IMPROVED PHAGE DISPLAY METHODS FOR THE SELECTION OF DNA-SHUFFLED AVIDIN MUTANTS

Master's thesis

Rolle Rahikainen

University of Tampere

Institute of Biomedical Technology

April 2012

ACKNOWLEDGEMENTS

This thesis was carried out in the Molecular Biotechnology group, Institute of Biomedical Technology, University of Tampere between May 2011 and April 2012. I am grateful for the group leader Markku Kulomaa for this opportunity to delve into the interesting world of biotin-binding proteins and to work as a part of his enthusiastic research group.

I also want to express my gratitude to docent Vesa Hytönen, the principal investigator of the Protein Dynamics group and one of my supervisors, for creating an inspiring atmosphere and helping at all stages of the project. My thanks also belong to the other supervisors of this study, MSc Soili Hiltunen and MSc Barbara Niederhauser for interesting and valuable discussions and support in the laboratory work. Moreover, the DNA-shuffled libraries created by MSc Barbara Niederhauser served as the technical basis of this study and made my work possible. I also want to thank for the excellent practical assistance given by Ulla Kiiskinen, Laura Kananen and Jukka Lehtonen. Your work made my day-to-day laboratory work enjoyable.

Lastly, I want to express my deepest gratitude to my girlfriend Iida, my family and my friends who have been constantly supporting me not only during this master's thesis project, but also during all my studies.

Tampere, April 2012

Rolle Rahikainen

PRO GRADU -TUTKIELMA

Paikka:	Tampereen yliopisto Biolääketieteellisen teknologian yksikkö
Tekijä :	RAHIKAINEN, ROLLE MIKAEL
Otsikko:	DNA:n sekoituksella valmistettujen avidiinimutanttien faagiseulontamenetelmien kehittäminen.
Sivut:	86 + 7 s.
Ohjaajat:	Dosentti Vesa Hytönen, FM Soili Hiltunen ja FM Barbara Niederhauser
Tarkastajat:	Professori Markku Kulomaa ja Dosentti Vesa Hytönen
Aika:	Huhtikuu 2012

TIIVISTELMÄ

Tutkielman tausta ja tavoitteet: Faagiseulonta on proteiinikirjastojen seulonnassa käytetty menetelmä, jonka avulla proteiineja ja peptidejä voidaan valikoida niiden ligandinsitomiskyvyn tai muiden biokemiallisten ominaisuuksien perusteella. Faagien epäspesifinen sitoutuminen ja faagien epätasaisen eluoitumisen aiheuttama vääristymä saattavat kuitenkin estää erittäin korkealla affiniteetilla ligandiinsa sitoutuvien proteiinien seulonnan. Tämän tutkimuksen tarkoituksena oli arvioida seitsemän erilaisen faagiseulontamenetelmän soveltuvuutta DNA:n sekoituksella valmistettujen avidiinimutanttien seulontaan.

Materiaalit ja menetelmät: DNA:n sekoituksella valmistettuja AVD/AVR2, AVD/BBP-A ja AVR2/BBP-A kirjastoja seulottiin 3 - 5 kertaa erilaisilla faagiseulontamenetelmillä. Kaikissa testatuissa menetelmissä käytettiin disulfidisidoksen sisältävää biotiinikonjugaattia, joka mahdollistaa biotiiniin sitoutuneiden faagien ei-selektiivisen eluoimisen DTT-käsittelyn avulla. Lisäksi erilaisia apuaineita ja esikäsittelyjä testattiin faagien epäspesifisen sitoutumisen vähentämiseksi. Faagiseulonnassa rikastuneiden avidiiniproteiinien ominaisuuksia tutkittiin faagi-ELISA-menetelmällä ja avidiinimutanttien dissosiaationopeusvakioiden vertailun mahdollistavalla testillä.

Tulokset: Tässä työssä kuvattujen uusien faagiseulontamenetelmien käyttäminen johti joidenkin aikaisemmin tuntemattomien avidiinimutanttien rikastumiseen. Alustavissa toiminnallisuutta arvioivissa tutkimuksissa monet näistä mutanteista dissosioituivat suhteellisen hitaasti alustaan sidotusta biotiinista. Toisaalta esimerkiksi villityypin avidiinia, AVR2:ta tai BBP-A:ta esittelevien faagien rikastumista ei havaittu. Lisäksi tässä työssä kehitetty kuoppalevymenetelmä osoittautui tehokkaaksi ja toistettavaksi menetelmäksi avidiinimutanttien valikointiin niiden dissosiaationopeuksien perusteella.

Johtopäätökset: Muokattuja faagiseulontamenetelmiä käytettiin tässä työssä onnistuneesti uusien, ominaisuuksiltaan lupaavien avidiinimutanttien seulonnassa. Erittäin tiukasti biotiiniin sitoutuvien avidiinimutanttien rikastumista ei kuitenkaan havaittu, mikä saattaa johtua faagien epäspesifisestä sitoutumisesta ja erilaisia avidiinimutantteja esittelevien faagien erilaisista tuottotasoista bakteerisoluissa. Tästä huolimatta tässä työssä esitetyt menetelmät ja tulokset antavat vankan pohjan avidiinimutanttien faagiseulonnassa käytettyjen menetelmien kehittämiselle.

Avainsanat: Faagiseulonta, avidiini, AVR2, BBP-A, D-biotiini, DNA:n sekoitus

MASTER'S THESIS

Place:	University of Tampere Institute of Biomedical Technology
Author:	RAHIKAINEN, ROLLE MIKAEL
Title:	Improved phage display methods for the selection of DNA-shuffled avidin mutants.
Pages:	86 + 7 p.
Supervisors:	Docent Vesa Hytönen, MSc Soili Hiltunen and MSc Barbara Niederhauser
Reviewed by:	Professor Markku Kulomaa and Docent Vesa Hytönen
Date:	April 2012

ABSTRACT

Background and aims: Phage display is a library screening method that allows efficient selection of proteins or peptides according to their ligand-binding affinities or other properties. However, non-specific binding of the phages and selection bias in the elution step may prevent the selection of proteins that have extremely high ligand-binding affinities. The aim of this study was to evaluate the suitability of seven different phage panning protocols for the phage display selection of DNA-shuffled avidin mutants.

Materials and methods: DNA-shuffled AVD/AVR2, AVD/BBP-A and AVR2/BBP-A libraries were panned for 3 - 5 rounds according to various phage display protocols. In all of these protocols, disulphide-containing linkers were used in the immobilization of biotin to allow dissociation-independent elution of the ligand-bound phages by a DTT treatment. In addition, the effects of various sample treatments and phage pre-elutions were evaluated in order to reduce the level of non-specific binding. The avidin mutants that were enriched during the phage panning experiments were functionally analyzed with phage-ELISA and with a microplate assay developed in this study.

Results: The use of the modified phage display protocols resulted in the enrichment of some previously unknown avidin mutants. In the preliminary functional analysis, some of these novel avidin mutants showed slow dissociation from immobilized biotin. However, phages displaying wild-type avidin, AVR2 or BBP-A were not enriched during phage panning. To evaluate the functional properties of the avidin mutants, a novel microplate assay was developed. This assay was proved to be an efficient and reproducible method for the screening of biotin-binding proteins according to the dissociation rate constants.

Conclusions: The new phage display protocols evaluated in this study were successfully used in the selection of novel avidin mutants with promising functional properties. However, as avidin mutants with extremely high biotin-binding affinities were not found, it is possible that non-specific binding and differences in the expression levels of the displayed proteins may have affected the results. Nevertheless, the work presented here provides a solid foundation for future development and optimization of the methods used in the phage display selection of avidin mutants.

Keywords: Phage display, avidin, AVR2, BBP-A, D-biotin, DNA-shuffling

CONTENTS

1. INTRODUCTION 9

2. REVIEW OF THE LITERATURE 11
   2.1 Avidin-biotin technology 11
      2.1.1 Chicken avidin and streptavidin 11
      2.1.2 The structural basis of the tight interaction between avidin and biotin 12
   2.2 Avidin-like proteins and the genes encoding them 15
      2.2.1 Avidin-like proteins 15
      2.2.2 Members of the avidin gene family in chicken 16
         2.2.2.1 The chicken avidin protein and the *AVD* gene 17
         2.2.2.2 Avidin-related proteins (AVRs) and the genes encoding them 18
         2.2.2.3 Biotin-binding proteins (BBPs) and the genes encoding them 20
      2.2.3 Other avidin-like proteins 21
         2.2.3.1 Streptavidin 21
         2.2.3.2 Recently discovered avidin-like proteins 21
   2.3 Phage display in protein engineering 23
      2.3.1 Rational mutagenesis and directed evolution 23
      2.3.2 Phage display in library screening 25
         2.3.2.1 Overview of the method 25
         2.3.2.2 The structure and life cycle of M13 bacteriophages 26
         2.3.2.3 The workflow of typical phage display experiments 29
         2.3.2.4 Phage elution strategies 30
         2.3.2.5 Phage display vectors and fusion protein 31

3. AIMS OF THE STUDY 34

4. MATERIALS AND METHODS 35
   4.1 DNA-shuffling and cloning of the mutant libraries 35
   4.2 Production and precipitation of the phages 35
   4.3 Phage panning 36
      4.3.1 Phage panning protocol 1 36
      4.3.2 Phage panning protocol 2 38
      4.3.3 Phage panning protocols A, B, C, D and E 38
   4.4 Phage titers 40
   4.5 Master plates 41

4.6 DNA Sequencing 42
4.6.1 Phagemid purification 42
4.6.2 Sequencing PCR and ethanol precipitation 42
4.7 Functional analyses 43
4.7.1 Phage-ELISA 44
4.7.2 The avidin-biotin displacement assay 44

5. RESULTS 47
5.1 Sequence analysis of the DNA-shuffled mutants 47
5.2 The libraries panned according to the protocols 1 and 2 48
5.2.1 Phage titers during the panning process 48
5.2.2 Sequence analysis of the phage display enriched mutants 50
5.2.2.1 Libraries panned according to the protocol 1 51
5.2.2.2 Libraries panned according to the protocol 2 52
5.2.3. Functional analysis of the phage display enriched mutants 53
5.2.3.1 Analysis of the ligand binding properties by phage-ELISA 53
5.2.3.2 Analysis of the dissociation kinetics by ABD-assay 54
5.3 The libraries panned according to the protocols A - E 55
5.3.1 Phage titers during the panning process 55
5.3.2 Functional analysis of the mutants 57
5.3.3 Sequence analysis of the mutants 58
5.4 Properties of the avidin-biotin displacement assay 59
5.4.1 Duration of the biotin treatment 60
5.4.2 Reproducibility of the assay 60

6. DISCUSSION 62
6.1 The phage display technology in the selection of novel avidin mutants 62
6.2 Dissociation-independent phage elution 64
6.2.1 Sequence analysis of the avidin mutants 65
6.2.1 The functional properties of the avidin mutants 68
6.3 Methods for reducing non-specific phage binding 69
6.4 The ABD-assay in the functional analysis of novel avidin proteins 74
6.5 Future objectives 76

7. CONCLUSIONS 78

REFERENCES 79

APPENDICES 87
Appendix 1: Solutions and reagents 87
Appendix 2 A: The non-redundant sequences of the mutants panned according to the phage panning protocol 1 89
Appendix 2 B: The non-redundant sequences of the mutants panned according to the phage panning protocol 2 90
Appendix 3: The non-redundant sequences of the mutants panned according to the phage panning protocols A-E 91
Appendix 4: The results of the ABD-assay for the mutants panned according to the protocols 1 and 2 92
Appendix 5: The results of the ABD-assay for the mutants panned according to the protocols A - E..... 93

ABBREVIATIONS

A/A2-library	DNA-shuffled avidin/AVR2 library
A/B-library	DNA-shuffled avidin/BBP-A library
A2/B-library	DNA-shuffled AVR2/BBP-A library
ABD-assay	Avidin-biotin displacement assay
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
AP	Alkaline phosphatase
AVD / AVD	Avidin (Gene / Protein)
AVR / AVR	Avidin-related protein (Gene / Protein)
BBP / BBP	Biotin-binding protein (Gene / Protein)
BSA	Bovine serum albumin
BTN	D(+)-Biotin
cDNA	Complementary DNA
CFU	Colony-forming unit
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
ETS	Expressed sequence tag
HRP	Horseradish peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
OD	Optical density at the indicated wavelength
PBS	Phosphate Buffered Saline
PNPP	para-Nitrophenylphosphate
RT	Room temperature (~21-23 °C)
SPR	Surface plasmon resonance

1. INTRODUCTION

Many of the methods used in life sciences are based on proteins that are capable of binding specifically and tightly other biomolecules (Uhlén, 2008). For example, immunohistochemistry, ELISA methods and many of the methods used in affinity chromatography are dependent on the availability of suitable target-specific binders. Moreover, such binders are also commonly exploited in the pharmaceutical industry and in disease diagnostics (Hey et al., 2005; Khan & Akhtar, 2011). The constant need for specific and sensitive binders has led to a rapid increase in the number of protein engineering studies aiming to the development and isolation of proteins with modified ligand specificities or improved properties (Thie et al., 2008 A). For example, increased thermal and chemical stability, higher protein expression levels and increased ligand specificity are all common goals in protein engineering studies.

If the exact three-dimensional structure of the studied protein has not been solved, it is often extremely difficult to rationally design the mutations needed for a specific modification in the protein structure or function (Khan & Akhtar, 2011). Hence, many of the recently published small-molecule binders have been generated through a process called directed evolution (Thie et al., 2008 A). In this method, the genes encoding the studied proteins are either randomly mutated or combined with fragments of homologous genes. This kind of randomized mutagenesis can be used to produce libraries of up to 10^{13} unique protein encoding genes (Kazlauskas & Bornscheuer, 2009). Although the large majority of the randomly modified genes encode non-functional proteins, some of the genes may encode proteins with improved properties. These protein mutants may themselves be suitable for various applications, or they can be used as the starting material for further mutagenesis studies. Moreover, structural analysis of the proteins with modified properties can help to identify the structural components important for a given property or function (Khan & Akhtar, 2011)

One of the challenges of directed evolution is identifying the genes that encode proteins with desired properties (Kazlauskas & Bornscheuer, 2009). Because the proteins mutants of the library should be selected according to their biochemical properties, analysis of the protein coding DNA sequence is not enough. On the other hand, isolation of the protein coding genes is required for the production and analysis of the selected

protein mutants. Several different library screening strategies have been designed for this purpose (Uhlén, 2008). Phage display is one of the oldest and the most commonly used library screening methods. In phage display, the protein mutants are displayed on the surface of phage particles, while the genes encoding the displayed proteins are packed inside the phage capsids. This creates a physical linkage between the genotype and the phenotype of the studied proteins and allows efficient selection and analysis of them (Hust et al., 2008).

During the past 25 years, phage display has proven to be an extremely versatile and efficient method for the screening of mutant libraries (Bratkovič, 2011). On the other hand, the high versatility of this method has its drawbacks too. Because the aims and the laboratory methods used in phage display can vary significantly from an experiment to experiment, it may not be possible to directly adopt phage display techniques for the screening of other mutant libraries. Moreover, so far most of the phage display studies have been focusing on the selection of modified antibodies or antibody fragments (Hust et al., 2008). Hence, some of the methods commonly used in phage display may not be optimal for the selection of mutants based on other protein scaffolds (Ebersbach et al., 2007).

The main aim of this study was to evaluate the suitability of various phage display methods for the selection of DNA-shuffled avidin mutants with high biotin-binding affinities. Most importantly, disulphide-containing linkers were used in biotin immobilization to allow efficient and non-selective elution of the phages displaying high affinity proteins. In addition, modified sample compositions were tested to reduce the degree of non-specific background binding. Also, novel microplate assays were developed to facilitate the initial screening of the phage display enriched avidin mutants.

2. REVIEW OF THE LITERATURE

2.1 Avidin-biotin technology

2.1.1 Chicken avidin and streptavidin

Avidin is a small tetrameric protein that was originally isolated from chicken egg-white (Eakin, Snell & Williams, 1941). Avidin binds its natural ligand biotin, also known as vitamin H or vitamin B7 with exceptionally high affinity (K_d ~10^{-15} M) (Green, 1975). Both avidin and its bacterial analog streptavidin are relatively small, resistant to many proteases and in their ligand-bound state extremely stable in a wide variety of conditions (Gonzalez, 1999; Green, 1975). In addition to these advantageous properties of avidin and streptavidin, the water solubility of biotin and its easy conjugation with other molecules have made so-called (strept)avidin-biotin technology an important tool in many fields of life-sciences (Laitinen et al., 2007).

Although wild-type avidin is as such suitable for many applications, both avidin and streptavidin have been extensively modified (Laitinen et al., 2006). Many of these modifications have been designed to either improve one or more of the properties of avidin or to shed light on the biochemical background of the tight avidin-biotin interaction (Laitinen et al., 2006). For example, the glycosylation state of avidin, its quaternary structure, isoelectric point, stability and many other structural and functional properties of avidin have been modified to make it better suitable for different applications (Marttila et al., 2000; Nordlund et al., 2004; Lim et al., 2011,). Moreover, the ligand specificity of avidin has been modified to allow specific binding of small molecules other than biotin (Määttä et al., 2008; Riihimäki et al. 2011 A). As a result, the current applications of the (strept)avidin-biotin technology range from protein purifications to diagnostics and tissue targeted drug delivery, and novel applications are constantly being developed (Laitinen et al., 2007). Besides to its value as a laboratory tool, the avidin-biotin interaction also offers biochemists an interesting system to study the atomic-level mechanisms behind extremely tight protein-ligand interactions (Laitinen et al., 2006).

2.1.2 The structural basis of the tight interaction between avidin and biotin

Avidin belongs to the calycin superfamily with lipocalins and fatty-acid binding proteins and shares structural features with the other members of this superfamily (Flower et al., 1993). The avidin tetramer consists of four identical subunits, each of which contain one biotin-binding site (See Figure 1). One avidin subunit consists of 128 amino acid residues and has a molecular mass of 15.5 – 18.0 kDa, depending on its glycosylation state (DeLange & Huang, 1971; Green, 1975). Wild-type avidin is glycosylated at one asparagine residue, which increases the molecular mass of the monomers by 10% (Flower et al., 1993). However, glycosylation is not necessary for biotin-binding, which allows the expression of functional chicken avidin in many other organisms, such as bacterial cells (Hiller et al., 1987).

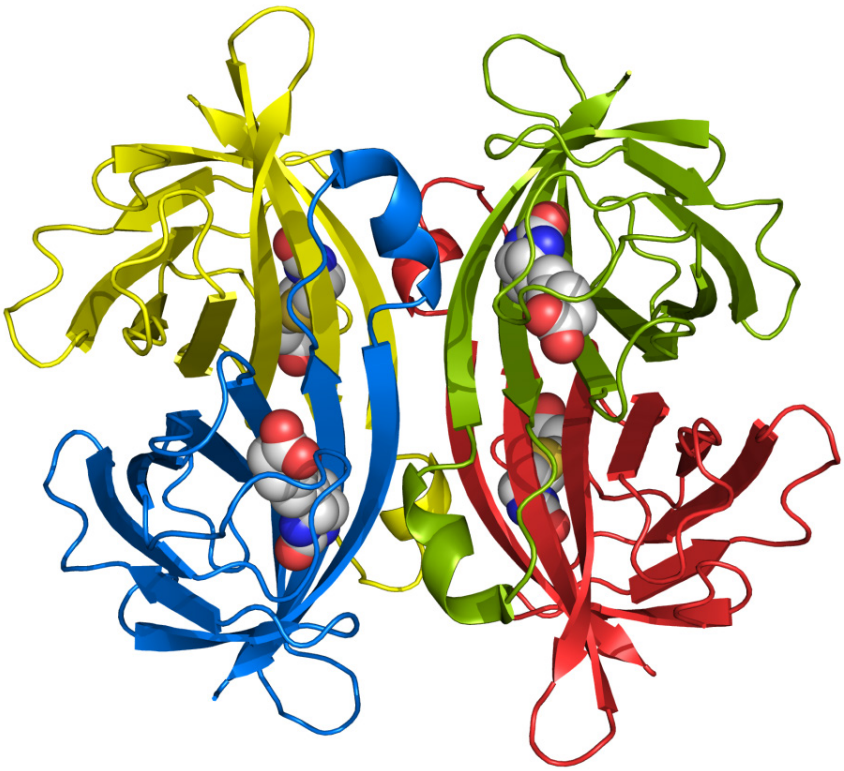


Figure 1. A structural model of tetrameric chicken avidin. The avidin subunits shown as cartoon models are colored to illuminate the tetrameric structure. Biotin molecules bound to each of the biotin-binding pockets are presented with space-filling models. The avidin subunits colored in blue and green or red and yellow are usually called functional pairs, as these avidin subunits offer each other amino acid residues that participate in biotin-binding. PDB ID: 2AVI (Livnah et al., 1993).

The tight interaction between (strept)avidin and biotin was utilized in numerous applications even before the structural basis of this interaction was known. The three dimensional structure of the biotin-binding pocket of streptavidin was first resolved in

1989 by Hendrickson et al., and the structure of a complete avidin tetramer was resolved in 1993 by Livnah et al. (Hendrickson et al., 1989; Livnah et al., 1993). Like the other proteins of the calycin superfamily, avidin monomers are eight-stranded antiparallel β-barrels with internal ligand binding site (See Figure 2) (Flower et al., 1993; Livnah et al., 1993). Although the amino acid sequences of avidins, lipocalins and fatty-acid binding proteins share only limited sequence identity, the three-dimensional β-barrel structure of these proteins has been well conserved in evolution (Flower et al., 1993).

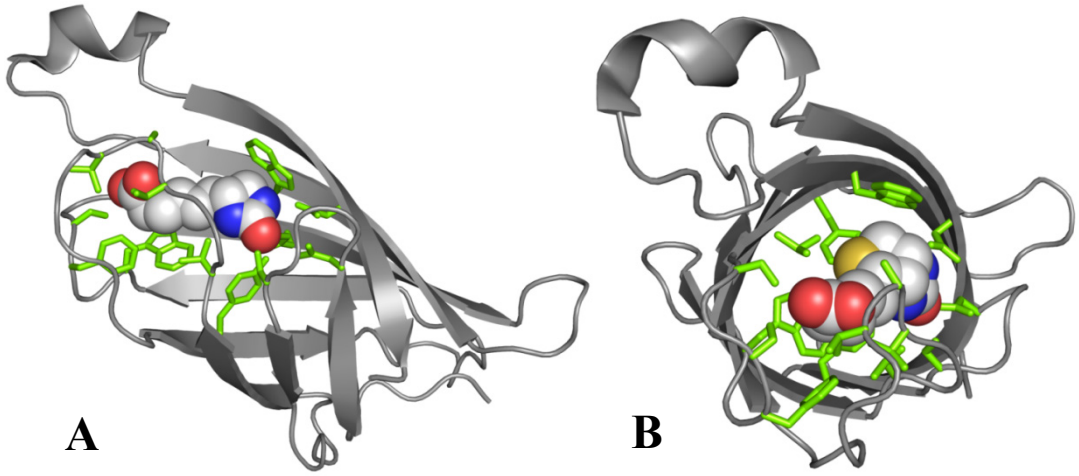


Figure 2. **The biotin-binding pocket of an avidin subunit**. The structure of the subunit is presented with grey cartoon model, while the sidechains of the biotin-binding residues are shown as green sticks. Tryptophan 110 residue participating to the ligand recognition from the adjacent subunit is not shown. A biotin molecule bound to the biotin-binding pocket is presented with a space-filling model. (A) Organization of the biotin-binding amino acid residues. (B) The biotin-binding site of the avidin subunits is located at one end of the β-barrel structure. PDB ID: 2AVI (Livnah et al., 1993).

The amino acid residues participating in biotin-binding have been studied with various methods, including X-ray crystallography and avidin mutagenesis (Repo et al., 2006; Laitinen et al., 2006). The biotin-binding site is located at one end of the barrel-like structure of the subunit and it is surrounded by amino acid residues pointing towards the center of the barrel (See Figure 2) (Livnah et al., 1993; Pugliese et al., 1993). In wild-type avidin, in total 16 amino acid residues (See Figure 3) directly participate in biotin-binding (Livnah et al., 1993). These residues include both polar and hydrophobic amino acid residues and they create a binding pocket that fits precisely to the shape and electrical properties of a biotin molecule (Livnah et al., 1993; Pugliese et al., 1993). Once a biotin molecule has bound to an avidin tetramer, it is held firmly in place by multiple hydrogen bonds between the amino acid residues of the binding site and the polar atoms of biotin (Livnah et al., 1993). Moreover, the hydrophobic parts of the

biotin molecule are lined with five hydrophobic amino acid residues, which stabilizes biotin to the binding site and accounts for the extremely high biotin-binding affinity of avidin (Livnah et al., 1993). Interestingly, one of these hydrophobic residues, tryptophan 110, is provided by another avidin subunit (Livnah et al., 1993). The inter-subunit interactions between Trp110 (Trp120 in streptavidin) and biotin are important for the high biotin-binding affinity of avidin tetramers and mutations affecting this residue substantially reduce the biotin-binding affinity (Sano et al., 1995). Because of their essential contribution to biotin-binding, the avidin subunits providing these residues to each other are usually called functional pairs (See Figure 1). In addition to increasing the biotin-binding affinity of the avidin subunits, the interaction between tryptophan 110 and biotin has been shown to significantly increase the thermal stability of the avidin tetramers (Sano et al., 1995; Gonzalez et al., 1997).

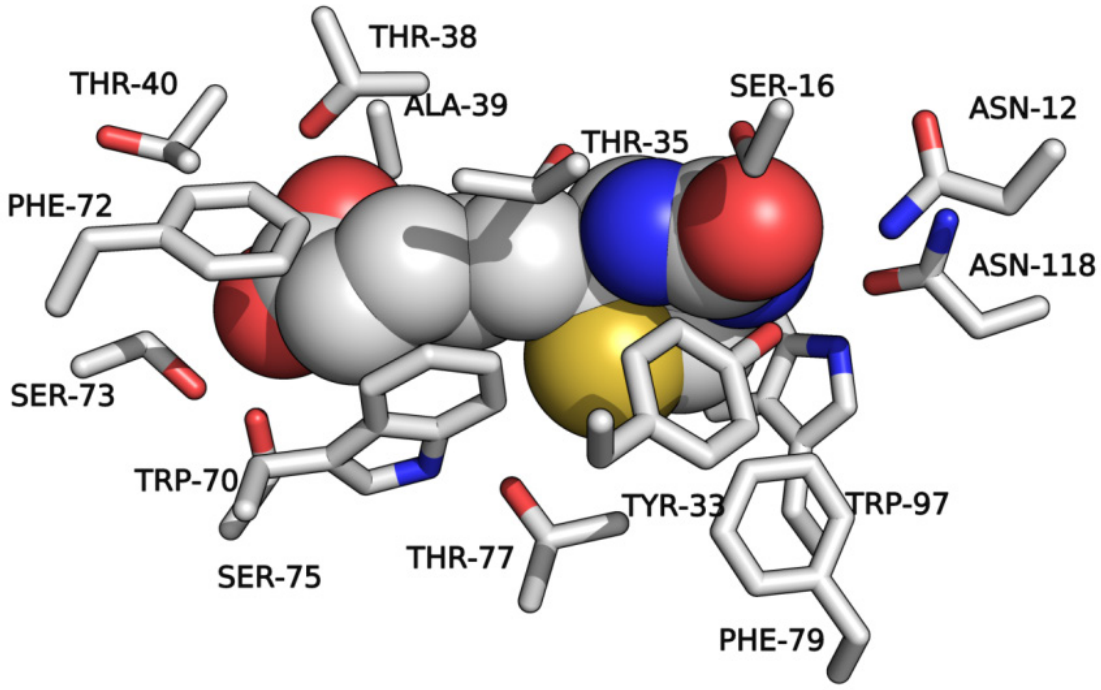


Figure 3. The amino acid residues participating in biotin-binding in chicken avidin. A biotin molecule bound to the biotin-binding site is presented as a space-filling model. Tryptophan 110 residue provided by the functional pair is not shown. The atoms are colored as follows: carbon, white; nitrogen, blue; oxygen, red; sulphur, yellow. PDB ID: 2AVI (Livnah et al., 1993).

2.2 Avidin-like proteins and the genes encoding them

2.2.1 Avidin-like proteins

Because of the exceptionally high biotin-binding affinities of chicken avidin and its bacterial homolog streptavidin, many avidin-like proteins have been studied extensively during the past few decades. However, these proteins and the genes encoding them are still nowadays under active research. The main driving force of this research has been the search for previously unknown proteins that are capable of binding biotin specifically and tightly. Different biotin-binding proteins vary in their physicochemical properties, e.g. ligand-binding affinity, thermal stability, isoelectric point and protease resistance and hence, they may be suitable for different purposes (Hytönen et al., 2005 A; Takakura et al., 2009; Sardo et al., 2001). Furthermore, immunologically different avidin-like proteins may prove to be important in therapeutic applications (Helppolainen et al., 2008). In addition to novel applications of the (strept)avidin-biotin technology, previously unknown biotin-binding proteins may also provide interesting insights into the structure-function relationship of biotin-binding proteins. By comparing the structural and functional properties of different avidin-like proteins, it has been possible to identify the structural components that are important for a specific function, e.g. for tetramerization or biotin-binding (Helppolanen et al., 2007).

Although many of the avidin-encoding genes were originally discovered from the genome of chicken, homologous genes have recently been found from many other species as well (Helppolainen et al., 2007; Helppolainen et al., 2008; Takakura et al., 2009). These species include both eukaryotes and bacteria, and the number of taxons known to contain avidin-encoding genes is steadily increasing (Sardo et al., 2011). It is clear that avidin-like proteins and the genes encoding them are widespread among many kingdoms of life. In many cases the exact biological function of the avidin-like proteins is unknown. Nevertheless, the ability of these proteins to bind biotin has been well conserved in evolution (Sardo et al., 2011). This suggests that biotin-binding is an important feature of many avidin-like proteins. On the other hand, some proteins containing avidin-like domains do not bind biotin at all (Yanai et al., 2005). Instead, these proteins may have adopted the tetrameric structure for multimerization rather than for ligand binding (Yanai et al., 2005).

2.2.2 Members of the avidin gene family in chicken

The chicken avidin gene family is not only the oldest, but also the largest known family of genes encoding avidin-like proteins. In addition to the avidin gene (*AVD*), the first member of the chicken avidin gene family, the chicken genome contains from 3 to 7 homologous genes encoding avidin-related proteins (*AVR1* - *AVR7*) and two genes encoding so-called biotin-binding proteins (*BBP-A* and *BBP-B*) (Niskanen et al., 2005; Ahlroth et al., 2000). These genes, the proteins they encode (See Figure 4) and their most important properties are discussed in the following paragraphs.

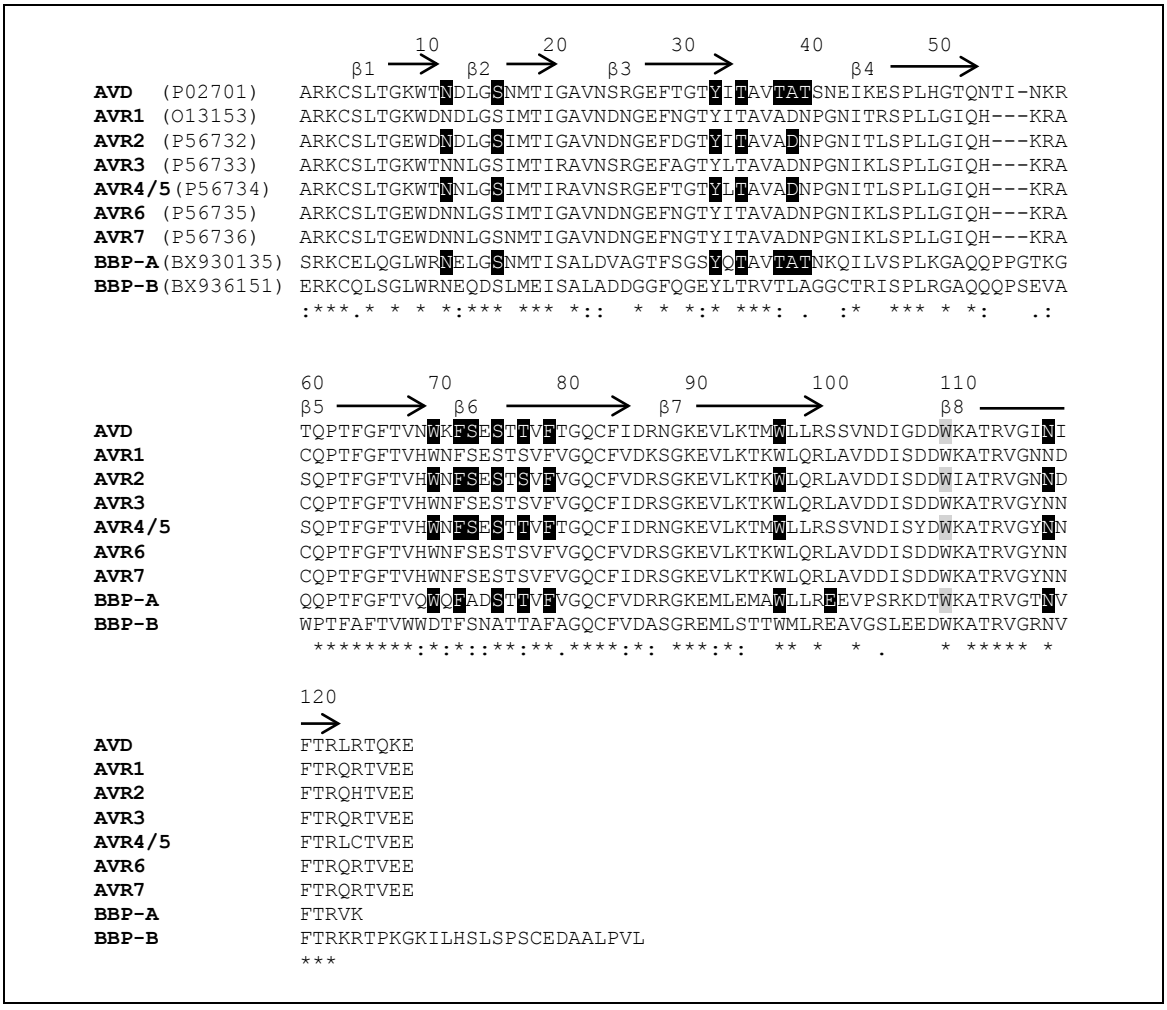


Figure 4. The amino acid sequences of chicken avidin and avidin-like proteins. The amino acids that have been conserved in all aligned sequences are marked with an asterisk, strong amino acid group similarity is indicated by a colon and weaker group similarity is indicated by a single dot. Residues that have been shown to participate in biotin-binding in structures determined by X-ray crystallography are shaded in black, apart from the interface tryptophans (W110) that are shaded in gray. The beta-strands of the avidin fold are marked with black arrows above the sequences. Because the proteins encoded by *AVR4* and *AVR5* are identical, only one sequence is shown. The GenBank accession numbers of the aligned sequences are presented after the name of each sequence.

2.2.2.1 The chicken avidin protein and the *AVD* gene

The first member of the avidin protein family, named avidin, was found from chicken egg-white already in the 1940's (Eakin et al., 1941). The first hints of the presence of avidin came from studies in which test animals fed with raw chicken egg-white suffered from various symptoms and injuries typical of biotin deficiency (Eakin et al., 1940 A). However, as added biotin did not relieve these symptoms, it was concluded that a constituent of the egg-white must have been sequestering biotin and rendering it unavailable for the animals (Eakin et al., 1940 B; Williams, 1943 and references therein). Although the exact structure of the avidin tetramer remained unknown for almost fifty years, the value of the extremely tight interaction between avidin and biotin was realized already in the 1970's (Warnke et al., 1980; Bayer & Wilchek, 1980).

The 128 amino-acid-long sequence of the avidin protein was first determined in 1971 and the respective cDNA was cloned in 1987 (DeLange & Huang, 1971; Gope et al., 1987). However, the full length *AVD* gene was not cloned until the year 1995 (Wallén et al., 1995). As described by Wallén et al. in 1995, this gene consists of four exons interrupted by three introns. The *AVD* gene is located on band q21 of the Z-chromosome of the chicken genome (Ahlroth et al., 2000). Interestingly, the *AVD* gene is accompanied by from 3 to 7 *AVR* genes that are located in the same gene cluster (See Figure 5) (Ahlroth et al., 2000; Niskanen et al., 2005). The total length of this gene cluster is around 27 kb and the intergenic regions vary in size between 2.5 and 9 kb (Ahlroth et al., 2000). As apparent from the Figure 5, the *AVD* gene is located at one end of the cluster, separated from the other genes by a 9-kb spacer (Ahlroth et al., 2000).

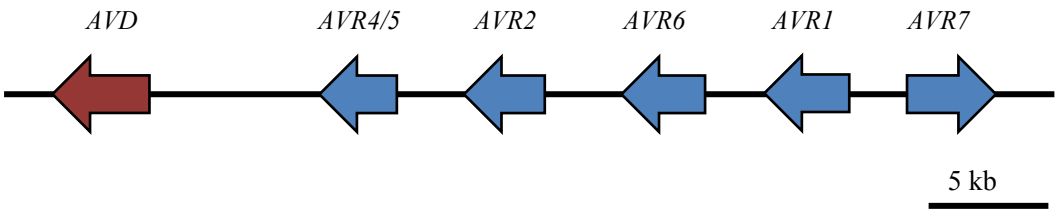


Figure 5. The *Avd* gene cluster in the chicken genome. The genes encoding avidin and all avidin-like proteins (AVR1-7) are located within the same gene cluster. All of these genes are located on the same strand, except for AVR7, which is located on the complementary strand. As the number of the *AVR* genes varies between different individuals, the organization of the gene cluster is not similar in all cases (Niskanen et al., 2005). Figure adapted from Ahlroth et al., 2000.

Despite of the active research in the field of biotin-binding proteins, the biological function of chicken avidin is still under debate (Stevens et al., 1991). The *AVD* gene is known to be expressed predominantly in the female reproductive track, from where the produced avidin proteins are secreted to the egg-white of chicken eggs (O'Malley et al., 1967). Nevertheless, the *AVD* gene can be expressed in almost any chicken tissue during microbial infections or in tissue trauma (Heinonen et al., 1978; Elo et al., 1980 A). Although the expression of the *AVD* gene can be induced either by a progesterone-dependent or by an inflammation-mediated mechanism, the biological functions of the produced avidin proteins are probably similar in both cases; avidin acts as an antimicrobial agent by sequestering biotin from the tissue (Tuohimaa et al., 1989 and references there-in; Green, 1975). By doing so, avidin probably inhibits the growth of biotin-requiring bacteria and yeasts and hence acts as a part of the innate immunity system in chicken.

2.2.2.2 Avidin-related proteins (AVRs) and the genes encoding them

The first 5 genes (*AVR1-5*) encoding avidin-related proteins, AVRs, were cloned in the late 1980's and early 1990's by Keinänen et al., after which *AVR6* and *AVR7* were cloned in the year 2000 by Ahlroth et al. (Keinänen et al., 1988; Keinänen et al., 1994; Ahlroth et al., 2000). As their name suggests, these genes are closely related to the chicken *AVD* gene. The sequence identities of the *AVR* and *AVD* genes range from 91% to 95%, with *AVR1* having the lowest and *AVR4/5* having the highest sequence identity with the *AVD* gene (Wallén et al., 1995). As mentioned above, all known *AVR* genes are located in the same gene cluster together with the *AVD* gene (See Figure 5). Intriguingly, the number of the AVR-encoding genes in this cluster differs between individuals and even between the cells of a single individual (Ahlroth et al., 2000). While the chicken genome project indicated that there are only 3 different *AVR* genes, many more genes were found in previous studies by Ahlroth et al. (Niskanen et al., 2005; Ahlroth et al., 2001 A). In these studies, the most frequent number of *AVR* genes was shown to be 4 or 5, and some of the analyzed cells contained 7 or more *AVR* genes (Ahlroth et al., 2001 A). In some cases, the whole gene cluster was duplicated or triplicated (Alhroth et al., 2001 A). The mechanisms behind the fluctuations in the number of the *AVR* genes are believed to be unequal crossing over or unequal sister chromatid exchange (Ahlroth et al., 2001 B). This hypothesis is supported by the sub-

telomeric location of the *AVD/AVR* complex and the finding that the intronic regions of the *AVR* genes are more conserved than the exonic regions (Ahlroth et al., 2001 A; Wallén et al., 1995).

Most of the differences between the *AVD* and *AVR* sequences are nonrandomly distributed nonsynonymous point mutations (Keinänen et al., 1994; Wallén et al., 1995). Out of the 16 amino acid residues known to directly participate in biotin-binding in avidin, 13 are conserved in all AVR proteins (Keinänen et al., 1994). The three substituted biotin-binding residues (T38A, A39D, T40N) are all located one after another between the amino acid positions 38 and 40 in a loop connecting the β-strands 3 and 4 (Keinänen et al., 1994). Although these residues are in direct contact with biotin, they are also substituted in many other biotin-binding proteins, which suggests that they are not critical for biotin-binding (Helppolainen et al., 2008; Laitinen et al., 2002). Despite of the high sequence identity of the *AVR* genes, the physicochemical properties of the encoded proteins vary significantly (Laitinen et al., 2002; Hytönen et al., 2005 B). For example, the biotin-binding affinities of the avidin-related proteins vary over a wide range of values, AVR4 having the highest and AVR2 having the lowest biotin-binding affinity (Laitinen et al., 2002). The AVR proteins also differ in their glycosylation patterns and isoelectric points. Similarly to avidin, the theoretical pI values for AVR3 and AVR4/5 are basic, while AVR1, AVR6 and AVR7 are neutral and AVR2 is acidic (Laitinen et al., 2002).

It is still not known whether the AVR proteins are expressed in the chicken tissues or not and the possible biological functions of these proteins are still poorly understood (Hytönen et al., 2005 B; Ahlroth et al., 2001 B). However, the mRNAs of *AVR2* and *AVR3* have been shown to be present in the chicken oviduct during inflammation (Kunnas et al., 1993). In addition, because all known *AVR* genes contain intact splicing sites, uninterrupted coding sequences and possibly functional promoter regions, it is unlikely that these genes would be pseudogenes (Wallén et al., 1995). On this background, it is not surprising that if these genes are expressed in an insect cell line (*Spodoptera frugiperda*), all *AVR* genes are able to encode functional, biotin-binding proteins (Laitinen et al., 2002).

2.2.2.3 Biotin-binding proteins (BBPs) and the genes encoding them

Biotin-binding proteins (BBPs) are the third class of the chicken proteins that are capable of binding biotin with high affinities. Unlike avidin, biotin-binding proteins are not secreted to egg-white. Instead, these proteins were discovered from the yolk of chicken eggs (György et al., 1942; White et al., 1976). It has been hypothesized that by binding biotin in the egg yolk, these proteins can prevent the biotin molecules from leaking to the egg-white. Thus biotin-binding proteins may act as a biotin storage system of the developing embryo (White & Whitehead, 1987). On the other hand, these proteins may also transport biotin between chicken plasma and the developing chicken oocytes (White & Whitehead, 1987).

The two currently known chicken BBP proteins, formerly called BBP-I and BBP-II, were originally hypothesized to be encoded by a single gene containing four tandemly repeated biotin-binding domains (Bush & White, 1989). Nevertheless, when the chicken genome was sequenced in 2004, two separate genes were found (Niskanen et al., 2005). Moreover, these genes do not encode four but only single biotin-binding domains. Because the results were partly conflicting with previous research, these genes were renamed as *BBP-A* and *BBP-B* (Niskanen et al., 2005). Apart from a long C-terminal extension in BBP-B, the proteins encoded by these genes (See Figure 4) have a high sequence identity with chicken avidin, which suggests that they may have a common ancestor (Niskanen et al., 2005). In addition, almost all of the biotin-binding residues and many of the interface residues are conserved in both BBP-proteins (Niskanen et al., 2005). Interestingly, one of the biotin-binding residues of BBP-A, glutamate 102, does not have biotin-binding counterparts in the chicken avidin or AVR proteins (See Figure 4) (Hytönen et al., 2007). This residue has been shown to form a hydrogen bond with one of the oxygen atoms in biotin, but the equivalent serine residue in avidin is too far to hydrogen-bond to biotin (Hytönen et al., 2007). Despite of these differences, if the *BBP-A* gene is expressed in insect or bacterial cells, the produced proteins form functional, biotin-binding tetramers (Hytönen et al., 2007). The biotin-binding affinity of the BBP-A tetramer is lower than that of chicken avidin, mostly because of a higher biotin dissociation rate. It is possible, that this results from the decreased structural stability of the BBP-A tetramers as compared to avidin (Hytönen et al., 2007).

2.2.3 Other avidin-like proteins

2.2.3.1 Streptavidin

Streptavidin, an avidin-like protein originally found from *Streptomyces avidinii*, is among the most studied avidin proteins. Because of its exceptional biotin-binding affinity (K_d ~10^{-14} M) and low non-specific binding, streptavidin has become one of the most commonly exploited small-molecule binders in the fields of biotechnology and medical diagnostics. To date, over 200 streptavidin mutants have been published (Laitinen et al., 2006). Each of these mutants has their unique properties, which makes the (strept)avidin-biotin technology suitable for a wide range of applications. For example, streptavidin mutants with modified ligand-binding specificities, increased stability and tailored ligand-binding properties have been designed (Laitinen et al., 2006; Chivers et al., 2010).

The gene encoding streptavidin was first cloned and sequenced in the year 1986 by Aragana et al. (Aragana et al., 1986). The most notable differences between the amino acid sequences of avidin and streptavidin are the 11 and 30 amino-acid-long N- and C-terminal extensions in streptavidin (Aragana et al., 1986). However, these extensions are cleaved off post-translationally, and the length of a mature streptavidin protein is similar to that of avidin (Weber et al., 1989). Despite of some differences in the amino acid sequences of avidin and streptavidin, the amino acid residues participating in biotin-binding have been conserved remarkably well (Aragana et al., 1986). Moreover, it has been hypothesized that the biological function of streptavidin could be similar to that of chicken avidin, i.e. it may act as an anti-microbial agent, hence giving a growth advantage to the *Streptomyces* species (Chaiet & Wolf, 1964).

2.2.3.2 Recently discovered avidin-like proteins

The avidin-like proteins discussed above are only a small fraction of all characterized or putative avidin-like proteins. In addition to these proteins, also many other avidin-like proteins are known. Most importantly, avidin-like proteins have been postulated to exist throughout the oviparous vertebrates and some experimental evidence supporting this hypothesis has been presented (Korpela et al., 1981; Elo, 1980 B). For example, avidin-

like proteins have been discovered from many avian species, including domestic duck (*Anas platyrhynchos*), ostrich (*Struthio camelus*), common turkey (*Meleagris gallopavo*) and zebra finch (*Taeniopygia guttata*) (Long et al., 2003; Hytönen et al., 2003). In addition, avidin-like proteins from non-avian species are known as well. Among these proteins are rhizavidin and bradavidins I and II from symbiotic nitrogen fixing bacteria (*Rhizobium etli* and *Bradyrhizobium japonicum*), xenavidin from Western clawed frog (*Xenopus tropicalis*) and zebavidin from zebrafish (*Danio rerio*) (Helppolainen et al., 2007; Nordlund et al., 2005; Helppolainen et al., 2008; Määttä et al., 2009; Niederhauser et al., unpublished results). More recently, new avidin-like proteins have been discovered from tamogitake mushroom *Pleurotus cornucopiae* (tamavidins I and II) and from a pathogenic proteobacterium *Burkholderia pseudomallei* (burkavidin) (Takakura et al., 2009; Sardo et al., 2011). Although the most distant members of the avidin superfamily share only limited sequence identity, the biotin-binding residues have been generally well conserved in evolution. Moreover, many of these proteins are known to bind biotin, which suggests that they either have or have recently had biological functions. Except for rhizavidin, which forms dimers, also the tetrameric structure of the avidin-like proteins has been well conserved (Helppolainen et al., 2007).

Several different methods have been used in the identification of the aforementioned avidin-like proteins and the genes encoding them. As already mentioned, avidin was first detected in functional studies, and its biochemical characteristics were well known before its mRNA or gene was cloned (Gope et al., 1987; Wallén et al., 1995). However, this is not always the case. Most of the avidin-like proteins presented above have been functionally characterized only after the genes encoding them have been cloned and sequenced. For example, some of the genes encoding the avidin-like proteins described above were originally identified from EST-databases (Veneskoski et al., unpublished results; Hytönen et al., 2007; Määttä et al., 2009). After the recent advances in DNA sequencing technologies, whole genome sequencing has become the most important tool for identifying genes encoding novel avidin-like proteins.

2.3 Phage display in protein engineering

2.3.1 Rational mutagenesis and directed evolution

The methods used in protein engineering are usually divided into two separate classes; rational mutagenesis and directed evolution (Bornscheuer & Pohl, 2001). As its name suggests, the mutations and substitutions made with rational mutagenesis are rationally designed to modify one or more of the biochemical properties of the studied protein. Rational mutagenesis usually requires detailed knowledge of the function and three dimensional structure of the protein (Khan & Akhtar, 2011). This often limits the usability of this mutagenesis method to proteins that have already been characterized in detail (Khan & Akhtar, 2011). To minimize the amount of laboratory work needed, the effects of rationally designed, site-directed mutations are usually predicted by using molecular modeling and other computational approaches (Kazlauskas, 2000). Nevertheless, the results of rational mutagenesis are often more or less unpredictable (Harayama, 1998).

If the three dimensional structure of the studied protein, or that of a close homolog is not known, it may be extremely difficult to successfully modify the protein with rational mutagenesis. If this is the case, directed evolution can be used (Khan & Akhtar, 2011). Directed evolution mimics the natural evolution of proteins, although the pace of this evolution is greatly accelerated (Arnold, 1998). This method is based on the selection of mutant proteins with desired properties from libraries consisting of a vast number of different mutants (Arnold, 1998). The construction of the mutant libraries and the selection of the functional mutants are the most important steps in directed evolution. Several different methods for both steps exist and the success of the experiment is largely dependent on the selection of suitable methods.

Mutant libraries can be constructed by either random mutagenesis or by *in-vitro* recombination (Ling & Robinson, 1997; Stemmer et al., 2002). The most commonly used methods for random mutagenesis are error-prone PCR, degenerated PCR primers, UV-light, chemical mutagens and so-called mutator bacterial strains (Arnold, 1998; Khan & Akhtar, 2011; Low et al., 1996). Although random mutagenesis results in mutant libraries with high diversity, usually a large majority of the mutants have poor

functionality (Stemmer et al., 2002). This is mainly because most of the randomly generated mutations have adverse effects on the protein structure and function (Harayama, 1998). Moreover, if the context of a beneficial mutation is wrong, e.g. the same sequence contains also deleterious mutations, the clone containing these mutations may not be selected for further analysis (Harayama, 1998).

On the contrary to random mutagenesis, *in-vitro* recombination utilizes the genetic material existing in nature. The methods used in *in-vitro* recombination, also known as molecular breeding, typically mimic the processes that rearrange and modify gene alleles and paralogous genes in living cells (Stemmer et al., 2002). DNA family shuffling, a method first presented in the year 1994 by Stemmer et al., is one of the classical examples of *in-vitro* recombination (Stemmer et al., 1994 A). In this method, two or more homologous genes are randomly cleaved with restriction enzymes to fragments of 10 to 50 nucleotides in length (Stemmer et al., 1994 A). When such fragments are randomly reassembled in a primerless PCR reaction, the resulting genes consist of parts from many parental genes (Stemmer et al., 1994 B). Such chimeric genes may exhibit properties that are not observed in any of the parental genes (Stemmer et al., 2002). Because the chimeric genes are composed of DNA sequences of the functional parental genes, the overall functionality of DNA-shuffled libraries is usually high. This, in turn, means that mutant proteins with desired properties can be found from smaller libraries.

DNA-shuffling has been used for many different applications, including enhancing enzyme activity, increasing the biological activity of cytokines and altering the ligand-binding specificities of antibodies and enzymes (Stemmer et al., 1994 B; Chang et al., 1999; Crameri et al., 1998). The parental genes used in DNA-shuffling can be either homologous genes existing in nature, chimeric genes that have been modified with other methods or both (Crameri et al., 1998). However, because of the PCR methods used in DNA-shuffling, the sequence identity of the parental genes has to be relatively high to ensure efficient recombination between the genes (Moore et al., 2000).

2.3.2 Phage display in library screening

2.3.2.1 Overview of the method

Protein engineering through directed evolution is largely dependent on efficient library screening methods. As randomly generated libraries are extremely large, characterization of all individual mutants is usually not possible (Diaz-Neto et al., 2009). The main prerequisite for high-throughput library screening is the coupling of the genotype and the phenotype of each mutant (Azzazy & Highsmith, 2002). Although the screening of the mutant libraries is based on the biochemical properties of the studied proteins, knowledge of the DNA sequences encoding them is required for further modifications, production of the screened proteins and characterization of them (Diaz-Neto et al., 2009). Several different *in-vivo* and *in-vitro* methods have been developed for coupling the studied protein mutants and the genes encoding them together (Uhlen, 2008). Different two-hybrid-techniques and the methods used in ribosome display, mRNA display and phage display all link genotype and phenotype together and provide tools for efficient library screening (O'Neil & Hoess, 1995; Hanes & Plückthun, 1997; Roberts & Szostak, 1997). All these methods have their strengths and weaknesses that should be taken into consideration when selecting a library screening method for a given project (Uhlen, 2008). For example, the maximum library sizes, use of host organisms and the need for special laboratory equipment differ between these methods (Uhlen, 2008).

Phage display is among the most utilized library screening methods (Uhlen, 2008). Introduced in 1985, phage display was one of the first methods coupling genes and the encoded proteins together (Smith, 1985). As its name suggests, phage display is based on displaying the studied proteins or peptides on the surface of bacteriophages, typically Ff-class (M13, f1 and fd) filamentous phages (O'Neil & Hoess 1995; Hust et al., 2008). Importantly, the genes encoding each of the displayed mutants are packaged within the same phage virion with the encoded proteins, which physically links the genotype and the phenotype together. This allows specific affinity purification and amplification of the bacteriophages that are displaying proteins with the desired properties (Hust et al., 2008).

Although phage display was originally designed to facilitate the identification and cloning of unknown genes, its potential in protein engineering was soon discovered (Smith, 1985; McCafferty et al., 1990). To date, phage display has been utilized in a multitude of different applications. For example, receptor binding sites, protein-ligand interactions and protein-DNA interactions have all been studied by using phage display (Azzazy & Highsmith, 2002; Gommans et al., 2005). Moreover, the potential applications of the phage display technology are not limited to screening for high-affinity binders. For example, it has been possible to identify aggregation resistant antibodies, proteins with increased thermal stability and proteins that show catalytic activity by modifying the conditions applied during phage panning (Jespers et al., 2004; Forrer et al., 1999; Sieber et al., 1998). Nevertheless, the most important application of phage display has been the selection of antibodies that recognize specific antigens (Azzazy & Highsmith, 2002). Because of the large size and complex structure of complete wild-type antibodies, usually only the antigen-binding site is displayed on the surface of the phage virions (Bruin et al., 1999). Libraries consisting of these so-called Fab- (Fragment antigen binding) or scFv- (Single-chain fragment variable) fragments have been successfully screened with phage display to isolate specific binders for numerous different antigens (Hust et al., 2008).

In the following paragraphs the workflow of a typical phage display experiment, the methods used in phage elution and the vectors most commonly used in phage display experiments are discussed. Because some background knowledge of phage biology is required to fully understand the different methods used in phage display, the structure and lifecycle of filamentous phages are discussed first.

2.3.2.2 The structure and life cycle of M13 bacteriophages

The Ff-class filamentous bacteriophages are ssDNA viruses that infect a variety of Gram-negative bacteria by using bacterial pili as receptors (Azzazy & Highsmith, 2002). The phages of this class do not lyse their host cells, although the constant production of new phage particles usually substantially increases the generation time of the cells infected by these viruses (Barbas et al., 2001). This is possible, as unlike most bacteriophages, the Ff-class phages are continuously secreted across the bacterial membranes as they are assembled (Russel et al., 1997).

The genomes of the characterized Ff-class bacteriophages are circular, approximately 6400 nucleotides in length and contain 11 different genes (Barbas et al., 2001). Apart from three proteins needed in the replication of the phage genome, all of the proteins encoded by these genes are integral membrane proteins (Russel, 1995). Five of the genes encode capsid proteins, three encode proteins needed in genome replication and the rest encode proteins that are important in phage assembly and secretion (Russel et al., 1997). A complete M13 phage virion is a cylinder-shaped protein assembly with a diameter of 65 Å and length of 9000 Å (Russel et al., 1997). Each of these phage virions contains one single-stranded copy of the circular DNA genome. The most important structural component of the phage capsid is the p8 protein, also known as the major coat protein. In addition to approximately 2700 copies of the p8 protein, the capsid of M13 bacteriophage also contains 5 copies of each of the five minor coat proteins (See Figure 8 for a schematic presentation) (Barbas et al., 2001). Apart from acting as structural components, these proteins are also required for host cell infection and phage capsid assembly (Barbas et al., 2001). For example, the p3 proteins are needed in the recognition of F-pili during the host cell infection, while the p7 and the p9 proteins are important in the assembly and secretion of the nascent phage particles (Marvin, 1998).

Figure 8. A schematic presentation of the cylindrical structure of the M13 bacteriophages. The capsid proteins (p3 and p6) that recognize host cell F-pili are shown on the right. The assembly of a phage particle is initiated when the capsid proteins at the opposite end (p7 and p9) interact with a DNA packing signal. Multimerization of the p8 proteins proceeds from left to right during the phage assembly. Figure adapted from Barbas et al., 2001.

The infection process of the M13 bacteriophages is initiated by the recognition of a host cell F-pilus by the p3 proteins of the phage capsid (Azzazy & Highsmith, 2002). After the phage has bound to the F-pilus, the pilus retracts, drawing the phage to a close contact with the host cell periplasm (Barbas et al., 2001). Only the circular genome of the phage is translocated into the host cytoplasm, while most of the coat proteins are disassembled back to monomers that remain on the inner cell membrane (Barbas et al., 2001). After entering a cell, the circular phage genome is first converted

into double-stranded form known as the parental replicative form, or RF, which is then replicated by the host cell DNA replication machinery (Azzazy & Highsmith, 2002). The replication of the RF DNA proceeds through a mechanism called rolling-circle replication, which produces ssDNA molecules that serve as templates for the synthesis of additional double-stranded phage genomes (Marvin, 1998). However, when the concentration of the p5 proteins reaches critical level, the p5 molecules co-operatively bind to the single-stranded phage genomes and prevent their conversion to RF DNA (Barbas et al., 2001). Eventually, the single-stranded phage genomes are almost completely covered by the p5 proteins. Together the phage genomes and the p5 proteins form long, helical structures that are about 8000 Å long and 80 Å in diameter (Marvin, 1998).

The assembly of the phage particles is initiated when a DNA segment called packing signal is recognized by the p1, p7 and p9 phage proteins (Russel et al., 1997). This process takes place at sites at which the inner and outer membranes of the host cell are brought into a close contact by membrane-spanning protein complexes (Barbas et al., 2001). During the phage assembly, the p5 proteins covering the phage genome are replaced by the p8 proteins, and the nascent phage particle is extruded out of the host cell through channels formed by protein complexes (Russel, 1995). This process is dependent on multiple proteins, some of which are provided by the host cell. After all p5 proteins have been replaced by the p8 proteins and the phage genome is almost completely extruded out of the cell, the assembly is terminated by the addition of the p3 and the p6 proteins to the cytosolic end of the phage particle (See Figure 8) (Barbas et al., 2001). After the addition of the p3 and p6 proteins, the complete phage particles are released from the host cell. Because this elegant mechanism allows the secretion of large phage particles without disrupting the host cell membranes and killing the cell, new phage particles are produced continuously even after the release of the first phages (Russel, 1995). It has been estimated that during the first phage generation, approximately 1000 new phages are released, after which 100-200 phages are produced per host cell generation (Barbas et al., 2001).

2.3.2.3 The workflow of typical phage display experiments

The workflow of a typical phage display experiment is presented in the Figure 6. This process, known as biopanning, consist of multiple sequential steps, in which the phages displaying recombinant proteins are produced in host cells, selected according to the ligand-binding affinities of the displayed proteins and amplified by infection of susceptible bacterial cells (Azzazy & Highsmith, 2002). These steps are usually repeated 2 - 5 times to enrich the phages that are displaying proteins with desired properties (Hust et al., 2008). The phages that are enriched in phage panning are determined by the antigen used and the conditions applied during the phage panning experiment. Because the best scFv- and Fab-libraries may contain up to 10^{11} unique DNA sequences, it is possible to isolate antibodies against multiple antigens from a single library (Sidhu, 2001).

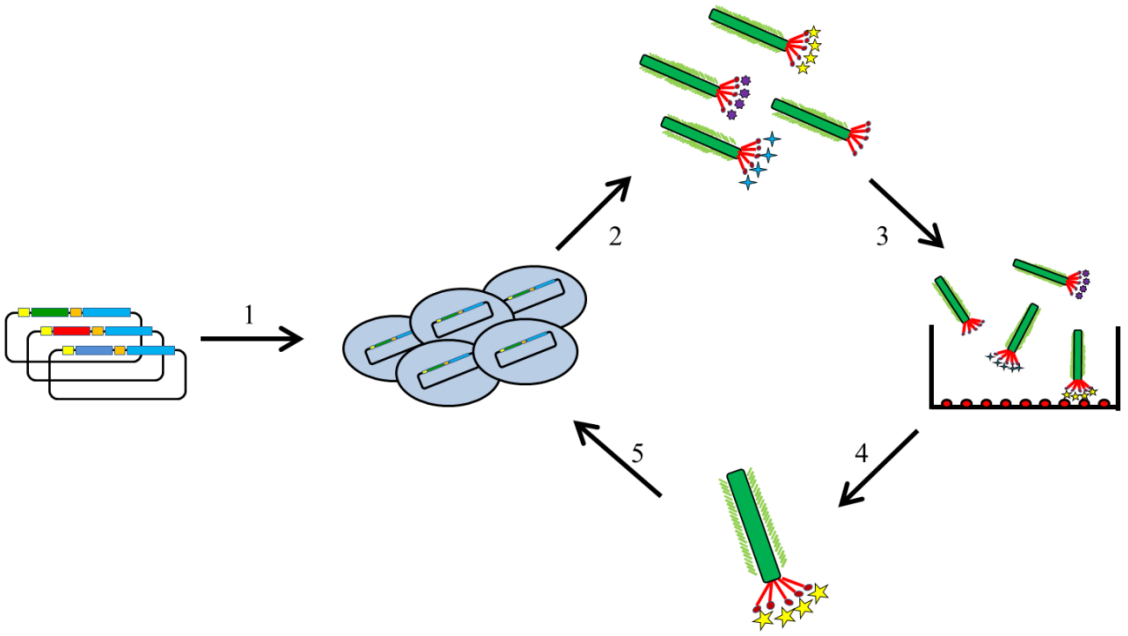


Figure 6. A simplified illustration of the phage display process. (1) Phagemid vectors are transformed into suitable host cells. These vectors contain chimeric genes encoding fusion proteins between a phage capsid protein and the members of the screened library. (2) The production of phage particles is initiated when the same cells are infected with helper phages (Not shown). These helper phages contain genes encoding proteins that are required in the replication of the phagemid genome. (3) The phages displaying the studied proteins are selected according to their antigen-binding affinities. The antigens used in this step are usually immobilized to a solid surface, such as microtiter plate wells or magnetic beads. (4) After washing off the unbound phages, the antigen-bound phages are eluted to improve their infectivity. (5) The eluted phages are re-infected into fresh host cells for amplification. If further phage enrichment is desired, these cells can be used in the production of new phages, after which the phage panning process is repeated.

2.3.2.4 Phage elution strategies

The affinity selection of the phages is one of the most important steps of the phage panning process. During this step, the phages displaying the studied proteins, for example Fab or scFv fragments, are incubated with an immobilized antigen of interest. The phages displaying proteins that have a high affinity for the antigen are bound to the solid surface, while unbound phages are removed by washing the surface with a suitable buffer (Barbas et al., 2001). The phages that remain bound to the surface after washing are eluted to restore their infectivity and to allow their amplification. Many different methods for the elution of these phages exist. The simplest and the most commonly applied elution method is altering the pH of the solution by the addition of hydrochloric acid or triethylamine. However, this method may favor the elution of only weakly bound phages (Kobayashi et al., 2005). This results in recovery bias, which prevents the enrichment of the phages that display proteins with a high affinities for the antigen used (Thomas & Smith, 2010).

Several different dissociation-independent elution methods have been proposed to overcome the problems caused by selective phage elution and recovery bias (See Figure 7). Firstly, inserting a protease-cleavable linker between the displayed protein and the phage capsid protein makes it possible to release the surface-bound phages by a protease treatment (Ward et al., 1996). While the phage particles themselves can resist this treatment relatively well, the linkers between the displayed proteins and the phage capsid proteins are cleaved efficiently during this treatment, which results in non-selective elution of the phages (Thomas & Smith, 2010). Moreover, cleavage of the displayed proteins may in some cases increase the infectivity of the eluted phages, as the displayed proteins may impair the functionality of the phage capsid proteins (Thomas & Smith, 2010). Similarly, if the antigens are immobilized via a DNA linker or a linker containing disulphide bonds, the immobilized antigens and the phages bound to them can be released by treating the sample with a suitable DNase or with dithiothreitol (DTT), respectively (Santala & Saviranta, 2004; Kobayashi et al., 2005). These elution strategies are not dependent on the amino acid sequence of the displayed proteins, which simplifies library design and makes these strategies suitable for the screening of already existing libraries (Kobayashi et al., 2005). The use of these dissociation-independent elution methods has already been shown to be an efficient way

to reduce recovery bias when screening for antibodies with extremely high ligand-binding affinities (Boder et al., 2000; Kobayashi et al., 2005).

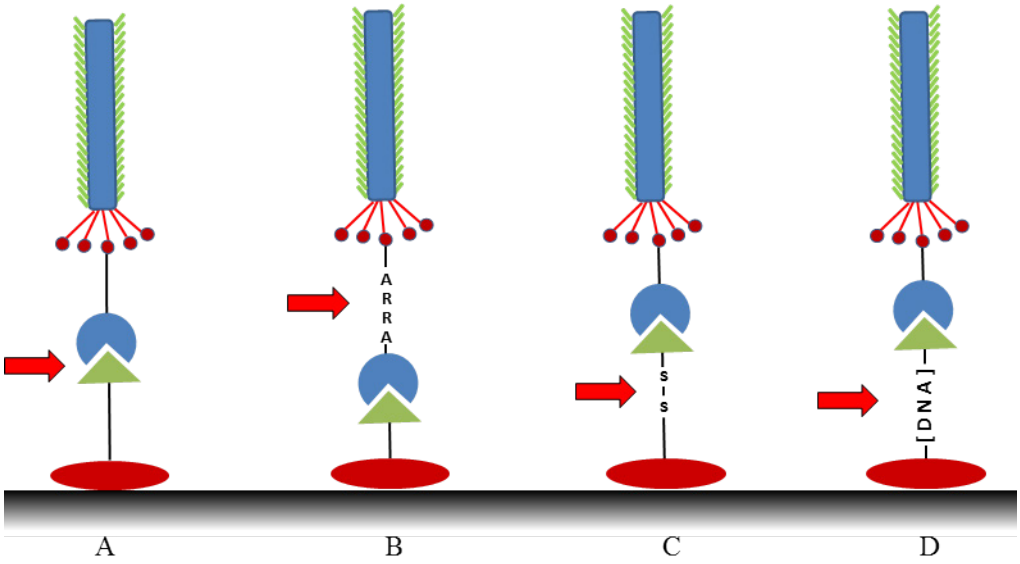


Figure 7. A schematic presentation of the different elution strategies used in phage display. The proteins displayed on the phage capsid are presented with blue three-quarter circles, the respective antigens are presented with green triangles and the carrier proteins or protein complexes used in antigen immobilization are presented with red ellipses. The sites at which the conjugates are cleaved in each case are indicated with red arrows. (A) The conventional elution method. In this method the antigen-bound phages are eluted by disrupting the antibody-antigen interactions with either low or high pH or chaotropic agents. (B) Phage elution by proteolytic cleavage. If a protease-susceptible linker sequence (Ala-Arg-Arg-Ala) is inserted between the displayed proteins and their fusion partners in the phage capsid, the surface-bound phages can be released by treating the sample with trypsin, for example. (C) and (D) Phage elution by DTT reduction or DNase treatment. If the antigens are immobilized with linkers containing either disulphide bonds (-S-S-) or specific DNA sequences, the antigen-bound phages can be eluted by treating the sample with reductive agents or with a suitable DNase, respectively.

2.3.2.5 Phage display vectors and fusion protein

Two different strategies have been used in the expression of the fusion proteins in phage display: complete phage genomes and phagemid vectors (Pasche, 2006). When complete phage genomes are used, the fusion genes consisting of the genes encoding the displayed protein and their fusion partners are inserted directly into the phage genomes (Smith & Petrenko, 1997). This strategy is simple, but it suffers from some drawbacks. For example, this strategy does not allow strict regulation of the phage production, which prevents long-term storage of the phage-infected host cells. Moreover, many of the phage display vectors based on wild-type phage genomes are genetically instable, and hence the genes encoding the displayed proteins are frequently lost during phage panning (Barbas et al., 2001; Hust, 2008).

To overcome these issues, various phagemid vectors have been developed. As their name indicates, phagemids are a distinct class of plasmids. A typical phagemid vector contains both bacterial and phage-derived *ori*-sequences for replication, an antibiotic resistance gene for selection, a packing signal for phage assembly and a multiple cloning site for cloning the genes encoding the displayed proteins (Smith & Petrenko, 1997). However, these vectors do not contain the genes encoding the proteins needed in phage replication and assembly (Azzazy & Highsmith, 2002). Hence, phagemid vectors can be propagated in bacterial cells in a double-stranded form like any other plasmid. The assembly of the phage particles is initiated when the cells harboring phagemid vectors are infected with helper phages. These bacteriophages contain the genes encoding the proteins needed in phage replication, assembly and extrusion, but mutations in their packing-signal effectively prevent the production of complete phage particles containing helper phage genomes (Barbas et al., 2001). Hence, the proteins encoded by the helper phage genome act in *trans* on the phagemid vectors, which results in the production of phage particles containing the respective phagemid vectors in a single-stranded form. This strategy has many advantages over the use of complete phage genomes, and thus it has largely superseded the other phage display strategies (Azzazy & Highsmith, 2002).

The fusion partners and the valency of the displayed proteins are yet another important aspect to consider when setting up a phage display experiment. In principle, foreign proteins can be fused with any of the five coat proteins of a typical filamentous phage (Barbas et al., 2001). However, the members of the screened library are usually expressed as fusion proteins with the p3 protein of the phage capsid, although also the major coat protein p8 has been used as the fusion partner when libraries consisting of short peptides have been screened (Hust, 2008). The displayed foreign peptides or proteins often interfere with the biological functions of their fusion partners, which is why only some of the phage capsid proteins are commonly used in phage display (Barbas et al., 2001). In addition, the effects of the displayed proteins are also dependent on the design of the fusion gene and on the valency of the displayed protein (Smith & Petrenko, 1997). Multivalent display (multiple proteins displayed per phage) of the studied proteins may in some cases reduce the number of produced phage virions or impair their infectivity. On the contrary, in monovalent phage display (one protein displayed per phage), most of the phage capsid proteins are expressed in wild-type

form, which effectively minimizes the adverse effects caused by the displayed proteins (Barbas et al., 2001). Moreover, monovalent display can be used to avoid avidity effects when the enrichment of proteins with high ligand-binding affinities is desired (Smith & Petrenko, 1997).

3. AIMS OF THE STUDY

The main aim of this study was to optimize the methods used in phage display to allow efficient and specific selection of novel DNA-shuffled avidin proteins with high biotin-binding affinities. Various phage display protocols were evaluated in order to ensure non-selective elution of the phages and to minimize non-specific binding of the phage particles. This study was divided into three parts, the aims of which were:

1. To evaluate the use of disulphide-containing linkers in the immobilization of biotin in phage display.
2. To investigate different methods for preventing non-specific binding of the phage particles.
3. To develop a new microplate assay for the functional analysis of the DNA-shuffled avidin mutants.

4. MATERIALS AND METHODS

4.1 DNA-shuffling and cloning of the mutant libraries

The phagemid libraries used in this study were constructed by Barbara Niederhauser. The parental genes and the methods used in DNA-shuffling, SLIC-cloning and electroporation are described in detail in Niederhauser et al. (2012). Briefly, the parental cDNA-sequences encoding chicken avidin, AVR2 and BBP-A were amplified in PCR reactions with gene specific primers containing suitable restriction sites as 5' overhangs. The PCR-products were gel purified and digested in pairs with the other parental sequences (AVD + AVR2, AVD + BBP-A and AVR2 + BBP-A) with DNaseI. The resulting fragments were then reassembled in a primerless PCR-reaction and amplified by using the same primers as in the amplification of the parental cDNAs. The reassembled genes were SLIC-cloned into a phagemid vector based on the VTT Fab vector (VTT Biotechnology, Espoo) and transformed into electrocompetent XL1-Blue cells (Stratagene) XL1-Blue is an amber suppressive *E. coli* cell line, which also contains the F' episome needed in the infection process of filamentous phages. The bacterial clones containing the phagemid vectors were selected by plating the transformation reactions on LB_{amp} plates (50 µg/ml ampicillin). The positive clones were collected by flooding the plates with sterile glycerol and the resulting mutant libraries were stored in -70 ºC freezer as glycerol stocks.

4.2 Production and precipitation of the phages

The mutant libraries were propagated by inoculating 10 ml of SB medium containing 10 µg/ml tetracycline, 50 µg/ml ampicillin and 0.1% (v/v) glucose with 50 µl of the library glycerol stocks. The cells were grown overnight at 28 °C with constant shaking (225 rpm). The libraries were diluted by transferring 200 µl of the overnight cultures to 10 ml of fresh SB medium containing 10 µg/ml tetracycline and 50 µg/ml ampicillin. The diluted cultures were incubated at 37 °C with constant shaking (225 rpm) until OD_{600} (Optical density at 600 nm wavelength) of 0.8 was reached. The cells were superinfected with 1 ml of helper phage (VSC-M13, 10^{11} cfu/ml, Stratagene) for 30 minutes at 37 °C without shaking to initiate the production of complete phage particles. After the helper phage infection, the cultures were diluted to 100 ml with pre-warmed

SB medium containing 10 µg/ml tetracycline and 50 µg/ml ampicillin. The cells were cultured for one hour at 37 °C and for another hour at 30 °C and 225 rpm. Kanamycin was added to a final concentration of 70 µg/ml and the cells were cultured overnight at 28 °C with constant shaking (225 rpm).

The phages produced to the culture medium were purified by polyethylene glycol (PEG-6000) precipitation. The bacterial cells were removed by a 15-minute centrifugation at 4000 g and 4 °C, and 25 ml of ice-cold 20% PEG-6000 in 2.5 M NaCl solution was added to the supernatant. The precipitation reactions were incubated on ice for 30 minutes, after which the precipitated phages were pelleted by centrifugation (13,200 g for 20 minutes at 4 °C). The precipitate was resuspended into 2 ml of PBS, transferred to a microcentrifuge tube and centrifuged for 10 min at 16,000 g and 4 °C. The supernatant was divided to two new microcentrifuge tubes and 250 µl of 20% PEG-6000 in 2.5 M NaCl solution was added to each tube. The phages were precipitated for 30 minutes on ice and collected by centrifugation at 16,000 g for 10 minutes. The precipitates were resuspended into 1 ml of PBS and stored at 4 °C.

4.3 Phage panning

In total seven different phage panning protocols were evaluated to find the optimal phage panning methods for the phage display screening of the DNA-shuffled avidin mutants described above. These phage panning protocols were named with either numbers (Protocols 1 and 2) or letters (Protocols A, B, C, D and E) depending on the step at which they differ from the protocols previously used in the phage panning of the same libraries (described in detail in Niederhauser et al., 2012). The phage panning protocol 1 was used as the reference protocol. Apart from the differences in the elution step (elution with dithiothreitol), this protocol is similar to the phage display protocol previously used by Niederhauser et al. (2012).

4.3.1 Phage panning protocol 1

Both bovine serum albumin (BSA) and milk casein were used as carrier proteins to immobilize biotin molecules to the surface of microtiter plates. These BSA-BTN and casein-BTN conjugates are based on the amine-reactive Sulfo-NHS-SS-BTN reagent

(Thermo Fisher Scientific) and they contain a DTT-cleavable disulphide bond. The BSA-BTN conjugates were used at panning rounds 1 and 3, while the casein-BTN conjugates were used on the second panning round. Both conjugates contained on average 4 biotin molecules per one protein molecule. These BSA-BTN and casein-BTN conjugates were produced and tested by Barbara Niederhauser according to the manufacturer's instructions.

Maxisorp Immuno F96 Microwell plates (Nunc A/S) were coated with either 200 ng of BSA or 200 ng of biotinylated BSA (BSA-BTN) overnight at 4 °C. As mentioned above, on the second panning round casein and biotinylated casein (casein-BTN) were used instead of BSA and BSA-BTN. The wells were washed 3 times with 300 µl of PBS and blocked with 300 µl of 1% (w/v) milk in PBS for one hour at room temperature (RT). The wells were washed again 3 times with 300 µl of PBS. To minimize the non-specific background binding of the phages during the phage panning, 100 µl of purified phages in PBS were first incubated in the BSA coated wells for 30 minutes. The phages were then transferred to the BSA-BTN coated wells and incubated for 60 minutes at RT with constant shaking (700 rpm). Unbound phages were removed by washing the plate 5 times with both 0.05% (v/v) PBS-Tween (PBST) and PBS, followed by a 15-minute wash with 300 µl of PBS at 37 °C and shaking at 700 rpm. The phages were eluted by adding 100 µl of 50 mM DTT in PBS and incubating the plate for one hour at 37 °C and shaking at 700 rpm. The eluted phages were transferred to 3 ml of fresh XL1-Blue cells at OD_{600} = 0.8 and the cells were incubated for 30 minutes at 37 °C for infection. The bacterial culture was diluted with 7 ml of SB medium containing tetracycline, ampicillin and glucose so that the final concentrations were 10 µg/ml for tetracycline, 10 µg/ml for ampicillin and 0.1% (v/v) for glucose. The cells were incubated for 30 minutes at 37 °C with constant shaking (225 rpm). The concentration of ampicillin was increased to 50 µg/ml and the cells were incubated for one hour at 37 °C with constant shaking (225 rpm). The cells were superinfected with 1 ml of VSC-M13 helper phage (Stratagene, 10^{11} cfu/ml) and incubated for 30 minutes at 37 °C without shaking. After the superinfection, the bacterial cultures were diluted to 100 ml with SB medium containing 10 µg/ml tetracycline and 50 µg/ml ampicillin and incubated for 15 minutes at 37 °C with shaking at 225 rpm. The temperature was lowered to 30 °C and the cells were incubated for one hour and 15 minutes with constant shaking (225 rpm). Kanamycin was added to a final concentration of 70 µg/ml and the cells were incubated overnight at

28 °C with constant shaking (225 rpm). The produced phages were precipitated by PEG-6000 precipitation as described above. The protocols used on the second and third panning rounds were mostly similar. On the second panning round casein was used as the biotin-carrier molecule and on the third round the number of washes after phage binding was increased from 5 to 10.

4.3.2 Phage panning protocol 2

The protocol 2 was mainly identical to the protocol 1 described above. However, in this protocol the phages were produced in the presence of glycosylated wild-type avidin (Belovo). Avidin was added at the same time with the helper phages to bind endogenous biotin present in the culture medium. The final concentration of avidin during the phage production was 100 µg/ml. Apart from the addition of avidin, the phage panning protocol 2 was identical with the protocol 1, and the libraries panned according to these two protocols were handled in parallel.

4.3.3 Phage panning protocols A, B, C, D and E

The phage panning protocols A - E were developed to study the effects of different sample treatments and additives on the non-ligand-specific background binding of the phages. Protocols A - E differ from the protocol 1 in the composition of the phage samples, phage pre-incubations and washing steps. However, many of the other steps of the phage panning process were identical to those of the protocol 1. The AVD/AVR2 library was panned in parallel according to each of these protocols for four phage panning rounds, after which randomly selected clones were analyzed.

In the protocols A - E, the precipitated phages were diluted 1:10 on the first panning round and 1:2 on the second, third and fourth panning round. To minimize the non-ligand-specific binding of the phages, all phage samples were diluted in either 1% BSA (the first and the third round) or 1% casein (the second and the fourth round) in PBS. The diluted phage samples were incubated overnight at 4 °C to allow the BSA- or casein-binding phages to bind their respective ligands. In addition, 5% of sucrose or 10% of glycerol were added in the protocols D and E, respectively, to reduce the nonspecific plastic binding of the phages. The compositions of the phage samples in

each case are shown in the Table 1 together with the dilution factors used in phage panning.

Table 1. The dilution factors and the compositions of the phage samples

Panning protocol	Phage dilution factor on the first round	Phage dilution factors on the other rounds	Composition of the diluted phage samples
Protocol 1	–	–	(Phages not diluted)
Protocol 2	–	–	(Phages not diluted)
Protocol A	1:10	1:2	1% BSA/Casein
Protocol B	1:10	1:2	1% BSA/Casein
Protocol C	1:10	1:2	1% BSA/Casein
Protocol D	1:10	1:2	1% BSA/Casein + 5% Sucrose
Protocol E	1:10	1:2	1% BSA/Casein + 10% Glycerol

The phage samples panned according to the protocols 1 and 2 were not diluted before phage panning. On the contrary, the samples panned according to the protocols A – E were diluted 1:10 on the first panning round and 1:2 on all successive rounds. On the first and the third panning round the diluted samples contained a final concentration of 1% (w/v) BSA in PBS, while on the second and the fourth round 1% (w/v) casein was used instead. In addition to this, the phage samples panned according to the protocols D and E contained 5% (w/v) sucrose and 10% (v/v) glycerol, respectively.

As described above, Maxisorp Immuno F96 Microwell plates (Nunc A/S) were coated with 200 ng of DTT-cleavable BSA-BTN (or casein-BTN) conjugates, washed 3 times with PBS and blocked with 300 µl of 1% milk in PBS. However, because the phages panned according to the protocols A – E were pre-incubated with either 1% BSA or 1% casein before the affinity selection, the pre-incubation with immobilized BSA or casein was omitted. Hence, aliquots (100 µl) of the diluted phage samples were transferred directly to the wells coated with BSA-BTN (or casein-BTN on the second and the fourth panning round) and incubated for 90 minutes at room temperature with constant shaking (600 rpm). Unbound phages were removed by washing the plate 5 times with both 0.05% PBS-Tween and PBS. As before, the wells were also filled with PBS and incubated for 15 minutes at 37 °C and 600 rpm to ensure efficient elution of the non-specifically bound phages. In the protocol B, the effect of this additional PBS wash was

further increased by the addition of 2 µg of non-conjugated biotin. In the protocol C, 100 µl of 100 mM HCl was added instead of PBS. After this 15-minute incubation, the wells were washed 5 times with PBS and the phages were eluted by the addition of 100 µl of 50 mM DTT, followed by a one-hour incubation at 37 °C with vigorous shaking (700 rpm). The washes and the sample treatments in each case are presented in the Table 2. Host cell infection by the eluted phages and all subsequent steps were done as in the protocol 1 described above.

Table 2. The pre-incubations and washes used in each of the phage panning protocols

Panning protocol	Pre-incubation	Affinity selection	1st wash	Additional 15-minute wash at 37 °C	2nd wash	Elution, 60 min at 37 °C
Protocol 1	30 min, immobilized BSA/Casein	60 min	5x PBS-Tween 5x PBS	PBS	1x PBS	100 µl 50 mM DTT
Protocol 2	30 min, immobilized BSA/Casein	60 min	5x PBS-Tween 5x PBS	PBS	1x PBS	100 µl 50 mM DTT
Protocol A	Overnight, 1% BSA/Casein solution	90 min	5x PBS-Tween 5x PBS	PBS	5x PBS	100 µl 50 mM DTT
Protocol B	Overnight, 1% BSA/Casein solution	90 min	5x PBS-Tween 5x PBS	PBS + 2 µg biotin	5x PBS	100 µl 50 mM DTT
Protocol C	Overnight, 1% BSA/Casein solution	90 min	5x PBS-Tween 5x PBS	100 mM HCl	5x PBS	100 µl 50 mM DTT
Protocol D	Overnight, 1% BSA/Casein solution	90 min	5x PBS-Tween 5x PBS	PBS	5x PBS	100 µl 50 mM DTT
Protocol E	Overnight, 1% BSA/Casein solution	90 min	5x PBS-Tween 5x PBS	PBS	5x PBS	100 µl 50 mM DTT

In the protocols 1 and 2, the phage samples were pre-incubated with immobilized BSA/Casein for 30 minutes, while in the protocols A - E, the phages were incubated in 1% BSA/Casein solution overnight at 4°C. In the protocol B, 2 µg of biotin (final concentration of 6.7 µg/ml) was added during the 15-minute wash with PBS. In the protocol C, 100 µl of 100 mM HCl was used in this step.

4.4 Phage titers

The phage titers of the samples were determined before and after each panning round. The phage titers before phage panning (Input titer) were assessed by serially diluting the resuspended phages 10^7 and 10^8 –fold and infecting 900 µl of XL1-Blue cells at OD_{600} = 0.8 with 111 µl of the two phage dilutions. After a 15-minute incubation at 37 °C, 10 µl of the phage-infected cells were transferred to 100 µl of fresh SB medium,

followed by plating the samples on LB $_{amp+tet}$ plates. The plates were incubated at 37 °C overnight, after which the numbers of colonies on each plate were counted.

After each phage panning round, the phage titers (Output titer) were determined by transferring 10 µl of the infection reactions to 100 µl of SB medium containing 20 µg/ml ampicillin and 10 µg/ml tetracycline. These samples were serially diluted 10^2-10^5 fold, depending on the expected phage titer. For each sample, two different dilutions were plated on LB $_{amp+tet}$ plates. After an overnight-incubation at 37 °C, the numbers of colonies on these plates were counted. If the numbers of colonies on both plates were countable (less than 2000 colonies per plate), the plate with the higher number of colonies was used in the analysis.

4.5 Master plates

Before the production of the phage was initiated (see section 4.2), aliquots of each library were diluted 1:200 in fresh SB medium, plated on LB plates containing appropriate antibiotics and incubated overnight at 37 °C. In total 48 randomly selected colonies were picked from each of the three library plates. The selected colonies were inoculated into 100 µl of freezing medium (See Appendix 1) containing 1% (v/v) glucose, 10 µg/ml tetracycline and 50 µg/ml ampicillin. These samples were incubated on a deep microtiter plate (MegaBlock-96, Sarstedt) for six hours at 37 °C with constant shaking (600 rpm). After this incubation, 24 µl of a solution containing freezing medium and glycerol in 1:1 ratio was added to the wells to increase the total glycerol concentration to 15%. The plates were shaken at 600 rpm for 15 minutes to mix the solutions and they were stored at −70 °C.

Similarly, after each phage panning round, master plates were created to allow the analysis of the same mutants with different methods. The analyzed colonies were randomly picked from the plates used in the determination of the output phage titers. In total 24 - 48 colonies from each plate were picked, depending on the panning protocol used and on the number of the panning round. The colonies were transferred to master plates as described above and stored at −70 °C.

4.6 DNA Sequencing

4.6.1 Phagemid purification

The avidin encoding genes of the clones stored on the master plates were sequenced both before and after the phage panning experiments. Before DNA sequencing, the phagemid vectors harboring the avidin encoding genes were propagated in host cells and purified to allow efficient DNA sequencing. The phagemid vectors were propagated by inoculating 1.5 ml of fresh SB medium containing 10 µg/ml tetracycline and 100 µg/ml ampicillin with 10 µl of the samples stored on the master plates. These samples were incubated overnight at 37 °C with constant shaking (420 rpm). The cells were pelleted by centrifugation for 10 minutes at 1000 g and 4 °C, after which the supernatants were discarded. The phagemid vectors were purified in 96-well format by using NucleoSpin® Robot-96 (Macherey Nagel) plasmid purification kit and Genesis RSP-100 (Tecan) liquid-handling robot. The purified phagemid vectors were eluted in 100 µl of the elution buffer provided with the plasmid purification kit.

4.6.2 Sequencing PCR and ethanol precipitation

The genes encoding the DNA-shuffled avidin mutants were sequenced by using BigDye® Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) and ABI-3130xl genetic analyzer (Applied Biosystems). Each sequencing reaction contained 100-300 ng of purified phagemid DNA, 6 pmol of the sequencing primer, 1.7 µl of 5x sequencing buffer and 0.7 µl of ready-made BigDye enzyme mix (BigDye Terminator v3.1 RR-100, Applied Biosystems). The sequences of the sequencing primers used are listed in the Table 3. The volume of the reaction was adjusted to 10 µl by using sterile deionized water. The sequencing reactions were denatured at 96 °C for 1 minute, followed by 30 cycles of 10 s at 96 °C, 10 s at 50 °C and 4 min at 60 °C. The PCR products were ethanol precipitated by the addition of 25 µl of 96% ethanol and 1 µl of 3 M sodium acetate (pH 5.2). The solutions were mixed by gentle shaking and incubated for 10 minutes at room temperature. The samples were centrifuged at 6000 g for 30 minutes. The supernatants were discarded and the DNA pellets were washed with 100 µl of 70% ethanol. The PCR products were pelleted again by centrifugation at 6000 g for 10 minutes. The resulting pellets were dried at 37 °C and resuspended into 12 µl of

ice cold HiDi-formamide (Applied Biosystems). The resuspended samples were denatured by incubating them for 3 minutes at 95 °C and cooled on ice. After sequencing, the resulting chromatograms and nucleotide sequences were analyzed with various software including Chromas Lite V2.01 (Technelysium Pty Ltd) and DNAman 4.11 (Lynnon Biosoft). The raw nucleotide sequences were filtered and translated with scripts written in Python programming language by Sampo Kukkurainen.

Table 3. The primers used in phagemid sequencing

Primer	Sequence	Tm
Phagemid 886	5'-ATTATCACCGTCACCGACTTCAGC-3'	58.8 °C
RR_pIII_rev2	5'- CGGTCATAGCCCCCTTATTAGC -3'	57.5 °C

The sequences of the primers used in phagemid sequencing are presented in the table. The DNA melting temperatures (Tm) calculated with the nearest-neighbor method are shown on the right hand side column.

4.7 Functional analyses

Two different methods were used to analyze the biotin-binding properties of the avidin mutants that were enriched in phage panning; phage-ELISA and an avidin-biotin displacement assay (ABD-assay). In phage-ELISA, the studied mutant proteins are produced as p3-fusion proteins and displayed on the surface of complete phage particles. The phages that bind to the immobilized antigens are detected with HRP-conjugated p8-antibodies. On the contrary, in the ABD-assay, the studied avidin mutants are not displayed on the surface of phage particles, but produced as soluble tetramers. However, as the studied proteins are produced in an amber suppressor cell line, some of the avidin mutants are produced as p3-fusion proteins. In this assay, the ligand-bound proteins are detected with biotinylated alkaline phosphatase (AP-BTN). The mutants that were enriched when protocols 1 and 2 were used were analyzed with both methods, while the mutants panned according to the protocols A – E were analyzed only with the ABD-assay.

4.7.1 Phage-ELISA

The phages were produced on a small scale in XL1-Blue *E. coli* cells (Stratagene) by inoculating 500 µl of SB medium containing 10 µg/ml tetracycline and 50 µg/ml ampicillin with 7 µl of the bacterial clones stored on the master plates. The cells were grown on a deep microtiter plate (MegaBlock-96 2.2 ml, Sarstedt) for six hours at 37 °C with constant shaking (500 rpm), after which 3 µl of helper phage (VSC-M13, 10^{11} cfu/ml, Stratagene) was added per well. The plates were incubated for two hours at 28 °C and 500 rpm. Kanamycin was added to a final concentration of 7 µg/ml and the plates were incubated overnight at 28 °C with constant shaking (500 rpm). The cells were pelleted by centrifugation at 1000 g for 10 minutes and the resulting supernatants were analyzed during the same day.

Maxisorp Immuno F96 Microwell plates (Nunc A/S) were coated overnight at 4 °C with 500 ng of either BSA or BSA-BTN. The plates were washed 3 times with PBS and blocked with 100 µl of 3% milk in PBS for one hour at room temperature. After blocking, 50 µl of the phage supernatants was added into the blocking solution, and the plates were incubated for one hour at 37 °C. The wells were washed 3 times with 0.05% PBS-Tween and 6 times with PBS. After the washes, 100 µl of HRP-conjugated anti-M13 antibody (GE Healthcare) diluted 1:5000 in 3% milk in PBS was added to each well and the plates were incubated for one hour at room temperature, shaking at 600 rpm. The plates were washed as before and 200 µl of ABTS substrate (Sigma Aldrich) containing 0.05% (v/v) H_2O_2 was added per well. The absorbances (A_{405}) of the wells were measured at 5-minute intervals for one hour using Model 680 XR microplate reader (Bio-Rad).

4.7.2 The avidin-biotin displacement assay

The studied avidin mutants were produced in XL1-Blue *E. coli* cells (Stratagene) by inoculating 200 µl of SB medium containing 10 µg/ml tetracycline, 100 µg/ml ampicillin and 0.1% (v/v) glucose with 7 µl of the bacterial clones stored on the master plates. The cells were grown on a deep microtiter plate (MegaBlock-96 2.2 ml, Sarstedt) for six hours at 37 °C, after which protein production was induced by the addition of 50 µl of SB medium containing 5 mM IPTG, 10 µg/ml tetracycline, 50 µg/ml ampicillin

and 0.1% (v/v) glucose. The final concentration of IPTG was 1 mM. The cells were grown overnight at 28 °C with constant shaking (600 rpm) and harvested by centrifugation at 1000 g for 10 minutes. The cell pellets were frozen at – 20 °C. The frozen cells were lysed with EasyLyse reagent (Epicentre Biotechnologies) immediately before the analysis. The cell pellets were resuspended into 40 µl of EasyLyse reagent by vigorous shaking for 10 minutes. The resuspended samples were diluted with 200 µl of 10 mM Tris-HCl buffer (pH 7.5). Cell debris was pelleted by centrifugation at 4500 rpm for 5 minutes at 4 °C.

Maxisorp Immuno F96 Microwell plates (Nunc A/S) were coated with either 200 ng of BSA or with a mixture of 50 ng of BSA and 150 ng of BSA-BTN. For each protein sample to be analyzed, one well was coated with BSA and two wells were coated with BSA + BSA-BTN. The wells were coated overnight at 4 °C and washed 3 times with PBS-Tween. The wells were blocked with 300 µl of 1% (w/v) milk in PBS-Tween for one hour at room temperature and washed 3 times with PBS-Tween. After the washing step, 70 µl of the cell lysate was added per well and the plates were shaken at 500 rpm for one hour at room temperature. Unbound proteins were washed off by washing the wells 3 times with PBS-Tween. In order to detect the bound proteins, 100 µl of biotinylated alkaline phosphatase (Vector Laboratories) diluted (1:5000) in 1% milk in PBS-Tween was added to each well. The plates were incubated for one hour at room temperature with constant shaking (500 rpm) and washed as before.

Before the addition of the phosphatase substrate (1 mg/ml PNPP in DEA buffer), some of the wells were treated with excess biotin to wash off the weakly bound avidin mutants. For each mutant, 185 µl of PBS was added to all three wells (one coated with BSA and two coated with BSA + BSA-BTN), followed by the addition of 12 µl of 170 µg/ml biotin in ITC buffer to one of the two BSA + BSA-BTN coated wells. The plates were incubated for 30 minutes at room temperature with shaking at 500 rpm. The wells were washed 3 times with PBS-Tween and 3 times with PBS. 100 µl of 1 mg/ml PNPP (Thermo Fisher Scientific) in DEA buffer was added to each well and the absorbance at 405 nm wavelength was measured at 10-minute intervals for one hour by using Model 680 XR microplate reader (Bio-rad).

Wild-type genes encoding avidin and BBP-A were used as control samples in the ABD-assay. Phagemid vectors containing the cDNA genes encoding avidin and BBP-A were transformed into XL1-Blue cells (Stratagene) and the encoded proteins were expressed in parallel with the DNA-shuffled mutants. In addition, purified samples of different biotin-binding proteins were used to evaluate the effects the biotin treatment has on purified avidin samples. Two of these samples contained purified egg-white avidin (5 µg/ml and 100 µg/ml) (Belovo) and the third sample consisted of chimeric avidin (5 µg/ml) (Hytönen et al., 2005 A) purified with affinity chromatography.

5. RESULTS

5.1 Sequence analysis of the DNA-shuffled mutants

To ensure that the diversity and the other properties of the DNA-shuffled libraries were similar to those of the same libraries in the previous study (Niederhauser et al., 2012), 35 - 40 randomly selected avidin mutants from each library were sequenced. Out of the three libraries analyzed, the AVD/AVR2 library had the best diversity, with over 95% (38 out of 40) of the sequenced mutants being unique (See Table 4). For the AVD/BBP-A and AVR2/BBP-A libraries, the percentages of unique sequences were 51% (18 out of 35) and 68% (27 out of 40), respectively. While almost all (over 99%) of the sequenced avidin mutants were DNA-shuffled, the average number of cross-overs varied between the libraries. The AVD/AVR2 library had the best diversity with an average of 3.1 crossovers per mutant. The AVD/BBP-A and the AVR2/BBP-A libraries had on average 1.1 and 1.4 crossovers per mutant, respectively. The number of crossovers was determined by using protein sequences only.

Table 4. The diversity of the avidin mutant libraries before phage panning

Library	No. of sequenced clones	No. of shuffled sequences	Percentage of unique sequences	Average number of cross-overs
AVD / AVR2	40	39*	95% (38 out of 40)	3.1
AVD / BBP-A	35	35	51% (18 out of 35)	1.1
AVR2 / BBP-A	40	40	68% (27 out of 40)	1.4

The numbers of sequenced and DNA-shuffled mutants and the percentages of unique sequences are presented in the table. All sequence analysis was done at the amino acid level, hence synonymous point mutations were not taken into account when the sequences were compared. The average numbers of cross-overs are presented on the right-hand column. Only one mutant did not contain any cross-overs (marked with an asterisk). However, because of the primers used in the construction of the libraries, this mutant is not identical to wild-type AVR2, but contains four C-terminal residues from avidin.

Similarly to the study described in Niederhauser et al. (2012), some of the sequenced avidin mutants contained DNA contaminations, point mutations, or both. The AVD/BBP-A library had the highest number of mutants containing contaminating sequences. Four out of the 35 sequenced avidin mutants of this library contained AVR2 sequences and two mutants had long 3' extensions containing parts of the VJL-region of

the gene encoding human immunoglobulin light chain *κ*. Similar extensions were also present in three mutants of the AVR2/BBP-A library. In addition, one mutant of the AVD/AVR2 library contained sequences characteristic of an avidin mutant known as sbAVD-1 (See Riihimäki et al., 2011).

5.2 The libraries panned according to the protocols 1 and 2

5.2.1 Phage titers during the panning process

During the phage panning process the phage titers of the libraries were analyzed before and after each panning round. Ratios between the output (sample of the eluted phages) and input (sample of the library before each phage panning round) titers were used to follow the enrichment of the ligand binding phages. The phage titers and the ratios between the output and input titers of each sample are presented in the Table 5. These ratios represent the fractions of the phages that were eluted after phage panning. For example, an output/input ratio of 1.0E-6 denotes that for every one million phages added during the phage panning, one infective phage is eluted in the elution step. The output/input ratios listed in the Table 5 are also presented graphically in the Figure 9. As apparent from the figure, the ratios between output and input titers increased during the panning process and they were highest on the last panning round.

The use of the phage panning protocol 2 resulted in almost ten-fold higher output/input-ratios when compared to those of the libraries panned according to the protocol 1. Although both input and output titers were higher on all panning rounds, the relative increase in input titers ranged from 1.2 to 7.7-fold, while many of the output titers were over 30-fold higher when compared to the phage titers of the libraries panned according to the protocol 1. As evident from the Figure 9, the difference in the phage titers between the libraries panned according to these two protocols was the most prominent on the first panning round and the smallest on the third panning round.

Table 5 A. The phage titers of the libraries panned according to the protocol 1

	Protocol 1								
	AVD/AVR2			AVD/BBP-A			AVR2/BBP-A		
Panning round	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input
1	2.0E+16	2.6E+07	1.3E-09	1.7E+16	4.7E+07	2.8E-09	1.2E+16	2.3E+07	1.8E-09
2	4.9E+15	4.4E+07	9.1E-09	9.6E+15	1.6E+07	1.7E-09	6.9E+15	2.4E+07	3.5E-09
3	5.7E+15	6.9E+08	1.2E-07	3.8E+15	2.9E+08	7.5E-08	6.6E+15	2.8E+07	4.2E-09

Table 5 B. The phage titers of the libraries panned according to the protocol 2

	Protocol 2								
	AVD/AVR2			AVD/BBP-A			AVR2/BBP-A		
Panning round	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input	Input titer (Cfu/ml)	Output titer (Cfu/ml)	Output/ Input
1	2.3E+16	8.3E+08	3.6E-08	2.1E+16	8.3E+08	4.0E-08	3.6E+16	6.9E+08	1.9E-08
2	3.0E+16	1.1E+09	3.7E-08	2.4E+16	6.5E+08	2.8E-08	4.8E+16	7.2E+08	1.5E-08
3	1.1E+16	1.5E+09	1.4E-07	3.0E+16	1.4E+09	4.6E-08	1.7E+16	9.6E+08	5.5E-08

The phage titers of the mutant libraries panned according to the protocols 1 and 2 are shown in the Tables 5 A and B, respectively. The input and output titers have been corrected to represent the absolute number of phages in each sample (i.e. the total volume of the sample is 1 ml), to allow the comparison of the output/input ratios between different phage panning experiments.

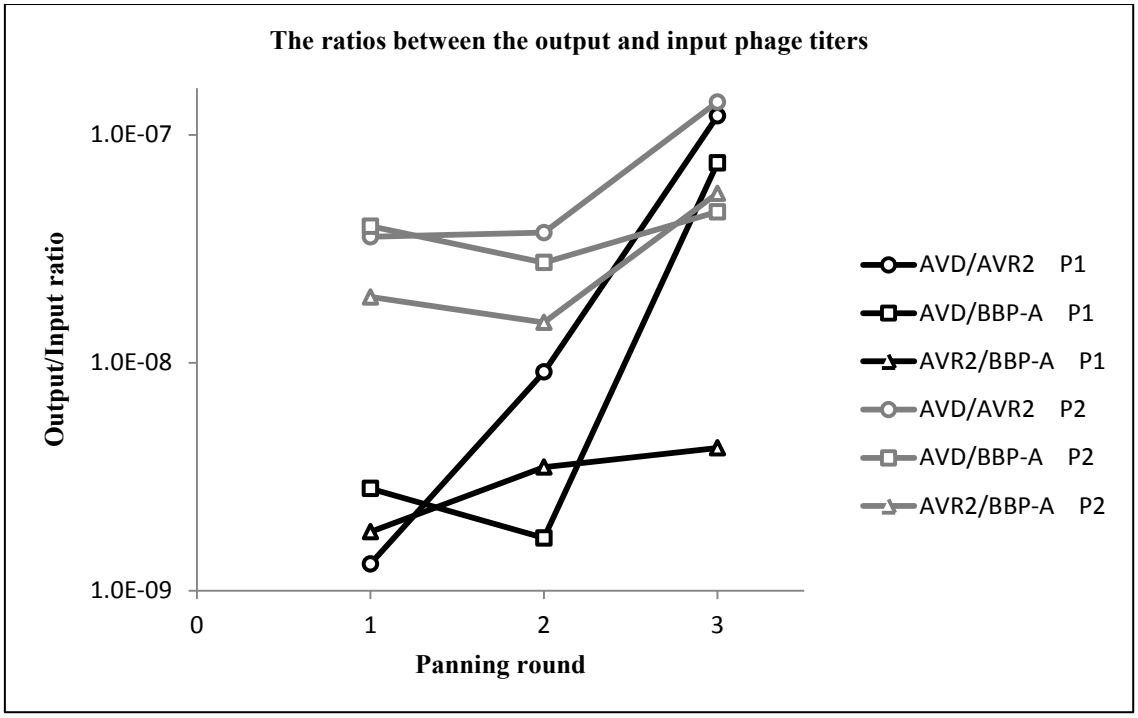


Figure 9. The recovery ratios for the libraries panned according to the protocols 1 and 2. These ratios represent the fractions of the phages that are eluted in the elution step. In the legend, P1 and P2 denote phage panning protocols 1 and 2, respectively.

5.2.2 Sequence analysis of the phage display enriched mutants

After three rounds of phage panning according to the protocols 1 and 2, the avidin mutants were analyzed by sequencing 32 clones of each library. The number of successfully sequenced clones varied between 19 and 29 depending on the library. In general, the clones of the libraries that were panned according to the protocol 2 were more likely to have sequencing errors than the clones panned according to the protocol 1. This difference in the sequence quality may be due to the very high confluency of the output titer plates of the libraries panned according to the protocol 2, and therefore it may reflect the properties of the DNA-shuffled genes themselves. On the other hand, it is possible that some of the sequences were more prone to sequencing errors than others, which may have led to a bias in the results. For example, the codons encoding amino acids Ser5 and Phe29 of avidin often had sequencing errors because of technical reasons, and therefore the sequences containing these codons were more likely to be discarded than sequences with different codons at these positions.

5.2.2.1 Libraries panned according to the protocol 1

The resulting non-redundant sequences of the mutants panned according to the protocol 1 are listed in the Appendix 2 A and the numbers of identical sequences are presented in the Table 6. The presence of multiple identical sequences indicates that some of the sequences have been enriched during the phage panning against immobilized biotin. Most of the sequenced mutants of the AVD/AVR2 and AVD/BBP-A libraries were identical with the other members of these libraries and only a few sequences were unique. The most highly enriched mutant of the AVD/AVR2 library, here called A/A2_P1_1, is identical to a mutant previously known as sbAVD-1 (described in detail in Riihimäki et al., 2011). This mutant alone covers 45% (13 out of 29) of the sequenced clones of the AVD/AVR2 library. Similarly, a mutant called A/B_P1_11 covers 79% (23 out of 29) of the sequenced clones of the AVD/BBP-A library. Unlike with the AVD/AVR2 and AVD/BBP-A libraries, phage panning of the AVR2/BBP-A library did not result in the enrichment of only one or two sequences. Instead, half (16 out of 32) of the sequenced mutants of this library were unique. The rest of the analyzed mutants belonged to five different groups, each of which consisted of 2 - 4 identical sequences. In total five of the analyzed genes of the AVR2/BBP-A library had long stretches of contaminating avidin sequences. However, none of the sequenced mutants of this library had deletions or sequences encoding immunoglobulin VJL regions.

Table 6. The avidin mutants analyzed after the phage panning

Library	Panning protocol	No. of sequenced mutants	No. of shuffled sequences	No. of unique sequences	Groups of identical sequences
AVD / AVR2	1	29	16	6	13*, 3, 3, 2, 2
AVD / BBP-A	1	29	29	4	23, 2
AVR2 / BBP-A	1	32	32	16	4, 4, 3, 3, 2
AVD / AVR2	2	19	19	19	-
AVD / BBP-A	2	18	18	2	11, 3, 2
AVR2 / BBP-A	2	24	24	9	7, 3, 3, 2

The numbers of unique sequences among the sequenced clones of each library are presented in the fifth column, followed by the numbers of identical mutants among the sequenced clones. The mutant group marked with an asterisk consisted of mutants previously known as sbAVD-1. This mutant is not DNA-shuffled and it contains DNA sequences not present in any of the parental genes.

If the genes containing DNA contaminations (For example sbAVD-1 sequences) are omitted, in total 54 out of the 90 sequenced mutant genes contained point mutations (See Appendix 2 A). Out of these point mutations, 87% (47 out of 54) affected amino acid residues that are known to directly participate in biotin-binding. Moreover, some of the mutated residues not directly participating in biotin-binding are close to the biotin-binding residues. Out of the 16 unique point mutations found, only two mutations (V26A and Q55R) are located outside the biotin-binding pocket. The locations of the point mutations that affect the biotin-binding residues of wild-type avidin are presented in the Figure 10.

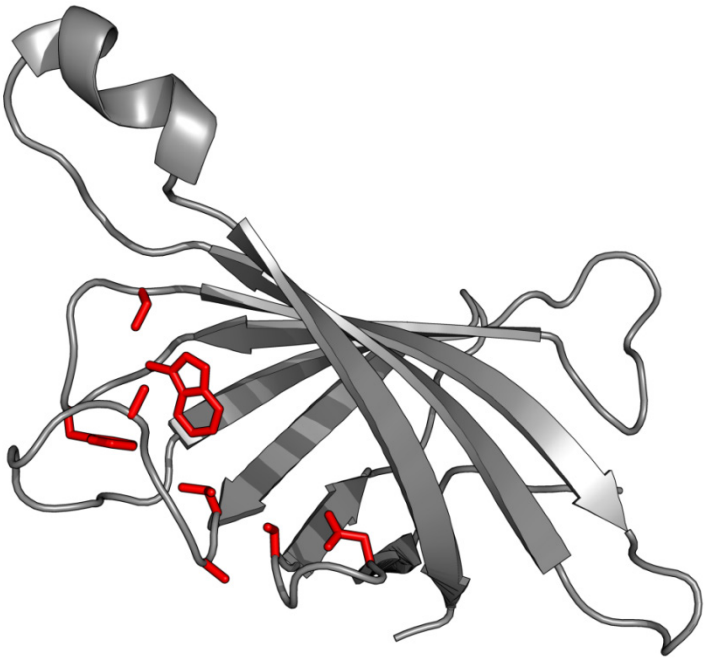


Figure 10. The point mutations present in the avidin mutants panned according to the protocol 1. The amino acid residues affected by point mutations are shown as sticks and colored in red. Only the mutations that affect residues present in wild-type avidin are shown (point mutations V26A, A75P and Q55R in BBP-A sequences and E72K in AVR2 sequences are not shown). PDB ID: 2AVI (Livnah et al., 1993).

5.2.2.2 Libraries panned according to the protocol 2

A similar sequence analysis was also performed for the avidin mutants that had been panned according to the phage panning protocol 2 (wild-type avidin added during phage production) (See Table 6 and Appendix 2 B). As expected, the mutants produced in the presence of wild-type avidin differed from the mutants panned according to the phage panning protocol 1. For example, all of the sequenced mutants of the AVD/AVR2

library were unique, and none of them was identical to the mutants that had been panned according to the protocol 1. Similarly, only 1 out of 18 and 3 out of 24 mutants of the AVD/BBP-A and AVR2/BBP-A libraries, respectively, were identical to the mutants panned according to the phage panning protocol 1. Interestingly, out of the 61 mutants sequenced, only 1 contained a point mutation (See Appendix 2 B). Moreover, this mutation (Q61K) did not affect any of the biotin-binding residues.

5.2.3. Functional analysis of the phage display enriched mutants

5.2.3.1 Analysis of the ligand binding properties by phage-ELISA

A Phage-ELISA experiment was conducted to evaluate the ligand-binding properties and production levels of the sequenced avidin mutants. In this assay the studied mutants were produced as p3-fusion proteins and displayed on the surface of complete phage particles. The absorbance readings of the BSA-coated wells were subtracted from the values measured for the BSA-BTN-coated wells to minimize the bias resulting from the background binding of the phages.

Table 7. The results of the phage-ELISA for the mutants panned according to the protocols 1 and 2

Name of the Library	Panning protocol	No. of mutants analyzed	No. of mutants with phage ELISA signals significantly above the background level	Average phage ELISA signals ± SD
AVD/AVD2	1	29	29 (100%)	0.251 (± 0.075)
AVD/BBP-A	1	29	28 (97%)	0.244 (± 0.079)
AVR2/BBP-A	1	32	27 (84%)	0.118 (± 0.094)
AVD/AVD2	2	19	4 (21%)	0.036 (± 0.089)
AVD/BBP-A	2	18	8 (44 %)	0.017 (± 0.012)
AVR2/BBP-A	2	24	4 (17 %)	0.003 (± 0.009)

The numbers and the percentages of the avidin mutants with phage-ELISA signals significantly above the background level (3 x SD of the background) are shown on the fourth column. The average phage-ELISA signals and the standard deviations of all analyzed mutants are presented on the right-hand column.

The number of absorbance signals significantly (3 times the standard deviation (SD) of the average background signal) above the background level and the average background-subtracted absorbance values are presented in the Table 7. Importantly, the

mutants panned according to the phage panning protocol 1 gave results strikingly different from those of the mutants panned according to the protocol 2. For the AVD/AVR2, AVD/BBP-A and AVR2/BBP-A libraries panned according to the protocol 1, the numbers of phage-ELISA signals significantly above the background level were 29 out of 29, 28 out of 29 and 27 out of 32, respectively. On the contrary, only a few of the mutants panned according to the protocol 2 gave signals above this threshold level.

5.2.3.2 Analysis of the dissociation kinetics by ABD-assay

To further increase the understanding of the phage panning process, an ABD-assay was performed with 41 mutants. These mutants were selected to cover a wide variety of avidin mutants with different sequences and they included both samples panned according to the panning protocol 1 and according to the protocol 2. To evaluate the reproducibility of the assay, duplicate samples were used in all steps of the process, including the bacterial culture and the protein expression steps. The signals measured from the BSA-BTN coated wells treated with excess biotin were compared to the signals measured from the untreated wells with similar coating. The ratio between these values was used to estimate the relative dissociation kinetics of the studied mutants. Because the ratios between biotin-treated and untreated wells were calculated for each mutant individually, the expression levels of the avidin mutants should not affect the results. Because of technical problems, we were unable to measure the BSA-binding signals for 18 mutants of the AVR2/BBP-A library panned according to the protocol 2. For these mutants, the BSA-binding signal was estimated by using the average signal of the other samples of this library.

The values measured from the BSA-coated wells were not subtracted from the values measured from the other wells. Instead, the average absorbance reading of empty wells (i.e. background) was subtracted from the raw absorbance readings of the analyzed avidin mutants. The number of mutants with absorbance readings significantly above the background level (absorbance reading > (average background reading + 3 x SD)) was counted. The numbers of such mutants and the average background-subtracted absorbance readings of all mutants presented are in the Table 8. The background-subtracted absorbance readings of all individual mutants are presented in the

Appendix 4. For two mutants, A2/B_P1_1 and A/A2_P1_5, the absorbance readings at the 60-minute timepoint exceeded the detection limit of the micro plate reader, and the actual results are likely to be greater than the values shown in the Appendix 4.

Table 8. The results of the ABD-assay for the mutants panned according to the protocols 1 and 2

Panning protocol	No. of samples	BSA coated wells	BSA-BTN coated wells	BSA-BTN coated wells + BTN treatment
1	n = 26	22 (85%) 0.106 (± 0.190)	23 (88%) 0.748 (± 1.049)	12 (46%) 0.076 (± 0.137)
2	n = 15	0 (0%) 0.000 (± 0.002)	4 (27%) 0.007 (± 0.016)	3 (20%) 0.003 (± 0. 007)

The numbers of absorbance readings significantly above the background level, their average values and the SDs are presented in the table. For example, out of the 26 analyzed mutants panned according to the protocol 1, 22 gave absorbance readings significantly above the background level when incubated on wells coated with BSA. The average absorbance reading and the standard deviation of the absorbance readings of all 26 mutants were 0.106 and 0.190, respectively.

To investigate the relative dissociation kinetics of the studied mutants, the ratio between the absorbance readings measured from the BSA-BTN-coated wells either treated or not treated with excess biotin was calculated (See Appendix 4). This ratio was calculated only for the mutants that showed absorbance readings significantly above the average background level. The average ratio of all DNA-shuffled mutants was 0.130 (± 0.147). This denotes that on average, the absorbance values measured from the biotin-treated wells were only 13% of the signal measured from the untreated wells, indicating a rapid dissociation of most of the studied DNA-shuffled avidin mutants.

5.3 The libraries panned according to the protocols A - E

5.3.1 Phage titers during the panning process

Five different phage panning protocols were tested to evaluate the effects different sample pre-treatments and additives have on the results of the phage panning experiment. Because of the high diversity of the AVD/AVR2 library, this library was used in these tests. During the panning process, the phage titers were determined with the methods described above. As before, the ratios between the output and input titers were used to follow the enrichment of the biotin-binding phages.

Table 9. The phage titers of the mutant libraries before and after each panning round

	Protocol A			Protocol B			Protocol C			Protocol D			Protocol E		
Panning round	Input titer Cfu/ml	Output titer Cfu/ml	Output/ Input	Input titer Cfu/ml	Output titer Cfu/ml	Output/ Input	Input titer Cfu/ml	Output titer Cfu/ml	Output/ Input	Input titer Cfu/ml	Output titer Cfu/ml	Output/ Input	Input titer Cfu/ml	Output titer Cfu/ml	Output/ Input
1	4.6E+15	5.7E+06	1.2E-09	4.6E+15	2.4E+06	5.3E-10	4.6E+15	1.3E+06	2.9E-10	4.6E+15	4.7E+06	1.0E-09	4.6E+15	3.8E+06	8.2E-10
2	6.6E+14	4.0E+06	6.0E-09	3.7E+14	8.0E+05	2.2E-09	1.2E+14	0.0E+00	–	1.2E+15	2.4E+06	2.0E-09	9.9E+14	8.3E+05	8.3E-10
3	2.9E+14	2.0E+07	5.7E-08	3.2E+13	7.9E+04	2.5E-09	1.5E+11	0.0E+00	–	2.7E+14	9.1E+06	3.3E-08	1.3E+14	4.4E+06	3.4E-08
4	1.6E+15	2.0E+08	1.2E-07	2.1E+13	1.9E+05	8.9E-09	2.7E+09	0.0E+00	–	2.3E+15	1.1E+08	4.6E-08	8.7E+14	2.1E+07	2.4E-08

The titers of the mutant libraries panned according to the protocols A - E are shown in the table. The input and output titers have been corrected to represent the absolute number of phages in each sample (i.e. the total volume of the sample is 1 ml), to allow the comparison of the output/input ratios between different phage panning experiments.

The phage titers and the output/input ratios are presented in the Table 9. The output/input ratios are also presented graphically in the Figure 11. As apparent from the Figure 11, all of the pre-elutions and additives tested decreased the output/input ratios when compared to those of the reference protocol. The decrease in the output/input-titers was mostly due to a decrease in the number of eluted phages (See table 9). The effects of added sucrose and glycerol (Protocols D and E) were smaller than those of the tested pre-elutions (Protocols B and C). In fact, the pre-elution with 100 mM HCl (Protocol C) was so harsh, that almost all of the phages panned according to this protocol were lost after the second panning round. Pre-elution with excess biotin had less striking effects, but the output/input ratios were still lower than those of the libraries panned according to the reference protocol. The ratios between the output and input titers increased during the panning process, which may indicate the enrichment of biotin-binding phages. The increase in the output/input titers was largest (100-fold) for the library panned according to the protocol A and smallest (16-fold) for the library panned according to the protocol B.

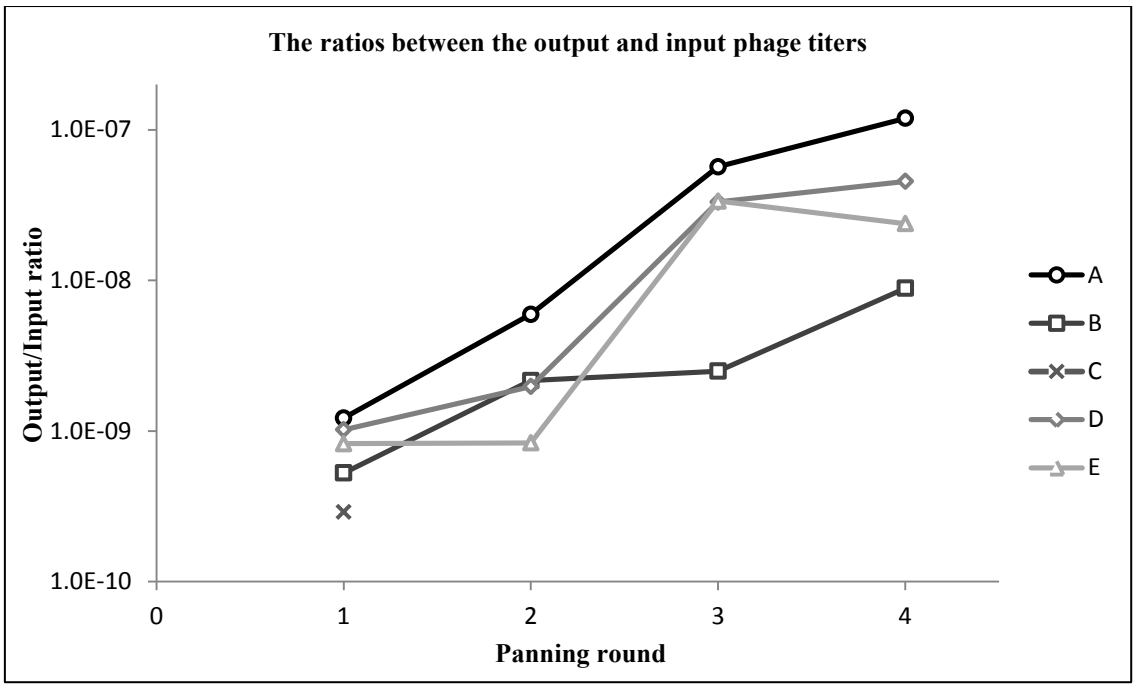


Figure 11. A graphical presentation of the output/input ratios presented in the Table 9. The values on the X-axis represent the number of the panning round and the values on the Y-axis represent the output/input ratios. Because of the loss of the phages panned according to the protocol C, the output/input ratio for this sample was determined only after the first phage panning round.

5.3.2 Functional analysis of the mutants

To evaluate whether the use of different phage panning protocols resulted in the enrichment of different mutants, 24 clones of each phage pool were analyzed. These clones were randomly picked from the LB $_{amp+tet}$ plates used in the determination of the output titers of the fourth panning round. Because the phages that were panned according to the protocol C were lost after the second panning round, only mutants panned according to the protocols A, B, D and E were analyzed. These mutants were analyzed with the ABD-assay as described above. The raw absorbance data consisted of A_{405} readings measured at the 60-minute time point. The average of the absorbance readings measured from empty microtiter plate wells (0.073) was subtracted from the absorbance readings measured from the BSA-BTN and BSA coated wells.

The numbers of the samples that gave absorbance readings significantly (> 3 x SD) above the background level and their average absorbance readings are presented in the Table 10. The average absorbance readings of all studied avidin mutants are presented in the Appendix 5 according to their amino acid sequences. As apparent from the

Table 10, panning with the protocol B resulted in the highest average absorbance reading in both BSA-BTN and BSA coated wells. Moreover, these mutants had the highest average absorbance readings after the biotin treatment. The phages panned according to the protocol A had the second highest absorbance readings in the BSA-BTN coated wells both before and after the biotin treatment. As previously, the ratios between the absorbance readings measured from the biotin-treated and untreated wells were calculated. These ratios are presented in the Appendix 5. In the case of multiple identical mutants, the average absorbance readings of all identical mutants were used to calculate the ratios.

Table 10. The results of the ABD-assay for the mutants panned according the protocols A - E

Panning protocol	No. of samples	BSA coated wells	BSA-BTN coated wells	BSA-BTN coated wells + BTN treatment
A	n = 24	2 (8%) 0.023 (± 0.008)	16 (67%) 0.154 (± 0.290)	4 (17%) 0.056 (± 0.059)
B	n = 24	9 (38%) 0.051 (± 0.024)	14 (58%) 0.819 (± 0.756)	11 (46%) 0.169 (± 0.087)
D	n = 24	3 (13%) 0.017 (±0.001)	16 (67%) 0.063 (± 0.044)	3 (13%) 0.018 (± 0.002)
E	n = 24	1 (4%) 0.014	10 (42%) 0.058 (± 0.046)	0 (0%) -

The numbers of absorbance readings significantly above the background level are presented in the table together with the average background-subtracted absorbance readings of these samples. Because the phages panned according to the protocol C were lost before the fourth panning round, none of these phages was analyzed with the ABD-assay.

5.3.3 Sequence analysis of the mutants

Based on the results of the ABD-assay, in total 20 mutants were selected for sequence analysis. The mutants with the highest absorbance readings in the BSA-BTN coated wells were selected to represent each of the four panning protocols (Protocols A, B, D and E). 18 out of 20 mutants were successfully sequenced. The list of non-redundant sequences and the numbers of identical sequences are presented in the Appendix 3. Over 61% (11 out of 18) of the sequenced mutants were identical to each other. One of these identical mutants had been panned according to the phage panning protocol A

(reference protocol) and all the others according to the protocol B (Pre-elution with excess biotin). Moreover, these mutants were identical to a mutant previously known as A/A1-3, which has been described in detail in Niederhauser et al. (2012). This suggests that especially the use of the protocol B favored the enrichment of this mutant. As previously presented by Niederhauser et al. (2012), this mutant contains a point mutation (S16Y) affecting one of the residues directly participating in biotin-binding.

Other identical avidin mutants belonged to two groups. Each of these groups consisted of two mutants panned according to the protocols B (pre-elution with excess biotin) or protocol D (Phages diluted in 10% sucrose solution). Interestingly, also the mutants of both of these groups contained missense mutations affecting residues involved in biotin-binding. The groups represented by the mutants called B4_15 and D4_7 contained T77I and T35A missense mutations, respectively. In addition, one of the unique sequences contained a nonsense mutation (TGG→TAG) truncating the protein after the amino acid N69. Because these avidin mutants were expressed in an amber suppressor cell line, some of the proteins may have been expressed as full-length protein products. In this case the TAG codon does not act as a stop codon. Instead, it encodes glutamine, and the possible full length products are likely to contain W70Q missense mutations.

5.4 Properties of the avidin-biotin displacement assay

One of the aims of this study was to develop and optimize an assay for functional screening of a large number of phage display selected avidin mutants. In order to accomplish this aim, the ABD-assay was developed. This assay was designed to minimize the effects protein expression levels have on the absorbance signals. This, in turn, allows the comparison of the biotin-binding properties of the mutants without a bias caused by unequal expression levels. As previously explained, ratios between the absorbances of the biotin-treated and untreated wells were calculated. Because these ratios are not dependent on the expression levels of the mutants, they can be used to compare the relative dissociation rate constants of avidin mutants with unequal expression levels.

5.4.1 Duration of the biotin treatment

The optimal duration of the treatment with excess biotin was determined by analyzing the effects treatments of different durations have on the absorbance readings of samples consisting of chicken egg-white avidin and an avidin mutant known as BN_A/A2_p1 (See Niederhauser et al., 2012). The results of this experiment are presented in the Figure 12. As apparent from the figure, a 15-minute incubation was long enough to decrease the absorbance signal of the BN A/A2 p1 sample down to background levels, while wild-type avidin did not show any changes even after a 60-minute incubation. Based on these results, the duration of the incubation time was set to 15 minutes, as a 15-minute incubation is likely to be enough for the dissociation of most of the DNA-shuffled avidin mutants with relatively low biotin-binding affinities.

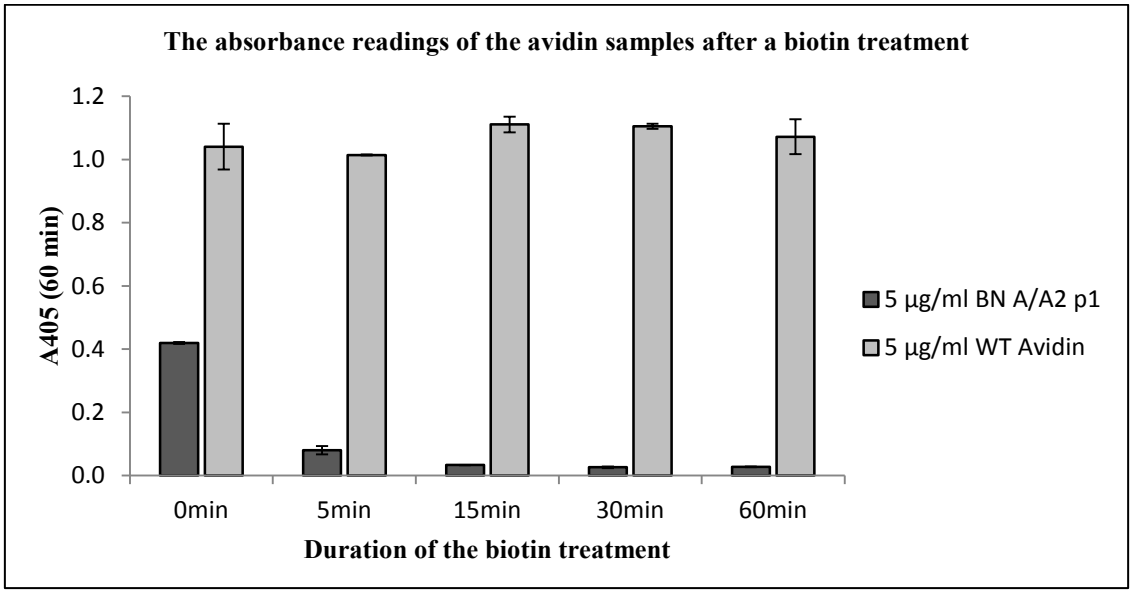
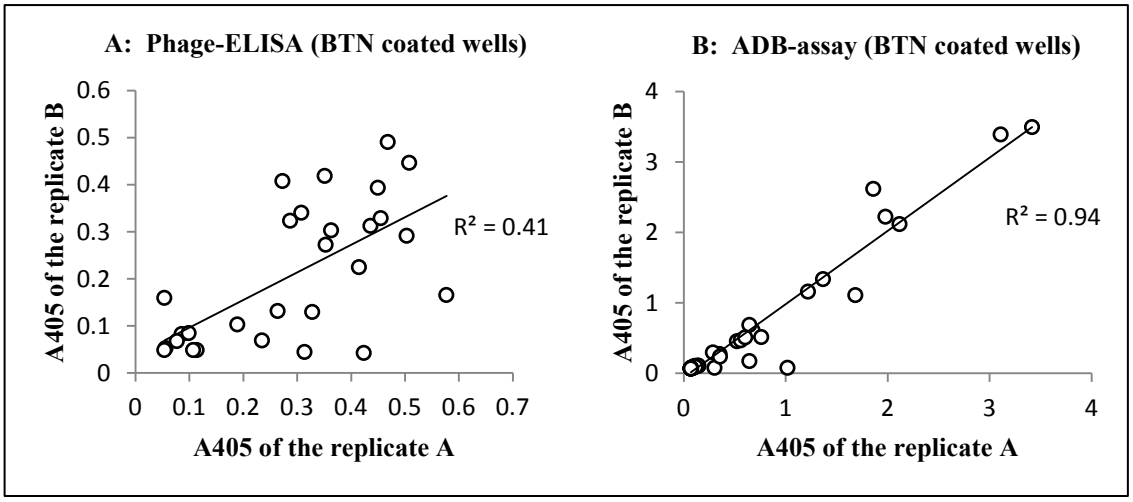


Figure 12. The effects of added biotin on the absorbance readings of the ABD-assay. The background-subtracted absorbance readings of BN A/A2 p1 and wild-type (WT) avidin samples are presented with dark grey and light gray bars, respectively. In each case, the duration of the biotin treatment is shown below the bars. Error bars represent the standard deviations of replicate samples.

5.4.2 Reproducibility of the assay

The reproducibility of the ABD-assay was expected to be higher than that of the assays based on the expression of complete phage particles (e.g. phage-ELISA discussed in Barbas et al., 2001). To evaluate the reproducibilities of the ABD-assay and phage-ELISA, the absorbance readings of the replicate samples of both assays were

plotted against each other. These plots are presented in the Figures 13 A and 13 B together with the respective linear regression lines. The coefficients of determination (R^2) for these linear regression lines were 0.936 (n = 48) and 0.413 (n = 28) for the ABD-assay and phage-ELISA, respectively. By comparing the coefficients of determination, it is apparent that the reproducibility of the ABD-assay was better than that of a conventional phage-ELISA experiment. In both assays, the replicate protein samples were also produced in parallel and therefore, some of the variance between the replicate samples may be due to differences in bacterial growth rates and protein expression levels.



Figures 13 A and 13 B. The reproducibilities of phage-ELISA (A) and ABD-assay (B). The background-subtracted absorbance readings of the replicate samples were plotted against each other (black circles) and linear regression lines (black line) were fitted to the values. The values of the coefficients of determination (R^2) for these lines are shown on the right.

6. DISCUSSION

6.1 The phage display technology in the selection of novel avidin mutants

During the past almost three decades, phage display has been proved to be an extremely efficient and versatile method for studying protein interactions and for identifying proteins with specific functions and properties (Azzazy & Highsmith, 2002). The most important application of this technology has been the selection of proteins, typically antibodies, with high affinity towards their ligands (Azzazy & Highsmith, 2002). The power of the phage display technology in antibody engineering has been demonstrated in many studies. For example, many clinically important recombinant antibodies have been identified by using phage display (reviewed by Thie et al., 2008 A). In addition to antibody engineering, phage display has been successfully used in a plethora of other applications. Importantly, proteins with high ligand-binding affinities have been successfully isolated from libraries based on non-antibody scaffolds. Among such alternative scaffolds are affibodies based on protein-A, affilins based on human γ-B-crystallin and anticalins based on the lipocalin fold (See Hey et al., 2005 for a review; Friedman et al., 2007; Ebersbach et al., 2007; Beste et al., 1999). Interestingly, also the chicken avidin has been used as an alternative scaffold in protein engineering (Riihimäki et al., 2011; Niederhauser et al., 2012; Hiltunen et al., unpublished results). The high chemical and thermal stability of the avidin fold, its deep ligand-binding site and its tetrameric structure make it an interesting alternative for the immunoglobulin fold (Riihimäki et al., 2011).

Although the use of suitable scaffold proteins and the sufficient diversity of the mutant libraries are the prerequisite for a successful phage display experiment, the protocols used and the panning conditions applied will eventually determine the phages that are enriched in phage panning (Brakovič, 2010). Therefore, it is essential to be aware of the factors that affect the results of the experiment. The ligands used in phage display are generally considered to be the most important factors determining the mutants that are enriched in phage panning. However, also many other factors affect the results. For example, the phage display technology itself has an intrinsic selection for mutants that fulfill certain requirements. Most importantly, only the mutants that are expressed and correctly folded in the host cells can be selected with the conventional phage display

methods (Forrer et al., 1999). Moreover, toxicity of the displayed proteins, their inefficient incorporation into the viral capsids and reduced infectivity of the phages displaying them can either cause selection bias or completely prevent the selection of functional mutants (Willats, 2002). In order to avoid these issues in phage panning, it is important to understand the biological and molecular mechanisms governing the panning process. However, because the methods used in phage panning vary from experiment to experiment, it may not be a trivial task to form a comprehensive picture of these mechanisms or to find the optimal phage panning conditions for a given project (Bratkovič, 2011; Thie et al., 2008 B).

The phage display technology has previously been used in the selection of chicken avidin mutants in two different studies, described in Riihimäki et al. (2011) and in Niederhauser et al. (2012). The encouraging results of these studies have already proven the feasibility of this approach in the selection of avidin mutants. Nevertheless, the phage panning methods used in these studies may not be optimal for the selection of mutants with high ligand-binding affinities. Importantly, in both studies, the phages were eluted by treating the sample with 100 mM hydrochloric acid. As described above, this may result in selection bias and prevent the enrichment of the mutants with the highest ligand-binding affinities (Ward et al., 1996; Kobayashi 2005). A second issue regarding the selection of avidin proteins with high biotin-binding affinities is directly related to the use of biotin as the immobilized ligand. As the SB medium used in phage production contains small amounts of biotin, it is possible that some of the avidin mutants bind this endogenous biotin instead of the immobilized ligands. The avidin mutants with the highest biotin-binding affinities bind this endogenous biotin most readily, which may result in selection bias. In addition, it is possible that the phage pre-incubations and washes used by Niederhauser et al. (2012) were not efficient enough to prevent the enrichment of non-ligand-specific phages (Niederhauser et al., 2012).

The main aim of this study was to evaluate the suitability of seven different phage panning methods for the selection of the avidin mutants with high biotin-binding affinities from the DNA-shuffled avidin libraries described in Niederhauser et al. (2012). Several different modifications were made to the previously used phage panning protocol to facilitate the selection of high-affinity binders. Most importantly, the biotin ligands were immobilized via DTT-cleavable linkers, which should allow non-selective

elution of all ligand-bound phages. In addition, in the phage panning protocol 2, wild-type avidin was added to the culture medium to sequester endogenous biotin before the production of the phages. To evaluate different methods for preventing non-ligand-specific phage binding, five different protocols with modified sample treatments and pre-elution methods were tested.

6.2 Dissociation-independent phage elution

Before the phage panning experiments were initiated, the diversity and the other properties of the libraries were analyzed by sequencing 30 – 40 randomly selected clones from each library. In general, the properties of the libraries were well in line with those described by Niederhauser et al. (2012). For example, in the current study the percentages of unique sequences in each library were 95%, 51% and 68% for the libraries AVD/AVR2, AVD/BBP-A and AVR2/BBP-A, respectively, while in the previous study the corresponding percentages were 96%, 43% and 55% (Niederhauser et al., 2012). Moreover, the average number of cross-overs was similar in both studies. It can be concluded, that despite of the one-year storage of the libraries between the two studies, the properties of all of the three libraries were similar in both studies before phage panning. Therefore, the differences in the results are likely to be due to the differences in the phage panning methods used in each study.

In this study, the three DNA-shuffled libraries (AVD/AVR2, AVD/BBP-A and AVR2/BBP-A) were panned for three rounds against biotin immobilized via a DDT-cleavable linker. Two different phage panning protocols were tested to evaluate the effects a non-selective elution by a DTT-treatment has on the phage enrichment. Except for the elution method, the protocol 1 is similar to the panning method used by Niederhauser et al. (2012). As previously described, in the protocol 2 the phages were produced in the presence of wild-type avidin.

During the phage panning, the level of phage enrichment was estimated by comparing the ratios between the output (after the phage panning) and input (before the phage panning) phage titers. All of these so-called recovery ratios increased during the phage panning. This indicates that on each subsequent panning round, a larger fraction of the phages bound to the surface of the microtiter plate well. Although this kind of increase

in the recovery ratio is usually considered to be a sign of phage enrichment, it does not necessarily indicate that the enriched phages were biotin-binding. Instead, it is possible that phages displaying avidin mutants with high affinities to BSA, casein or polystyrene were enriched during the phage panning. Surprisingly, the wild-type avidin added during the phage production greatly increased the output/input-ratios on the first two panning rounds. These differences in the recovery ratios seem to be mostly due to the differences in the output-titers. A small increase in the output-titers was expected, as by sequestering endogenous biotin, the added wild-type avidin allows more phages to bind immobilized biotin. However, on the first round the difference in the output titers was almost ten-fold, suggesting that other, still poorly understood mechanisms may have affected the results. Interestingly, the addition of wild-type avidin also increased the growth rates of the bacterial cultures. This observation is surprising, as wild-type avidin is usually considered to be toxic to biotin-requiring bacterial cells. Thus, further studies are needed to completely understand the mechanisms behind these effects caused by the added wild-type avidin.

6.2.1 Sequence analysis of the avidin mutants

The amino acid sequences of the phage display enriched avidin mutants were analyzed by sequencing the avidin-encoding genes of randomly selected colonies. As expected, after phage panning the number of unique sequences was smaller than in the original mutant libraries, indicating that some of the avidin mutants were enriched during phage panning. The most highly enriched mutant of the AVD/AVR2 library panned according to the protocol 1 (A/A2_P1_1) was identical to an avidin mutant previously known as sbAVD-1 (Riihimäki et al., 2011). This mutant alone covers almost half of all sequenced avidin mutants of the AVD/AVR2 library. However, this mutant is not DNA-shuffled and it contains sequences not present in any of the parental genes. Hence, it is likely that the presence of this mutant results from a DNA contamination that has occurred during the library construction. As described in Riihimäki et al. (2011), the sbAVD-1 mutant is binding biotin weakly. Because strong enrichment of this mutant was observed, it is probable that properties other than the biotin-binding affinities of the mutants (e.g. expression levels) affected the results of the experiment.

Despite of the differences in the phage panning protocols, avidin mutants identical to three (A/A2-1, A/A2-3 and A/B-2) of the five mutants described in detail in Niederhauser et al. (2012) were among the mutants sequenced in this study. Although the level of enrichment differed between the two studies, the presence of these mutants suggests that the panning methods used in both studies favored the enrichment of avidin mutants with similar properties. Moreover, similarly to the clones analyzed in Niederhauser et al. (2012), a majority of the mutants panned according to the protocol 1 contained point mutations affecting the amino acid residues directly participating in biotin-binding. Because the sequence context of these mutations differs from that of the avidin proteins used in previous mutagenesis studies, it is difficult to predict the exact effects these mutations have on the structural and functional properties of the avidin mutants. However, as most of the mutations affecting biotin-binding residues have drastic effects on the ligand-binding affinities of the mutants, it is likely that the biotin-binding affinities of these mutants are lower than those of the wild-type proteins (Laitinen et al., 2006). Hence, it seems that the use of a dissociation-independent phage elution method did not completely abolish the tendency of the phage selection process to enrich avidin mutants with decreased biotin-binding affinities and increased expression levels. The most obvious explanation for this phenomenon is the toxicity of biotin-binding avidin proteins for *E. coli*. The avidin mutants with the highest biotin-binding affinities most readily bind biotin present in the periplasm of the host cell, which may prevent biotin from entering the cell. It is also possible that a small fraction of the expressed avidin proteins form functional, biotin-binding tetramers inside the host cell cytoplasm and sequester biotin synthesized by the cell. Therefore, during the phage production step, there may be negative selection of the avidin mutants with high biotin-binding affinities.

Surprisingly, only one of the mutants panned according to the protocol 2 contained a point mutation. Moreover, this mutation (G61K) is far from the biotin-binding site. This striking difference in the number of mutations between the clones panned according to either protocol 1 or protocol 2 suggests, that added wild-type avidin did indeed affect the phage selection process and led to the enrichment of different avidin mutants. However, the reasons for the differences in the numbers of mutations are only poorly understood. Smaller differences in the growth rates of the host cells might at least partly explain these results. As added wild-type avidin most probably bound a significant

fraction of the biotin present in the culture medium, the growth advantage of the bacterial cells expressing avidin mutants with low biotin-binding affinities may have been smaller than in the absence of wild-type avidin. On the other hand, this theory does not explain the high phage titers of the libraries panned according to the protocol 2.

It is probable that the variance in the expression levels of different avidin mutants prevented the enrichment of some high-affinity binders. The effects of this problem are most prominent in the case of the phages displaying wild-type avidin proteins. Although parental genes encoding wild-type avidin and AVR2 are known to be present in the libraries studied here, enrichment of these genes was not observed in either of the two studies (Niederhauser et al., 2012). Because the phage elution method used in this study was not dependent on the biotin-binding affinities of the displayed avidin mutants, the possibility of a selection bias in the elution step can be ruled out with a reasonable certainty. Hence, as hypothesized by Niederhauser et al. (2012), the low expression levels of the wild-type proteins are likely to be the reason for the absence of the genes encoding these proteins after the phage panning. In addition to the toxicity of the avidin proteins to the biotin-requiring bacterial cells, also many other factors may affect the phage display efficiencies of different proteins. For example, the use of a suitable secretion signal is important for successful expression and phage display selection of the studied proteins.

The *PelB* secretion signal used in this study to direct the periplasmic secretion of the fusion proteins was also used in the study described by Riihimäki et al. (2011). Surprisingly, in that study, wild-type chicken avidin was successfully displayed on the surface of phage particles. Although the functionality of phage-displayed wild-type avidin was confirmed, libraries containing such phages were not panned against immobilized biotin. Thus, the results of the study described by Riihimäki et al. (2011) cannot be directly compared with those of the studies presented here and in Niederhauser et al. (2012). Based on the results of these two studies, it can be concluded that the relatively low production levels of the phages displaying wild-type avidin and avidin-like proteins may be enough to prevent the enrichment of these phages.

In future studies, the use of other protein secretion signals might help to avoid the selection bias resulting from the variance in the protein expression levels. For example,

the *OmpA* secretion signal from *Bordetella avium* has been successfully used to promote the periplasmic secretion of non-glycosylated wild-type chicken avidin in Gram-negative bacterial cells (Hytönen et al., 2004). Moreover, the *OmpA* secretion signal has also been shown to efficiently promote the expression and phage display of various proteins fused with the phage p3 protein (Thie et al., 2008 B). Yet another option could be the use of completely different protein translocation pathways. In addition to the Sec pathway exploited by the proteins containing either *PelB* or *OmpA* secretion signals, also the SRP and TAT secretion pathways have been utilized in some phage display studies (Thie et al., 2008 B; Speck et al., 2005). On the other hand, the use of other secretion signals does not necessarily help to overcome the problems caused by the possible toxicity of the studied avidin proteins.

6.2.1 The functional properties of the avidin mutants

The functional properties of the phage display enriched avidin mutants were evaluated with both phage-ELISA and the ABD-assay. Because of the poor reproducibility of the results of the phage-ELISA experiment, phage-ELISA was used only in the initial screening of a large number of avidin mutants. Surprisingly, the phage-ELISA signals of the mutants panned according to the protocol 2 were on average significantly lower than those of the mutants panned according to the protocol 1. Hence, it seems that despite of the non-mutated biotin-binding residues, the mutants panned according to the protocol 2 are either only weakly biotin-binding or poorly expressed in bacterial cells. It is possible that during the PEG-precipitation of the phages, also the wild-type avidin added in the protocol 2 was precipitated. If this is the case, the added avidin may have bound to the immobilized biotin and thus prevented the biotin-binding of the phage displayed avidin mutants. In the absence of biotin-specific binding, the phages specific to BSA, casein or polystyrene may have been enriched, which could explain the virtually nonexistent biotin-binding affinities of the enriched mutants. However, this does not explain the high output-titers of the libraries that were panned according to the protocol 2.

To minimize the bias caused by the differences in the protein expression levels, selected mutants were also analyzed with the ABD-assay. Similarly to the phage-ELISA experiment, only a few of the mutants panned according to the protocol 2 showed

detectable biotin-binding, while most of the mutants panned according to the protocol 1 gave relatively high absorbance signals on the BSA-BTN-coated surfaces. However, some of the mutants panned according to the protocol 1 exhibited also noticeable BSA-binding. The BSA-binding may be mediated by the p3-fusion of these mutants, and therefore it does not necessarily indicate that the avidin mutants themselves are recognizing BSA.

The treatment with excess biotin decreased the absorbance readings of all DNA-shuffled avidin mutants. The only exceptions to this were the samples that did not show any biotin-binding at all, and therefore had absorbance readings close to the background level. The average ratio between the absorbance readings measured from the biotin-treated and the untreated wells was 0.13. However, some of the DNA-shuffled avidin mutants had substantially higher ratios. For example, mutants called A/A2_P2_30, A/B_P1_1 and A/A2_P2_24 had ratios of 0.54 and 0.35 and 0.31, respectively. Although the absolute absorbance values of these avidin mutants were low, the high ratios suggest that these avidin mutants are relatively slowly dissociating from the immobilized biotin molecules. Hence, these mutants could be interesting candidates for further functional analysis with more accurate methods. Surprisingly, also two of the avidin mutants studied in detail in Niederhauser et. al. (2012) were among the mutants with the highest ratios; a mutant previously known as A/A2-3 (here called A/A2_P1_5) gave a ratio of 0.16 and a mutant known as A/B-p2 (here called A/B_P1_10) gave a ratio of 0.18. As expected, both wild-type egg-white avidin and chimeric avidin resisted this biotin treatment well, and the changes in the absorbance readings were only negligible. This indicates that these avidin proteins are so tightly biotin-binding, that only a small fraction of the biotin-bound avidin proteins dissociated during the one-hour biotin treatment. This observation is consistent with the extremely low dissociation rate constants previously measured with other methods (Green, 1975; Hytönen et al., 2005 A).

6.3 Methods for reducing non-specific phage binding

As described above, some of the mutants selected with the dissociation-independent phage panning methods were expressed at levels higher than those of the wild-type avidin proteins and exhibited relatively slow dissociation from immobilized biotin.

Although these results are encouraging, most of the mutants that were enriched when the libraries were panned according to the protocols 1 and 2 were only weakly biotin-binding and readily displaced by excess biotin. Moreover, many of the mutants analyzed in this study were identical to the ones described in Niederhauser et al. (2012). This indicates that the use of a dissociation-independent elution method did not by itself result in the enrichment of avidin mutants with properties that considerably differ from those of the previously analyzed mutants.

In the study published by Niederhauser et al. (2012), it was hypothesized that the absence of avidin mutants with high biotin-binding affinity may have resulted from excessive non-specific binding. The use of DTT-cleavable linkers in the immobilization of biotin at least in theory helps to reduce the effects of this non-specific binding, as unlike phage elution with hydrochloric acid or triethylamine, a DTT treatment specifically elutes only the phages that have bound to the immobilized biotin molecules (Kobayashi et al., 2005). However, it is possible that during the one-hour treatment with 50 mM DTT, also some of the non-specifically bound phages were released from the polystyrene surface. This, in turn, may have been enough to prevent the enrichment of the phages displaying rare avidin mutants with high biotin-binding affinities and relatively low expression levels

To study different methods for preventing the enrichment of phages that are not biotin-binding but readily expressed in the host cells, five modified phage panning protocols (protocols A - E) were tested. The exact methods used in each of these protocols are presented in the section 4.3.2. Briefly, in the protocols B and C, some of the surface-bound phages were pre-eluted by treating the samples with soluble biotin or with 100 mM hydrochloric acid, respectively. The aim of these additional sample treatments was to remove the weakly bound phages before the actual phage elution with DTT. On the contrary, the methods used in the protocols D (phages diluted in 5% sucrose) and E (phages diluted in 10% glycerol) were designed to minimize the non-specific binding of the phages in the first place. In addition, in all of these protocols extended phage pre-incubations with BSA or casein were used to avoid the enrichment of BSA/casein-binding phages. The phage panning protocol A was used as a reference for the comparison of the other phage panning protocols.

The use of any of the modified phage panning protocols (protocols B – E) decreased the ratios between the output and input titers. This was not surprising, since the methods used in these phage panning protocols were expected to reduce the number of phages eluted in the elution step. However, the treatment with 100 mM hydrochloric acid (in the protocol C) was clearly too harsh, as almost all of the phages were lost on the second panning round. This pre-elution was designed to mimic the elution method used by Niederhauser et al. (2012), in order to analyze the mutants that remain ligand-bound after a conventional phage elution. However, as no phage enrichment was observed, it is possible that only a negligible fraction of the phages remained ligand-bound after the pre-elution with hydrochloric acid. This suggests that a treatment with 100 mM hydrochloric acid is most probably efficient enough to elute any of the avidin mutants present in the AVD/AVR2 library. On the other hand, it is also possible that some phages remained bound to biotin even after the acid treatment, but their infectivity was substantially lower than that of the phages not pre-eluted with hydrochloric acid. If this is the case, it is likely that the output phage titers of the samples panned according to the protocol C are biased because of the reduced infectivity of the phages treated with hydrochloric acid. The use of dissociation-independent elution methods helps to avoid this kind of selection bias and ensures nearly complete elution of the ligand-bound phages. Therefore, such elution methods are highly recommended for future studies.

Sequencing of the phage display selected mutants revealed that almost all of the mutants panned according to the protocol B were identical to the A/A2-3 mutant described by Niederhauser et al. (2012). It seems clear, that the pre-elution with excess biotin performed in the protocol B strongly favored the enrichment of this mutant. However, in the ABD-assay this mutant was readily displaced during the biotin-treatment, which was indicated by a treated/untreated ratio of only 0.15. On the other hand, the absolute absorbance signals given by these mutants were relatively high, suggesting that the expression levels of these mutants were above the average. Hence, it is possible that the enrichment of this mutant was mostly driven by the high expression levels, rather than being caused by slow dissociation from the immobilized biotin molecules. Nevertheless, in the ABD-assay, one of the other mutants panned according to the protocol B (a mutant called B4_9) exhibited both relatively high absolute absorbance readings and a high ratio of 0.79 between the biotin-treated and untreated samples. As this mutant

seems to be both highly expressed and slowly dissociating from biotin, it could be an interesting candidate for further analysis with other methods.

The selection strategy represented by the protocol B is usually called off-rate selection (Barbas et al., 2001; Laffly et al., 2008). This kind of phage panning strategies have been designed, because the methods based on washing off the weakly bound phages often fail at favoring the enrichment of proteins with ligand-binding affinities at or below the lower picomolar range (Zahnd et al., 2004). In off-rate selection, the vast excess of the non-immobilized ligands makes the surface-dissociation of the phages virtually irreversible. Hence, the relative dissociation rates of the displayed proteins and the length of the incubation time will determine the mutants that are enriched during the phage panning experiment. In some studies, extremely long incubation times have been used to favor the enrichment of antibodies with high ligand-binding affinities. For example, in a study presented by Laffly et al. (2008), an incubation time of 24 days was used to select for antibodies with picomolar ligand-binding affinities (Laffly et al., 2008). Therefore, it is possible that when off-rate selection (protocol B) was used for the selection of avidin mutants with high biotin-binding affinities, the incubation with excess biotin was not long enough to allow the dissociation of all, except the most tightly biotin-bound avidin mutants. If this is the case, it is likely that properties other than ligand-binding affinity were determinative to the results of the phage panning experiment.

According to the preliminary functional analyses, phage panning according to the protocols D and E did not result in the enrichment of avidin mutants with higher biotin-binding affinities than those of the mutants panned according to the reference protocol. Although the stringency of the binding reaction was increased in these protocols by the addition of 10% glycerol or 5% sucrose, it is possible that the conditions applied during the binding reaction still allowed excessive non-specific binding of the phages. Similarly, in the study described by Niederhauser et al. (2012), the phages were resuspended into a buffer containing 20% glycerol, 1 M NaCl and 1% BSA. Nevertheless, it seems that despite of the relatively high glycerol and salt concentrations, non-specific phage binding could not be avoided. These results are somewhat surprising, as it was shown in a previous study by Kjaer et al. (1998), that the addition of glycerol to a concentration of 10 – 20% may decrease non-ligand-specific

background binding of phages displaying scFv-fragments by three orders of magnitude (Kjaer et al., 1998).

The contrasting results of these two studies provide a concrete example of the problems discussed above. Although many phage display protocols with different phage panning conditions have been published, these protocols are usually optimized for a specific protein or application (Barbas et al., 2001). However, some general guidelines can be given. Increased concentrations of detergents and so-called protein folding aids, such as glycerol and DMSO, typically reduce non-specific hydrophobic interactions between the phages and the polystyrene surface (Barbas et al., 2001; Kjaer et al., 1998). On the contrary, high salt concentrations favor the formation of hydrophobic interactions, which may be beneficial if non-specific phage binding is mediated by ionic interactions (Barbas et al., 2001). Naturally, too harsh conditions will prevent the formation of the specific interactions between the displayed protein and its ligand. Therefore, some trial and error may be required to find the phage panning conditions that best suit to a given phage display project.

In addition to modifying the phage panning conditions, the use of alternative phage panning methods might also help to overcome the problems caused by the non-ligand-specific binding of the phages. For example, when the ligands are immobilized to the surface of paramagnetic beads instead of the microtiter plate wells, the enrichment of polystyrene-binding phages is avoided and less stringent washes are needed to wash off the phages with low ligand-binding affinities (Griffiths & Duncan, 1998). In addition, when compared to microtiter plates, the higher surface to volume ratio of magnetic beads allows the phage panning of a broader range of different mutants on the first panning rounds. The use of magnetic beads also allows an easy control over the surface density of the immobilized ligands, which is important when monovalent ligand binding is required.

If the enrichment of high-affinity binders is desired, yet another option could be the use of the SPR-based phage panning method first described by Malmborg et al. in the year 1996. In this phage panning method, the target ligands are immobilized to a biosensor chip and allowed to react with the phages displaying the studied proteins. The binding reaction is followed by a washing step, during which the number of the surface-bound

phages is monitored in real time with the SPR-instrument. This allows efficient selection of high-affinity binders in only one phage panning round (Yuan et al., 2006). This method is especially suitable when the enrichment of the rare high-affinity binders is wanted or when the target ligands are poorly soluble in water (Bratkovič, 2011; Yuan et al., 2006).

6.4 The ABD-assay in the functional analysis of novel avidin proteins

One of the aims of this study was to develop new methods for the functional analysis of the mutants that are enriched in phage display experiments. Although many powerful methods can be used in the analysis of different biotin-binding proteins, these methods generally require at least some milligrams of purified proteins. Hence, these methods are only applicable when the studied proteins have been produced in a liter-scale culture and purified with suitable methods, for example with affinity chromatography. Because the analysis of single mutant proteins is usually not enough for obtaining a comprehensive picture of the results of the phage display experiment, these methods are poorly suited for the initial analysis of the mutants that have been enriched by phage display. Therefore, other methods allowing the simultaneous analysis of potentially hundreds of mutants are needed for the preliminary analysis of the mutant proteins. Ideally, such methods would provide unbiased and quantitative information about the functional properties of the studied proteins, without the need to handle many samples individually.

Microplate assays, such as ELISA, offer an inexpensive, simple and readily available method for analyzing hundreds of samples in parallel. Even though the results of microplate assays are usually only semi-quantitative, the accuracy of these methods is high enough for an initial analysis of the studied proteins. In this study, two different microplate assays were employed to study the functional properties of the novel avidin mutants. The most straightforward method for the functional analysis of phage display enriched mutants was phage-ELISA, in which the studied proteins were displayed on the surface of complete phage particles. However, similarly to phage panning, in phage-ELISA the phage capsid may mediate non-specific binding to the polystyrene surface of the microtiter plates. Moreover, the expression levels of complete phage particles depend on the expression levels of multiple phage and host proteins. Both of

these factors can at least partly explain why the reproducibility of the phage-ELISA experiment conducted in this study was relatively poor.

Because of the issues related to the reproducibility of the phage-ELISA experiment, also another method was used to analyze the properties of the phage display enriched avidin mutants. In ABD-assay, the studied avidin proteins were not displayed on the surface of complete phage virions, but expressed as soluble tetramers. This probably helps to reduce the problems caused by non-specific binding and it may also enhance the reproducibility of the assay. The coefficient of determination for the ABD-assay performed in this study was high ($R^2 = 0.94$, n = 48), indicating a good reproducibility of the results.

An absolute absorbance reading given by a typical microplate assay is not only dependent on the ligand-binding affinity of the studied protein, but also on the concentration of the protein in the sample. Because the antibodies that are specific to the wild-type avidin proteins do not necessarily bind the DNA-shuffled avidin mutants at all, conventional sandwich ELISA cannot be used to evaluate the concentrations of the studied avidin mutants. Hence, in order to evaluate the relative dissociation rates of the studied proteins, the effect of the variance in the protein expression levels has to be minimized. In this study, one of the two duplicate samples was treated with excess biotin before the colorigenic phosphatase substrate was added. As previously explained, during this biotin treatment, some of the biotin-bound avidin mutants dissociated from the immobilized biotin conjugates. The presence of non-conjugated biotin made this dissociation virtually irreversible. Hence, a fraction of the alkaline phosphatase was lost during the washing step, which was indicated by lower sample absorbance readings after the addition of the phosphatase substrate. Ratios between the absorbance readings measured from the biotin-treated and untreated wells allow the comparison of the dissociation rates of different avidin mutants, even if the expression levels of the studied proteins are largely different. However, determination of the exact dissociation rate constants would require the use of other methods, such as methods based on radiolabeled biotin molecules or the SPR technology.

Although the use of the ABD-assay can help to minimize the bias caused by the differences in the protein expression levels, this method is only applicable when the

expression levels of the studied proteins are above the detection limit. This limitation applies to other microplate assays as well, and it should be taken into consideration when non-purified protein lysates are used in functional studies. For example, in this study, neither wild-type avidin nor wild-type BBP-A was expressed at measurable levels and the absorbance readings of these samples were near the background level. Therefore, low absorbance readings do not necessarily indicate low biotin-binding affinities, but may also result from low protein expression levels.

The ABD-assay developed in this study was proved to be a simple and reproducible method for the functional analysis of novel avidin mutants. As this assay makes it possible to semi-quantitatively evaluate the relative dissociation rate constants of hundreds of avidin mutants simultaneously, it is an efficient method for the initial functional analysis of phage display selected avidin mutants. Therefore, in future studies, the focus of the initial analysis of the selected avidin mutants can be shifted from sequence analysis to functional analysis, which streamlines the selection process and allows us to avoid unnecessary sequencing of avidin mutants with low ligand-binding affinities. Moreover, the applicability of this assay is not limited to the analysis of biotin-binding avidin mutants. With small modification, similar assays can be used to evaluate the ligand-binding properties of avidin mutants that recognize other water-soluble ligands.

6.5 Future objectives

The dissociation-independent phage elution methods described in this study were originally designed to facilitate the selection of antibodies with femtomolar antigen-binding affinities (Kobayashi et al., 2005). However, because of the extremely high biotin-binding affinities of wild-type avidin and streptavidin, the interactions mediated by these wild-type proteins are usually strong enough for most applications. Therefore, the focus of future studies is likely to be on the selection avidin mutants that have both high biotin-binding affinities and properties not present in any of the wild-type proteins, instead of selecting avidin mutants according to their ligand-binding affinities only. As described above, the phage display technology can be exploited in the selection of proteins with high expression levels, improved thermal and chemical stability, catalytic activity and increased protease resistance (Forrer et al., 1999; Jespers et al., 2004;

Sieber et al., 1998). With slight modifications, the phage display methods used in this study could be utilized in the selection of avidin mutants with increased stability, for example. As demonstrated by Riihimäki et al. (2011), also avidin mutants with modified ligand-binding specificities can be selected by using phage display. However, because only proteins that exist in the original library can be selected, it is possible that new libraries that are larger in size and have better diversity are required for such experiments. On the other hand, if these requirements are met, the phage display technology offers intriguing possibilities for the selection of avidin proteins with modified structural and functional properties.

7. CONCLUSIONS

In this study, we evaluated the suitability of seven different phage display protocols for the selection of DNA-shuffled avidin mutants with high biotin-binding affinities. The use of the dissociation-independent phage elution methods resulted in the enrichment of some previously unknown avidin mutants that had high expression levels and showed promising functional properties in the preliminary functional analysis. Closer analysis of these mutants could shed light to the structural basis of the tight avidin-biotin interaction. Moreover, some of the novel phage display selected avidin mutants may offer valuable tools for various applications in life sciences.

It seems clear that because of its important biological functions, biotin is a very challenging ligand for phage display experiments. The results of this study suggest that when avidin mutants are selected by using phage display, both variance in the protein expression levels and non-ligand-specific binding may have a significant impact on the outcome of the experiment. Further optimization of the methods used in the phage display selection of DNA-shuffled avidin mutants may be needed to overcome these issues. However, the phage display protocols evaluated in this study and the results presented here provide a solid foundation for future studies. Importantly, the results of this study suggest that off-rate selection (represented by the protocol B) might be a suitable method for the selection of novel avidin mutants with high biotin-binding affinities. In addition, the ABD-assay described here facilitates the functional analysis of the phage display selected avidin mutants and it will be utilized in future phage display studies.

REFERENCES

Ahlroth MK, Ahlroth P & Kulomaa MS. Copy-Number Fluctuation by Unequal Crossing-Overin the Chicken Avidin Gene Family. *Biochem Biophys Res Commun* 2001 A;288:400-406.

Ahlroth MK, Grapputo A, Laitinen OH & Kulomaa MS. Sequence Features and Evolutionary Mechanisms in the Chicken Avidin Gene Family. *Biochem Biophys Res Commun* 2001 B;285:734-741.

Ahlroth MK, Kola EH, Ewald D, Masabanda J, Sazanov A, Fries R, Kulomaa MS. Characterization and chromosomal localization of the chicken avidin gene family. *Anim Genet* 2000;31:367-375.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403-410.

Aragaña CE, Kuntz ID, Birken S, Axel R & Cantor CR. Molecular cloning and nucleotide sequence of the streptavidin gene. *Nucl Acid Res* 1986;14(4):1871-1882.

Arnold FH. Design by directed evolution. *Acc Chem Res* 1998;31:125-131.

Azzazy HME & Highsmith WE. Phage display technology: clinical applications and recent innovations. *Clin Biochem* 2002;35:425-445.

Barbas CF, Burton DR, Scott JK & Silverman GJ (eds). Phage display: A laboratory manual. *Cold Spring Harbor Library Press*, USA, 2001. ISBN 978-087969740-2.

Bayer EA, Wilchek M. The Use of the Avidin-Biotin Complex as a Tool in Molecular Biology. *Methods Biochem Anal* 1980;26:1-45.

Beste G, Schmidt FS, Stibora T & Skerra A. Small antibody-like proteins with prescribed ligand specificities derived from the lipocalin fold. *Proc Natl Acad Sci USA* 1999;96:1898-1903.

Boder ET, Midelfort KS & Wittrup KD. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc Natl Acad Sci USA* 2000;97(20):10701-10705.

Bornscheuer UT & Pohl M. Improved biocatalysts by directed evolution and rational protein design. *Biocatal Biotransformation* 2001;5:137-143.

Bratkovič T. Progress in phage display: evolution of the technique and its applications. *Cell Mol Life Sci* 2011;67:749-767.

Bruin R, Spelt K, Mol J, Koes R & Quattrochio F. Selection of high-affinity antibodies from phage display libraries. *Nat Biotechnol* 1999;17:397-399.

Bush L & White HB. Conversion of domains into subunits in the processing of egg yolk biotin-binding protein I. *J Biol Chem* 1989;264(10):5741-5745.

Chaiet L & Wolf FJ. The properties of streptavidin, a biotin-binding protein produced by *Streptomycetes*. *Arch Biochem Biophys* 1964;106:1-5.

Chang CCJ, Chen TT, Cow BW, Dawes GD, Stemmer WPC, Punnonen J & Patten PA. Evolution of a cytokine using DNA family shuffling. *Nat Biotechnol* 1999;17:793-797.

Chivers CE, Crozat E, Chu C, Moy VT, Sherratt DJ & Howarth M. A streptavidin variant with slower biotin dissociation and increased mechanostability. *Nat Methods* 2010;7(5):391-393.

Crameri A, Raillard SA, Bermudez E & Stemmer WPC. DNA-shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 1998;391:288-291.

Dale JW, Von Schantz M & Plant N. From genes to genomes: Concepts and applications of DNA technology. *John Wiley & Sons Ltd,* USA, 2002. ISBN 978-0471497837.

DeLange RJ & Huang TS. Egg White Avidin. *J Biol Chem* 1971;246(3):698-709.

Diaz-Neto E, Nunes DN, Giordano RJ, Sun J, Botz GH, Yang K, Setubal JC, Pasqualini R, Arap W. Next-generation phage display: integrating and comparing available molecular tools to enable cost- effective high-throughput analysis. *Plos One* 2009;4(12):1-11-

Eakin RE, McKinley WA, Williams RJ. Egg-white injury in chicks and its relationship to a deficiency of vitamin H (biotin). *Science* 1940 B;92:224-225.

Eakin RE, Snell EE & Williams RJ. The concentration and assay of avidin, the injury-producing protein in raw egg white. *J Biol Chem* 1941;140:535-543.

Eakin RE, Snell EE, Williams RJ. A constituent of raw egg white capable of inactivating biotin in vitro. *J Biol Chem* 1940 A;136:801-802.

Ebersbach H, Fiedler E, Scheuermann T, Fiedler M, Stubbs MT, Reimann C, Proetzel G, Rudolph R & Fiedler U. Affilin–Novel Binding Molecules Based on Human γ-B-Crystallin, an All β-Sheet Protein. *J Mol Biol* 2007;372:172-185.

Elo HA, Räisänen S & Tuohimaa P. Induction of an antimicrobial biotin-binding egg white protein (avidin) in chick tissues in septic *Escherichia coli* infection. *Cell Mol Life Sci* 1980 A;36(3):312-313.

Elo HA. Occurrence of avidin-like biotin-binding capacity in various vertebrate tissues and its induction by tissue injury. *Comp Biochem Physiol B* 1980 B;67(2):221-224.

Flower DR. Structural relationship of streptavidin to the calycin protein superfamily. *FEBS Lett* 1993;33(1-2):99-102.

Forrer P, Jung S & Plückthun A. Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins. *Curr Opin Struct Biol* 1999;9:514-520.

Fraser CM, Eisen JA & Salzberg SL. Microbial genome sequencing. *Nature* 2000;406:799-803.

Friedman M, Nordberg E, Höidén-Guthenberg I, Brismar H, Adams GB, Nilsson FY, Carlsson J & Ståhl S. *Protein Eng Des Sel* 2007;20(4):189-199.

Giri AV, Anishetty S & Gautam P. Functionally specified protein signatures distinctive for each of the different blue copper proteins. *BMC Bioinformatics* 2004;5(127):1-8.

Gommans WM, Haisma HJ & Rots MG. Engineering Zinc Finger Protein Transcription Factors: The Therapeutic Relevance of Switching Endogenous Gene Expression On or Off at Command. *J Mol Biol* 2005;354(3):507-519.

Gonzalez M, Argaraña CE & Fidelio GD. Extremely high thermal stability of streptavidin and avidin upon biotin binding. *Biomol Eng* 1999;16(1-4):67-72.

González M, Bagatolli LA, Echabe I, Arrondo JLR, Argaraña CE, Cantor CR & Fidelio GD. Interaction of biotin with streptavidin - thermostability and conformational changes upon binding. *J Biol Chem* 1997;272(17):11288-11294.

Gope ML, Keinänen RA, Kristo PA, Conneely OM, Beattie WG, Zarucki-Schulz T, O'Malley BW, Kulomaa MS. Molecular clonig of the chicken avidin cDNA. *Nucl Acid Res* 1987;15(8):3595-3606.

Green NM. Avidin. In: Anfinsen CB (ed), Advances in protein chemistry, Academic Press, New York, 1975, pp. 85-132.

Griffiths AD & Duncan AR. Strategies for selection of antibodies by phage display. *Curr Opin Biotechnol* 1998;9:102-108.

György P, Rose CS & Tomarelli R. Investigations on the tability of avidin. *J Biol Chem* 1942;144:169-173

Hanes J & Plückthun A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci USA* 1997;94:4937-4942.

Harayama S. Artificial evolution by DNA-shuffling. *Trends Biotechno*l 1998;16(2):76-82.

Heinonen PK, Metsä-Ketelä T & Tuohimaa P. Induction of avidin in the chick oviduct by tissue damage and prostaglandins. *J Reprod Fert* 1978;52:159-162.

Helppolainen SH, Määttä JAE, Halling KK, Slotte JP, Hytönen VP, Jänis J, Vainiotalo P, Kulomaa MS, Nordlund HR. Bradavidin II from Bradyrhizobium japonicum: A new avidin-like biotin-binding protein. *Biochim Biophys Acta* 2008;1784:1002-1010.

Helppolainen SH, Nurminen KP, Määttä JAE, Halling KK, Slotte JP, Huhtala T, Liimatainen T, Ylä-Herttuala S, Airanne KJ, Närvänen A, Jänis J, Vainiotalo P, Valjakka J, Kulomaa MS & Nordlund HR. Rhizavidin from Rhizobium etli: the first natural dimer in the avidin protein family. *Biochem J* 2007;405:397-405.

Hendrickson WA, Pähler A, Smith JL, Satow Y, Merritt EA & Phizackerley RP. Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proc Natl Acad Sci USA* 1987;86:2190-2194.

Hey T, Fiedler E, Rudolph R & Fiedler M. Artificial, non-antibody binding proteins for pharmaceutical and industrial applications. *Trends Biotechnol* 2005;23(10):514-521.

Hiller Y, Gershoni JM, Bayer EA & Wilchek M. Biotin binding to avidin: Oligosaccharide side chain not required for ligand association. *Biochem J* 1987;248:167-171.

Hust M, Thie H, Schirrmann T & Dübel S. Antibody phage display In: Zhiqiang A (ed), Therapeutic Monoclonal Antibodies: From Bench to Clinic. *John Wiley & Sons*, USA, 2008. ISBN 0470117915

Hytönen VP, Laitinen OH, Airanne TT, Kidron H, Meltola NJ, Porkka EJ, Hörhö J, Paldanius T, Määttä JAE, Nordlund HR, Johnson MS, Salminen TA, Airanne KJ, Ylä-Herttuala S & Kulomaa MS. Efficient production of active chicken avidin using a bacterial signal peptide in *Escherichia coli. Biochem J* 2004;384:385-390.

Hytönen VP, Laitinen OH, Grapputo A, Kettunen A, Savolainen J, Kalkkinen N, Marttila AT, Nordlund HR, Nyholm TKM, Paganelli G & Kulomaa MS. Characterization of poultry egg-white avidins and their potential as a tool in pretargeting cancer treatment. *Biochem* J 2003;372:219-225.

Hytönen VP, Määttä JAE, Kidron H, Halling KK, Hörhä J, Kulomaa T, Nyholm TKM, Johnson MS, Salminen TA, Kulomaa MS & Airanne TT. Avidin related protein 2 shows unique structural and functional features among the avidin protein family. *BMC Biotechnol* 2005 B;5(28):1-14.

Hytönen VP, Määttä JAE, Niskanen EA, Huuskonen J, Helttunen KJ, Halling KK, Nordlund HR, Rissanen K, Johnsonn MS, Salminen TA, Kulomaa MS, Laitinen OH & Airanne TT. Structure and characterization of a novel chicken biotin-binding protein A (BBP-A). *BMC Struct Biol* 2007;7(8):1-20

Hytönen VP, Määttä JAE, Nyholm TKM, Livnah O, Eisenberg-Domovich Y, Hyre D, Nordlund HR, Hörhä J, Niskanen EA, Paldanius T, Kulomaa T, Porkka EJ, Stayton PS, Laitinen OH & Kulomaa MS. Design and Construction of Highly Stable, Protease-resistant Chimeric Avidins. *J Biol Chem* 2005 A;280(11):10228-10233.

Jespers L, Schon O, Famm K & Winter G. Aggregation-resistant domain antibodies selected on phage by heat denaturation. *Nat Biotechnol* 2004;22(9):1161-1165.

Kazlauskas RJ & Bornscheuer UT. Finding better protein engineering strategies. *Nature Chem Biol* 2009;5(8):526-529.

Kazlauskas RJ. Molecular modeling and biocatalysis: explanations, predictions, limitations, and opportunities. *Curr Opin Chem Biol* 2000;4:81-88.

Keinänen RA, Laukkanen M-L & Kulomaa MS. Molecular cloning of three structurally related genes for chicken avidin. *J Steroid Biochem* 1988;30:17-21.

Keinänen RA, Wallén MJ, Kristo PA, Laukkanen MO, Toimela TA, Helenius MA & Kulomaa MS. Molecular cloning and nucleotide sequences of chicken avidin-related genes 1-5. *Eur J Biochem* 1994;220:615-621.

Kersey PJ, Staines DM, Lawson D, Kulesha E, Derwent P, Humphrey JC, Hughes DST, Keenan S, Kerhornou A, Koscielny G, Langridge N, McDowall MD, Megy K, Maheswari U, Nuhn M, Paulini M, Pedro H, Toneva I, Wilson D, Yates A & Birney E. Ensembl Genomes: an integrative resource for genome-scale data from non-vertebrate species. *Nucleic Acids Res* 2012;40:91-97.

Khan I & Akhtar MW. Different approaches for protein engineeringin industrial biotechnology. *Nature Precedings* 2011. Manuscript.

Kjaer S, Brian Stausbøl-Grøn, Wind T, Ravn P, Jensen KH, Kahns L, Clark BFC. Glycerol diversifies phage repertoire selections and lowers non-specific phage absorption. *FEBS Letters* 1998;431:448-452.

Kobayashi N, Karibe T & Goto J. Dissociation-independent selection of high-affnity anti-hapten phage antibodies using cleavable biotin-conjugated haptens. *Anal Biochem* 2005;347:287-296.

Korf I. Gene finding in novel genomes. *BMC Bioinformatics* 2004;5(59):1-9.

Korpela JK, Kulomaa MS, Elo HA & Tuohimaa PJ. Biotin-binding proteins in eggs of oviparous vertebrates. *Cell Mol Life Sci* 1981;37(10):1065-1066.

Kunnas TA, Wallén MJ & Kulomaa MS. Induction of chicken avidin and related mRNAs after bacterial infection. *Biochim Biophys Acta* 1993;1216(3):441-445.

Laffly E, Pelat T, Cédrone F, Blésa S, Bedouelle H & Thullier P. Improvement of an antibody neutralizing the anthrax toxin by simultaneous mutagenesis of its six hypervariable loops. *J Mol Biol* 2008;378:1094-1103.

Laitinen OH, Hytönen VP, Ahlroth MK, Pentikäinen OT, Gallagher C, Nordlund HR, Ovod V, Marttila AT, Porkka E, Heino S, Johnson MS, Airanne KJ & Kulomaa MS. Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin. *Biochem J* 2002;363:609-617.

Laitinen OH, Hytönen VP, Nordlund HR & Kulomaa MS. Genetically engineered avidins and streptavidins. *Cell Mol Life Sci* 2006;63:2992-3017.

Laitinen OH, Nordlund HR, Hytönen VP & Kulomaa MS. Brave new (strept)avidins in biotechnology. *Trends biotechnol* 2007;25(6):269-277.

Ling M & Robinson BH. Approaches to DNA Mutagenesis: An Overview. *Anal Biochem* 1997;254:157-178.

Livnah O, Bayer EA, Wilchek M & Sussman JL. Three-dimensional structures of avidin and avidin-biotin complex. *Proc Natl Acad Sci USA* 1993;90:5076-5080.

Long EL, Sonstegard TD, Long JA, Tassek CPV & Zuelke KA. Serial Analysis of Gene Expression in Turkey Sperm Storage Tubules in the Presence and Absence of Resident Sperm. *Biol Reprod* 2003;69(2):469-474.

Low NM, Holliger P & Winter G. Mimicking Somatic Hypermutation: Affinity Maturation of Antibodies Displayed on Bacteriophage Using a Bacterial Mutator Strain. *J Mol Biol* 1996;260:359-368.

Määttä JAE, Airenne TT, Nordlund HR, Jänis J, Paldanius TA, Vainiotalo P, Johnson MS, Kulomaa MS & Hytönen VP. Rational Modification of Ligand-Binding Preference of Avidin by Circular Permutation and Mutagenesis. *ChemBioChem* 2008;9(7):1124-1135.

Määttä JAE, Helppolainen SH, Hytönen VP, Johnson MS, Kulomaa MS, Airanne TT & Nordlund HR. Structural and functional characteristics of xenavidin, the first frog avidin from *Xenopus tropicalis*. *BMC Struct Biol* 2009;9(63):1-13.

Malmborg AC, Dueñas M, Ohlin M, Söderlind E & Borrebaeck CAK. Selection of binders from phage display libraries using the BIAcore biosensor. *J Immunol Methods* 1996;198:51-57.

Marttila AT, Laitinen OH, Airenne KJ, Kulik T, Bayer EA, Wilchek M & Kulomaa MS. Recombinant NeutraLite Avidin: a non-glycosylated, acidic mutant of chicken avidin that exhibits high affinity for biotin and low-specific binding properties. *FEBS Lett* 2000;467(1):31-36.

Marvin DA. Filamentous phage structure, infection and assembly. *Curr Opin Struct Biol* 1998;8:150-158.

McCafferty J, Griffiths AD, Winter G, Chiswell DJ. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature* 1990;348:552-554.

Moore GL, Maranas CD, Lutz S & Benkovic CJ. Predicting crossover generation in DNA-shuffling. *Proc Natl Acad Sci USA* 2000;98(6):3226-3231.

Nagaraj SH, Gasser RB & Ranganathan S. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Brief Bioinform* 2006;8(1):6-21.

Niederhauser B, Siivonen J, Määttä JAE, Jänis J, Kulomaa MS & Hytönen VP. DNA family shuffling within the chicken avidin protein family – A shortcut to more powerful protein tools. *J Biotechnol* 2012;157(1):38-49.

Niskanen EA, Hytönen VP, Grapputo A, Nordlund HR, Kulomaa MS & Laitinen OH. Chicken genome analysis reveals novel genes encoding biotin-binding proteins related to avidin family. *BMC Genomics* 2005;6:41-54.

Nordlund HR, Hytönen VP, Laitinen OH & Kulomaa MS. Novel Avidin-like Protein from a Root Nodule Symbiotic Bacterium, *Bradyrhizobium japonicum*. *J Biol Chem* 2005;280(14):13250-12255.

Nordlund HR, Laitinen OH, Hytönen VP, Uotila STH, Porkka E & Kulomaa MS. Construction of a Dual Chain Pseudotetrameric Chicken Avidin by Combining Two Circularly Permuted Avidins. *J Biol Chem* 2004;279(35):36715-36719.

O'Malley BW. In Vitro Hormonal Induction of a Specific Protein (Avidin) in Chick Oviduct. *Biochemistry* 1967;6(8):2546-2551.

O'Neil KT & Hoess RH. Phage display: protein engineering by directed evolution. *Curr Opin Struct Biol* 1995;5:443-449.

Paschke M. Phage display systems and their applications. *Appl Microbiol Biotechnol* 2006;70:2-11.

Pugliese L, Coda A, Malcovati M & Bolognesi M. Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution. *J Mol Biol* 1993;231(5):698-710.

Repo S, Paldanius TA, Hytönen VP, Nyholm TKM, Halling KK, Huuskonen J, Pentikäinen OT, Rissanen K, Slotte JP, Airanne TT, Salminen TA, Kulomaa MS & Johnson MS. Binding Properties of HABA-Type Azo Derivatives to Avidin and Avidin-Related Protein 4. *Chem Biol* 2006;13:1029-1039.

Riihimäki TA, Hiltunen S, Rangl M, Nordlund HR, Määttä JAE, Ebner A, Hinterdorfer P, Kulomaa MS, Takkinen K & Hytönen VP. Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein. *BMC Biotechnol* 2011;11:64-75.

Roberts RW & Szostak JW. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci USA* 1997;94(23):12297-12302.

Russel M, Linderoth NA & Šali A. Filamentous phage assembly: variation on a protein export theme. *Gene* 1997;192:23-32.

Russel M. Moving through the membranes with filamentous phages. *Trends Microbiol* 1995;3(6):223-228.

Sano T & Cantor CR. Intersubunit contacts made by tryptophan 120 with biotin are essential for both strong biotin binding and biotin-induced tighter subunit association of streptavidin. *Proc Natl Acad Sci USA* 1995;92:3180-3184.

Santala V & Saviranta P. Affinity-independent elution of antibody-displaying phages using cleavable DNA linker containing streptavidin beads. *J Immunol Methods* 2004;284:159-163.

Sardo A, Wohlschlaget T, Lo C, Zoller H, Ward TR & Creus M. Burkavidin: a novel secreted biotin-binding protein from the human pathogen *Burkholderia pseudomallei*. *Protein Expr Purif* 2011;77(2):131-139.

Sidhu SS. Engineering for M13 phage display. *Biomol Eng* 2001;18:56-63.

Sieber V, Plückthun A & Schmid FX. Selecting proteins with improved stability by a phage-based method. *Nat Biotechnol* 1998;16:955-960.

Smith GP & Petrenko VA. Phage Display. *Chem Rev* 1997;97:391-410.

Smith GP. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* 1985;228(4705):1315-1317.

Speck J, Arndt KM & Müller KM. Efficient phage display of intracellularly folded proteins mediated by the TAT pathway. *Protein Engin Des Sel* 2011. Manuscript.

Stemmer WPC. DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Proc Natl Acad Sci USA* 1994 A;91:10747-10751.

Stemmer WPC. Molecular breeding of genes, pathways and genomes by DNA shuffling. *Biotechnol Bioprocess Eng* 2002;7:121-129.

Stemmer WPC. Rapid evolution of a protein *In vitro* by DNA-shuffling. Nature 1994 B;370:389-391.

Stevens L. Egg white proteins. *Comp Biochem Physiol B* 1991;1:1-9.

Takakura Y, Tsunashima M, Suzuki J, Usami S, Kakuta Y, Oniko N, Ito M & Yamamoto T. Tamavidins – novel avidin-like biotin-binding proteins from the Tamogitake mushroom. *FEBS J* 2009;276:1383-1397.

Thie H, Meyer T, Schirrmann T, Hust M & Dübel S. Phage display derived therapeutic antibodies. *Curr Pharm Biotechnol* 2008 A;9:439-445.

Thie H, Schirrman T, Paschke M, Dübel S & Hust M. SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in *E. coli*. *New Biotechnol* 2008 B;25(1):49-54.

Thomas WD & Smith GP. The case for trypsin release of affinity-selected phages. *BioTechniques* 2010;48(3):651-653.

Tuohimaa P, Joensuu T, Isola J, Keinänen RA, Kunnas T, Niemelä A, Pekki A, Wallén MJ, Ylikomi T & Kulomaa MS. Development of progestin-specific response in the chicken oviduct. *Int J Dev Biol* 1989;33:125-134.

Uhlén M. Affinity as a tool in life science. *BioTechnoloques* 2008;44(5):649-655.

Wallén MJ, Laukkanen MO, Kulomaa MS. Cloning and sequencing of the chicken egg-white avidin-encoding gene and its relationship with the avidin-related genes *Avr1 -Avr5*. *Gene* 1995;161:205-209.

Ward RL, Clark MA, Lees J, Hawkins NJ. Retrieval of human antibodies from phage-display libraries using enzymatic cleavage. *J Immunol Methods* 1996;189:73-82.

Warnke R & Levy R. Detection of T and B cell antigens hybridoma monoclonal antibodies: a biotin-avidin-horseradish peroxidase method. *J Histochem Cytochem* 1980;28(8):771-776.

Weber PC, Ohlendorf DH, Wendoloski JJ & Salamme FR. Structural origins of high-affinity biotin binding to streptavidin. *Science* 1989;243(4887):85-88.

White HB & Whitehead CC. Role of avidin and other biotin-binding proteins in the deposition and distribution of biotin in chicken eggs. *Biochem J* 1987;241:677-684.

White HB, Dennison BA, Fera MAD, Whitney CJ, McGuire JC, Meslar HW & Sammelwitz PH. Biotin-binding protein from chicken egg-yolk – assay and relationship to egg-white avidin. *Biochem J* 1976;157:395-400.

Willats WGT. Phage display: practicalities and prospects. *Plant Mol Biol* 2002;50:837-854.

Williams RH. Clinical biotin deficiency. *N Engl J Med* 1943;228:247-251

Yanai I, Yu Y, Zhu X, Cantor CR & Weng Z. An avidin-like domain that does not bind biotin is adopted for oligomerization by the extracellular mosaic protein fibropellin. *Protein Sci* 2005;14:417-423.

Yuan B, Schulz P, Liu R & Sierks MR. Improved affinity selection using phage display technology and off-rate based selection.*Electron J Biotechnol* 2006;9(2):171-175.

Zahnd C, Spinelli S, Luginbühl B, Amstrutz P, Cambillaus C & Plückthun P. Directed in vitro evolution and crystallographic analysis of a peptide-binding single chain antibody fragment (scfv) with low picomolar affinity. *J Biol Chem* 2004;279(18):18870-18877.

APPENDICES

Appendix 1: Solutions and reagents

BSA	Lyophilized powder (Sigma Aldrich)
BSA-BTN	BSA biotinylated with Sulfo-NHS-SS-BTN (Thermo Fisher Scientific). On average 4 biotin molecules per BSA molecule. Barbara Niederhauser 2010.
Casein	Lyophilized skimmed-milk powder
Casein-BTN	Casein biotinylated with Sulfo-NHS-SS-BTN (Thermo Fisher Scientific). On average 4 biotin molecules per casein molecule. Barbara Niederhauser 2010.
DEA buffer	1 M Diethanolamine 0.5 mM $MgCl_2$ dH_2O (pH 9.8)
DTT	50 mM Dithiothreitol in PBS
Freezing medium	10 mg/ml NaCl 22 mg/ml Troptone 30 mg/ml Yeast extract 10 mg/ml MOPS (3-morpholinopropane-1-sulfonic acid) 48 µg/ml (400 µM) $MgSO_4$ 6.3 mg/ml (36 mM) K_2HPO_4 1.8 mg/ml (13.2 mM) KH_2PO_4 890 µg/ml (6.8 mM) $(NH_4)_2SO_4$ 363 µg/ml (1.7 mM) Na-Citrate 4.4% (v/v) Glycerol dH_2O
LB medium	10 mg/ml NaCl 10 mg/ml Tryptone 5 mg/ml Yeast extract dH_2O
LB $_{amp+tet}$ plates	20 ml LB medium (for 1 plate) 0.6 g Bactoagar (Becton Dickinson) 100 µg/ml Ampicillin 10 µg/ml Tetracycline

PBS	8.01 mg/ml (137 mM) NaCl 200 µg/ml (2.7 mM) KCl 1.78 mg/ml (10 mM) Na_2HPO_4 0.27 mg/ml (2 mM) KH_2PO_4 dH_2O (ph 7.4)
PBS-Tween (PBST)	0.05% (v/v) Tween 20 in 1x PBS
PNPP	5 mg tablets, Sigma Aldrich. Final concentration of 1 mg/ ml in DEA-buffer.
SB medium	10 mg/ml NaCl 22 mg/ml Troptone 30 mg/ml Yeast extract 10 mg/ml MOPS (3-morpholinopropane-1-sulfonic acid) dH_2O
5x Sequencing Buffer	200 mM Tris 5 mM $MgCl_2$ dH_2O
Wild-type avidin	Glycosylated egg-white avidin, Belovo

Appendix 2 A: The non-redundant sequences of the mutants panned according to the phage panning protocol 1

The residues originating from avidin, AVR2 and BBP-A are colored blue, green and yellow, respectively. Point mutations and DNA contaminations are colored red. On top, the amino acid sequence of wild-type avidin is shown together with the biotin binding residues (shaded in black) and the interface residues (shaded in gray). The black arrows above the avidin sequence indicate the eight beta-sheets of the avidin fold. Mutants from different libraries (AVD/AVR2, AVR/BBP-A and AVR2/BBP-A) are separated by thick horizontal lines. Sequences containing DNA contaminations are marked with an asterisk.

Library	Mutant	Clones	Mutation	Avidin encoding sequence
				β1 → β2 → β3 → β4 → β5 → β6 → β7 → β8 →
	WT AVIDIN			ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQKE
A/A2	A/A2_P1_01	13	* (sbAVD-1)	ARKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A/A2	A/A2_P1_21	2	* (sbAVD-1)	ARKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRQRTQK
A/A2	A/A2_P1_05	2	S16Y	ARKCSLTGKWTNDLGYTMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRSGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A/A2	A/A2_P1_04	3	T35A	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYIAAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A/A2	A/A2_P1_11	3	N12D	ARKCSLTGKWTLDLGSNMTIGAVNSRGEFTGTYITAVADNPGNITLSPLLGIQH...KRASQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDISDDWIATRVGINIFTRQRTQK
A/A2	A/A2_P1_17	1	* (sbAVD-1)	ARKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRQRTQK
A/A2	A/A2_P1_02	1	* (sbAVD-1)	ARKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A/A2	A/A2_P1_22	1	T35I	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYILAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A/A2	A/A2_P1_27	1	W70R	ARKCSLTGKWTNDLGSIMTIGAVNDNGEFDGTYITAVADNPGNITLSPLHGTQNTIN.KRTQPTFGFTVNRNFSESTSVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A/A2	A/A2_P1_09	1	T35K	ARKCSLTGKWDNDLGSNMTIGAVNSRGEFTGTYIKAVADNPGNITLSPLLGIQH...KRASQPTFGFTVHWKFSESTTVFTGQCFVDRSGKEVLKTKWLQRLAVDDISDDWKATRVGINIFTRQRTQK
A/A2	A/A2_P1_16	1	S16Y	ARKCSLTGKWTNDLGYNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVHWKFSESTSVFVGQCFIDRNGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A/B	A/B_P1_11	23	W70L	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNLKFSDSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A/B_P1_01	2	N12D	ARKCSLTGKWTDDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A/B_P1_02	1	N12D	ARKCSLTGKWTDDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A/B_P1_09	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A/B_P1_10	1	S16Y	ARKCSLTGKWTNDLGYNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A/B	A/B_P1_16	1	T35K	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYIKAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGTNVFTRV....
A2/B	A2/B_P1_10	4	A39P	SRKCELQGLWRNELGSNMTISALDVAGTFSGSYQTAVTPTNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_12	4	A36V	SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTVVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_16	3		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_20	3	F72Y	SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNYSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_03	2		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_15	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_13	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRRGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_09	1	V26A	SRKCELQGLWRNELGSIMTISALDAAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRRGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_14	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVAATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_24	1	S77F	SRKCELQGLWRNELGSNMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADFTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_01	1	P75A	SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_07	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_29	1	E76K	SRKCELQGLWRNDLGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2/B_P1_02	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSKSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_04	1	* (sbAVD-1)	SRKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A2/B	A2/B_P1_17	1	* (sbAVD-1)	SRKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A2/B	A2/B_P1_11	1	* (sbAVD-1)	SRKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A2/B	A2/B_P1_05	1	* (sbAVD-1)	ARKCSLTGKWTNRMNHNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A2/B	A2/B_P1_06	1	*	SRKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYILAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGIMIFTRLRTQK
A2/B	A2/B_P1_08	1	Q55R	SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQRPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2/B_P1_19	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK

Appendix 2 B: The non-redundant sequences of the mutants panned according to the phage panning protocol 2

The residues originating from avidin, AVR2 and BBP-A are colored blue, green and yellow, respectively. Point mutations and DNA contaminations are colored red. On top, the amino acid sequence of wild-type avidin is shown together with the biotin binding residues (shaded in black) and the interface residues (shaded in gray). The black arrows above the avidin sequence indicate the eight beta-sheets of the avidin fold. Mutants from different libraries (AVD/AVR2, AVR/BBP-A and AVR2/BBP-A) are separated by thick horizontal lines. Sequences containing DNA contaminations are marked with an asterisk.

Library	Mutant	Clones	Mutation	Avidin encoding sequence
				β1 → β2 → β3 → β4 → β5 → β6 → β7 → β8 →
	WT AVIDIN			ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQKE
A-A2	A-A2_P2_16	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDISDDWIATRVGNNDFTRLRTQK
A-A2	A-A2_P2_32	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRLRTQK
A-A2	A-A2_P2_09	1		ARKCSLTGEWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFVDRSGKEVLKTKWLLRSSVNDIGDDWKATRVGNNDFTRLRTQK
A-A2	A-A2_P2_11	1		ARKCSLTGEWDNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVHWNFSESTTVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRQRTQK
A-A2	A-A2_P2_01	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTSVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A-A2	A-A2_P2_12	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDISDDWIATRVGNNDFTRQRTQK
A-A2	A-A2_P2_06	1		ARKCSLTGKWTNDLGSIMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A-A2	A-A2_P2_02	1		ARKCSLTGKWDNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTSVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A-A2	A-A2_P2_05	1		ARKCSLTGEWDNDLGSIMTIGAVNDNGEFDGTYITAVADNPGNITLSPLLGIQH...KRASQPTFGFTVNWKFSESTTVFTGQCFVDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQK
A-A2	A-A2_P2_18	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A-A2	A-A2_P2_19	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNXFTRQRTQK
A-A2	A-A2_P2_20	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWNFSESTTVFTGQCFIDRNGKEVLKTKWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A-A2	A-A2_P2_22	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWIATRVGNNDFTRLRTQK
A-A2	A-A2_P2_26	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWIATRVGNNDFTRQRTQK
A-A2	A-A2_P2_28	1		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTSVFVGQCFIDRNGKEVLKTKWLQRLAVDDISDDWKATRVGNNDFTRQRTQK
A-A2	A-A2_P2_31	1		ARKCSLTGKWTNDLGSNMTIGAVNDNGEFDGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK
A-A2	A-A2_P2_24	1		ARKCSLTGKWTNDLGSIMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A-A2	A-A2_P2_21	1		ARKCSLTGKWDNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRQRTQK
A-A2	A-A2_P2_25	1		ARKCSLTGEWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRQRTQK
A/B	A-B_P2_06	11		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A-B_P2_12	3		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSESTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A-B_P2_23	2		ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTQPTFGFTVNWKFSDSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A-B_P2_19	1		ARKCSLTGKWTNDLGSNMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A/B	A-B_P2_21	1	G61K	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTIN.KRTKPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A2/B	A2-B_P2_02	7		SRKCELQGLWRNELGSNMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_23	3		SRKCELQGLWRNELGSNMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2-B_P2_25	3		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_10	2		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_08	1		SRKCELTGEWDNDLGSIMTIGAVNDNGEFDGTYITAVADNPGNITLSPLLGIQH...KRASQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2-B_P2_06	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVQWQFADSTTVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRLRTQK
A2/B	A2-B_P2_20	1	* (sbAVD-1)	SRKCSLTGKWTNDLGSIMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTI.NKRTQPTFGFNVNWKFSESTTVFTGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_05	1		SRKCSLTGEWDNDLGSIMTIGAVNDNGEFDGTYITAVADNPGNITLSPLLGIQH...KRASQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_22	1		SRKCELQGLWRNELGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTKWLQRLAVDDISDDWIATRVGNNDFTRQRTQK
A2/B	A2-B_P2_16	1		ARKCSLTGEWDNDLGSIMTISALDVAGTFSGSYQTAVTATNKQILVSPLKGAQQPPGTKGQQPTFGFTVHWNFSESTSVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A2/B	A2-B_P2_15	1		ARKCSLTGEWDNDLGSIMTIGAVNDNGEFDGTYITAVADNPGNITLSPLLGIQH...KRASQPTFGFTVQWQFADSTTVFVGQCFVDRRGKEMLEMAWLLREEVPSRKDTWKATRVGTNVFTRV....
A2/B	A2-B_P2_19	1	* (sbAVD-1)	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTI.NKRTQPTFGFTVNWKFSDSTTVFVGQCFVDRRGKEMLEMAWLLREEVVPSRKDTWKATRVGTNVFTRV...
A2/B	A2-B_P2_09	1	* (sbAVD-1)	SRKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTI.NKRTQPNFGFTVHWNFSESTSVFVGQCFVDRSGKEVLKTMWLLRSSVNDIGDDWKATRVGNLFTRQRTQK

Appendix 3: The non-redundant sequences of the mutants panned according to the phage panning protocols A-E

The residues originating from avidin, AVR2 and BBP-A are colored blue, green and yellow, respectively. Point mutations and DNA contaminations are colored red. On top, the amino acid sequence of wild-type avidin is shown together with the biotin binding residues (shaded in black) and the interface residues (shaded in gray). The black arrows above the avidin sequence indicate the eight beta-sheets of the avidin fold.

Panning protocol	Mutant	Previously known as	No. of clones	Mutation	Avidin encoding sequence
					β1 → β2 → β3 → β4 → β5 → β6 → β7 → β8 →
	AVIDIN				ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGINIFTRLRTQKE
A	A4_8		1	W70Q	ARKCSLTGKWTNDLGSNMTIGAVNDNGEFTGTYITAVADNPGNITLSPLHGTQNTINKRTQPTFGFTVHQNFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRLRTQK.
B	B4_2	BN p3	11	S16Y	ARKCSLTGKWTNDLGYTMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFIDRSGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK.
B	B4_15		2	T77I	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYITAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTIVFTGQCFVDRSGKEVLKTMWLLRSSVNDIGDDWKATRVGNNDFTRQRTQK.
B	B4_9		1		ARKCSLTGEWTNDLGSNMTIGAVNSRGEFTGTYITAVADNPGNITLSPLLGTQH..KRASQPTFGFTVNWKFSESTTVFTGQCFIDRNGKEVLKTMWLLRSSVNDIGDDWLATRVGNNDFTRLRTQK.
D	D4_7		2	T35A	ARKCSLTGKWTNDLGSNMTIGAVNSRGEFTGTYIAAVTATSNEIKESPLHGTQNTINKRTQPTFGFTVNWKFSESTTVFTGQCFVDRSGKEVLKTKWLQRLAVDDISDDWLATRVGNNDFTRLRTQK.
D	D4_15		1		ARKCSLTGEWTNDLGSNMTIGAVNSRGEFTGTYITAVTDNPGNITLSPLLGTQH..KRASQPTFGFTVHWNFSESTSVFVGQCFIDRNGKEVLKTKWLQRLAVDDISDDWLATRVGNNDFTRQRTQK.

Appendix 4: The results of the ABD-assay for the mutants panned according to the protocols 1 and 2

The background-subtracted absorbance values of each sample are presented in the table. The binding of the mutants was assayed on both BSA and BSA-BTN coated surfaces and after a treatment with excess biotin. The ratio of the biotin-treated samples to untreated samples is shown on the right hand column.

Name of the mutant	Previously known as	No. of clones	Panning protocol	BTN coating + BTN treatment	SD	BTN coating	SD	BSA coating	SD	Ratio between the biotin-treated and untreated samples
A/A2_P1_22		1	1	0.007	0.006	0.573	0.033	0.069	0.003	0.01
A/A2_P1_4		3	1	0.000	0.001	0.232	0.023	0.010	0.001	
A/A2_P1_27		1	1	0.003	0.001	0.596	0.021	0.186	0.028	0.00
A/A2_P1_25		N/A	1	0.000	0.001	0.024	0.005	0.074	0.003	
A/A2_P1_16		1	1	0.022	0.010	0.947	0.471	0.043	0.020	0.02
A/A2_P1_5	A/A2-3	2	1	0.554	0.062	3.427	0.000	0.661	0.071	0.16
A/A2_P1_11	A/A2-1	3	1	0.002	0.000	0.033	0.013	0.034	0.001	0.06
A/A2_P2_24		1	2	0.017	0.002	0.054	0.021	0.003	0.002	0.31
A/A2_P2_23		1	2	0.002	0.001	0.023	0.001	0.000	0.001	0.07
A/A2_P2_30		1	2	0.019	0.001	0.035	0.002	0.002	0.000	0.54
A/A2_P2_6		1	2	0.000	0.001	0.002	0.002	0.004	0.001	
A/A2_P2_19		1	2	0.000	0.002	0.000	0.001	0.001	0.003	
A/A2_P2_20		1	2	0.000	0.001	0.000	0.001	0.000	0.001	
A/B_P1_11		23	1	0.097	0.036	0.377	0.134	0.022	0.010	0.26
A/B_P1_1		2	1	0.175	0.016	0.505	0.084	0.031	0.010	0.35
A/B_P1_2		1	1	0.056	0.027	0.231	0.113	0.006	0.004	0.24
A/B_P1_16		1	1	0.269	0.016	3.180	0.143	0.449	0.014	0.08
A/B_P1_10	A/B-2	1	1	0.391	0.009	2.169	0.380	0.199	0.027	0.18
A/B_P2_6		11	2	0.000	0.001	0.000	0.001	0.002	0.000	
A/B_P2_12		3	2	0.001	0.002	0.001	0.001	0.004	0.004	
A/B_P2_23		2	2	0.000	0.002	0.004	0.001	0.002	0.001	
A/B_P2_19		1	2	0.013	0.003	0.008	0.002	0.001	0.003	
A2/B_P1_20		3	1	0.004	0.001	1.224	0.158	0.095	0.036	0.00
A2/B_P1_12		4	1	0.001	0.001	0.055	0.013	0.011	0.004	0.01
A2/B_P1_10		4	1	0.001	0.002	0.487	0.046	0.023	0.013	0.00
A2/B_P1_24		1	1	0.012	0.005	0.226	0.059	0.000	0.002	0.05
A2/B_P1_19		1	1	0.001	0.003	0.003	0.001	0.000	0.001	
A2/B_P1_29		1	1	0.000	0.003	0.017	0.007	0.014	0.004	
A2/B_P1_1		1	1	0.155	0.008	3.386	0.041	0.698	0.160	0.05
A2/B_P2_27		N/A	2	0.002	0.001	0.006	0.003	0.001	0.001	
A2/B_P2_2		7	2	0.000	0.001	0.000	0.001	0.000	0.000	
A2/B_P2_15		1	2	0.000	0.000	0.000	0.001	0.000	0.001	
A2/B_P2_8		1	2	0.000	0.001	0.000	0.000	0.000	0.001	
WT Avidin (Phagemid)				0.000	0.000	0.000	0.000	0.001	0.002	
WT BBP-A (Phagemid)				0.000	0.000	0.000	0.000	0.000	0.000	
5µg/ml ChiAVD (Purified protein)				1.827	0.012	2.048	0.001	0.000	0.001	0.89
5µg/ml WT Avidin (Purified protein, Belovo)				1.185	0.006	1.282	0.014	0.005	0.002	0.92
100µg/ml WT Avidin (Purified protein, Belovo)				2.274	0.282	2.031	0.123	0.004	0.001	1.12

Appendix 5: The results of the ABD-assay for the mutants panned according to the protocols A - E.

The average background subtracted results of the ABD-assay performed for the mutants that had been panned according to the panning protocols A, B, D or E are presented in the table below. The standard deviations (SD) of the absorbance values are shown only if multiple identical mutants were tested. The ratio of the biotin-treated samples to untreated samples is shown on the right hand column. The first letter of the names of the mutants refers to the panning protocol used in each case.

Name of the mutant	Previously known as	No. of clones	BTN coating + BTN treatment	SD	BTN coating	SD	BSA coating	SD	Ratio between the biotin-treated and untreated samples
A4_8		1	0.017		0.340		0.014		0.05
B4_2	A/A2-3	11	0.167	0.090	1.129	0.659	0.044	0.025	0.15
B4_15		2	0.016	0.011	0.016	0.005	0.001	0.001	1.00
B4_9		1	0.156		0.198		0.000		0.79
D4_7		2	0.016	0.004	0.092	0.079	0.000	0.003	0.17
D4_15		1	0.015		0.135		0.011		0.11