune response is not elicited, and randomization will generate novel characteristics, which may influence immunogenicity (Hosse 2006). Several examples of immunogenicity of partly or fully human proteins have been reported (Bender 2007; Casadevall 2005; Harding 2010; Hey 2005). Thus, immunogenicity is complex to predict and is affected by several factors, which also include route of administration, and it does not simply depend on the amino acid sequence alone (De Groot 2007; Mukovozov 2008). Even a single point mutation may render a protein immunogenic (Hosse 2006) and one may argue that there is a higher risk associated with using human proteins since they may evoke an autoimmune response (Binz 2005a). Protein therapeutics may be scanned for T-cell epitopes, which can then be removed by mutagenesis (De Groot 2005; Jones 2005; Tangri 2002). Such strategies have been applied to generate a de-immunized variant of ABD035, and proved successful based on a T-cell proliferation assay where this molecule gave very weak responses similar to recombinant HSA ((Feldwisch 2012; Frejd 2012), Affibody AB unpublished data). Immunogenicity is, at least in theory, closely related to protein size since a larger protein has a higher probability of including potentially immunogenic regions. Therefore, a small protein such as ABD would have an inherent advantage in this aspect. To conclude, only clinical trials will provide a reliable picture of a protein's immunogenic behavior in humans and this issue affects all proteins regardless of their origin.

5. PRESENT INVESTIGATION

5.1 Aim

The aim of the projects in this thesis was to explore the potential of a small albumin-binding domain (ABD) consisting of 46 amino acids as a scaffold for engineering of bispecific affinity proteins. The native binding site for albumin provides an interesting functionality, which can be exploited for affinity purification or *in vivo* half-life extension. By supplying this small protein with an additional binding site for a chosen target protein, bispecific domains with unique characteristics may be produced for use in several applications. The work in the first two papers in the thesis mainly concerns the technical challenges associated with this protein engineering task. The last two papers represent the first steps toward therapeutic applications where a small albumin-binding domain with an additional binding site for a suitable target, here two cancer-related members of the human epidermal growth factor receptor family, may provide a new strategy for therapeutic interventions.

5.2 ABD as a scaffold for protein engineering, design of a library for selection of bispecific domains and proof of principle

In some cases, evolution has converged on small protein domains with anti-parallel three-helix bundle structural motifs to perform several molecular recognition tasks (Frick 1992). Interestingly, very similar structures are found in the family of albumin-binding domains described in section 4 and among the immunoglobulin-binding domains derived from staphylococcal protein A (SPA). Stable hydrophobic cores contribute to the generally high stability of these small domain structures, which may be one reason that they have been conserved during evolution (Johansson 2002a). Using for example G148-ABD3 (PDB-entry 1GJT.pdb) as a template for so called structural alignment, i.e. querying the protein data bank for similar three dimensional structural motifs, returns a long list of proteins that contain a structural unit that is similar to this anti-parallel three helix bundle. This indicates that structure may sometimes be better conserved than sequence and that many naturally occurring proteins have adopted this favorable fold. The examples mentioned in the introductory sections of this thesis also demonstrate that both the Z-domain derived from SPA and several ABD molecules can tolerate a large number of substitutions and still preserve their overall structures. What is even more interesting for the work in this thesis is that some domains with a similar structure naturally possess two distinct binding sites on their relatively limited

surface areas, for example all five domains of SPA can bind to two distinct sites in either the Fc- or Fab-region of many immunoglobulins (Jansson 1998). Similarly, the immunoglobulin-binding domains of streptococcal protein G (Figure 12), although based on a structurally diverse fold, have such bispecific properties. The engineered proteins that could switch between two functionally distinct folds (He 2012), mentioned in the introduction, represent an additional example of bispecific protein domains. In summary, small and stable protein domains are promising templates for protein engineering and, given that some of them have a natural dual binding specificity, it would be attractive to investigate this characteristic for relevant pairs of target molecules by protein engineering.

When considering ABD as a potential scaffold for protein engineering, it is interesting to look into other scaffolds based on a α-helical structure. Apart from the Affibody molecules, a similar scaffold that utilizes randomization of surface exposed residues on helices is represented by the, so far unpublished, Alphabodies (www.complix.be). Recently, binders based on a three-helical Measles virus phosphoprotein P were published (Cyranka-Czaja 2012) and, excitingly, monospecific binding-proteins based on the same (but not stabilized) albumin-binding domain as described in this thesis (Ahmad 2012). These examples illustrate a common anticipation around the potential of these, small, three-helical alternative protein scaffolds, all of which seem to be inspired by the success of the Z-domain as a scaffold for the Affibody molecules. Figure 17 illustrates some available scaffolds with structures similar to ABD that have been published. As mentioned in the introduction, randomization of residues on a conformationally restricted surface rather than in loops has been used to obtain binding proteins, and such diversification of rigid structural elements has also been suggested to potentially reduce off-target interactions (Fleishman 2011). Moreover, indications that α -helical structures may potentially generate well-behaving proteins to a greater extent compared to other secondary structures, when random or semi-random libraries built from variable units of secondary structure have been screened for soluble proteins (Graziano 2008; Matsuura 2002), is another interesting feature.

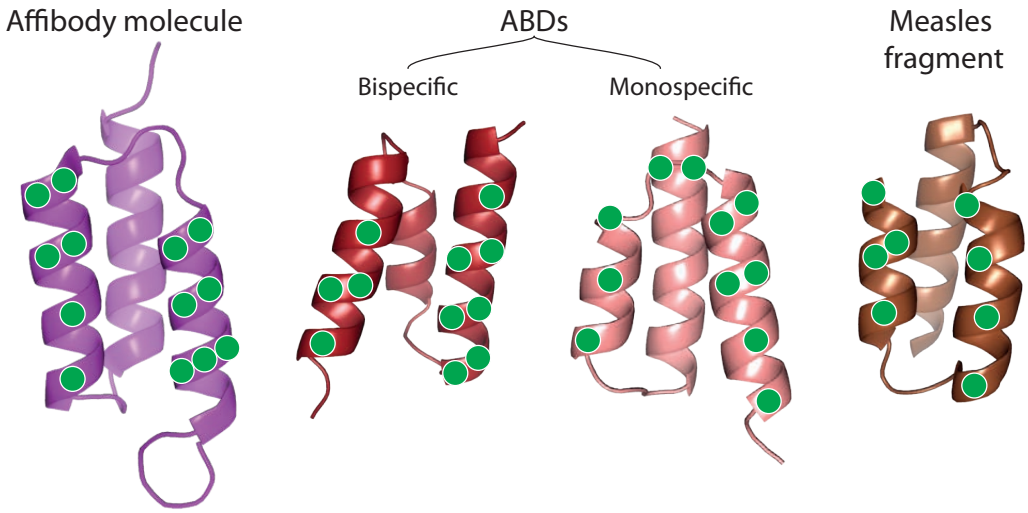


Figure 17. Comparison of small protein scaffolds with similar three-dimensional structures. A. A 58 amino acid Affibody molecule (generated from a structure of the Z-domain, 1Q2N.pdb (Zheng 2004)). **B.** The 46 amino acid albumin-binding domain used as a scaffold both for the generation of monospecific binding proteins (Ahmad 2012) and, in its stabilized form, for bispecific affinity proteins as described in this thesis (generated from 1GJT.pdb (Johansson 2002a)). **C.** A 49 amino acid fragment from a measles protein, which was recently reported for use as a scaffold protein (Cyranka-Czaja 2012). Here represented by one of several libraries constructed in this study (generated from 2K9D.pdb (Bernard 2009)).

Inspired by the naturally occurring three-helix bundle proteins, and particularly the natural bispecificity exhibited by the domains of SPA and the potential of a small domain with an inherent binding to albumin (see section 4.2), a library based on ABD was created (**Paper I**, Figures 18 and 19). A variant of G148-ABD3 called ABDstable, which was previously stabilized towards alkaline exposure (Gulich 2000) and thereby also gained an improved thermostability, was used as a template for this library construction. This molecule and variants derived from this library are referred to as ABD in the text in the remaining sections of this thesis, but it should be emphasized that this scaffold template contains a few substitutions as compared to G148-ABD3 (see Figure 20 for details). This figure also compares different library designs based on ABD and includes some variants mentioned in the text.

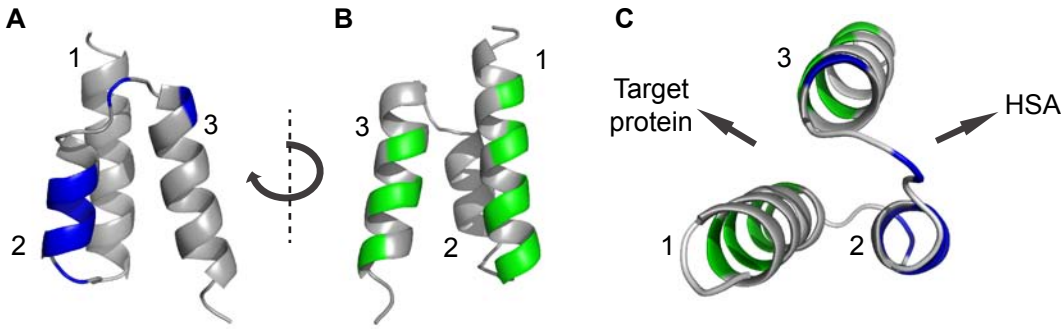


Figure 18. Strategy for engineering bispecificity into the albumin-binding domain. A. Residues mainly located in the second helix are responsible for albumin binding (blue, based on (Linhult 2002)). **B**. Eleven residues distributed mainly on the surfaces of helix 1 and helix 3 were randomized in the library (green) to generate a novel binding site. **C.** A top view illustrates how two spatially separated binding sites may be incorporated into the small ABD. The pictures were generated using PDB-entry 1GJT.pdb (Johansson 2002a).

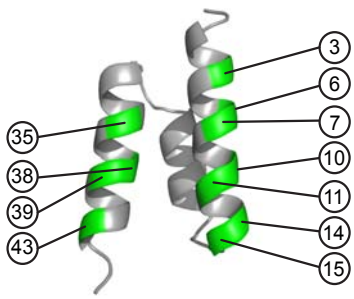


Figure 19. Library design. Eleven residues in helix 1 and helix 3, with positions indicated by numbers, where randomized (NNK) in the library. Based on the crystal complex of the related ALB8-GA module and HSA (figure 15), the flat albumin-binding area involves approximately one fourth of the total surface area of the domain (Lejon 2004). Consequently, a large area that does not take part in albumin binding is potentially available for protein engineering (which is also indicated in figure 18).

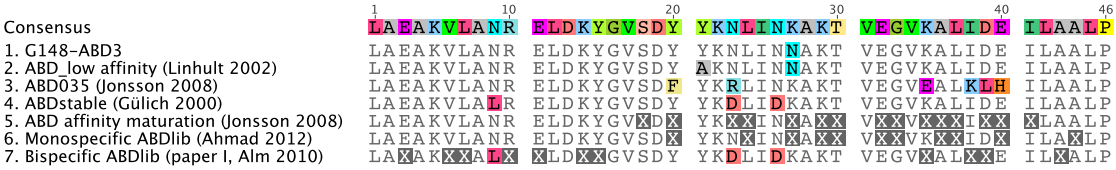


Figure 20. Sequence alignment of ABD-libraries and chosen ABD-variants. This alignment illustrates how the design of the ABD-library described in **paper I** (sequence 7 in the alignment) relates to a monospecific ABD-library (6, (Ahmad 2012)), the wild type G148-ABD3 (1), the affinity-matured ABD035 (3; (Jonsson 2008)), the ABD-library used for generation of ABD035 (5, (Jonsson 2008)), a low affinity (Y21A) variant (2; (Linhult 2002)), and the stabilized ABD molecule used as a scaffold for the bispecific library construction (4; (Gulich 2000)). A consensus sequence is shown at the top and differences between the sequences are highlighted. Randomized sequences in the combinatorial libraries are marked by X, regardless of which degenerate codons that were used in the library designs. Note that all sequences presented here are 46 amino acids and do not contain any gaps, which makes the numbering somewhat different as compared to in figure 14. Based on this sequence alignment, 27/46 amino acids (nearly 60 %) in ABD have been substituted (excluding the Y21A substitution in sequence 2 that is detrimental for albumin-binding, but not for structure), which illustrates how tolerant this domain can be for substitutions. The picture was generated in Geneious Pro version 5.5.7.

As a basis for the work included in **paper I**, the ABD-library was constructed from degenerate oligonucleotides where 11 surface-exposed residues on helices 1 and 3 were randomized (Figure 19) using NNK codons and cloned into a phagemid vector. These eleven residues were chosen for several reasons, i) they are all surface-exposed and point in an opposite direction as compared to the albumin-binding site in the structural complex of HSA and the homologous ALB8-GA (Lejon 2004), ii) they are not considered to be directly involved in albumin-binding (see section 4) and some of them have been substituted without any observable negative effects on overall secondary structure or albumin-binding ability in various alanine-substituted variants (Linhult 2002) and iii) none of them are highly conserved among homologous domains (Figure 14). Eleven residues randomized with a broad degenerate codon such as NNK generates a very large library (~$2 \cdot 10^{14}$ variants on the amino acid level) that generally cannot be fully covered experimentally. This is a common feature of many naïve libraries. In this case a larger area composed of several residues was included instead of a more restricted surface that would more easily be covered. In the alignment of homologs (sequences 1-16 in figure 14), all of the 11 randomized positions are represented by at least four different residues (and many by 7-8 different amino acids). The only residue that seems somewhat more conserved is E11, where 13/16 homologs have this glutamate. Regarding the randomization in helix 3, essentially overlapping patterns of randomization have

been used in the mono- (Ahmad 2012) and bispecific (**paper I**) library designs (Figure 20), respectively. This overlap agrees with structural data for how the residues point in the crystal complex (Lejon 2004), i.e. residues pointing toward HSA were randomized in the monospecific library and residues pointing away from HSA were diversified in the bispecific library (Figure 15). In addition, a comparison with the NMR-evaluation of the phage selected variant with a broad albumin-binding specificity (PSD-1, (Rozak 2006)) shows that none of the residues that were annotated as albumin-binding in that experiment overlaps with residues randomized in the bispecific library (He 2007). Several NMR-experiments (see section 4), and especially the two available crystal complexes (Lejon 2008; Lejon 2004), indicate that parts of the third helix are involved in albumin-binding, yet the opposite side of the same helix may still be subjected to substitutions without affecting this binding site (Linhult 2002). This observation is further verified by data obtained from bispecific ABD molecules presented in the papers included in this thesis.

The ABD-library was transformed into *E. coli* and, after infection with helper phages, functional display was demonstrated through selections using HSA as a target. Sequencing of 150 random clones from the non-selected library demonstrated that the design objectives were achieved and a Western blot on phage stocks using biotinylated HSA demonstrated that HSA-binding clones could be both displayed and enriched by phage display. After these initial validations, the Z-domain was chosen as a target to challenge the library for selection of binders to an antigen. Many commercially available matrices for affinity purification of antibodies or Fc-fusion proteins contain affinity ligands based on multimeric derivatives of the Z-domain (Hober 2007). Hence, binders selected against this target are expected to bind to these affinity matrices, which could be used for protein purification. Furthermore, ABD itself has been used as a purification tag before based on its albumin-binding property (Boström 2012; Nilsson 1987).

Four rounds of selection were performed using a dimeric form of the Z-domain as a target. The dimeric target was selected since it is more similar to the protein A-based affinity ligands used for affinity purification of antibodies and because it results in avidity in the selection. Sequencing of random clones enriched during selection revealed a clear dominance of one ABD-sequence, called ABDz1 (Figure 21). This clone constituted 100 % of the output from the fourth selection round and was also found in 84 % of the characterized clones originating from the third round. ABDz1 and two additional candidates (not shown here) were sub-cloned as His_6-tagged constructs in an expression vector and purified by immobilized metal ion affinity chromatography (IMAC). All variants were shown to have CD-spectra similar to the parental ABD and, excitingly, binding analysis using surface plasmon resonance (SPR) spectroscopy demonstrated that all three ABD-variants had retained HSA-binding. No positive selection

to enrich HSA-binding variants was performed in the selection against Z, so these clones retained their native binding function in spite of the additional substitutions. However, only ABDz1 bound the Z-domain used as target in the selection and with a relatively low affinity. ABDz1 expressed in *E. coli* could successfully be purified from cell lysates on a protein A-based matrix as well as on HSA Sepharose, which demonstrated that both the specificity engineered into this small domain and the inherent albumin binding could be used for affinity purification. To evaluate the general utility of these characteristics of ABDz1, it was utilized as an N-terminal fusion tag to three target proteins with different properties. The three fusion proteins, as well as the ABDz1 tag alone, were individually loaded on a HSA Sepharose column, which was then washed and eluted. Fractions constituting the eluted peak for each protein were pooled, adjusted to neutral pH and loaded on a protein A-based affinity column for an additional purification step. Samples taken before purification, after the first affinity purification step and after the full two-step procedure were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 22). The results clearly demonstrate that ABDz1 can be used as a general purification tag to obtain very pure fusion proteins using this orthogonal affinity purification strategy. As compared to classical orthogonal protein purification, where typically two different physical characteristics of a target protein are used for purification, the ABDz1-tag mediates such characteristics in a more general manner that does not require individual optimization for each target protein. Furthermore, a reversed order of the two purification steps was shown to give a comparable result (see **paper I**).

Figure 21. Sequences of ABDz1 and ABDz1_C6S. ABDz1 was identified when the ABD-library was panned against the Z-domain. A C6S-substituted variant of ABDz1 was used to address the importance of the cysteine residue. The library design is included at the top of the alignment for comparison. The picture was generated in Geneious Pro version 5.5.7.

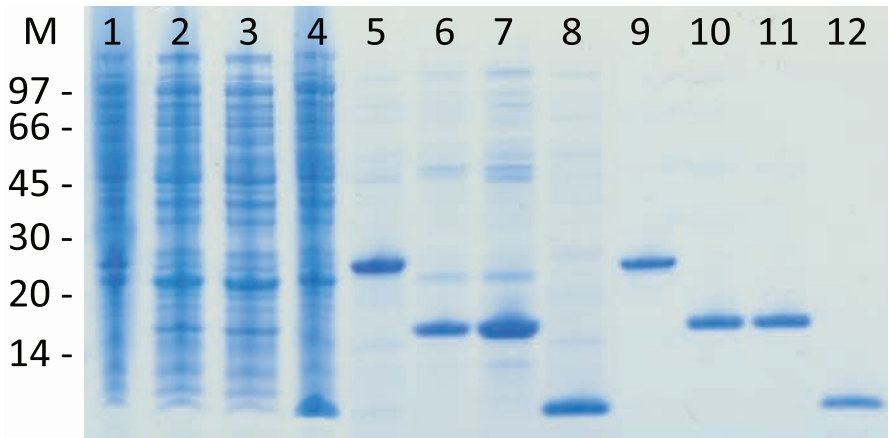


Figure 22. Orthogonal affinity purification using ABDz1 as a dual affinity tag. Lanes 1-4 on the SDS-PAGE gel represent *E. coli* lysates containing three different target proteins and ABDz1 alone before purification. Samples in lanes 5-8 have been purified on HSA Sepharose and samples in lanes 9-12 have also undergone a second purification step on a protein A-based matrix.

Since ABDz1 contained a cysteine in the first helix and the purified ABDz1-protein was shown to be present mainly in the form of homodimers, a C6S-substituted variant (Figure 21) was cloned, expressed and purified by IMAC. Interestingly, this variant lost its ability to bind the Z-domain, as measured by SPR. The indicated importance of the cysteine residue was further evaluated in an experiment where an ABDz1-sample, pre-treated with the reducing agent dithiothreitol (DTT), was loaded on a protein A-based column. This sample lost its ability to bind to the matrix upon addition of reducing agent, which demonstrates that the ABDz1-containing proteins need to associate in the form of disulfide-bridged dimers in order to be effectively purified using this method.

5.3 Engineering of high affinity bispecific binding proteins based on ABD

The encouraging results obtained using the bispecific properties of ABDz1 described in **paper I** led to the question whether such bispecific characteristics could be a common property of members of the ABD-library, as intended in the library design. The relatively small size of the library (about 10^7 variants as determined from plating after library transformation), and the moderate affinity of the single positive clone ADBz1 in the initial selection, also raised the question whether the current library was sufficient to generate binders to different target proteins. Another concern was whether the two binding sites in the

same domain might interfere with or somehow limit each other. The requirement for dimerization of ABDz1 could also indicate a limitation of the surface area that formed the new binding site since two molecules formed the Z-binding surface together. Therefore, selection of binders to TNF-α was undertaken (**paper II**). This protein is a validated target for several therapeutic proteins and is known to play an important role in inflammatory processes in rheumatoid arthritis and several other disorders. However, in our study TNF-α should rather be considered as a model target protein used to look into how higher affinities could be engineered into bispecific ABD molecules. The panning against the Z-domain implied that at least three rounds of selection were needed to obtain a phage pool enriched in binders (ABDz1 was present in less than 1 % of the clones after two rounds, **paper I**), therefore four rounds of selection against biotinylated TNF-α were performed. Sequencing of randomly picked clones identified 19 unique variants, two of which had sequence similarities (Figure 23) and were also shown to bind both HSA and TNF-α in binding analyses performed by SPR (Figure 23). Interestingly, the most common variant, ABD_{TNF1}, bound TNF-α with an acceptable affinity, but had instead lost much of its affinity for HSA. In contrast, the second most common candidate, ABD_{TNF2}, bound more weakly to TNF-α but retained much of its HSA-binding affinity. These results indicate a difficulty of incorporating two high-affinity interactions in the same domain and motivated the design of a secondary library for affinity maturation to further investigate if two high affinity interactions could be obtained using the same molecule.

To evaluate the feasibility of using cell-display on staphylococci (as schematically illustrated in Figure 24) for affinity maturation, ABD_{TNF1}, ABD_{TNF2} and non-randomized ABD (ABDstable) were cloned into a staphylococcal display vector designed for this purpose. Functional expression on the cell surface and binding to HSA as well as TNF-α could be demonstrated (not shown), which inspired the use of this platform for sorting of an affinity maturation library based on the sequences of the two first-generation binding molecules. This library was designed to cover sequences that were similar to ABD_{TNF1} and ABD_{TNF2} but still be of a limited size ($<10^7$) to more conveniently allow full coverage during transformation to *S. carnosus* and sorting by using FACS. Three of the originally randomized positions were locked to the amino acid present in the two initial candidates identified by phage display. Following sequence evaluation of the cloned and transformed library to verify that it fulfilled the design criteria, FACS-sorting was performed in three rounds, with alternating cycles of amplification by cell growth, to enrich binders to TNF-α using the strategy illustrated in figure 24. Six parallel sorting strategies were utilized, two of which aimed at, in addition to selecting binders to labeled TNF-α, also select library members with ability to bind both TNF-α and HSA, labeled by different fluorophores, using multi-

parameter cell-sorting. Importantly, since each staphylococcal cell displays several copies of its encoded ABD molecule, the dual selection strategy does not necessarily enrich variants that bind both TNF-α and HSA simultaneously. As shown in figure 24, all binding signals were normalized for expression level through monitoring of the binding of a reporter protein (IgG) to the reporter tag (a dimer of the Z-domain).

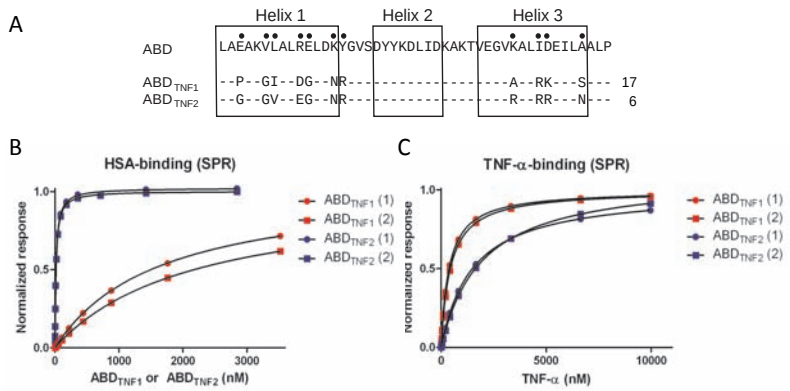


Figure 23. Characterization of ABD molecules targeting TNF-α selected by phage display. A. Two candidates, ABD_{TNF1} and ABD_{TNF2}, were observed several times after selection and they contained some sequence similarities. AND_{TNF1} was observed 17 times and ABD_{TNF2} 6 times among the sequenced clones after four rounds of phage display selection. Boxes indicate helical regions derived from the structure of G148-ABD3 and randomized positions in the library are labeled with dots. **B.** Both selected variants and the scaffold molecule (not shown in this figure) had retained binding to HSA, but the strongest TNF-α-binder, AND_{TNF1}, bound more weakly to HSA than AND_{TNF2}. **C.** Both variants bound TNF-α, a property that was absent in the non-randomized scaffold protein (not shown in this figure). The affinity for TNF-α of AND_{TNF1} was higher than for AND_{TNF2}. Normalized data from replicate measurements from an equilibrium binding analysis performed by SPR are shown. ABD_{TNF1} or ABD_{TNF2} were injected over immobilized HSA (B). In (C), TNF-α was injected over immobilized ADB_{TNF1} or ABD_{TNF2}.

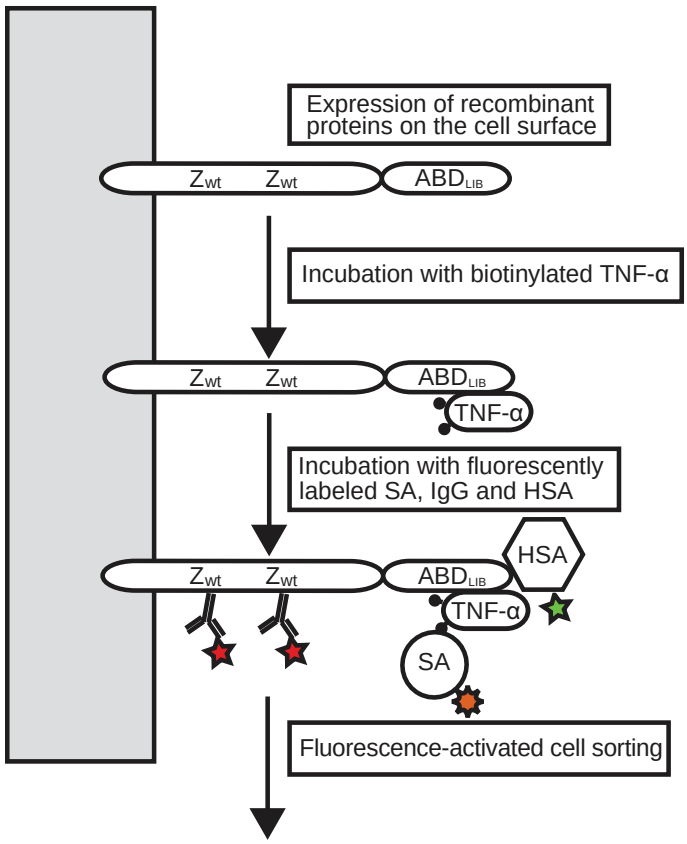


Figure 24. Schematic illustration of the display of an ABD-library on *S. carnosus* for selection of binders to TNF-α. The ABD-library was expressed as a fusion to a dimer of the IgG-binding Z-domain used for normalization of binding signal against expression level. In one sorting strategy, binders to TNF-α were isolated. A parallel strategy aimed at selecting cells that bound both TNF-α and HSA, which were detected using two different fluorophores. Bound biotinylated TNF-α was detected by labeled streptavidin (SA).

Affinity maturation by staphylococcal display identified several highly similar clones, which indicated that the selection was working as intended. Interestingly, selections targeting only TNF-α resulted in a set of clones that was distinct from the set identified when selection also monitored and enriched for HSA-binding. A few examples of commonly observed clones from these two selection strategies are shown in figure 25. Analysis of the binding characteristics of purified candidates from both sorting strategies by SPR demonstrated that low nanomolar affinities for TNF-α were obtained from both strategies, and that clones selected for retained HSA-binding did bind albumin with markedly higher affinities as compared to clones isolated only on the basis of TNF-α-binding (see **paper II**). In conclusion, affinity maturation by

cell-display could successfully isolate variants with improved binding characteristics, even when two such selection pressures were applied simultaneously. Improvements in the affinity for TNF-α as high as a 100-fold, as compared to the best first-generation binding molecule ABD_{TNF1}, were achieved. This demonstrates that two high-affinity binding sites can be accommodated on a small albumin-binding domain.

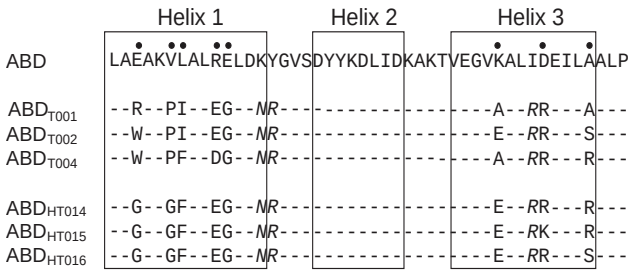


Figure 25. Common ABD-variants identified after cell display selection against TNF-α or TNF-α and HSA. Variants ABD_{T001}, ABD_{T002} and ABD_{T004} were derived from sorting targeting only TNF-α. The remaining candidates (ABD_{HT014}, ABD_{HT015} and ABD_{HT016}) were obtained from cells enriched for their capability to bind to both TNF-α and HSA in the same sorting cycle. However, since each cell displays the affinity proteins in a multivalent fashion, this sorting does not necessarily enrich for simultaneous binding to HSA and TNF-α. These six ABD-variants comprise a representative selection from a larger panel of candidates (**paper II**).

A closer examination of the sequences of the TNF-α-binders raises the question of the structural preservation of the parental three-helical fold in the obtained proteins. All isolated clones contain a high number of glycine and/or proline, especially in the region encoding the first helix of the ABD-scaffold, which may interfere with proper formation of this α-helix. HSA-binding probably requires a fold similar to what has been determined for ABD using NMR, or observed in the crystal complexes of its homolog ALB8-GA with HSA, to form the required intermolecular contacts. However, CD-spectroscopic characterization of several TNF-α-binding variants demonstrated that the spectra obtained (not shown) deviated from the typical appearance seen for helical proteins such as the original ABD. A possible interpretation of this lack of secondary structure is that these binding-molecules have an ability to adopt two different folds to engage in binding with either HSA or TNF-α. This hypothesis, although not verified by experiments, would also implicate that the selected molecules should not be able to simultaneously bind TNF-α and albumin, which was also shown using several SPR-based assays (**paper II**).

To extrapolate on the lack of simultaneous binding of the TNF-α-binding ABD molecules, a question of general importance is whether simultaneous binding is a desired property of a bispecific ABD molecule for a specific application. On one hand, simultaneous binding would increase the overall size of the bispecific molecule since a bound albumin molecule would be part of the overall molecular complex. This may be undesired for applications where a small size is of key importance, such as for efficient penetration into a solid tumor (as discussed further especially in **paper III**). On the other hand, for a more accessible target, for example a soluble protein present in the blood, simultaneous binding may be beneficial in order to decrease the free fraction of ABD molecule and thereby improve the half-life and limit undesired kidney elimination. Furthermore, the high concentration of albumin in the blood may potentially consume an entire administered dose of ABD molecules if it does not have a simultaneous binding capability, i.e. excess albumin may stockpile the ABD molecules and thereby make them unavailable for desired target-binding through their second binding interface. However, all such considerations will be highly target- and application dependent and this balance may potentially be modulated by the relative and absolute affinities for albumin and the second target, respectively. A functional balance of affinities may perhaps be determined from studies of ABD-variants where one affinity is adjusted while the second is kept constant. Such modifications have, so far, not been evaluated using bispecific ABD molecules. It is not yet known whether simultaneous binding to albumin would be feasible, and it would also demand that a suitable epitope on the target molecule can be recruited that does not interfere with the inherent geometry of albumin-binding. One could consider strategies to select specifically for molecules with capability for simultaneous binding. This topic is briefly discussed in the future outlook given in the final section.

5.4 ABD-based targeting of cancer-related members of the human epidermal growth factor receptor family

The ability to bind a target in the presence of albumin can be selected for by applying an excess of albumin during the encounter with the target to preferentially enrich variants with such characteristics. This strategy was applied in **paper III**, where binding molecules were selected both in the presence and absence of an excess of HSA. However, due to the very high concentration of HSA in the blood, experiments cannot easily mimic physiological conditions. The aim of the work included in **paper III** was to select high affinity binders to the human epidermal growth factor receptor 2 (ErbB2), which is a well-validated target for cancer therapy and thereby represents an interesting direction for further evaluation of the concept of bispecific ABD molecules. ErbB2 is a transmembrane protein that is over-expressed in

several cancer types, but present only at low levels (or not at all) in normal adult tissues. To generate ErbB2-binding ABD molecules, four rounds of phage display selection, using the same ABD-library as in **papers I** and **II**, were performed using biotinylated ErbB2 as a target. Excess HSA was not included in these initial selections but used as a second selection pressure during affinity maturation (see below). In line with the observations in **paper I**, the ErbB2-targeted selection returned one dominating candidate when random clones were sequenced. Interestingly, a repeated selection performed at another occasion from a separate aliquot of the same ABD-library gave this same candidate again. This could, however, be a result of cross-contamination or simply be due to the relatively small library size. Sequencing of nearly 1000 clones, after the second, third and fourth selection round, did not reveal any additional candidate with a distinguishable sequence similarity that would indicate an ErbB2-binding function. The commonly isolated variant, called $ABD_{ErbB2-1}$, was seen at a frequency of 20 % after two rounds and represented more than 80 % of the clones after the third and fourth rounds of selection. It bound ErbB2 with an affinity of 75 nM and demonstrated a somewhat improved affinity for HSA as compared to the parental ABD molecule. Improved HSA-binding has been observed several times in variants identified in **paper III** and **IV**. It is commonly the result of an increased association rate in the interaction with the negatively charged albumin (at physiological pH), which can be interpreted as an effect caused by the appearance of positively charged residues in the selected molecules.

In most applications, a higher affinity than 75 nM for ErbB2 would be desired, so affinity maturation was performed. Since only one ErbB2-binding candidate was available, library design for affinity maturation based on sequence similarity between several clones, analogous to the approach taken in **paper II**, was not possible. Instead, eleven residues in $ABD_{ErbB2-1}$ localized at the randomized sites in the ABD-library were individually substituted to alanine. The rational behind this method is that variants where an important residue has been replaced can be identified in an ErbB2-binding assay, and such data collected for all substituted variants may provide information for a library design. These substituted ABD-variants were expressed in *E. coli* and purified by HSA affinity chromatography, a strategy that seemed to be generally usable for purification of bispecific ABD-variants based on results from purification of ABDz1 and various TNF-α-binders described in **papers I** and **II**. The general applicability of this purification strategy indicates that the randomized positions normally do not interfere with albumin binding to such an extent that functional binding ability is lost.

All alanine-substituted variants derived from $ABD_{ErbB2-1}$ were subjected to a thorough binding analysis using SPR (**paper III**). Substitution of four of the residues to alanine led to a complete loss of binding.

This demonstrated that these residues were crucial for binding of $ABD_{ErbB2-1}$ to ErbB2. Interestingly, these residues are positioned close together on the engineered binding surface of ABD and, since both $ABD_{ErbB2-1}$ and its substituted derivatives were shown to give CD-spectra characteristic for molecules that fold like ABD, these residues probably form a local hot spot for ErbB2-binding. Based on these binding data, two affinity maturation libraries were designed (Figure 26A). In the first, more conservative library, the four critical residues were retained and the remaining seven positions were diversified using NNK-codons. In a second, semi-conservative approach, a mixture of degenerate oligonucleotides was applied in the library construction to preserve various combinations of important residues to be present at an average of 50 % in the sequences represented in the library (see **paper III** for more details). Selections from these libraries were performed in four rounds using phage display. In half of the selection tracks an excess of unlabeled HSA was present and in the remaining tracks selection was only aimed at improved ErbB2-binding. Notably, sequencing of a large number of clones (469) derived from these two selection strategies, and from the two libraries, revealed that the overall sequence output between the libraries and the selection strategies was very similar (Figure 26B and C). Many unique ABD-variants with a high degree of similarity among them were identified, which indicates that the new library designs included more ErbB2-binding sequences than the original, non-targeted, ABD-library. Moreover, the four residues identified as important for ErbB2-binding were present in all sequenced clones from both libraries and all selection strategies, even though they were allowed to vary in one of the library designs.

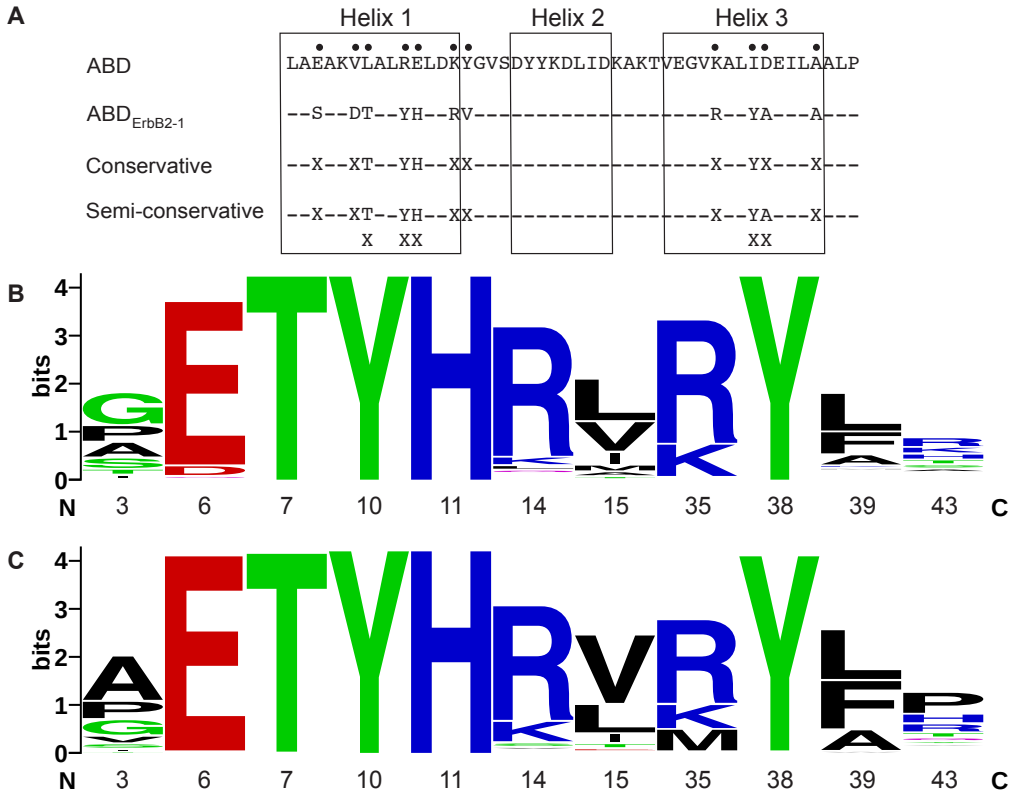


Figure 26. Affinity maturation of ErbB2-binding ABD molecules. A. Two affinity maturation libraries were designed based on the first generation candidate $ABD_{ErbB2-1}$. Four residues were retained in the conservative library design since they were crucial for ErbB2-binding (T7, Y10, H11 and Y28). In a semi-conservative library, all eleven residues were allowed to change, but the important residues and all possible combinations of them where retained to a high extent by using a mixture of degenerate oligonucleotides in the library construction. **B.** Sequence logotypes from sequenced colonies from selections without HSA present (271 sequences). The eleven randomized positions are shown from left (N-terminus) to right (C-terminus) and numbers indicate their location in the 46 amino acid ABD sequence. Logotypes were generated using Weblogo 3.3 and the overall height of each residue corresponds to its degree of conservation. The height within each stack relates to the relative frequency. **C.** Corresponding logotype for sequences from selections with HSA present (167 sequences). This strategy returned a surprisingly similar set of sequences compared to selections without HSA (logotype in B).

Expression and evaluation of a set of clones selected from the isolated repertoires revealed that several binders with low nanomolar affinity for both ErbB2 and HSA were represented, and that at least 80-fold affinity improvements, compared to $ABD_{ErbB2-1}$, had been achieved. Since the output from the phage display selections was too diverse to identify top candidates, staphylococcal display is anticipated to facilitate fine sorting of clones to isolate the best performing variants in the future. Thus, the selected repertoires were sub-cloned to the staphylococcal display vector (described in **paper II**), transformed to staphylococci and analyzed by flow cytometry. These results (**paper III**) clearly demonstrate that there exist differences between variants selected only to bind ErbB2 and variants selected for binding in the presence of HSA (Figure 27A). In general, ErbB2-binding signals of variants selected in the absence of albumin decreased significantly upon addition of an excess of HSA. In contrast, many clones selected with albumin present retained more of their ErbB2-binding. However, measurement of albumin-binding alone revealed that most clones selected only for ErbB2-binding had a retained albumin binding ability similar to $ABD_{ErbB2-1}$ (Figure 27B) whereas many clones selected in the presence of HSA apparently lost much of their albumin binding capacity. This behavior is indicated by the formation of a second population of cells with very low albumin-binding signals in repertoires selected using this strategy (Figure 27B). This dissimilarity may be expected from the two selection pressures that were applied, but considering the high degree of sequence similarity between the samples (Figure 26B and C) it was still somewhat surprising. Sorting by FACS and analysis of isolated variants from these distinct groups of clones is currently being undertaken to investigate these differences in more detail. The cell display data on affinity matured ErbB2-binding ABD molecules, as well as an apparent enrichment of variants with unintended substitutions in scaffold regions that may influence albumin binding (see **paper III**), show how challenging the incorporation of two binding sites may be and how these two binding sites may interplay.

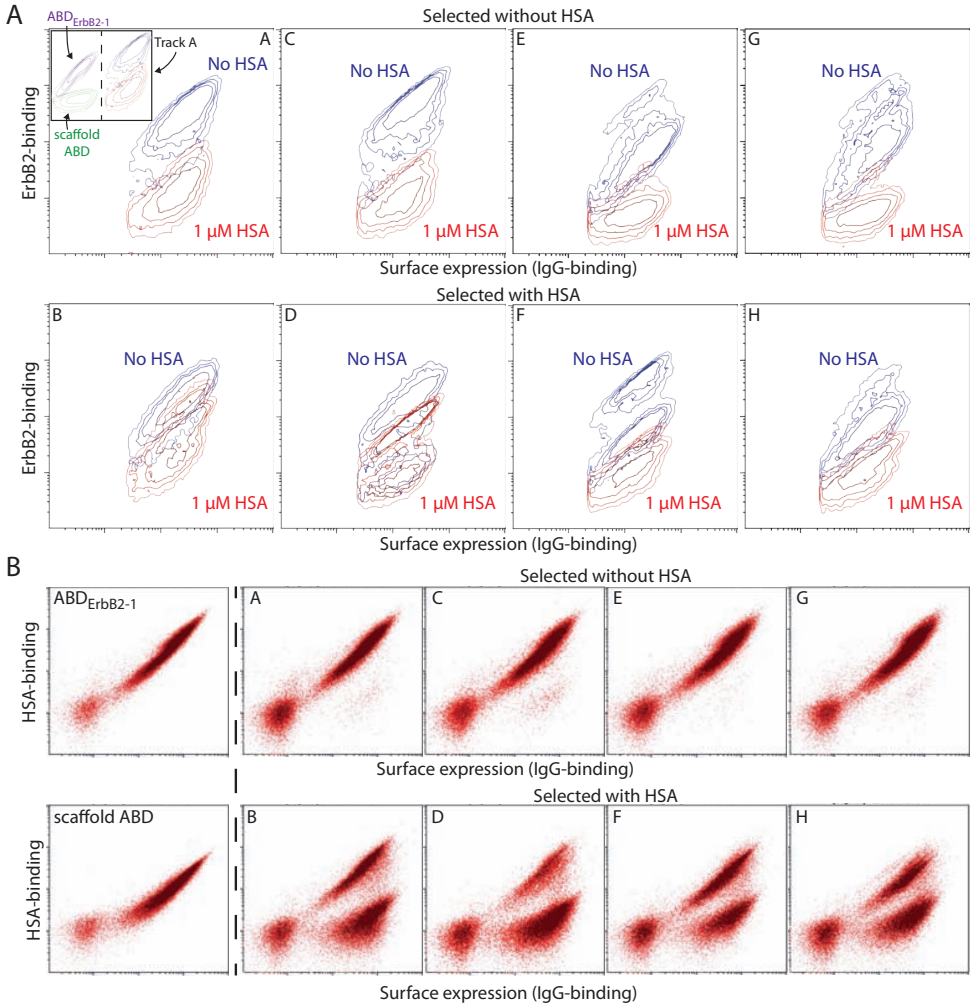


Figure 27. Cell display of clones pre-enriched by phage display. A. Cells expressing variants from tracks selected for ErbB2-binding without (A, C, E and G) or with (B, D, F and H) 1 µM HSA present were incubated with 50 nM ErbB2 +/- 1 µM HSA and analyzed by flow-cytometry. ErbB2-binding was detected through incubation with streptavidin-phycoerythrin (y-axis) and expression levels by IgG-Alexa Fluor 647 conjugate (x-axis), data are shown using logarithmic scales. Each diagram presents an overlay of contour plots of the expressing populations of cells from two separate samples from the same batch (A-H) incubated with only ErbB2 (blue) or ErbB2 and HSA (red). The contours represent cell density in the respective regions corresponding to 20, 40, 60 or 80 % of the maximum cell density observed. Each sample is represented by 25-30000 events. Corresponding data for $ABD_{ErbB2-1}$ and scaffold ABD are shown in the inlay of panel A. **B.** To analyze if the differences observed during cell display with and without unlabeled HSA present were due to overall differences in albumin-binding affinities, the samples were incubated with labeled HSA (50 nM) and compared to $ABD_{ErbB2-1}$ and scaffold ABD. Binding of HSA-Alexa Fluor 488 is shown on the y-axis and expression level on the x-axis.

In **paper IV** the ABD-library was screened for binders to the epidermal growth factor receptor 3, ErbB3, which is a cancer-associated receptor that depends on binding of a growth factor to become activated. ErbB3 deviates from ErbB2 since it requires a ligand to adopt an active conformation, whereas ErbB2 is constitutively in an active, ligand-independent conformation. Moreover, ErbB3 lacks an active tyrosine kinase domain while the one carried by ErbB2 is functional. As a consequence, ErbB3 lacks the ability to transduce a signal by itself, but heterodimerization generates potent signaling units that drive proliferation and differentiation. Interestingly, the selection against ErbB3 generated many more unique positive variants than the selections described in **papers I**, **II** and **III**. This indicates that the library, in spite of the more restricted output in previous selections and a relatively small size, has a capacity to deliver several binding clones for some targets. Seven of a total of 32 unique variants were expressed, purified and subjected to affinity measurements by SPR. All variants were shown to bind ErbB3 with affinities (K_D) in the order of 10-100 nM and all proteins bound strongly to HSA. Similar affinities were measured for the corresponding murine targets, which may indicate that a conserved epitope on ErbB3 was recognized. Most candidates had CD-spectra characteristic for folded helical proteins and were able to completely refold after heating to 90°C. Several variants did not reach complete unfolding and their melting temperatures where above 80°C.

Since ErbB3 contains a ligand-binding domain, a binding competition was performed with its ligand Neuregulin using SPR. This experiment demonstrated that all variants tested were capable of competing with Neuregulin for ErbB3-binding, a function that could potentially be antagonistic on the signaling that is activated through this ligand-binding event. Binding to ErbB3 and competition with Neuregulin was also demonstrated on a human cell line for the two variants with the highest ErbB3-binding affinities, referred to as ABD3-3 and ABD3-27 in the text. Interestingly, an Affibody molecule directed against an overlapping epitope on ErbB3 (Kronqvist 2011) that competes with these ABD molecules has been shown to have anti-proliferative effects on cancer cells *in vitro* (Gostring 2012), which implies that this shared binding site has a biological significance. In contrast to observations in previous selections, the highest affinity variants were not the most commonly observed candidates in the sequencing data after selection. Therefore, a functional assay to screen selected clones would be useful. In a way, the staphylococcal display platform can be used for such purposes and it also facilitates simultaneous monitoring of multiple parameters in a more flexible manner than a standard enzyme-linked immunosorbent assay. The utility of this approach is indicated by the analysis performed on ErbB2-binding ABD-variants pre-enriched by phage display (**paper III**) and a simple cloning step is the only requirement to transfer material from the phagemid vector to this staphylococcal display vector.

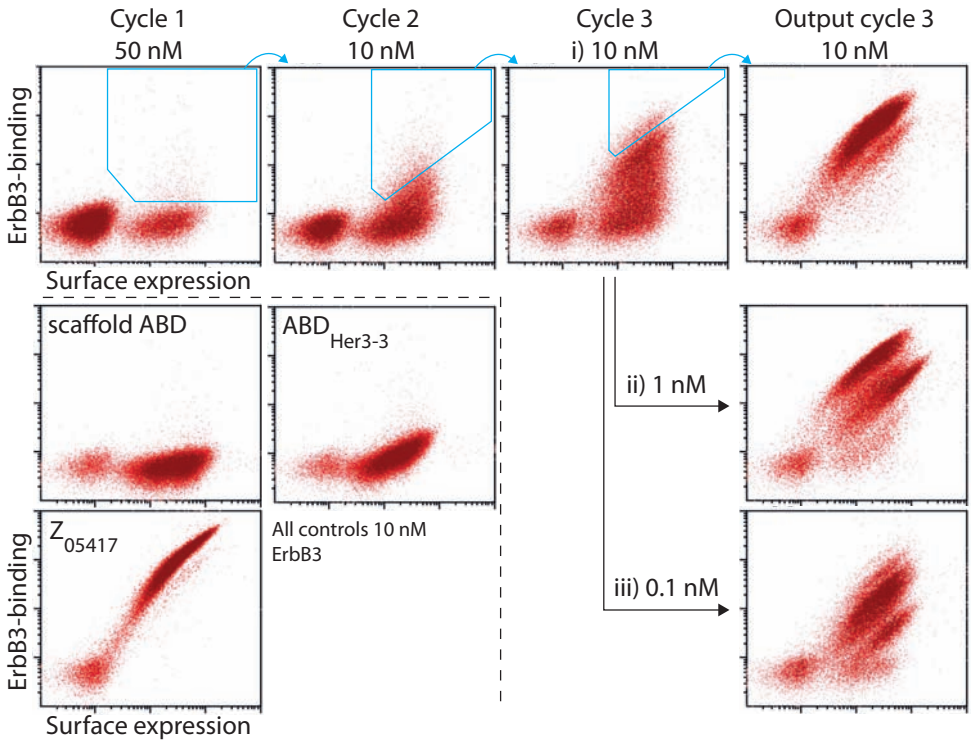


Figure 28. Preliminary data from affinity maturation of ErbB3-binding ABD molecules. An affinity maturation library for staphylococcal display was designed based on the first generation ErbB3-binding ABD molecules. This library was sorted in three cycles using decreasing concentrations of ErbB3. Promising enrichment of ErbB3-binding variants was observed over the three sorting rounds. Much stronger ErbB3-binding signals are obtained from candidates in the affinity maturation library as compared to ABD3-3, which is the first generation clone with the highest affinity. A comparison with an Affibody molecule with a sub-nanomolar ErbB3-binding affinity, Z05417 (Kronqvist 2011), used as a positive control, indicates that the selection is successful. Interestingly, more sub-populations can be distinguished in the output from selections performed with a lower target concentration in the third round.

Another interesting property of the ErbB3-binding variants, which were characterized in detail in **paper IV**, was the variation they exhibited in ErbB3-binding in the presence of different concentrations of albumin. In this experiment, ErbB3-binding was monitored by SPR for ABD molecules that had been pre-incubated with varying concentrations of albumin. Some variants lost the ability to bind ErbB3 already at a low molar ratio of albumin whereas others were remarkably tolerant to this challenge. Considering the overall similarity of the ABD-variants, these differences are a result of subtle sequence differences

between them. Of all ErbB2- and ErbB3-binding variants evaluated so far, neither has had the ability of simultaneous binding to albumin and the target molecule. In the intended application of tumor targeting, this may be beneficial if the equilibrium can be adjusted to a favorable balance. However, the expression level of ErbB3 is lower than for ErbB2, and therefore requires a higher affinity for an ErbB3-targeted ABD molecule. Consequently, work in progress aims to affinity mature the ErbB3-binding molecules by using the staphylococcal display platform. An affinity maturation library has been designed based on data from the first-generation variants described in **paper IV**. This library, which is not included in the paper, has been cloned and functionally expressed on staphylococcal cells. Initial sorting rounds show a promising enrichment of ErbB3-binding variants, which attain much stronger ErbB3-binding signals when compared to first-generation variants (Figure 28). Hopefully, this affinity maturation effort will generate an improved set of second-generation ErbB3-binding ABD molecules.

6. CONCLUSIONS AND FUTURE OUTLOOK

Based on the work summarized in this thesis, ABD is a promising scaffold for engineering of bispecific binding proteins. Selection of ABD-variants recognizing four different targets from a total of five combinatorial libraries, one naïve and four designed for affinity maturation, have been undertaken. Functional characterization of more than fifty ABD-variants derived from these libraries, either in the form of purified proteins or as individually expressed clones on staphylococcal cells, demonstrates that bispecific binding molecules based on ABD can be generated. Many more functional variants have been identified on the sequence- or batch level, respectively. ABD represents one of the smallest structured scaffolds reported and probably represents the smallest such bispecific affinity protein engineered to date. Together, the results obtained from selections using phage- and cell display against several target proteins bring ABD beyond a single initial study and demonstrates that the concept may be more broadly applied. ABD represents an interesting complement to the existing group of alternative non-immunoglobulin scaffold proteins and is unique with its inherent ability to bind serum albumin.

The small size and the albumin-binding ability of ABD may, as illustrated in figure 29, provide unique advantages of this molecule for tumor targeting. A small size is generally considered a desired property for fast penetration and accumulation in tumors. However, this advantage usually comes at the expense of rapid elimination via the kidneys. Since ABD has the ability to associate with albumin, it may have an advantage over other small and rapidly eliminated proteins in this respect. The studies included in this thesis demonstrate that it is possible to accommodate two high-affinity binding sites on the small and stable ABD molecule, but how these properties translate *in vivo* has not yet been investigated.

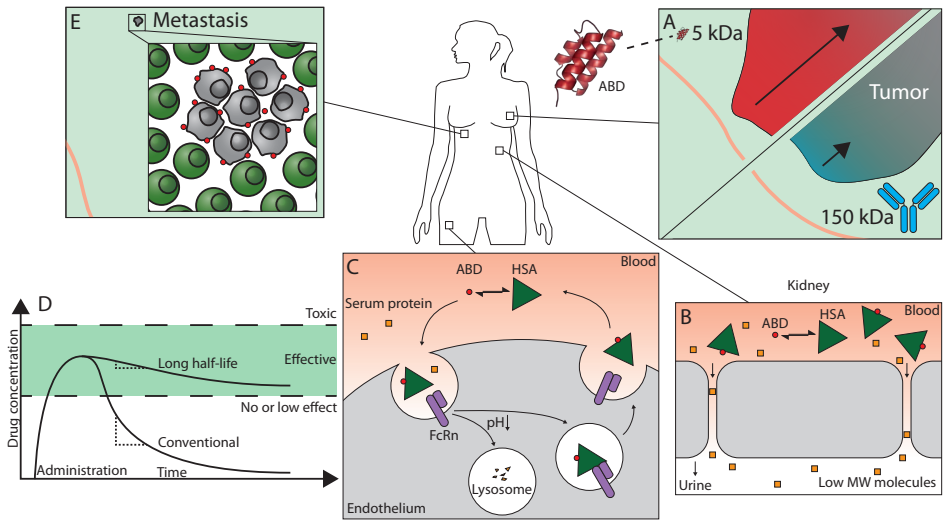


Figure 29. Hypothetical and speculative properties and use of bispecific ABD molecules for cancer therapy. A. Small ABD molecules are expected to penetrate tumor tissue more efficiently than much larger antibody molecules. **B.** Non-covalent association to albumin is known to decrease renal elimination of small proteins. In the illustration, the bispecific domains also utilize this ability of ABD. **C.** Albumin-binding can also be employed for FcRn-mediated rescue from endosomal degradation, provided binding is retained at lowered pH and does not interfere with FcRn-function. **D.** As a result of decreased renal elimination and FcRn-mediated recycling, a bispecific ABD molecule could possibly achieve a prolonged half-life *in vivo*. **E.** If armed with a suitable mechanism of action or payload, a small targeting agent in the form of a bispecific ABD molecule may provide a new tool for investigation in tumor therapy.

Albumin binding is a topic that has attracted much attention lately (section 4.1). Promising data for half-life extension of several proteins by association to albumin have been reported. A comparison of the new ABD-based affinity proteins to other available formats would be very interesting. One important question to answer is whether simultaneous binding is a prerequisite for efficacy *in vivo*. Considering the extremely high concentration of albumin, a very uneven competition for target binding may occur for a protein without ability for simultaneous binding. Local differences in target concentration may potentially shift the equilibrium in favor of target binding at particular locations, but manipulation of albumin-binding affinity would be a more straightforward means to optimize a molecule for specific targeting. Previous studies have demonstrated that even a weak albumin-binding affinity has a considerable effect on protein half-life *in vivo*. This could be investigated further by using a panel of ABD-variants where one binding surface is unaltered, for example for ErbB2 or ErbB3, while variants with different affinities for albumin are generated by site-directed mutagenesis. ABD could provide a useful model for such studies since the

effects of many substitutions may be predicted based on existing data and because several variants can be conveniently produced, purified and characterized using the methods developed in this thesis.

However, a higher affinity for albumin may potentially decrease undesired kidney accumulation and may also reduce the proteolysis of non-FcRn-associated protein in the endosomes, which could possibly cause immunogenicity mediated through presentation of ABD-derived peptides to the immune system. Moreover, experiences from the affinity matured- as well as the de-immunized variant of ABD may be incorporated into the bispecific format, especially if a higher albumin-binding affinity is desired. However, for applications demanding a strong albumin binding, simultaneous target binding may be necessary. Based on the orientation of the novel binding surface of ABD in relation to albumin in the available crystal complexes of ALB8-GA and HSA, simultaneous binding might still be possible. Perhaps this property mainly depends on the binding geometry that is allowed by the epitope recognized on the target molecule. Consequently, a system to preferentially select against epitopes that are compatible with simultaneous binding would be intriguing. In essence, a selection system for simultaneous binding has already been developed for other affinity proteins, i.e. the anchored periplasmic expression system where a binding protein is non-covalently captured on the cell surface via one of the desired binding capabilities for subsequent encounter with its target during selection. However, anchoring on cell-displayed albumin would be relatively complex and less complicated techniques would be more attractive. For example, capture of phages that monovalently express an ABD-library on albumin coupled to magnetic beads could be utilized. These beads may then be incubated with a fluorescently labeled target protein followed by FACS-sorting of the beads with associated fluorescence. This strategy has an attractive conceptual simplicity, but may require pre-enrichment of target-binding clones or signal amplification to achieve sufficient fluorescent signals for sorting. Naturally, use of multivalent display formats would be more complicated for selection of simultaneous binders since they may cause undesired additive effects from neighboring (identical) molecules that each carries out either one of the desired binding events. Several solutions should be possible for this kind of selection, but to be meaningful they request a therapeutically relevant target protein that is easily accessible in the circulatory system and has a small size or molecular geometry that may allow binding to a bispecific ABD already in complex with albumin. Small interleukins or cytokines could potentially provide one such group of target proteins.

An alternative future direction to technically improve the pipeline for generation of ABD molecules would be to generate a new, larger ABD-library. Library transformation and display level could perhaps be improved further by using another phagemid (Velappan 2010) in combination with a co-translational,

rather than a post-translational, display route. Possibly, a fully monovalent system could also be useful. For example, the signal recognition particle (SRP)-display system (Steiner 2008; Steiner 2006) was shown to improve the display level of the structurally similar measles protein derived scaffold mentioned before as compared to a more conventional system (Cyranka-Czaja 2012). The SRP-system was developed for fast-folding proteins and has, for example, successfully been used for display of DARPins (Steiner 2008; Steiner 2006) and Adnectins (Koide 2007). Another possibility to enable a higher flexibility and simplify the phage display selection protocols would be to explore the phage display systems that do not require helper phages (Chasteen 2006).

A hypothetical alteration of the bispecific approach would be to generate a library where the surface formed by helices one and two is randomized. In the albumin-bound ABD, this area appears to be more accessible than the current binding surface. However, since the second helix is shorter and known to be highly important for albumin binding, the current strategy may be more tolerable. A similarly unconventional option would be to incorporate the monospecific- library design (Ahmad 2012) in a high affinity bispecific ABD molecule. Since engineering of bispecific affinity proteins is an active area of research (Kontermann 2011a), the ABD molecule could potentially provide new combinations of binding specificities that do not necessarily include albumin. Moreover, since Fab-fragments fused to albumin-binding peptides have shown promising results as imaging agents (Dennis 2007), perhaps such applications should not immediately be ruled out. A similar format of bispecific proteins is the so-called Fcabs, which are Fc-fragments that incorporate an engineered binding site (Wozniak-Knopp 2010). However, both these immunoglobulin-derived formats are roughly 10-fold larger than the bispecific ABD molecules and also based on more complex molecular architectures.

ABD has been used as a half-life extending fusion partner to several other proteins. Perhaps an ABD molecule with an additional targeting ability could expand the utility of ABD in such applications. Our studies (**paper IV**) indicated that a fusion of ABD with a C-terminal Z-domain might interfere with target binding, at least for some ErbB3-binding ABD molecules. Interestingly, data from ABD-fused Affibody molecules indicate that, instead, an ABD molecule at the C-terminus of the promising ErbB2-binding Affibody molecule was preferred over an N-terminal position (Andersen 2011; Tolmachev 2007). Therefore, at least based on this pair of molecules, it seems like a format with a C-terminal bispecific (ErbB3-binding) ABD molecule fused to an N-terminal ErbB2-binding Affibody molecule to generate a very small tri-functional fusion protein may be feasible. A comparison of such fusion proteins, assembled from available ABD and Affibody molecules, with Affibody molecules fused to wild-type ABD, the

affinity matured ABD (ABD035) or single ABD or Affibody domains, would be exciting. Maybe an added functionality incorporated into the ABD-unit can provide improved homing or tumor retention of the fusion protein, but the risk of causing undesired agonistic effects mediated through receptor dimerization needs to be considered.

Several characteristics of the bispecific ABD molecules isolated so far raise interesting structural questions. For example, how can the binding geometry of ABDz1 depend on homodimerization through a disulfide-bridge? How does a flexible TNF-α-binding ABD molecule interact with the two target proteins and are there any differences between the molecules that were enriched partly based on HSA-binding as compared to candidates enriched only based on TNF-α-binding? Perhaps these proteins undergo a fold switching similar to the designed proteins shown in figure 3. They definitely share some characteristics with these proteins since they seem to be flexible and have a diminished stability that could potentially allow large conformational changes. Which epitopes on ErbB2 and ErbB3 the affinity proteins recognize, how they compare to known epitopes of available antibodies and how ABD-binding inhibits the binding of Neuregulin to ErbB3 are other interesting subjects?

Hopefully future experiments will be able to answer some of these questions and help finding an appropriate niche for further development of bispecific affinity proteins based on the ABD scaffold.

Acknowledgements

Tack **Sophia** för allt spännande under dessa år, för den frihet jag har fått, för att du är en förebild när det gäller att få ihop livspusslet och för att du alltid har något klokt att säga både om privata saker och när det gäller forskning! Hoppas att du lyckas lösgöra mer tid för forskning framöver.

Min bihandledare **John**, tack för allt stöd och alla tips som du har delat med dig av från din oerhörda kunskap om bibliotek, scaffolds och selektioner! Flera projekt fick ordentlig snurr när du och dina staffar gjorde entré. Tack också för trevligt rese- och konferenssällskap.

Tack **Stefan S** för din inställning och ditt trevliga sätt! Jag tvekade aldrig på att du var en grym bihandledare trots att vi aldrig hade något möte innan John successivt tog över.

Tove A, tack för en fenomenal introduktion och allt du har lärt mig. Tack inte minst för att du skapat ett grymt ABD-bibliotek! Extra kul var det att hälsa på lite på SciLife och bli filmstjärnor tillsammans.

Alla institutioner behöver en **Per-Åke** som med sin fantastiska kunskap och sitt engagemang i alla doktoranders projekt bidrar med något alldeles speciellt! Extra tack för hjälp med granskning av avhandling och alla tankar kring ABD, tack vare din hjälp vet jag att flera luckor i resonemangen och kunskapsluckor hos mig kunde täppas till.

Tack till alla som har läst och bidragit till hela eller delar av denna text: **Sophia**, **John**, **Per-Åke**, **Stefan N** och **Bitte**. Ett särskilt tack till **Mattias B**, som bidragit med majoriteten av det innehåll som en läsare tillgodogör sig, för din skickliga assistans med figurerna!

Tack till alla exjobbare för ert engagemang, för allt som jag har lärt mig om mig själv, om er och om handledning! **Åsa**, som tappert kastade sig in i CD44v6-djungeln för att fiska Affibodies under min spelade men egentligen obefintliga handledarpondus. **Stephan** som kom ända från Tyskland för att ge sig på dessa Affibodies och som hann med fantastiskt mycket på kort tid! Tack **Anneli** för tålmodig kloning och proteinrening med ABDz1, tänk vilken grym metod det visade sig bli. **Tove B** som inte bara skapade ordning kring CD44v6 utan senare också bidrog enormt till vårt bokkapitel. Tack för roligt sällskap på proteinkursen i Lissabon också. **Micke**, tänk om jag kunnat lika mycket som du när jag var ny. Du blev övergiven några veckor in på ditt exjobb och passade då på att lägga grunden för hela HER2/HER3-

historien, wow! ABD-bebisen överlämnas i trygga armar. Tack **Mariya**, som tålmodigt via google translate tolkade alla våra labbprotokoll för att hjälpa oss få ordning på vår ensamma stackars HER2-bindare. **Mattias B**, om du är så här grym på bioteknik och egentligen är en nanokille kan jag knappt föreställa mig vad du kommer att kunna åstadkomma inom det fältet i framtiden. Lycka till!

Tack till alla nya och gamla medlemmar i och omkring Sophias grupp för allt kul: **Tove A**, **Cilla**, **Hanna**, **Jenny**, **Karin L**, **Johanna**, **Anna**, **Tove B**, **Micke**, **Sara**, **Mattias**, **Cajsa**, **Henrik W**. Särskilt tack till **Louise** för hennes introduktion till ABD och fagdisplay och till **Sara** som såg till att vårt Zloop-projekt kickade igång efter att i ett par år legat i malpåse. **Anna**, jag kommer alltid att vara ditt fan! Tack **My** för en intressant inblick i Zbasic (och spindeltråd).

Charlotte, **Johanna, Henrik** och **Sebastian** för luncher och allt trevligt utanför jobbet! Det ska vi fortsätta med. Till sist blev även jag klar…

Nina K, **Per-Åke L** och **Andreas**, tack för hjälp med de första stapplande stegen i arbetet med TNF. **Nina**, **Magdalena**, **Filippa** och **Hanna**, tack för alla stalltips kring staffar. **Basia**, kan du inte komma och konsulta lite? Alla bandylirare, snart väntar en mer seriös comeback än det futtiga gästspel jag hunnit med på sistone. **Magdalena** och **Micke** för allt pillande med SPR-maskinerna och **Anna P** för att du håller ordning. **Håkan**, skönt att inte vara ensam om att ha hjärtat på det rätta stället! **Jochen** för en och annat hot pot i Peking, trots skönt SMS-missförstånd (you don't need to hang out if you don't want to), och din härliga peppande inställning. Alla i Sörgården, det har varit lite hårdrock på sistone, ska försöka labba lite och rätta till det snart. Tack **Emma**, **Inger**, **Kristina** och **Fredrik** för ert viktiga jobb!

Alla nya och gamla rumskompisar, **Emma L**, **Pawel**, **Tove A**, **Charlotte**, **Anna K**, **Mårten**, **Julia**, **Filippa**, **Magdalena**, **Chi**, **Andreas**, **Pavan**, **Ali** och **Amrita**.

Tack till övriga PIs på plan 3 **Mathias**, **Amelie**, **Tobbe**, **Helene** och precis alla andra både på plan 3 och på SciLife som bidrar till en så bra miljö!

Tack till **Marika**, **Hanna**, **Gamze**, **Hans** och **Karl** i Uppsala och **Ronny** för arbete med CD44v6. **Michael** och **Marie**, det ska bli spännande att se hur ni kan trolla med våra molekyler. **Adnande** och **Tatiana**, lycka till nu blir det spännande!

Tack Affibody, särskilt till **Elin** som trots viss tveksamhet engagerade sig i det knepiga CD44v6-arbetet och till **Fredrik** för hans smittande engagemang. Tack till **Niklas** och **Joanna** på AWA, det var spännande att få en inblick i er värld.

Stort tack till KVA, Stiftelsen PE Wahlbergs minnesfond, Ahlstrands testamentesfond, Hierta-Rezius fond för vetenskaplig forskning, Olof Ahlöfs Fond, Kemistsamfundet, Gålöstiftelsen, Wallenbergfonden, Elisabeth och Alfred Ahlqvists stiftelse, Stiftelsen Sigurd och Elsa Gojes Minne, Signeuls fond, Letterstedts resestipendiefond, Magnus Bergvalls stiftelse och Stiftelsen Tornspiran för stöd och de möjligheter detta har gett mig.

Tack till Vetenskapsrådet och Knut och Alice Wallenbergs Stiftelse för finansiering som har gjort detta arbete möjligt!

Tack till gamla kursare och alla vänner utanför KTH, nu är det slut på tryckandet i avhandlingsgrottan! Särkilt tack till min och Camillas familjer för allt stöd och tålamod. Tack **Bitte**, **Nicke**, **Emil**, **Britt-Mari**, **Jürgen**, **Caroline** och **Niclas**.

Tack mormor **Ulwa** för ditt stöd genom livet. Du såg fram emot alla festligheter men fick tyvärr inte möjlighet att delta, min avhandling tillägnas dig så att du kan vara med här ändå.

Camilla, tack för att du alltid tar hand om mig, för allt roligt vi har gjort under åren med allt ifrån lägenhetsrenoveringar till resor. Ofattbart vad tiden har gått fort, nu är det nästan exakt nio år sedan den där ölmässan och festen på KI. Jag älskar dig, beundrar dig för alla dina fina egenskaper och längtar till januari!

7. REFERENCES

Achari, A., S. P. Hale, et al. (1992). "1.67-A X-ray structure of the B2 immunoglobulin-binding domain of streptococcal protein G and comparison to the NMR structure of the B1 domain." Biochemistry 31(43): 10449-57.

Ackerman, M., D. Levary, et al. (2009). "Highly avid magnetic bead capture: an efficient selection method for de novo protein engineering utilizing yeast surface display." Biotechnol Prog 25(3): 774-83.

Ahlgren, S. and V. Tolmachev (2010). "Radionuclide molecular imaging using Affibody molecules." Curr Pharm Biotechnol 11(6): 581-9.

Ahmad, J. N., J. Li, et al. (2012). "Novel high-affinity binders of human interferon gamma derived from albumin-binding domain of protein G." Proteins 80(3): 774-89.

Akerstrom, B., E. Nielsen, et al. (1987). "Definition of IgG- and albumin-binding regions of streptococcal protein G." J Biol Chem 262(28): 13388-91.

Albericio, F. (2004). "Developments in peptide and amide synthesis." Curr Opin Chem Biol 8(3): 211-21.

Alexander, P. A., Y. He, et al. (2007). "The design and characterization of two proteins with 88% sequence identity but different structure and function." Proc Natl Acad Sci U S A 104(29): 11963-8.

Alexander, P. A., Y. He, et al. (2009). "A minimal sequence code for switching protein structure and function." Proc Natl Acad Sci U S A 106(50): 21149-54.

Almagro, J. C. and J. Fransson (2008). "Humanization of antibodies." Front Biosci 13: 1619-33.

Amstutz, P., H. Koch, et al. (2006). "Rapid selection of specific MAP kinase-binders from designed ankyrin repeat protein libraries." Protein Eng Des Sel 19(5): 219-29.

An, Z. (2010). "Monoclonal antibodies - a proven and rapidly expanding therapeutic modality for human diseases." Protein Cell 1(4): 319-30.

Andersen, J. T., B. Dalhus, et al. (2012). "Structure-based mutagenesis reveals the albumin-binding site of the neonatal Fc receptor." Nat Commun 3: 610.

Andersen, J. T., R. Pehrson, et al. (2011). "Extending half-life by indirect targeting of the neonatal Fc receptor (FcRn) using a minimal albumin binding domain." J Biol Chem 286(7): 5234-41.

Andersen, J. T. and I. Sandlie (2009). "The versatile MHC class I-related FcRn protects IgG and albumin from degradation: implications for development of new diagnostics and therapeutics." Drug Metab Pharmacokinet 24(4): 318-32.

Anderson, C. L., C. Chaudhury, et al. (2006). "Perspective-- FcRn transports albumin: relevance to immunology and medicine." Trends Immunol 27(7): 343-8.

Andreoni, C., L. Goetsch, et al. (1997). "Flow cytometric quantification of surface-displayed recombinant receptors on staphylococci." Biotechniques 23(4): 696-702, 704.

Angelini, A., J. Morales-Sanfrutos, et al. (2012). "Bicyclization and tethering to albumin yields long-acting peptide antagonists." J Med Chem.

Arber, W. and S. Linn (1969). "DNA modification and restriction." Annu Rev Biochem 38: 467-500.

Arora, P., T. G. Oas, et al. (2004). "Fast and faster: a designed variant of the B-domain of protein A folds in 3 microsec." Protein Sci 13(4): 847-53.

Barbas, C. F., 3rd, J. D. Bain, et al. (1992). "Semisynthetic combinatorial antibody libraries: a chemical solution to the diversity problem." Proc Natl Acad Sci U S A 89(10): 4457-61.

Barelle, C., D. S. Gill, et al. (2009). "Shark novel antigen receptors--the next generation of biologic therapeutics?" Adv Exp Med Biol 655: 49-62.

Bass, S., R. Greene, et al. (1990). "Hormone phage: an enrichment method for variant proteins with altered binding properties." Proteins 8(4): 309-14.

Baum, R. P., V. Prasad, et al. (2010). "Molecular imaging of HER2-expressing malignant tumors in breast cancer patients using synthetic 111In- or 68Ga-labeled affibody molecules." J Nucl Med 51(6): 892-7.

Beck, A. and J. M. Reichert (2011a). "The amazing, multipurpose antibody." MAbs 3(3): 221-2.

Beck, A. and J. M. Reichert (2011b). "Therapeutic Fc-fusion proteins and peptides as successful alternatives to antibodies." MAbs 3(5): 415-6.

Bender, N. K., C. E. Heilig, et al. (2007). "Immunogenicity, efficacy and adverse events of adalimumab in RA patients." Rheumatol Int 27(3): 269-74.

Benhar, I. (2001). "Biotechnological applications of phage and cell display." Biotechnol Adv 19(1): 1-33.

Berman, H., K. Henrick, et al. (2003). "Announcing the worldwide Protein Data Bank." Nat Struct Biol 10(12): 980.

Berman, H., K. Henrick, et al. (2007). "The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data." Nucleic Acids Res 35(Database issue): D301-3.

Bernard, C., S. Gely, et al. (2009). "Interaction between the C-terminal domains of N and P proteins of measles virus investigated by NMR." FEBS Lett 583(7): 1084-9.

Bernath, K., S. Magdassi, et al. (2005). "Directed evolution of protein inhibitors of DNA-nucleases by in vitro compartmentalization (IVC) and nano-droplet delivery." J Mol Biol 345(5): 1015-26.

Bertschinger, J., D. Grabulovski, et al. (2007). "Selection of single domain binding proteins by covalent DNA display." Protein Eng Des Sel 20(2): 57-68.

Beste, G., F. S. Schmidt, et al. (1999). "Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold." Proc Natl Acad Sci U S A 96(5): 1898-903.

Better, M., C. P. Chang, et al. (1988). "Escherichia coli secretion of an active chimeric antibody fragment." Science 240(4855): 1041-3.

Bianchi, E., A. Folgori, et al. (1995). "A conformationally homogeneous combinatorial peptide library." J Mol Biol 247(2): 154-60.

Binder, U., G. Matschiner, et al. (2010). "High-throughput sorting of an Anticalin library via EspP-mediated functional display on the Escherichia coli cell surface." J Mol Biol 400(4): 783-802.

Binz, H. K., P. Amstutz, et al. (2004). "High-affinity binders selected from designed ankyrin repeat protein libraries." Nat Biotechnol 22(5): 575-82.

Binz, H. K., P. Amstutz, et al. (2005a). "Engineering novel binding proteins from nonimmunoglobulin domains." Nat Biotechnol 23(10): 1257-68.

Binz, H. K. and A. Pluckthun (2005b). "Engineered proteins as specific binding reagents." Curr Opin Biotechnol 16(4): 459-69.

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

Bird, R. E., K. D. Hardman, et al. (1988). "Single-chain antigen-binding proteins." Science 242(4877): 423-6.

Birtalan, S., Y. Zhang, et al. (2008). "The intrinsic contributions of tyrosine, serine, glycine and arginine to the affinity and specificity of antibodies." J Mol Biol 377(5): 1518-28.

Bitonti, A. J., J. A. Dumont, et al. (2004). "Pulmonary delivery of an erythropoietin Fc fusion protein in non-human primates through an immunoglobulin transport pathway." Proc Natl Acad Sci U S A 101(26): 9763-8.

Bjorck, L., W. Kastern, et al. (1987). "Streptococcal protein G, expressed by streptococci or by Escherichia coli, has separate binding sites for human albumin and IgG." Mol Immunol 24(10): 1113-22.

Björck, L. and G. Kronvall (1984). "Purification and some properties of streptococcal protein G, a novel IgG-binding reagent." The Journal of Immunology 133(2): 969-974.

Blagodatski, A. and V. L. Katanaev (2012). "Technologies of directed protein evolution in vivo." Cell Mol Life Sci 68(7): 1207-14.

Bloom, J. D., S. T. Labthavikul, et al. (2006). "Protein stability promotes evolvability." Proc Natl Acad Sci U S A 103(15): 5869-74.

Bloom, J. D., M. M. Meyer, et al. (2005). "Evolving strategies for enzyme engineering." Curr Opin Struct Biol 15(4): 447-52.

Boder, E. T., K. S. Midelfort, et al. (2000). "Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity." Proc Natl Acad Sci U S A 97(20): 10701-5.

Boder, E. T. and K. D. Wittrup (1997). "Yeast surface display for screening combinatorial polypeptide libraries." Nat Biotechnol 15(6): 553-7.

Boder, E. T. and K. D. Wittrup (1998). "Optimal screening of surface-displayed polypeptide libraries." Biotechnol Prog 14(1): 55-62.

Boehm, T. (2011). "Design principles of adaptive immune systems." Nat Rev Immunol 11(5): 307-17.

Boersma, Y. L. and A. Pluckthun (2011). "DARPins and other repeat protein scaffolds: advances in engineering and applications." Curr Opin Biotechnol 22(6): 849-57.

Bostrom, J., S. F. Yu, et al. (2009). "Variants of the antibody herceptin that interact with HER2 and VEGF at the antigen binding site." Science 323(5921): 1610-4.

Boström, T., J. Nilvebrant, et al. (2012). Purification Systems Based on Bacterial Surface Proteins. Protein Purificaiton. R. Ahmad, InTech.

Bowley, D. R., A. F. Labrijn, et al. (2007). "Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage." Protein Eng Des Sel 20(2): 81-90.

Bradbury, A. R. and J. D. Marks (2004). "Antibodies from phage antibody libraries." J Immunol Methods 290(1-2): 29-49.

Bradbury, A. R., S. Sidhu, et al. (2011). "Beyond natural antibodies: the power of in vitro display technologies." Nat Biotechnol 29(3): 245-54.

Braisted, A. C. and J. A. Wells (1996). "Minimizing a binding domain from protein A." Proc Natl Acad Sci U S A 93(12): 5688-92.

Brannigan, J. A. and A. J. Wilkinson (2002). "Protein engineering 20 years on." Nat Rev Mol Cell Biol 3(12): 964-70.

Bratkovic, T., A. Berlec, et al. (2006). "Engineered staphylococcal protein A's IgG-binding domain with cathepsin L inhibitory activity." Biochem Biophys Res Commun. 349: 449-453.

Brouzes, E., M. Medkova, et al. (2009). "Droplet microfluidic technology for single-cell high-throughput screening." Proc Natl Acad Sci U S A 106(34): 14195-200.

Brown, S. (1992). "Engineered iron oxide-adhesion mutants of the Escherichia coli phage lambda receptor." Proc Natl Acad Sci U S A 89(18): 8651-5.

Bryan, P. N. (2000). "Protein engineering of subtilisin." Biochim Biophys Acta 1543(2): 203-222.

Buchanan, A., F. Ferraro, et al. (2012). "Improved drug-like properties of therapeutic proteins by directed evolution." Protein Eng Des Sel 25(10): 631-8.

Burmeister, W. P., L. N. Gastinel, et al. (1994a). "Crystal structure at 2.2 A resolution of the MHC-related neonatal Fc receptor." Nature 372(6504): 336-43.

Burmeister, W. P., A. H. Huber, et al. (1994b). "Crystal structure of the complex of rat neonatal Fc receptor with Fc." Nature 372(6504): 379-83.

Cabantous, S., T. C. Terwilliger, et al. (2005). "Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein." Nat Biotechnol 23(1): 102-7.

Cadwell, C. and G. Joyce (1994). Mutagenic PCR. PCR Methods and Applications, Cold Sprig HArbor Laboratory: S136-S140.

Casadevall, N., K. U. Eckardt, et al. (2005). "Epoetin-induced autoimmune pure red cell aplasia." J Am Soc Nephrol 16 Suppl 1: S67-9.

Cedergren, L., R. Andersson, et al. (1993). "Mutational analysis of the interaction between staphylococcal protein A and human IgG1." Protein Eng 6(4): 441-8.

Chames, P., M. Van Regenmortel, et al. (2009). "Therapeutic antibodies: successes, limitations and hopes for the future." Br J Pharmacol 157(2): 220-33.

Chao, G., W. L. Lau, et al. (2006). "Isolating and engineering human antibodies using yeast surface display." Nat Protoc 1(2): 755-68.

Charbit, A., J. C. Boulain, et al. (1986). "Probing the topology of a bacterial membrane protein by genetic insertion of a foreign epitope; expression at the cell surface." EMBO J 5(11): 3029-37.

Chasteen, L., J. Ayriss, et al. (2006). "Eliminating helper phage from phage display." Nucleic Acids Res 34(21): e145.

Chaudhury, C., C. L. Brooks, et al. (2006). "Albumin binding to FcRn: distinct from the FcRn-IgG interaction." Biochemistry 45(15): 4983-90.

Chaudhury, C., S. Mehnaz, et al. (2003). "The major histocompatibility complex-related Fc receptor for IgG (FcRn) binds albumin and prolongs its lifespan." J Exp Med 197(3): 315-22.

Chen, G., A. Hayhurst, et al. (2001). "Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS)." Nat Biotechnol 19(6): 537-42.

Christianson, G. J., V. Z. Sun, et al. (2012). "Monoclonal antibodies directed against human FcRn and their applications." MAbs 4(2).

Cirino, P., K. Mayer, et al. (2003). "Generating mutant libraries using error-prone PCR." Methods Mol Biol. 231: 3-9.

Cochran, J. R., Y. S. Kim, et al. (2006). "Improved mutants from directed evolution are biased to orthologous substitutions." Protein Eng Des Sel 19(6): 245-53.

Cole, M. F. and E. A. Gaucher (2011). "Utilizing natural diversity to evolve protein function: applications towards thermostability." Curr Opin Chem Biol 15(3): 399-406.

Constantinou, A., C. Chen, et al. (2010). "Modulating the pharmacokinetics of therapeutic antibodies." Biotechnol Lett 32(5): 609-22.

Coppieters, K., T. Dreier, et al. (2006). "Formatted anti-tumor necrosis factor alpha VHH proteins derived from camelids show superior potency and targeting to inflamed joints in a murine model of collagen-induced arthritis." Arthritis Rheum 54(6): 1856-66.

Couchman, J. R. (2009). "Commercial antibodies: the good, bad, and really ugly." J Histochem Cytochem 57(1): 7-8.

Craig, R. A., J. Lu, et al. (2010). "Optimizing nucleotide sequence ensembles for combinatorial protein libraries using a genetic algorithm." Nucleic Acids Res 38(2): e10.

Cramer, J. F., P. A. Nordberg, et al. (2007). "Crystal structure of a bacterial albumin-binding domain at 1.4 A resolution." FEBS Lett 581(17): 3178-82.

Crawford, F., K. Jordan, et al. (2006). "Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands." Immunological Reviews 210: 156-170.

Cull, M. G., J. F. Miller, et al. (1992). "Screening for receptor ligands using large libraries of peptides linked to the C terminus of the lac repressor." Proc Natl Acad Sci U S A 89(5): 1865-9.

Cunningham, B. C., P. Jhurani, et al. (1989). "Receptor and antibody epitopes in human growth hormone identified by homolog-scanning mutagenesis." Science 243(4896): 1330-6.

Curry, S., H. Mandelkow, et al. (1998). "Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites." Nat Struct Biol 5(9): 827-35.

Cwirla, S. E., E. A. Peters, et al. (1990). "Peptides on phage: a vast library of peptides for identifying ligands." Proc Natl Acad Sci U S A 87(16): 6378-82.

Cyranka-Czaja, A. and J. Otewski (2012). "A novel, stable, helical scaffold as an alternative binder — construction of phage display libraries." Acta Biochimica Polonica 59(3): 383-390.

Dahiyat, B. I. and S. L. Mayo (1997). "Probing the role of packing specificity in protein design." Proc Natl Acad Sci U S A 94(19): 10172-7.

Dalbadie-McFarland, G., L. W. Cohen, et al. (1982). "Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function." Proc Natl Acad Sci U S A 79(21): 6409-13.

Danilova, N. (2012). THE EVOLUTION OF ADAPTIVE IMMUNITY. Self and Nonself. C. López-Larrea, Landes Bioscience and Springer Science+Business Media: 218-235.

Datta-Mannan, A., D. R. Witcher, et al. (2007). "Monoclonal antibody clearance. Impact of modulating the interaction of IgG with the neonatal Fc receptor." J Biol Chem 282(3): 1709-17.

Daugherty, P. S. (2007). "Protein engineering with bacterial display." Curr Opin Struct Biol 17(4): 474-80.

Daugherty, P. S., G. Chen, et al. (1998). "Antibody affinity maturation using bacterial surface display." Protein Eng 11(9): 825-32.

Dawson, P. E. and S. B. Kent (2000). "Synthesis of native proteins by chemical ligation." Annu Rev Biochem 69: 923-60.

de Chateau, M. and L. Bjorck (1994). "Protein PAB, a mosaic albumin-binding bacterial protein representing the first contemporary example of module shuffling." J Biol Chem 269(16): 12147-51.

de Chateau, M., E. Holst, et al. (1996). "Protein PAB, an albumin-binding bacterial surface protein promoting growth and virulence." J Biol Chem 271(43): 26609-15.

De Groot, A. S., P. M. Knopp, et al. (2005). "De-immunization of therapeutic proteins by T-cell epitope modification." Dev Biol (Basel) 122: 171-94.

De Groot, A. S. and D. W. Scott (2007). "Immunogenicity of protein therapeutics." Trends Immunol 28(11): 482-90.

de Marco, A. (2011). "Biotechnological applications of recombinant single-domain antibody fragments." Microb Cell Fact 10: 44.

Deisenhofer, J. (1981). "Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution." Biochemistry 20(9): 2361-70.

Dennis, M. S., H. Jin, et al. (2007). "Imaging tumors with an albumin-binding Fab, a novel tumor-targeting agent." Cancer Res 67(1): 254-61.

Dennis, M. S., M. Zhang, et al. (2002). "Albumin binding as a general strategy for improving the pharmacokinetics of proteins." J Biol Chem 277(38): 35035-43.

Doi, N. and H. Yanagawa (1999). "STABLE: protein-DNA fusion system for screening of combinatorial protein libraries in vitro." FEBS Lett 457(2): 227-30.

Dooley, H., M. F. Flajnik, et al. (2003). "Selection and characterization of naturally occurring single-domain (IgNAR) antibody fragments from immunized sharks by phage display." Mol Immunol 40(1): 25-33.

Dryden, D., A. Thomson, et al. (2008). "How much of protein sequence space has been explored by life on Earth?" J. R. Soc. Interface 5: 953-956.

Dubel, S., O. Stoevesandt, et al. (2010). "Generating recombinant antibodies to the complete human proteome." Trends Biotechnol 28(7): 333-9.

Duenas, M. and C. A. Borrebaeck (1994). "Clonal selection and amplification of phage displayed antibodies by linking antigen recognition and phage replication." Biotechnology (N Y) 12(10): 999-1002.

Edwards, B. M., S. C. Barash, et al. (2003). "The remarkable flexibility of the human antibody repertoire; isolation of over one thousand different antibodies to a single protein, BLyS." J Mol Biol 334(1): 103-18.

Eigenbrot, C., M. Ultsch, et al. (2010). "Structural basis for high-affinity HER2 receptor binding by an engineered protein." Proc Natl Acad Sci U S A 107(34): 15039-44.

Ekblad, T., V. Tolmachev, et al. (2009). "Synthesis and chemoselective intramolecular crosslinking of a HER2-binding affibody." Biopolymers 92(2): 116-23.

Ellington, A. D. and J. W. Szostak (1990). "In vitro selection of RNA molecules that bind specific ligands." Nature 346(6287): 818-22.

Enever, C., T. Batuwangala, et al. (2009). "Next generation immunotherapeutics--honing the magic bullet." Curr Opin Biotechnol 20(4): 405-11.

Ernst, W., R. Grabherr, et al. (1998). "Baculovirus surface display: construction and screening of a eukaryotic epitope library." Nucleic Acids Res 26(7): 1718-23.

Esvelt, K. M., J. C. Carlson, et al. (2011). "A system for the continuous directed evolution of biomolecules." Nature 472(7344): 499-503.

Evans, T. W. (2002). "Review article: albumin as a drug--biological effects of albumin unrelated to oncotic pressure." Aliment Pharmacol Ther 16 Suppl 5: 6-11.

Ewert, S., T. Huber, et al. (2003). "Biophysical properties of human antibody variable domains." J Mol Biol 325(3): 531-53.

Fagete, S., U. Ravn, et al. (2009). "Specificity tuning of antibody fragments to neutralize two human chemokines with a single agent." MAbs 1(3): 288-96.

Fahnestock, S. R., P. Alexander, et al. (1986). "Gene for an immunoglobulin-binding protein from a group G streptococcus." J Bacteriol 167(3): 870-80.

Fedorova, A., K. Zobel, et al. (2011). "The development of peptide-based tools for the analysis of angiogenesis." Chem Biol 18(7): 839-45.

Feldhaus, M. J. and R. W. Siegel (2004). "Yeast display of antibody fragments: a discovery and characterization platform." J Immunol Methods 290(1-2): 69-80.

Feldhaus, M. J., R. W. Siegel, et al. (2003). "Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library." Nat Biotechnol 21(2): 163-70.

Feldwisch, J. and V. Tolmachev (2012). "Engineering of affibody molecules for therapy and diagnostics." Methods Mol Biol 899: 103-26.

Feldwisch, J., V. Tolmachev, et al. (2010). "Design of an optimized scaffold for affibody molecules." J Mol Biol 398(2): 232-47.

Fellouse, F. A., K. Esaki, et al. (2007). "High-throughput generation of synthetic antibodies from highly functional minimalist phage-displayed libraries." J Mol Biol 373(4): 924-40.

Fellouse, F. A., B. Li, et al. (2005). "Molecular recognition by a binary code." J Mol Biol 348(5): 1153-62.

Fellouse, F. A., C. Wiesmann, et al. (2004). "Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition." Proc Natl Acad Sci U S A 101(34): 12467-72.

Fernandez, L. A. (2004). "Prokaryotic expression of antibodies and affibodies." Curr Opin Biotechnol 15(4): 364-73.

Ferrer, M., J. Maiolo, et al. (2005). "Directed evolution of PDZ variants to generate high-affinity detection reagents." Protein Eng Des Sel 18(4): 165-73.

Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature 340(6230): 245-6.

Firth, A. E. and W. M. Patrick (2008). "GLUE-IT and PEDEL-AA: new programmes for analyzing protein diversity in randomized libraries." Nucleic Acids Res 36(Web Server issue): W281-5.

Fishwild, D. M., S. L. O'Donnell, et al. (1996). "High-avidity human IgG kappa monoclonal antibodies from a novel strain of minilocus transgenic mice." Nat Biotechnol 14(7): 845-51.

Fleetwood, F., N. Devoogdt, et al. (2012). "Surface display of a single-domain antibody library on Gram-positive bacteria." Cell Mol Life Sci.

Fleishman, S. J., J. E. Corn, et al. (2011). "Hotspot-centric de novo design of protein binders." J Mol Biol 413(5): 1047-62.

Flisiak, R. and I. Flisiak (2010). "Albinterferon-alpha 2b: a new treatment option for hepatitis C." Expert Opin Biol Ther 10(10): 1509-15.

Floudas, C., S. Fung, et al. (2006). "Advances in protein structure prediction and de novo protein design: A review." Chemical Engineering Science 61: 966–988.

Forrer, P., S. Jung, et al. (1999). "Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins." Curr Opin Struct Biol 9(4): 514-20.

Francisco, J. A., R. Campbell, et al. (1993). "Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface." Proc Natl Acad Sci U S A 90(22): 10444-8.

Frejd, F. (2012). Half-Life Extension by Binding to Albumin through an Albumin Binding Domain. Therapeutic Proteins: Strategies to Modulate Their Plasma Half- Lives. R. Kontermann. Weinheim, Wiley-VCH Verlag GmbH & Co.

Freudl, R., S. MacIntyre, et al. (1986). "Cell surface exposure of the outer membrane protein OmpA of Escherichia coli K-12." J Mol Biol 188(3): 491-4.

Frick, I. M., M. Wikstrom, et al. (1992). "Convergent evolution among immunoglobulin G-binding bacterial proteins." Proc Natl Acad Sci U S A 89(18): 8532-6.

Fujii, R., M. Kitaoka, et al. (2004). "One-step random mutagenesis by error-prone rolling circle amplification." Nucleic Acids Res 32(19): e145.

Fujii, R., M. Kitaoka, et al. (2006). "Error-prone rolling circle amplification: the simplest random mutagenesis protocol." Nat Protoc 1(5): 2493-7.

Gaberc-Porekar, V., I. Zore, et al. (2008). "Obstacles and pitfalls in the PEGylation of therapeutic proteins." Curr Opin Drug Discov Devel 11(2): 242-50.

Gai, S. A. and K. D. Wittrup (2007). "Yeast surface display for protein engineering and characterization." Curr Opin Struct Biol 17(4): 467-73.

Gao, J., S. S. Sidhu, et al. (2009). "Two-state selection of conformation-specific antibodies." Proc Natl Acad Sci U S A 106(9): 3071-6.

Gebauer, M. and A. Skerra (2009). "Engineered protein scaffolds as next-generation antibody therapeutics." Curr Opin Chem Biol 13(3): 245-55.

Gebauer, M. and A. Skerra (2012). "Anticalins small engineered binding proteins based on the lipocalin scaffold." Methods Enzymol 503: 157-88.

Georgiou, G., C. Stathopoulos, et al. (1997). "Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines." Nat Biotechnol 15(1): 29-34.

Gera, N., M. Hussain, et al. (2012). "Protein selection using yeast surface display." Methods.

Getz, J. A., T. D. Schoep, et al. (2012). "Peptide discovery using bacterial display and flow cytometry." Methods Enzymol 503: 75-97.

Geyer, C. R., J. McCafferty, et al. (2012). "Recombinant antibodies and in vitro selection technologies." Methods Mol Biol 901: 11-32.

Gilbreth, R. N., K. Esaki, et al. (2008). "A dominant conformational role for amino acid diversity in minimalist protein-protein interfaces." J Mol Biol 381(2): 407-18.

Gilbreth, R. N. and S. Koide (2012). "Structural insights for engineering binding proteins based on non-antibody scaffolds." Curr Opin Struct Biol 22(4): 413-20.

Gill, D. S. and N. K. Damle (2006). "Biopharmaceutical drug discovery using novel protein scaffolds." Curr Opin Biotechnol 17(6): 653-8.

Glaser, S. M., D. E. Yelton, et al. (1992). "Antibody engineering by codon-based mutagenesis in a filamentous phage vector system." J Immunol 149(12): 3903-13.

Goetsch, L., J. F. Haeuw, et al. (2003). "Identification of B- and T-cell epitopes of BB, a carrier protein derived from the G protein of Streptococcus strain G148." Clin Diagn Lab Immunol 10(1): 125-32.

Gold, L. (2001). "mRNA display: diversity matters during in vitro selection." Proc Natl Acad Sci U S A 98(9): 4825-6.

Golynskiy, M. V. and B. Seelig (2010). "De novo enzymes: from computational design to mRNA display." Trends Biotechnol 28(7): 340-5.

Gostring, L., M. Malm, et al. (2012). "Cellular effects of HER3-specific affibody molecules." PLoS One 7(6): e40023.

Graille, M., E. A. Stura, et al. (2000). "Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity." Proc Natl Acad Sci U S A 97(10): 5399-404.

Gramatikoff, K., O. Georgiev, et al. (1994). "Direct interaction rescue, a novel filamentous phage technique to study protein-protein interactions." Nucleic Acids Res 22(25): 5761-2.

Graziano, J. J., W. Liu, et al. (2008). "Selecting folded proteins from a library of secondary structural elements." J Am Chem Soc 130(1): 176-85.

Greenberg, A. S., A. L. Hughes, et al. (1996). "A novel "chimeric" antibody class in cartilaginous fish: IgM may not be the primordial immunoglobulin." Eur J Immunol 26(5): 1123-9.

Greener, A., M. Callahan, et al. (1996). "An efficient random mutagenesis technique using an E. coli mutator strain." Methods Mol Biol 57: 375-85.

Gregoret, L. M. and R. T. Sauer (1993). "Additivity of mutant effects assessed by binomial mutagenesis." Proc Natl Acad Sci U S A 90(9): 4246-50.

Grodzki, A. C. and E. Berenstein (2010). "Antibody purification: affinity chromatography - protein A and protein G Sepharose." Methods Mol Biol 588: 33-41.

Gronwall, C. and S. Stahl (2009). "Engineered affinity proteins--generation and applications." J Biotechnol 140(3-4): 254-69.

Gulich, S., M. Linhult, et al. (2000). "Stability towards alkaline conditions can be engineered into a protein ligand." J Biotechnol 80(2): 169-78.

Gunneriusson, E., P. Samuelson, et al. (1996). "Surface display of a functional single-chain Fv antibody on staphylococci." J Bacteriol 178(5): 1341-6.

Guo, H. H., J. Choe, et al. (2004). "Protein tolerance to random amino acid change." Proc Natl Acad Sci U S A 101(25): 9205-10.

Guss, B., M. Eliasson, et al. (1986). "Structure of the IgG-binding regions of streptococcal protein G." EMBO J 5(7): 1567-75.

Hackel, B. J., A. Kapila, et al. (2008). "Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling." J Mol Biol 381(5): 1238-52.

Hackel, B. J. and K. D. Wittrup (2010). "The full amino acid repertoire is superior to serine/tyrosine for selection of high affinity immunoglobulin G binders from the fibronectin scaffold." Protein Eng Des Sel 23(4): 211-9.

Hamers-Casterman, C., T. Atarhouch, et al. (1993). "Naturally occurring antibodies devoid of light chains." Nature 363(6428): 446-8.

Han, T., J. Sui, et al. (2011). "Fine epitope mapping of monoclonal antibodies against hemagglutinin of a highly pathogenic H5N1 influenza virus using yeast surface display." Biochem Biophys Res Commun 409(2): 253-9.

Hanes, J., L. Jermutus, et al. (1998). "Ribosome display efficiently selects and evolves high-affinity antibodies in vitro from immune libraries." Proc Natl Acad Sci U S A 95(24): 14130-5.

Hanes, J. and A. Pluckthun (1997). "In vitro selection and evolution of functional proteins by using ribosome display." Proc Natl Acad Sci U S A 94(10): 4937-42.

Hanes, J., C. Schaffitzel, et al. (2000). "Picomolar affinity antibodies from a fully synthetic naive library selected and evolved by ribosome display." Nat Biotechnol 18(12): 1287-92.

Haraldsson, B., J. Nystrom, et al. (2008). "Properties of the glomerular barrier and mechanisms of proteinuria." Physiol Rev 88(2): 451-87.

Harayama, S. (1998). "Artificial evolution by DNA shuffling." Trends Biotechnol 16(2): 76-82.

Harding, F. A., M. M. Stickler, et al. (2010). "The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions." MAbs 2(3): 256-65.

Harmsen, M. M. and H. J. De Haard (2007). "Properties, production, and applications of camelid single-domain antibody fragments." Appl Microbiol Biotechnol 77(1): 13-22.

Harris, L. J., S. B. Larson, et al. (1997). "Refined structure of an intact IgG2a monoclonal antibody." Biochemistry 36(7): 1581-97.

Harvey, B. R., G. Georgiou, et al. (2004). "Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries." Proc Natl Acad Sci U S A 101(25): 9193-8.

Harvey, B. R., A. B. Shanafelt, et al. (2006). "Engineering of recombinant antibody fragments to methamphetamine by anchored periplasmic expression." J Immunol Methods 308(1-2): 43-52.

Hawkins, R. E., S. J. Russell, et al. (1992). "Selection of phage antibodies by binding affinity. Mimicking affinity maturation." J Mol Biol 226(3): 889-96.

He, M. and M. J. Taussig (1997). "Antibody-ribosome-mRNA (ARM) complexes as efficient selection particles for in vitro display and evolution of antibody combining sites." Nucleic Acids Res 25(24): 5132-4.

He, Y., Y. Chen, et al. (2012). "Mutational tipping points for switching protein folds and functions." Structure 20(2): 283-91.

He, Y., Y. Chen, et al. (2007). "An artificially evolved albumin binding module facilitates chemical shift epitope mapping of GA domain interactions with phylogenetically diverse albumins." Protein Sci 16(7): 1490-4.

He, Y., D. A. Rozak, et al. (2006). "Structure, dynamics, and stability variation in bacterial albumin binding modules: implications for species specificity." Biochemistry 45(33): 10102-9.

Hecht, M. H., A. Das, et al. (2004). "De novo proteins from designed combinatorial libraries." Protein Sci 13(7): 1711-23.

Hedhammar, M., T. Gräslund, et al. (2005). "Protein Engineering Strategies for Selective Protein Purification." Chem. Eng. 28(11): 1315-1325.

Heinis, C., T. Rutherford, et al. (2009). "Phage-encoded combinatorial chemical libraries based on bicyclic peptides." Nat Chem Biol 5(7): 502-7.

Heitz, A., D. Le-Nguyen, et al. (1999). "Min-21 and min-23, the smallest peptides that fold like a cystine-stabilized beta-sheet motif: design, solution structure, and thermal stability." Biochemistry 38(32): 10615-25.

Hey, T., E. Fiedler, et al. (2005). "Artificial, non-antibody binding proteins for pharmaceutical and industrial applications." Trends Biotechnol 23(10): 514-22.

Ho, M., S. Nagata, et al. (2006). "Isolation of anti-CD22 Fv with high affinity by Fv display on human cells." Proc Natl Acad Sci U S A 103(25): 9637-42.

Hober, S., K. Nord, et al. (2007). "Protein A chromatography for antibody purification." J Chromatogr B Analyt Technol Biomed Life Sci 848(1): 40-7.

Hoess, R. H. (2001). "Protein design and phage display." Chem Rev 101(10): 3205-18.

Hoffhines, A. J., E. Damoc, et al. (2006). "Detection and purification of tyrosine-sulfated proteins using a novel anti-sulfotyrosine monoclonal antibody." J Biol Chem 281(49): 37877-87.

Hogbom, M., M. Eklund, et al. (2003). "Structural basis for recognition by an in vitro evolved affibody." Proc Natl Acad Sci U S A 100(6): 3191-6.

Holliger, P. and P. J. Hudson (2005). "Engineered antibody fragments and the rise of single domains." Nat Biotechnol 23(9): 1126-36.

Holliger, P., T. Prospero, et al. (1993). ""Diabodies": small bivalent and bispecific antibody fragments." Proc Natl Acad Sci U S A 90(14): 6444-8.

Holliger, P. and G. Winter (1997). "Diabodies: small bispecific antibody fragments." Cancer Immunol Immunother 45(3-4): 128-30.

Holt, L. J., A. Basran, et al. (2008). "Anti-serum albumin domain antibodies for extending the half-lives of short lived drugs." Protein Eng Des Sel 21(5): 283-8.

Holt, L. J., C. Herring, et al. (2003). "Domain antibodies: proteins for therapy." Trends Biotechnol 21(11): 484-90.

Hoogenboom, H. R. (2005). "Selecting and screening recombinant antibody libraries." Nat Biotechnol 23(9): 1105-16.

Hoogenboom, H. R., A. D. Griffiths, et al. (1991). "Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains." Nucleic Acids Res 19(15): 4133-7.

Hopp, J., N. Hornig, et al. (2010). "The effects of affinity and valency of an albumin-binding domain (ABD) on the half-life of a single-chain diabody-ABD fusion protein." Protein Eng Des Sel 23(11): 827-34.

Hoppmann, S., Z. Miao, et al. (2011). "Radiolabeled affibody-albumin bioconjugates for HER2-positive cancer targeting." Bioconjug Chem 22(3): 413-21.

Hosse, R. J., A. Rothe, et al. (2006). "A new generation of protein display scaffolds for molecular recognition." Protein Sci 15(1): 14-27.

Hosseinimehr, S. J., V. Tolmachev, et al. (2012). "Liver uptake of radiolabeled targeting proteins and peptides: considerations for targeting peptide conjugate design." Drug Discov Today 17(21-22): 1224-32.

Hoyer, W., C. Gronwall, et al. (2008). "Stabilization of a beta-hairpin in monomeric Alzheimer's amyloid-beta peptide inhibits amyloid formation." Proc Natl Acad Sci U S A 105(13): 5099-104.

Hudson, P. J. and A. A. Kortt (1999). "High avidity scFv multimers; diabodies and triabodies." J Immunol Methods 231(1-2): 177-89.

Hughes, M. D., D. A. Nagel, et al. (2003). "Removing the redundancy from randomised gene libraries." J Mol Biol 331(5): 973-9.

Hulme, E. C. and M. A. Trevethick (2010). "Ligand binding assays at equilibrium: validation and interpretation." Br J Pharmacol 161(6): 1219-37.

Huston, J. S., D. Levinson, et al. (1988). "Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli." Proc Natl Acad Sci U S A 85(16): 5879-83.

Hutchison, C. A., 3rd, S. Phillips, et al. (1978). "Mutagenesis at a specific position in a DNA sequence." J Biol Chem 253(18): 6551-60.

Hutt, M., A. Farber-Schwarz, et al. (2011). "Plasma half-life extension of small recombinant antibodies by fusion to immunoglobulin-binding domains." J Biol Chem 287(7): 4462-9.

Igawa, T., H. Tsunoda, et al. (2010). "Reduced elimination of IgG antibodies by engineering the variable region." Protein Eng Des Sel 23(5): 385-92.

Imai, K. and A. Takaoka (2006). "Comparing antibody and small-molecule therapies for cancer." Nat Rev Cancer 6(9): 714-27.

Jakobovits, A., R. G. Amado, et al. (2007). "From XenoMouse technology to panitumumab, the first fully human antibody product from transgenic mice." Nat Biotechnol 25(10): 1134-43.

Jansson, B., M. Uhlen, et al. (1998). "All individual domains of staphylococcal protein A show Fab binding." FEMS Immunol Med Microbiol 20(1): 69-78.

Jarver, P., C. Mikaelsson, et al. (2010). "Chemical synthesis and evaluation of a backbone-cyclized minimized 2-helix Z-domain." J Pept Sci 17(6): 463-9.

Jendeberg, L., B. Persson, et al. (1995). "Kinetic analysis of the interaction between protein A domain variants and human Fc using plasmon resonance detection." J Mol Recognit 8(4): 270-8.

Jendeberg, L., M. Tashiro, et al. (1996). "The mechanism of binding staphylococcal protein A to immunoglobin G does not involve helix unwinding." Biochemistry 35(1): 22-31.

Jeong, K. J., M. J. Seo, et al. (2007). "APEx 2-hybrid, a quantitative protein-protein interaction assay for antibody discovery and engineering." Proc Natl Acad Sci U S A 104(20): 8247-52.

Jespers, L., O. Schon, et al. (2004). "Crystal structure of HEL4, a soluble, refoldable human V(H) single domain with a germ-line scaffold." J Mol Biol 337(4): 893-903.

Jin, M., G. Song, et al. (2006). "Directed evolution to probe protein allostery and integrin I domains of 200,000-fold higher affinity." Proc Natl Acad Sci U S A 103(15): 5758-63.

Joern, J. M. (2003). "DNA shuffling." Methods Mol Biol 231: 85-9.

Johansson, M. U., M. de Chateau, et al. (1995). "The GA module, a mobile albumin-binding bacterial domain, adopts a three-helix-bundle structure." FEBS Lett 374(2): 257-61.

Johansson, M. U., M. de Chateau, et al. (1997). "Solution structure of the albumin-binding GA module: a versatile bacterial protein domain." J Mol Biol 266(5): 859-65.

Johansson, M. U., I. M. Frick, et al. (2002a). "Structure, specificity, and mode of interaction for bacterial albumin-binding modules." J Biol Chem 277(10): 8114-20.

Johansson, M. U., H. Nilsson, et al. (2002b). "Differences in backbone dynamics of two homologous bacterial albumin-binding modules: implications for binding specificity and bacterial adaptation." J Mol Biol 316(5): 1083-99.

Jones, P. T., P. H. Dear, et al. (1986). "Replacing the complementarity-determining regions in a human antibody with those from a mouse." Nature 321(6069): 522-5.

Jones, T. D., W. J. Phillips, et al. (2005). "Identification and removal of a promiscuous CD4+ T cell epitope from the C1 domain of factor VIII." J Thromb Haemost 3(5): 991-1000.

Jonsson, A., J. Dogan, et al. (2008). "Engineering of a femtomolar affinity binding protein to human serum albumin." Protein Eng Des Sel 21(8): 515-27.

Jonsson, A., H. Wallberg, et al. (2009). "Generation of tumour-necrosis-factor-alpha-specific affibody molecules capable of blocking receptor binding in vitro." Biotechnol Appl Biochem 54(2): 93-103.

Jostock, T. and S. Dubel (2005). "Screening of molecular repertoires by microbial surface display." Comb Chem High Throughput Screen 8(2): 127-33.

Juraja, S. M., T. D. Mulhern, et al. (2006). "Engineering of the Escherichia coli Im7 immunity protein as a loop display scaffold." Protein Eng Des Sel 19(5): 231-44.

Kaltenbach, M. and F. Hollfelder (2012). "SNAP display: in vitro protein evolution in microdroplets." Methods Mol Biol 805: 101-11.

Kamtekar, S., J. M. Schiffer, et al. (1993). "Protein design by binary patterning of polar and nonpolar amino acids." Science 262(5140): 1680-5.

Kaneko, T., S. S. Sidhu, et al. (2011). "Evolving specificity from variability for protein interaction domains." Trends Biochem Sci 36(4): 183-90.

Kazlauskas, R. J. and U. T. Bornscheuer (2009). "Finding better protein engineering strategies." Nat Chem Biol 5(8): 526-9.

Keefe, A. D. and J. W. Szostak (2001). "Functional proteins from a random-sequence library." Nature 410(6829): 715-8.

Kehoe, J. W., N. Velappan, et al. (2006). "Using phage display to select antibodies recognizing post-translational modifications independently of sequence context." Mol Cell Proteomics 5(12): 2350-63.

Kim, H. J., A. Eichinger, et al. (2009). "High-affinity recognition of lanthanide(III) chelate complexes by a reprogrammed human lipocalin 2." J Am Chem Soc 131(10): 3565-76.

Kim, J., W. L. Hayton, et al. (2007). "Kinetics of FcRn-mediated recycling of IgG and albumin in human: pathophysiology and therapeutic implications using a simplified mechanism-based model." Clin Immunol 122(2): 146-55.

Koch, H., N. Grafe, et al. (2006). "Direct selection of antibodies from complex libraries with the protein fragment complementation assay." J Mol Biol 357(2): 427-41.

Koder, R. L., J. L. Anderson, et al. (2009). "Design and engineering of an O(2) transport protein." Nature 458(7236): 305-9.

Kohler, G. and C. Milstein (1975). "Continuous cultures of fused cells secreting antibody of predefined specificity." Nature 256(5517): 495-7.

Koide, A., C. W. Bailey, et al. (1998). "The fibronectin type III domain as a scaffold for novel binding proteins." J Mol Biol 284(4): 1141-51.

Koide, A., R. N. Gilbreth, et al. (2007). "High-affinity single-domain binding proteins with a binary-code interface." Proc Natl Acad Sci U S A 104(16): 6632-7.

Koide, A., J. Wojcik, et al. (2012a). "Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold." J Mol Biol 415(2): 393-405.

Koide, S. (2009a). "Generation of new protein functions by nonhomologous combinations and rearrangements of domains and modules." Curr Opin Biotechnol 20(4): 398-404.

Koide, S., A. Koide, et al. (2012b). "Target-binding proteins based on the 10th human fibronectin type III domain ((1)(0)Fn3)." Methods Enzymol 503: 135-56.

Koide, S. and S. Sidhu (2009b). "The Importance of Being Tyrosine: Lessons in Molecular Recognition from Minimalist Synthetic Binding Proteins." ACS CHEMICAL BIOLOGY 4(5): 325-334.

Kolmar, H. and A. Skerra (2008). "Alternative binding proteins get mature: rivalling antibodies." FEBS J 275(11): 2667.

Kontermann, R. (2011a). "Bispecific Antibodies."

Kontermann, R. E. (2009). "Strategies to extend plasma half-lives of recombinant antibodies." BioDrugs 23(2): 93-109.

Kontermann, R. E. (2011b). "Strategies for extended serum half-life of protein therapeutics." Curr Opin Biotechnol 22(6): 868-76.
Korndorfer, I. P., G. Beste, et al. (2003). "Crystallographic analysis of an "anticalin" with tailored specificity for fluorescein reveals high structural plasticity of the lipocalin loop region." Proteins 53(1): 121-9.
Kouadio, J. L., J. R. Horn, et al. (2005). "Shotgun alanine scanning shows that growth hormone can bind productively to its receptor through a drastically minimized interface." J Biol Chem 280(27): 25524-32.
Krag, D. N., G. S. Shukla, et al. (2006). "Selection of tumor-binding ligands in cancer patients with phage display libraries." Cancer Res 66(15): 7724-33.
Kraulis, P. J., P. Jonasson, et al. (1996). "The serum albumin-binding domain of streptococcal protein G is a three-helical bundle: a heteronuclear NMR study." FEBS Lett 378(2): 190-4.
Kronqvist, N., J. Lofblom, et al. (2008a). "A novel affinity protein selection system based on staphylococcal cell surface display and flow cytometry." Protein Eng Des Sel 21(4): 247-55.
Kronqvist, N., J. Lofblom, et al. (2008b). "Simplified characterization through site-specific protease-mediated release of affinity proteins selected by staphylococcal display." FEMS Microbiol Lett 278(1): 128-36.
Kronqvist, N., M. Malm, et al. (2011). "Combining phage and staphylococcal surface display for generation of ErbB3-specific Affibody molecules." Protein Eng Des Sel 24(4): 385-96.
Kronqvist, N., M. Malm, et al. (2010). "Staphylococcal surface display in combinatorial protein engineering and epitope mapping of antibodies." Recent Pat Biotechnol 4(3): 171-82.
Ku, J. and P. G. Schultz (1995). "Alternate protein frameworks for molecular recognition." Proc Natl Acad Sci U S A 92(14): 6552-6.
Kurz, M., K. Gu, et al. (2001). "cDNA - protein fusions: covalent protein - gene conjugates for the in vitro selection of peptides and proteins." Chembiochem 2(9): 666-72.
Ladner, R. C. (1995). "Constrained peptides as binding entities." Trends Biotechnol 13(10): 426-30.
Lahr, S. J., A. Broadwater, et al. (1999). "Patterned library analysis: a method for the quantitative assessment of hypotheses concerning the determinants of protein structure." Proc Natl Acad Sci U S A 96(26): 14860-5.
Lancet, D., E. Sadovsky, et al. (1993). "Probability model for molecular recognition in biological receptor repertoires: significance to the olfactory system." Proc Natl Acad Sci U S A 90(8): 3715-9.
Lappe, M., G. Bagler, et al. (2009). "Designing evolvable libraries using multi-body potentials." Curr Opin Biotechnol 20(4): 437-46.
Lee, C. V., S. S. Sidhu, et al. (2004). "Bivalent antibody phage display mimics natural immunoglobulin." J Immunol Methods 284(1-2): 119-32.
Lee, S. B., M. Hassan, et al. (2008). "Affibody molecules for in vivo characterization of HER2-positive tumors by near-infrared imaging." Clin Cancer Res 14(12): 3840-9.
Lejon, S., J. F. Cramer, et al. (2008). "Structural basis for the binding of naproxen to human serum albumin in the presence of fatty acids and the GA module." Acta Crystallogr Sect F Struct Biol Cryst Commun 64(Pt 2): 64-9.
Lejon, S., I. M. Frick, et al. (2004). "Crystal structure and biological implications of a bacterial albumin binding module in complex with human serum albumin." J Biol Chem 279(41): 42924-8.
Lendel, C., V. Dincbas-Renqvist, et al. (2004). "Biophysical characterization of Z(SPA-1)--a phage-display selected binder to protein A." Protein Sci 13(8): 2078-88.
Lendel, C., J. Dogan, et al. (2006). "Structural basis for molecular recognition in an affibody:affibody complex." J Mol Biol 359(5): 1293-304.
Leung, D., E. Chen, et al. (1989). "A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction." Technique 1: 11-15.
Levin, A. M. and G. A. Weiss (2006). "Optimizing the affinity and specificity of proteins with molecular display." Mol Biosyst 2(1): 49-57.
Levin, K. B., O. Dym, et al. (2009). "Following evolutionary paths to protein-protein interactions with high affinity and selectivity." Nat Struct Mol Biol 16(10): 1049-55.
Li, M. (2000). "Applications of display technology in protein analysis." Nat Biotechnol 18(12): 1251-6.
Li, S., S. Millward, et al. (2002). "In vitro selection of mRNA display libraries containing an unnatural amino acid." J Am Chem Soc 124(34): 9972-3.
Libon, C., N. Corvaia, et al. (1999). "The serum albumin-binding region of streptococcal protein G (BB) potentiates the immunogenicity of the G130-230 RSV-A protein." Vaccine 17(5): 406-14.
Lindgren, J., C. Ekblad, et al. (2012). "A native chemical ligation approach for combinatorial assembly of affibody molecules." Chembiochem 13(7): 1024-31.
Lindman, S., A. Hernandez-Garcia, et al. (2010). "In vivo protein stabilization based on fragment complementation and a split GFP system." Proc Natl Acad Sci U S A 107(46): 19826-31.
Linhult, M., H. K. Binz, et al. (2002). "Mutational analysis of the interaction between albumin-binding domain from streptococcal protein G and human serum albumin." Protein Sci 11(2): 206-13.
Linhult, M., S. Gulich, et al. (2004). "Improving the tolerance of a protein a analogue to repeated alkaline exposures using a bypass mutagenesis approach." Proteins 55(2): 407-16.
Lipovsek, D. (2011). "Adnectins: engineered target-binding protein therapeutics." Protein Eng Des Sel 24(1-2): 3-9.

Lipovsek, D., S. M. Lippow, et al. (2007). "Evolution of an interloop disulfide bond in high-affinity antibody mimics based on fibronectin type III domain and selected by yeast surface display: molecular convergence with single-domain camelid and shark antibodies." J Mol Biol 368(4): 1024-41.

Lipovsek, D. and A. Pluckthun (2004). "In-vitro protein evolution by ribosome display and mRNA display." J Immunol Methods 290(1-2): 51-67.

Liu, C. C., A. V. Mack, et al. (2008). "Protein evolution with an expanded genetic code." Proc Natl Acad Sci U S A 105(46): 17688-93.

Liu, C. C. and P. G. Schultz (2010). "Adding new chemistries to the genetic code." Annu Rev Biochem 79: 413-44.

Ljungberg, U. K., B. Jansson, et al. (1993). "The interaction between different domains of staphylococcal protein A and human polyclonal IgG, IgA, IgM and F(ab')2: separation of affinity from specificity." Mol Immunol 30(14): 1279-85.

Lobo, E. D., R. J. Hansen, et al. (2004). "Antibody pharmacokinetics and pharmacodynamics." J Pharm Sci 93(11): 2645-68.

Lofblom, J. (2011a). "Bacterial display in combinatorial protein engineering." Biotechnol J 6(9): 1115-29.

Lofblom, J., J. Feldwisch, et al. (2010). "Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications." FEBS Lett 584(12): 2670-80.

Lofblom, J., F. Y. Frejd, et al. (2011b). "Non-immunoglobulin based protein scaffolds." Curr Opin Biotechnol 22(6): 843-8.

Lofblom, J., J. Sandberg, et al. (2007). "Evaluation of staphylococcal cell surface display and flow cytometry for postselectional characterization of affinity proteins in combinatorial protein engineering applications." Appl Environ Microbiol 73(21): 6714-21.

Lofblom, J., H. Wernerus, et al. (2005). "Fine affinity discrimination by normalized fluorescence activated cell sorting in staphylococcal surface display." FEMS Microbiol Lett 248(2): 189-98.

Lofdahl, P. A., O. Nord, et al. (2009). "Selection of TNF-alpha binding affibody molecules using a beta-lactamase protein fragment complementation assay." N Biotechnol 26(5): 251-9.

Lofdahl, P. A. and P. A. Nygren (2010). "Affinity maturation of a TNFalpha-binding affibody molecule by Darwinian survival selection." Biotechnol Appl Biochem 55(3): 111-20.

Lonberg, N. (2005). "Human antibodies from transgenic animals." Nat Biotechnol 23(9): 1117-25.

Low, N. M., P. H. Holliger, et al. (1996). "Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain." J Mol Biol 260(3): 359-68.

Lowman, H. B., S. H. Bass, et al. (1991). "Selecting high-affinity binding proteins by monovalent phage display." Biochemistry 30(45): 10832-8.

Magliery, T. J. and L. Regan (2004). "Library approaches to biophysical problems." Eur J Biochem 271(9): 1593-4.

Main, A. L., T. S. Harvey, et al. (1992). "The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions." Cell 71(4): 671-8.

Makrides, S. C., P. A. Nygren, et al. (1996). "Extended in vivo half-life of human soluble complement receptor type 1 fused to a serum albumin-binding receptor." J Pharmacol Exp Ther 277(1): 534-42.

Mandal, K., M. Uppalapati, et al. (2012). "Chemical synthesis and X-ray structure of a heterochiral {D-protein antagonist plus vascular endothelial growth factor} protein complex by racemic crystallography." Proc Natl Acad Sci U S A 109(37): 14779-84.

Mandecki, W. (1998). "The game of chess and searches in protein sequence space." TIBTECH 16: 200-202.

Mannocci, L., M. Leimbacher, et al. (2011). "20 years of DNA-encoded chemical libraries." Chem Commun (Camb) 47(48): 12747-53.

Mardis, E. R. (2008). "Next-generation DNA sequencing methods." Annu Rev Genomics Hum Genet 9: 387-402.

Marshall, G. (1992). "Three-dimensional structure of peptide-protein complexes: implications for recognition." Current Opinion in Structural Biology 2: 904-919.

Marvin, J. S. and H. B. Lowman (2003). "Redesigning an antibody fragment for faster association with its antigen." Biochemistry 42(23): 7077-83.

Matouschek, A., J. T. Kellis, Jr., et al. (1990). "Transient folding intermediates characterized by protein engineering." Nature 346(6283): 440-5.

Matsumura, M., W. J. Becktel, et al. (1989). "Stabilization of phage T4 lysozyme by engineered disulfide bonds." Proc Natl Acad Sci U S A 86(17): 6562-6.

Matsuura, T., A. Ernst, et al. (2002). "Construction and characterization of protein libraries composed of secondary structure modules." Protein Sci 11(11): 2631-43.

Mattheakis, L. C., R. R. Bhatt, et al. (1994). "An in vitro polysome display system for identifying ligands from very large peptide libraries." Proc Natl Acad Sci U S A 91(19): 9022-6.

Mazor, Y., R. Noy, et al. (2007a). "chFRP5-ZZ-PE38, a large IgG-toxin immunoconjugate outperforms the corresponding smaller FRP5(Fv)-ETA immunotoxin in eradicating ErbB2-expressing tumor xenografts." Cancer Lett 257(1): 124-35.

Mazor, Y., T. Van Blarcom, et al. (2010). "Selection of full-length IgGs by tandem display on filamentous phage particles and Escherichia coli fluorescence-activated cell sorting screening." FEBS J 277(10): 2291-303.

Mazor, Y., T. Van Blarcom, et al. (2008). "E-clonal antibodies: selection of full-length IgG antibodies using bacterial periplasmic display." Nat Protoc 3(11): 1766-77.

Mazor, Y., T. Van Blarcom, et al. (2007b). "Isolation of engineered, full-length antibodies from libraries expressed in Escherichia coli." Nat Biotechnol 25(5): 563-5.

McCafferty, J., A. D. Griffiths, et al. (1990). "Phage antibodies: filamentous phage displaying antibody variable domains." Nature 348(6301): 552-4.

McCullum, E. O., B. A. Williams, et al. (2010). "Random mutagenesis by error-prone PCR." Methods Mol Biol 634: 103-9.

Metzner, H. J., T. Weimer, et al. (2009). "Genetic fusion to albumin improves the pharmacokinetic properties of factor IX." Thromb Haemost 102(4): 634-44.

Miao, Z., J. Levi, et al. (2010). "Protein scaffold-based molecular probes for cancer molecular imaging." Amino Acids 41(5): 1037-47.

Michaud, G. A., M. Salcius, et al. (2003). "Analyzing antibody specificity with whole proteome microarrays." Nat Biotechnol 21(12): 1509-12.

Miersch, S. and S. S. Sidhu (2012). "Synthetic antibodies: Concepts, potential and practical considerations." Methods 57(4): 486-98.

Miller, O. J., K. Bernath, et al. (2006). "Directed evolution by in vitro compartmentalization." Nat Methods 3(7): 561-70.

Moore, S., C. Leung, et al. (2012). "Knottins: disulfide-bonded therapeutic and diagnostic peptides." Drug Discovery Today: Technologies 9(1): e3-e11.

Morrison, K. L. and G. A. Weiss (2001). "Combinatorial alanine-scanning." Curr Opin Chem Biol 5(3): 302-7.

Mossner, E., H. Koch, et al. (2001). "Fast selection of antibodies without antigen purification: adaptation of the protein fragment complementation assay to select antigen-antibody pairs." J Mol Biol 308(2): 115-22.

Mukovozov, I., T. Sabljic, et al. (2008). "Factors that contribute to the immmunogenicity of therapeutic recombinant human proteins." Thromb Haemost 99(5): 874-82.

Müller, M., K. Saunders, et al. (2012). "Improving the pharmacokinetic properties of biologics by fusion to an anti-HSA shark VNAR domain." mAbs 4(6): 673-685.

Mullis, K., F. Faloona, et al. (1986). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction." Cold Spring Harb Symp Quant Biol 51 Pt 1: 263-73.

Munoz, E. and M. W. Deem (2008). "Amino acid alphabet size in protein evolution experiments: better to search a small library thoroughly or a large library sparsely?" Protein Eng Des Sel 21(5): 311-7.

Munson, M., S. Balasubramanian, et al. (1996). "What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties." Protein Sci 5(8): 1584-93.

Murakami, H., T. Hohsaka, et al. (2002). "Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs." Nat Biotechnol 20(1): 76-81.

Murakami, H., T. Hohsaka, et al. (2012). Random Insertion and Deletion Mutagenesis. Methods in Molecular Biology, vol. 231: Directed Evolution Library Creation: Methods and Protocols. F. Arnold and G. Gregoriou, Humana Press: 53-64.

Myhre, E. B. and G. Kronvall (1980). "Demonstration of a new type of immunoglobulin G receptor in Streptococcus zooepidemicus strains." Infect Immun 27(3): 808-16.

Nakashima, N. and T. Tamura (2009). "Conditional gene silencing of multiple genes with antisense RNAs and generation of a mutator strain of Escherichia coli." Nucleic Acids Res 37(15): e103.

Nelson, A. L., E. Dhimolea, et al. (2010). "Development trends for human monoclonal antibody therapeutics." Nat Rev Drug Discov 9(10): 767-74.

Nemoto, N., E. Miyamoto-Sato, et al. (1997). "In vitro virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro." FEBS Lett 414(2): 405-8.

Neylon, C. (2004). "Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution." Nucleic Acids Res 32(4): 1448-59.

Nguyen, A. and P. Daugherty (2012). Production of Randomly Mutated Plasmid Libraries Using Mutator Strains. Methods in Molecular Biology, Directed Evolution Library Creation: Methods and Protocols. F. Arnold and G. Georgiou, Humana Press.

Nguyen, A., A. E. Reyes, 2nd, et al. (2006). "The pharmacokinetics of an albumin-binding Fab (AB.Fab) can be modulated as a function of affinity for albumin." Protein Eng Des Sel 19(7): 291-7.

Nilsson, B., T. Moks, et al. (1987). "A synthetic IgG-binding domain based on staphylococcal protein A." Protein Eng 1(2): 107-13.

Nizak, C., S. Monier, et al. (2003). "Recombinant antibodies to the small GTPase Rab6 as conformation sensors." Science 300(5621): 984-7.

Nord, K., E. Gunneriusson, et al. (1997). "Binding proteins selected from combinatorial libraries of an alpha-helical bacterial receptor domain." Nat Biotechnol 15(8): 772-7.

Nord, K., J. Nilsson, et al. (1995). "A combinatorial library of an alpha-helical bacterial receptor domain." Protein Eng 8(6): 601-8.

Nord, O., M. Uhlen, et al. (2003). "Microbead display of proteins by cell-free expression of anchored DNA." J Biotechnol 106(1): 1-13.

Northrup, S. H. and H. P. Erickson (1992). "Kinetics of protein-protein association explained by Brownian dynamics computer simulation." Proc Natl Acad Sci U S A 89(8): 3338-42.

Nov, Y. (2011). "When Second Best is Good Enough: Another Probabilistic Look at Saturation Mutagenesis." Appl. Environ. Microbiol.

Nuttall, S. D., U. V. Krishnan, et al. (2001). "Isolation of the new antigen receptor from wobbegong sharks, and use as a scaffold for the display of protein loop libraries." Mol Immunol 38(4): 313-26.

Nuttall, S. D. and R. B. Walsh (2008). "Display scaffolds: protein engineering for novel therapeutics." Curr Opin Pharmacol 8(5): 609-15.

Nygren, P., M. Uhlén, et al. (1991). In vivo stabilization of a human recombinant CD4 derivative by fusion to a serum-albumin-binding receptor. Vaccines, Cold Spring Harbor Press. 91: 363-368.

Nygren, P. A. (2008). "Alternative binding proteins: affibody binding proteins developed from a small three-helix bundle scaffold." FEBS J 275(11): 2668-76.

Nygren, P. A., M. Eliasson, et al. (1988). "Analysis and use of the serum albumin binding domains of streptococcal protein G." J Mol Recognit 1(2): 69-74.

Nygren, P. A., C. Ljungquist, et al. (1990). "Species-dependent binding of serum albumins to the streptococcal receptor protein G." Eur J Biochem 193(1): 143-8.

Nygren, P. A. and A. Skerra (2004). "Binding proteins from alternative scaffolds." J Immunol Methods 290(1-2): 3-28.

Nygren, P. A. and M. Uhlen (1997). "Scaffolds for engineering novel binding sites in proteins." Curr Opin Struct Biol 7(4): 463-9.

O'Connell, D., B. Becerril, et al. (2002). "Phage versus phagemid libraries for generation of human monoclonal antibodies." J Mol Biol 321(1): 49-56.

O'Fagain, C. (2011). "Engineering protein stability." Methods Mol Biol 681: 103-36.

Obeidat, M. and B. J. Ballermann (2012). "Glomerular endothelium: a porous sieve and formidable barrier." Exp Cell Res 318(9): 964-72.

Odegrip, R., D. Coomber, et al. (2004). "CIS display: In vitro selection of peptides from libraries of protein-DNA complexes." Proc Natl Acad Sci U S A 101(9): 2806-10.

Olivera, B. M., J. Rivier, et al. (1990). "Diversity of Conus neuropeptides." Science 249(4966): 257-63.

Orlandi, R., D. H. Gussow, et al. (1989). "Cloning immunoglobulin variable domains for expression by the polymerase chain reaction." Proc Natl Acad Sci U S A 86(10): 3833-7.

Orlova, A., M. Magnusson, et al. (2006). "Tumor imaging using a picomolar affinity HER2 binding affibody molecule." Cancer Res 66(8): 4339-48.

Osbourn, J., M. Groves, et al. (2005). "From rodent reagents to human therapeutics using antibody guided selection." Methods 36(1): 61-8.

Paegel, B. M. (2010). "Microfluidic landscapes for evolution." Curr Opin Chem Biol 14(5): 568-73.

Pal, G., J. L. Kouadio, et al. (2006). "Comprehensive and quantitative mapping of energy landscapes for protein-protein interactions by rapid combinatorial scanning." J Biol Chem 281(31): 22378-85.

Pancer, Z., C. T. Amemiya, et al. (2004). "Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey." Nature 430(6996): 174-80.

Pantazes, R. J., M. J. Grisewood, et al. (2011). "Recent advances in computational protein design." Curr Opin Struct Biol 21(4): 467-72.

Parrish, J. R., K. D. Gulyas, et al. (2006). "Yeast two-hybrid contributions to interactome mapping." Curr Opin Biotechnol 17(4): 387-93.

Pelletier, J. and S. Sidhu (2001). "Mapping protein-protein interactions with combinatorial biology methods." Curr Opin Biotechnol 12(4): 340-7.

Pelletier, J. N., K. M. Arndt, et al. (1999). "An in vivo library-versus-library selection of optimized protein-protein interactions." Nat Biotechnol 17(7): 683-90.

Pelletier, J. N., F. X. Campbell-Valois, et al. (1998). "Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments." Proc Natl Acad Sci U S A 95(21): 12141-6.

Pepper, L. R., Y. K. Cho, et al. (2008). "A decade of yeast surface display technology: where are we now?" Comb Chem High Throughput Screen 11(2): 127-34.

Perelson, A. S. and G. F. Oster (1979). "Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination." J Theor Biol 81(4): 645-70.

Pluckthun, A. (2012). "Ribosome display: a perspective." Methods Mol Biol 805: 3-28.

Plückthun, A. (2009). Alternative scaffolds: expanding the options of antibodies. Recombinant antibodies for immunotherapy. M. Little. New York, Cambridge University Press: 244-271.

Prassler, J., S. Thiel, et al. (2011). "HuCAL PLATINUM, a synthetic Fab library optimized for sequence diversity and superior performance in mammalian expression systems." J Mol Biol 413(1): 261-78.

Predki, P. F., V. Agrawal, et al. (1996). "Amino-acid substitutions in a surface turn modulate protein stability." Nat Struct Biol 3(1): 54-8.

Qi, H., H. Lu, et al. (2012). "Phagemid vectors for phage display: properties, characteristics and construction." J Mol Biol 417(3): 129-43.

Qian, Z. and S. Lutz (2005). "Improving the catalytic activity of Candida antarctica lipase B by circular permutation." J Am Chem Soc 127(39): 13466-7.

Qiu, J. K., S. T. Jung, et al. (2010). "Enrichment of Escherichia coli spheroplasts displaying scFv antibodies specific for antigens expressed on the human cell surface." Appl Microbiol Biotechnol 88(6): 1385-91.

Qiu, X. Q., H. Wang, et al. (2007). "Small antibody mimetics comprising two complementarity-determining regions and a framework region for tumor targeting." Nat Biotechnol 25(8): 921-9.

Quin, M. B. and C. Schmidt-Dannert (2011). "Engineering of biocatalysts - from evolution to creation." ACS Catal 1(9): 1017-1021.

Rader, C. and C. F. Barbas, 3rd (1997). "Phage display of combinatorial antibody libraries." Curr Opin Biotechnol 8(4): 503-8.

Raghavan, M., M. Y. Chen, et al. (1994). "Investigation of the interaction between the class I MHC-related Fc receptor and its immunoglobulin G ligand." Immunity 1(4): 303-15.

Ramesh, B., V. G. Sendra, et al. (2012). "Single-cell characterization of autotransporter mediated Escherichia coli surface display of disulfide-bond containing proteins." J Biol Chem.

Reetz, M. T., D. Kahakeaw, et al. (2008). "Addressing the numbers problem in directed evolution." Chembiochem 9(11): 1797-804.

Regan, L. and W. F. DeGrado (1988). "Characterization of a helical protein designed from first principles." Science 241(4868): 976-8.

Reichert, J. M. and E. Dhimolea (2012). "The future of antibodies as cancer drugs." Drug Discov Today 17(17-18): 954-63.

Reiersen, H., I. Lobersli, et al. (2005). "Covalent antibody display--an in vitro antibody-DNA library selection system." Nucleic Acids Res 33(1): e10.

Reiersen, H. and A. R. Rees (1999). "An engineered minidomain containing an elastin turn exhibits a reversible temperature-induced IgG binding." Biochemistry 38(45): 14897-905.

Reiersen, H. and A. R. Rees (2000). "Sodium sulphate reactivates a protein A minidomain with a short elastin beta-turn." Biochem Biophys Res Commun 276(3): 899-904.

Reis, K. J., E. M. Ayoub, et al. (1984). "Streptococcal Fc receptors. I. Isolation and partial characterization of the receptor from a group C streptococcus." J Immunol 132(6): 3091-7.

Ren, G., J. M. Webster, et al. (2012). "In vivo targeting of HER2-positive tumor using 2-helix affibody molecules." Amino Acids 43(1): 405-13.

Ren, G., R. Zhang, et al. (2009). "A 2-helix small protein labeled with 68Ga for PET imaging of HER2 expression." J Nucl Med 50(9): 1492-9.

Roberts, R. W. and J. W. Szostak (1997). "RNA-peptide fusions for the in vitro selection of peptides and proteins." Proc Natl Acad Sci U S A 94(23): 12297-302.

Rockberg, J., J. Lofblom, et al. (2008). "Epitope mapping of antibodies using bacterial surface display." Nat Methods 5(12): 1039-45.

Romero, P. A. and F. H. Arnold (2009). "Exploring protein fitness landscapes by directed evolution." Nat Rev Mol Cell Biol 10(12): 866-76.

Roopenian, D. C. and S. Akilesh (2007). "FcRn: the neonatal Fc receptor comes of age." Nat Rev Immunol 7(9): 715-25.

Rosik, D., A. Orlova, et al. (2011). "Direct comparison of 111In-labelled two-helix and three-helix Affibody molecules for in vivo molecular imaging." Eur J Nucl Med Mol Imaging 39(4): 693-702.

Rozak, D. A., P. A. Alexander, et al. (2006). "Using offset recombinant polymerase chain reaction to identify functional determinants in a common family of bacterial albumin binding domains." Biochemistry 45(10): 3263-71.

Rozak, D. A., J. Orban, et al. (2005). "G148-GA3: a streptococcal virulence module with atypical thermodynamics of folding optimally binds human serum albumin at physiological temperatures." Biochim Biophys Acta 1753(2): 226-33.

Saerens, D., G. H. Ghassabeh, et al. (2008). "Single-domain antibodies as building blocks for novel therapeutics." Curr Opin Pharmacol 8(5): 600-8.

Saiki, R. K., S. Scharf, et al. (1985). "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science 230(4732): 1350-4.

Samuelson, P., E. Gunneriusson, et al. (2002). "Display of proteins on bacteria." J Biotechnol 96(2): 129-54.

Samuelson, P., M. Hansson, et al. (1995). "Cell surface display of recombinant proteins on Staphylococcus carnosus." J Bacteriol 177(6): 1470-6.

Sandberg, D., A. Wennborg, et al. (2011). "Abstract nr P2-09-01: First-in-Human Whole-Body HER2–Receptor Mapping Using Affibody Molecular Imaging." Cancer Res 71(24).

Sanger, F., S. Nicklen, et al. (1977). "DNA sequencing with chain-terminating inhibitors." Proc Natl Acad Sci U S A 74(12): 5463-7.

Sauer-Eriksson, A. E., G. J. Kleywegt, et al. (1995). "Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG." Structure 3(3): 265-78.

Schatz, P. J. (1993). "Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli." Biotechnology (N Y) 11(10): 1138-43.

Schellenberger, V., C. W. Wang, et al. (2009). "A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner." Nat Biotechnol 27(12): 1186-90.

Schimmele, B. and A. Pluckthun (2005). "Identification of a functional epitope of the Nogo receptor by a combinatorial approach using ribosome display." J Mol Biol 352(1): 229-41.

Schlapschy, M., I. Theobald, et al. (2007). "Fusion of a recombinant antibody fragment with a homo-amino-acid polymer: effects on biophysical properties and prolonged plasma half-life." Protein Eng Des Sel 20(6): 273-84.

Schlehuber, S. and A. Skerra (2001). "Duocalins: engineered ligand-binding proteins with dual specificity derived from the lipocalin fold." Biol Chem 382(9): 1335-42.

Schonfeld, D., G. Matschiner, et al. (2009). "An engineered lipocalin specific for CTLA-4 reveals a combining site with structural and conformational features similar to antibodies." Proc Natl Acad Sci U S A 106(20): 8198-203.

Schreiber, G. (2002). "Kinetic studies of protein-protein interactions." Curr Opin Struct Biol 12(1): 41-7.

Schreiber, G. and A. R. Fersht (1993). "Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering." Biochemistry 32(19): 5145-50.

Scott, J. K. and G. P. Smith (1990). "Searching for peptide ligands with an epitope library." Science 249(4967): 386-90.

Secco, P., E. D'Agostini, et al. (2009). "Antibody library selection by the {beta}-lactamase protein fragment complementation assay." Protein Eng Des Sel 22(3): 149-58.

Seeburg, P. H., W. W. Colby, et al. (1984). "Biological properties of human c-Ha-ras1 genes mutated at codon 12." Nature 312(5989): 71-5.

Seelig, B. (2011). "mRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries." Nat Protoc 6(4): 540-52.

Seelig, B. and J. W. Szostak (2007). "Selection and evolution of enzymes from a partially randomized non-catalytic scaffold." Nature 448(7155): 828-31.

Selzer, T., S. Albeck, et al. (2000). "Rational design of faster associating and tighter binding protein complexes." Nat Struct Biol 7(7): 537-41.

Sepp, A., D. S. Tawfik, et al. (2002). "Microbead display by in vitro compartmentalisation: selection for binding using flow cytometry." FEBS Lett 532(3): 455-8.

Sergeeva, A., M. G. Kolonin, et al. (2006). "Display technologies: application for the discovery of drug and gene delivery agents." Adv Drug Deliv Rev 58(15): 1622-54.

Shekhawat, S. S. and I. Ghosh (2011). "Split-protein systems: beyond binary protein-protein interactions." Curr Opin Chem Biol 15(6): 789-97.

Sheridan, C. (2007). "Pharma consolidates its grip on post-antibody landscape." Nat Biotechnol 25(4): 365-6.

Shivange, A. V., J. Marienhagen, et al. (2009). "Advances in generating functional diversity for directed protein evolution." Curr Opin Chem Biol 13(1): 19-25.

Shusta, E. V., M. C. Kieke, et al. (1999). "Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency." J Mol Biol 292(5): 949-56.

Sidhu, S. S. (2007). "Full-length antibodies on display." Nat Biotechnol 25(5): 537-8.

Sidhu, S. S. (2012). "Antibodies for all: The case for genome-wide affinity reagents." FEBS Lett 586(17): 2778-9.

Simister, N. E. and K. E. Mostov (1989). "An Fc receptor structurally related to MHC class I antigens." Nature 337(6203): 184-7.

Sjobring, U., C. Falkenberg, et al. (1988). "Isolation and characterization of a 14-kDa albumin-binding fragment of streptococcal protein G." J Immunol 140(5): 1595-9.

Sjolander, A., P. A. Nygren, et al. (1997). "The serum albumin-binding region of streptococcal protein G: a bacterial fusion partner with carrier-related properties." J Immunol Methods 201(1): 115-23.

Sjölander, A., S. Ståhl, et al. (1993). "Bacterial Expression Systems Based on a Protein A and Protein G Designed for the Production of Immunogens: Applications to Plasmodium falciparum Malaria Antigens." ImmunoMethods 2(1): 79-92.

Skerra, A. (2007). "Alternative non-antibody scaffolds for molecular recognition." Curr Opin Biotechnol 18(4): 295-304.

Skerra, A. (2008). "Alternative binding proteins: anticalins - harnessing the structural plasticity of the lipocalin ligand pocket to engineer novel binding activities." FEBS J 275(11): 2677-83.

Skerra, A. and A. Pluckthun (1988). "Assembly of a functional immunoglobulin Fv fragment in Escherichia coli." Science 240(4855): 1038-41.

Smith, G. P. (1985). "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface." Science 228(4705): 1315-7.

Smith, G. P. and V. A. Petrenko (1997). "Phage Display." Chem Rev 97(2): 391-410.

Smith, J. M. (1970). "Natural selection and the concept of a protein space." Nature 225(5232): 563-4.

Smothers, J. F., S. Henikoff, et al. (2002). "Tech.Sight. Phage display. Affinity selection from biological libraries." Science 298(5593): 621-2.

Sockolosky, J. T., M. R. Tiffany, et al. (2012). "Engineering neonatal Fc receptor-mediated recycling and transcytosis in recombinant proteins by short terminal peptide extensions." Proc Natl Acad Sci U S A 109(40): 16095-100.

Sollod, B., D. Wilson, et al. (2005). "Were arachnids the first to use combinatorial peptide libraries." Peptides 26: 131-139.

Souriau, C., L. Chiche, et al. (2005). "New binding specificities derived from Min-23, a small cystine-stabilized peptidic scaffold." Biochemistry 44(19): 7143-55.

Spada, S. and A. Pluckthun (1997). "Selectively infective phage (SIP) technology: a novel method for in vivo selection of interacting protein-ligand pairs." Nat Med 3(6): 694-6.

Ståhl, S., N. Kronqvist, et al. (2012). "Affinity proteins and their generation." J Chem Technol Biotechnol.

Starovasnik, M. A., A. C. Braisted, et al. (1997). "Structural mimicry of a native protein by a minimized binding domain." Proc Natl Acad Sci U S A 94(19): 10080-5.

Steiner, D., P. Forrer, et al. (2008). "Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display." J Mol Biol 382(5): 1211-27.

Steiner, D., P. Forrer, et al. (2006). "Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display." Nat Biotechnol 24(7): 823-31.

Steipe, B. (1999). "Evolutionary approaches to protein engineering." Curr Top Microbiol Immunol 243: 55-86.

Stemmer, W. P. (1994). "Rapid evolution of a protein in vitro by DNA shuffling." Nature 370(6488): 389-91.
Stohl, W. and D. M. Hilbert (2012). "The discovery and development of belimumab: the anti-BLyS-lupus connection." Nat Biotechnol 30(1): 69-77.
Stork, R., E. Campigna, et al. (2009). "Biodistribution of a bispecific single-chain diabody and its half-life extended derivatives." J Biol Chem 284(38): 25612-9.
Stork, R., D. Muller, et al. (2007). "A novel tri-functional antibody fusion protein with improved pharmacokinetic properties generated by fusing a bispecific single-chain diabody with an albumin-binding domain from streptococcal protein G." Protein Eng Des Sel 20(11): 569-76.
Strohl, W. R. (2009). "Optimization of Fc-mediated effector functions of monoclonal antibodies." Curr Opin Biotechnol 20(6): 685-91.
Stumpp, M. T., H. K. Binz, et al. (2008). "DARPins: a new generation of protein therapeutics." Drug Discov Today 13(15-16): 695-701.
Suzuki, T., A. Ishii-Watabe, et al. (2010). "Importance of neonatal FcR in regulating the serum half-life of therapeutic proteins containing the Fc domain of human IgG1: a comparative study of the affinity of monoclonal antibodies and Fc-fusion proteins to human neonatal FcR." J Immunol 184(4): 1968-76.
Tabuchi, I., S. Soramoto, et al. (2001). "An in vitro DNA virus for in vitro protein evolution." FEBS Lett 508(3): 309-12.
Tamaskovic, R., M. Simon, et al. (2012). "Designed ankyrin repeat proteins (DARPins) from research to therapy." Methods Enzymol 503: 101-34.
Tang, L., H. Gao, et al. (2012). "Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers." Biotechniques 52(3): 149-58.
Tangri, S., C. LiCalsi, et al. (2002). "Rationally engineered proteins or antibodies with absent or reduced immunogenicity." Curr Med Chem 9(24): 2191-9.
Tashiro, M., R. Tejero, et al. (1997). "High-resolution solution NMR structure of the Z domain of staphylococcal protein A." J Mol Biol 272(4): 573-90.
Tawfik, D. S. and A. D. Griffiths (1998). "Man-made cell-like compartments for molecular evolution." Nat Biotechnol 16(7): 652-6.
Thomas, W. D. and G. P. Smith (2010). "The case for trypsin release of affinity-selected phages." Biotechniques 49(3): 651-4.
Thompson, C. A. (2010). "FDA approves kallikrein inhibitor to treat hereditary angioedema." Am J Health Syst Pharm 67(2): 93.
Tian, F., M. L. Tsao, et al. (2004). "A phage display system with unnatural amino acids." J Am Chem Soc 126(49): 15962-3.
Tijink, B. M., T. Laeremans, et al. (2008). "Improved tumor targeting of anti-epidermal growth factor receptor Nanobodies through albumin binding: taking advantage of modular Nanobody technology." Mol Cancer Ther 7(8): 2288-97.
Tokuriki, N. and D. S. Tawfik (2009). "Protein dynamism and evolvability." Science 324(5924): 203-7.
Tolmachev, V. (2008). "Imaging of HER-2 overexpression in tumors for guiding therapy." Curr Pharm Des 14(28): 2999-3019.
Tolmachev, V., A. Orlova, et al. (2007). "Radionuclide therapy of HER2-positive microxenografts using a 177Lu-labeled HER2-specific Affibody molecule." Cancer Res 67(6): 2773-82.
Tolmachev, V., T. A. Tran, et al. (2012). "Tumor targeting using affibody molecules: interplay of affinity, target expression level, and binding site composition." J Nucl Med 53(6): 953-60.
Tolmachev, V., H. Wallberg, et al. (2009). "The influence of Bz-DOTA and CHX-A''-DTPA on the biodistribution of ABD-fused anti-HER2 Affibody molecules: implications for (114m)In-mediated targeting therapy." Eur J Nucl Med Mol Imaging 36(9): 1460-8.
Tuerk, C. and L. Gold (1990). "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science 249(4968): 505-10.
Ueno, S. and N. Nemoto (2012). "cDNA display: rapid stabilization of mRNA display." Methods Mol Biol 805: 113-35.
Uhlén, M. (2008). "Affinity as a tool in life science." 44: 649-654.
Uhlen, M., B. Nilsson, et al. (1983). "Gene fusion vectors based on the gene for staphylococcal protein A." Gene 23(3): 369-78.
Uhlen, M., P. Oksvold, et al. (2010). "Towards a knowledge-based Human Protein Atlas." Nat Biotechnol 28(12): 1248-50.
Ulmer, K. M. (1983). "Protein engineering." Science 219(4585): 666-71.
Unverdorben, F., A. Farber-Schwarz, et al. (2012). "Half-life extension of a single-chain diabody by fusion to domain B of staphylococcal protein A." Protein Eng Des Sel 25(2): 81-8.
Urban, J. H., R. M. Schneider, et al. (2005). "Selection of functional human antibodies from retroviral display libraries." Nucleic Acids Res 33(4): e35.
Vajo, Z. and W. C. Duckworth (2000). "Genetically engineered insulin analogs: diabetes in the new millenium." Pharmacol Rev 52(1): 1-9.
van Bloois, E., R. T. Winter, et al. (2011). "Decorating microbes: surface display of proteins on Escherichia coli." Trends Biotechnol 29(2): 79-86.
Van der Brulle, J., M. Fischer, et al. (2008). "A novel solid phase technology for high-throughput gene synthesis." BioTechniques 45(3): 340-343.
VanAntwerp, J. J. and K. D. Wittrup (2000). "Fine affinity discrimination by yeast surface display and flow cytometry." Biotechnol Prog 16(1): 31-7.
Vanhercke, T., C. Ampe, et al. (2005). "Reducing mutational bias in random protein libraries." Anal Biochem 339(1): 9-14.

Varadamsetty, G., D. Tremmel, et al. (2012). "Designed Armadillo Repeat Proteins: Library Generation, Characterization and Selection of Peptide Binders with High Specificity." J Mol Biol.

Velappan, N., H. E. Fisher, et al. (2010). "A comprehensive analysis of filamentous phage display vectors for cytoplasmic proteins: an analysis with different fluorescent proteins." Nucleic Acids Res 38(4): e22.

Vincke, C., R. Loris, et al. (2009). "General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold." J Biol Chem 284(5): 3273-84.

Virnekas, B., L. Ge, et al. (1994). "Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis." Nucleic Acids Res 22(25): 5600-7.

Visintin, M., E. Tse, et al. (1999). "Selection of antibodies for intracellular function using a two-hybrid in vivo system." Proc Natl Acad Sci U S A 96(21): 11723-8.

Vita, C., C. Roumestand, et al. (1995). "Scorpion toxins as natural scaffolds for protein engineering." Proc Natl Acad Sci U S A 92(14): 6404-8.

Voigt, C. A., C. Martinez, et al. (2002). "Protein building blocks preserved by recombination." Nat Struct Biol 9(7): 553-8.

Wahlberg, E. and T. Hard (2006). "Conformational stabilization of an engineered binding protein." J Am Chem Soc 128(23): 7651-60.

Wahlberg, E., C. Lendel, et al. (2003). "An affibody in complex with a target protein: structure and coupled folding." Proc Natl Acad Sci U S A 100(6): 3185-90.

Walker, A., G. Dunlevy, et al. (2010). "Anti-serum albumin domain antibodies in the development of highly potent, efficacious and long-acting interferon." Protein Eng Des Sel 23(4): 271-8.

Wang, R., X. Fang, et al. (2004). "The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures." J Med Chem 47(12): 2977-80.

Ward, E. S., D. Gussow, et al. (1989). "Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli." Nature 341(6242): 544-6.

Watts, R. E. and A. C. Forster (2012). "Update on pure translation display with unnatural amino acid incorporation." Methods Mol Biol 805: 349-65.

Weaver-Feldhaus, J. M., J. Lou, et al. (2004). "Yeast mating for combinatorial Fab library generation and surface display." FEBS Lett 564(1-2): 24-34.

Webster, J. M., R. Zhang, et al. (2009). "Engineered two-helix small proteins for molecular recognition." Chembiochem 10(8): 1293-6.

Wehrman, T., B. Kleaveland, et al. (2002). "Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments." Proc Natl Acad Sci U S A 99(6): 3469-74.

Weiss, B. and C. C. Richardson (1967). "Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from Escherichia coli infected with T4 bacteriophage." Proc Natl Acad Sci U S A 57(4): 1021-8.

Weiss, G. A., C. K. Watanabe, et al. (2000). "Rapid mapping of protein functional epitopes by combinatorial alanine scanning." Proc Natl Acad Sci U S A 97(16): 8950-4.

Wells, J., B. Cunningham, et al. (1999). From big molecules to smaller ones. Peptide Science – Present and Future. Y. Shimonishi: 1-5.

Wells, J. A. (1990). "Additivity of mutational effects in proteins." Biochemistry 29(37): 8509-17.

Wells, J. A., M. Vasser, et al. (1985). "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites." Gene 34(2-3): 315-23.

Wernerus, H., J. Lehtio, et al. (2002). "Engineering of staphylococcal surfaces for biotechnological applications." J Biotechnol 96(1): 67-78.

Wernerus, H., P. Samuelson, et al. (2003). "Fluorescence-activated cell sorting of specific affibody-displaying staphylococci." Appl Environ Microbiol 69(9): 5328-35.

Wernerus, H. and S. Stahl (2004). "Biotechnological applications for surface-engineered bacteria." Biotechnol Appl Biochem 40(Pt 3): 209-28.

Wiederstein, M. and M. J. Sippl (2005). "Protein sequence randomization: efficient estimation of protein stability using knowledge-based potentials." J Mol Biol 345(5): 1199-212.

Wilchek, M., E. A. Bayer, et al. (2006). "Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction." Immunol Lett 103(1): 27-32.

Wittrup, K. D. (2001). "Protein engineering by cell-surface display." Curr Opin Biotechnol 12(4): 395-9.

Wojcik, J., O. Hantschel, et al. (2010). "A potent and highly specific FN3 monobody inhibitor of the Abl SH2 domain." Nat Struct Mol Biol 17(4): 519-27.

Wolkowicz, R., G. C. Jager, et al. (2005). "A random peptide library fused to CCR5 for selection of mimetopes expressed on the mammalian cell surface via retroviral vectors." J Biol Chem 280(15): 15195-201.

Wong, T. S., D. Roccatano, et al. (2007). "Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries." Environ Microbiol 9(11): 2645-59.

Wozniak-Knopp, G., S. Bartl, et al. (2010). "Introducing antigen-binding sites in structural loops of immunoglobulin constant domains: Fc fragments with engineered HER2/neu-binding sites and antibody properties." Protein Eng Des Sel 23(4): 289-97.

Wright, M. C. and G. F. Joyce (1997). "Continuous in vitro evolution of catalytic function." Science 276(5312): 614-7.

Wu, A. M. and P. D. Senter (2005). "Arming antibodies: prospects and challenges for immunoconjugates." Nat Biotechnol 23(9): 1137-46.

Wurch, T., A. Pierré, et al. (2012). "Novel protein scaffolds as emerging therapeutic proteins: from discovery to clinical proof-of-concept." Trends in Biotechnology 30(11): 575-582.

Yeung, Y. A. and K. D. Wittrup (2002). "Quantitative screening of yeast surface-displayed polypeptide libraries by magnetic bead capture." Biotechnol Prog 18(2): 212-20.

Zahnd, C., C. A. Sarkar, et al. (2010). "Computational analysis of off-rate selection experiments to optimize affinity maturation by directed evolution." Protein Eng Des Sel 23(4): 175-84.

Zahnd, C., E. Wyler, et al. (2007). "A designed ankyrin repeat protein evolved to picomolar affinity to Her2." J Mol Biol 369(4): 1015-28.

Zalevsky, J., A. K. Chamberlain, et al. (2010). "Enhanced antibody half-life improves in vivo activity." Nat Biotechnol 28(2): 157-9.

Zheng, D., J. M. Aramini, et al. (2004). "Validation of helical tilt angles in the solution NMR structure of the Z domain of Staphylococcal protein A by combined analysis of residual dipolar coupling and NOE data." Protein Sci 13(2): 549-54.

Zhou, C., F. W. Jacobsen, et al. (2010). "Development of a novel mammalian cell surface antibody display platform." MAbs 2(5): 508-18.

Zoller, F., U. Haberkorn, et al. (2011). "Miniproteins as phage display-scaffolds for clinical applications." Molecules 16(3): 2467-85.