BARBARA TASKINEN

Protein Engineering of Avidin by
Rational Design and DNA Family Shuffling

■

UNIVERSITY OF TAMPERE

UNIVERSITY
OF TAMPERE

Cover design by
Mikko Reinikka

441 729
Painotuote

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following published articles and a submitted manuscript, which are referred to in the text by their Roman numerals I-IV, and on unpublished data.

I. **Niederhauser, B.**, Siivonen, J., Määttä, J.A., Jänis, J., Kulomaa, M.S. & Hytönen, V.P. 2012, "DNA family shuffling within the chicken avidin protein family - A shortcut to more powerful protein tools", *Journal of Biotechnology,* vol. 157, no. 1, pp. 38-49.

II. **Taskinen, B.**, Airenne, T.T., Jänis, J., Rahikainen, R., Johnson, M.S., Kulomaa, M.S. & Hytönen, V.P., "A novel chimeric avidin with increased thermal stability using DNA shuffling", *Submitted manuscript.*

III. Pollheimer, P., **Taskinen, B.**, Scherfler, A., Gusenkov, S., Creus, M., Wiesauer, P., Zauner, D., Schöfberger, W., Schwarzinger, C., Ebner, A., Tampé, R., Stutz, E., Hytönen, V.P. & Gruber, H. 2013, "Reversible biofunctionalization of surfaces with a switchable mutant of avidin", *Bioconjugate Chemistry,* vol. 24, no. 10, pp.1656-1668.

IV. **Taskinen, B.**, Zmurko, J., Ojanen, M., Kukkurainen, S., Parthiban, M., Määttä, J.A.E., Leppiniemi, J., Jänis, J., Parikka, M., Turpeinen, H., Rämet, M., Pesu, M., Johnson, M.S., Kulomaa, M.S., Airenne T.T. & Hytönen, V.P., 2013, "Zebavidin - An avidin-like protein from zebrafish", *PLOS ONE,* vol. 8, no.10, e77207

ABBREVIATIONS

ABD	avidin-biotin displacement assay
AVD	chicken avidin
att	site-specific attachment site for recombination
AVR	avidin related protein
BBP-A	biotin-binding protein A
BSA	bovine serum albumin
Da	Dalton (mass unit)
DLS	dynamic light scattering
DSC	differential scanning calorimetry
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	ethylenediaminetetraacetic acid
EF1a	elongation factor 1-alpha
ESI	electrospray ionization
FT-ICR	Fourier transform ion cyclotron resonance
ΔG	change in Gibbs free energy
ΔH	change in enthalpy
HBS	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffered saline
dNTP	deoxynucleotide
ITC	isothermal titration calorimetry
ITCHY	incremental truncation for the creation of hybrid enzymes
K_a	association equilibrium constant
K_d	dissociation equilibrium constant
k_{diss}	dissociation rate constant
M	molar concentration (mol/l)
MS	mass spectrometry
NHS	N-hydroxysuccinimide
pIII	minor coat protein of bacteriophage M13
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol

pI	isoelectric point
qRT-PCR	quantitative real time polymerase chain reaction
R	gas constant (8.314 J/molK)
RACHITT	random chimeragenesis on transient templates
ΔS	change in entropy
SAM	self-assembly monolayer
SCRATCHY	combination of ITCHY and DNA shuffling
SDS	sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SHIPREC	sequence homology-independent protein recombination
SLS	static light scattering
SLIC	sequence and ligation independent cloning
SPR	surface plasmon resonance
StEP	staggered extension process
T	temperature

YHTEENVETO

Kanan avidiini ja bakteeriperäinen streptavidiini ovat suosittuja työkaluja bioteknologian, lääketieteen ja teknologian sovelluksissa, koska ne sitoutuvat poikkeuksellisen tiukasti pieneen biotiini-vitamiiniin. Näiden proteiinien toimintaa ja rakennetta on tutkittu laaja-alaisesti jo vuosikymmeniä. Useita uusia avidiinin ja streptavidiinin muotoja on kehitetty käyttäen kohdennettua eli rakenteeseen perustuvaa proteiinien muokkausta. Näin on pyritty vastaamaan jatkuvasti kasvavaan uusien proteiinityökalujen tarpeeseen. Samalla on opittu ymmärtämään näiden korkeita lämpötiloja kestävien proteiinien poikkeuksellisia ominaisuuksia. Lisäksi useista eri eliöistä eristettyjen uusien avidiinin kaltaisten proteiinien rakenteiden ja toiminnan tutkiminen on lisännyt tietoa avidiiniproteiiniperheen ominaisuuksista.

Tässä tutkimuksessa hyödynnettiin avidiiniproteiiniperheen monimuotoisuutta valmistamalla eri avidiinien osia yhdistämällä ns. kimeerinen geenikirjasto. Työssä käytettiin satunnaismutageneesimenetelmää, jossa useiden eri geenien DNA:ta sekoitetaan pilkkomalla ne ensin osiin, ja yhdistämällä syntyneitä palasia sattumanvaraisesti. Toiminnalliset biotiinia sitovat muunnokset valikoitiin valmistetusta geenikirjastosta ns. "faagidisplay" menetelmällä, joka perustuu erään bakteeriviruksen pintaproteiinien hyödyntämiseen. Lupaavimpien, tiukasti biotiinia sitovien muotojen ominaisuuksia tutkittiin monipuolisesti biokemiallisten ja biofysikaalisten menetelmien avulla. Tutkimuksen tuloksena oli kahden uuden lämpökestävän kimeerisen avidiinimutantin, A/A2-1 ja A/A2-B, ominaisuuksien selvittäminen.

Tutkimuksessa sovellettiin aikaisempaa tietämystä avidiiniproteiinien kolmiulotteisesta rakenteesta ja toiminnasta käyttämällä aiemmin tuotettua avidiinin mutatoitua muotoa (AVD M96H) uudelleenkäytettävän biologisia molekyylejä tunnistavan biosensorin valmistamiseen. Tämän muokatun avidiinin käytettävyyttä biosensoreissa parannettiin uusien mutaatioiden avulla. Näin voitiin vähentää epäspesifistä sitoutumista pintoihin sekä hidastaa johonkin toiseen molekyyliin liitetyn biotiinin irtoamista eli dissosiaatiota.

Avidiinien biologista toimintaa ei täysin tunneta. Jotta ymmärrettäisiin paremmin avidiiniproteiinien fysiologiaan liittyviä toimintoja, tutkimuksessa tuotettiin ja tutkittiin seeprakalan avidiinin kaltaista proteiinia, zebavidiinia. Seeprankalan

avidiinin, zebavidiinin, tuotannon keskittyminen kalan sukurauhasiin vahvistaa aiemmin esitettyä hypoteesia avidiinin merkityksestä osana eliön ja kehittyvän alkion puolustussysteemiä. Lisäksi tutkimustulokset osoittivat, että seeprakala on lupaava malliorganismi munivien selkärankaisten eläinten tuottamien avidiinien fysiologisen merkityksen tutkimisessa.

ABSTRACT

Avidin and streptavidin are popular tools used in biotechnological applications due to their high affinity towards the small molecule biotin. The function and structure of these remarkable proteins have been extensively studied in the past decades. Many useful avidin and streptavidin mutants have been created by rational protein design in order to study their extraordinary properties and to fulfill the demand for more specialized protein tools. Furthermore, the characterization of avidin-like proteins in a diverse set of organisms has improved our understanding of the functional and structural properties of members of the avidin protein family, collectively known as avidins.

In this study, we took advantage of the sequence diversity within the avidin protein family and used it to create chimeric mutant libraries constructed by the random mutagenesis method DNA shuffling. Functional biotin-binding mutants were selected from the chimeric mutant library using phage display. Promising candidates were extensively characterized by biophysical methods. This led to the discovery of two thermostable chimeric avidin mutants, A/A2-1 and A/A2-B.

In addition, we made use of available information about the structure and function of avidin, and found that the previously designed avidin mutant AVD M96H was suitable for the design of reusable biosensors based on the avidin-biotin-technology. The mutant was further improved by rational protein design to reduce the observed non-specific binding in biosensor experiments and to lower the dissociation rate of conjugated biotin.

In order to gain a deeper understanding of the physiological function of the avidin proteins, we produced and characterized zebavidin, an avidin-like protein from zebrafish. An analysis of zebavidin expression levels in zebrafish supported the earlier hypothesis that avidin plays a role in the defense system of the developing zebrafish embryo. Furthermore, our results confirmed that zebrafish is a promising organism for studying the physiological function of zebavidin.

1. INTRODUCTION

Evolution is based on the random occurrence of mutations that change the properties of a protein, which in turn increases the fitness of the organism and its chance of survival and propagation. Throughout billions of years of evolution, nature has created a vast variety of protein folds, protein structures and protein functions. Even though evolution is a random process, with most of the occurring mutations being deleterious or in the best case neutral, nature has been able to create diverse and complicated functional protein networks that are able to build, organize and maintain complex multicellular organisms (Darwin 1859, Dawkins 1985).

Many naturally occurring protein functions are suitable for industrial processes, pharmacy and medicine, but the new environment the proteins are transferred into often compromises their structural stability and function. For example, the proteases used in laundry detergents need to have high activity at low temperatures and high stability under alkaline conditions (Kirk, Borchert & Fuglsang 2002); or, the α-amylases used in the production of glucose from starch need to be active at low pH (Shaw, Bott & Day 1999). In other cases however, completely novel functions may be desired (Shao, Arnold 1996, Park et al. 2006). Therefore, scientists have found ways to harness evolution in the laboratory, by guiding the evolutionary process into the desired direction, and reducing the time it takes to create a given property (Arnold 1998a, Jackel, Kast & Hilvert 2008). By developing sophisticated methods to introduce mutations into the amino acid sequence and by designing elaborate screening and selection methods to search for a specific function in the target protein, the time it takes for a protein to evolve has decreased to a mere few weeks or months (Neylon 2004, Lin, Cornish 2002).

Recreating protein evolution in the laboratory is not the only way to alter the properties of a protein. With knowledge of the function and structure of proteins and their homologous sequences, proteins can be manipulated in a rational manner by selectively altering their sequences (Cedrone, Menez & Quemeneur 2000, Marshall et al. 2003, Eijsink et al. 2004).

The avidin protein family consists of a group of proteins that share a highly similar three-dimensional structure composed of four identical eight stranded β-barrels (Livnah et al. 1993). Members of the family can be found in bacteria, fungi and vertebrates and a number of new family members have been found and characterized recently (Nordlund et al. 2005b, Helppolainen et al. 2007, Määttä et al. 2009, Meir, Bayer & Livnah 2012, Takakura et al. 2009). Although their structures are highly similar, members of the avidin protein family show low sequence identities and display modified structural and functional properties. A common property of the family members is that they bind the vitamin D-biotin with high affinity. This phenomenon has made avidins important tools for many biotechnological applications, where the strong immobilization or capture of analyte proteins or other molecules is required (Diamandis, Christopoulos 1991). The need of high biotin-binding properties in a broad range of different applications has increased the demand for avidins with structural and functional characteristics that are not found in the natural proteins. The diverse avidin protein family therefore serves as an excellent starting material for directed evolution experiments.

In this study, we used rational and random mutagenesis to engineer the structural and functional properties of avidin proteins. DNA shuffling was used to create stable chimeric proteins combining sequences of avidin and avidin-related protein 2 (AVR2). Then, to extend the toolbox of evolvable avidin proteins, we studied a previously uncharacterized avidin family member from zebrafish. Finally, we found the avidin mutant AVD M96H, developed previously using rational design (Nordlund et al. 2003), to be suitable for the reversible biofunctionalization of biosensor surfaces.

2. REVIEW OF THE LITERATURE

2.1 Protein engineering by directed protein evolution

In natural evolution, proteins evolve because randomly occurring mutations take place and mutants survive due to the fitness advantage they give to the organism. The same principle is applied when directed protein evolution is used to engineer proteins. Genes encoding proteins are subjected to random mutagenesis and the individual carrying the desired function is detected using a specifically designed selection or screening method. The selected individuals are subsequently amplified to be used in the next round of mutagenesis and selection (Figure 1) (Jackel, Kast & Hilvert 2008). In contrast to natural evolution, where selection happens randomly, in directed evolution the fittest strain can be selected and is therefore directed (Arnold 1998a).

The manipulation of a protein's function by altering its sequence is commonly referred to as protein engineering. Protein engineering is an important tool in different fields of research and industries. It has been used, for example, in the biocatalyst industry to improve the stability and activity of the serine endoprotease substilisin, a common additive in laundry detergents (Ness et al. 1999); in medicine to improve the antiviral and antiproliferation activities of human alpha-interferon in murine cells (Chang et al. 1999); in agriculture to increase the expression levels of the *Bacillus thuringiensis* endotoxin Cry, and to widen its range of specificity towards different insects. (Lassner, McElroy 2002). In basic research protein engineering is used to study protein evolution (Bershtein, Goldin & Tawfik 2008, Cochran et al. 2006, Fasan et al. 2008).
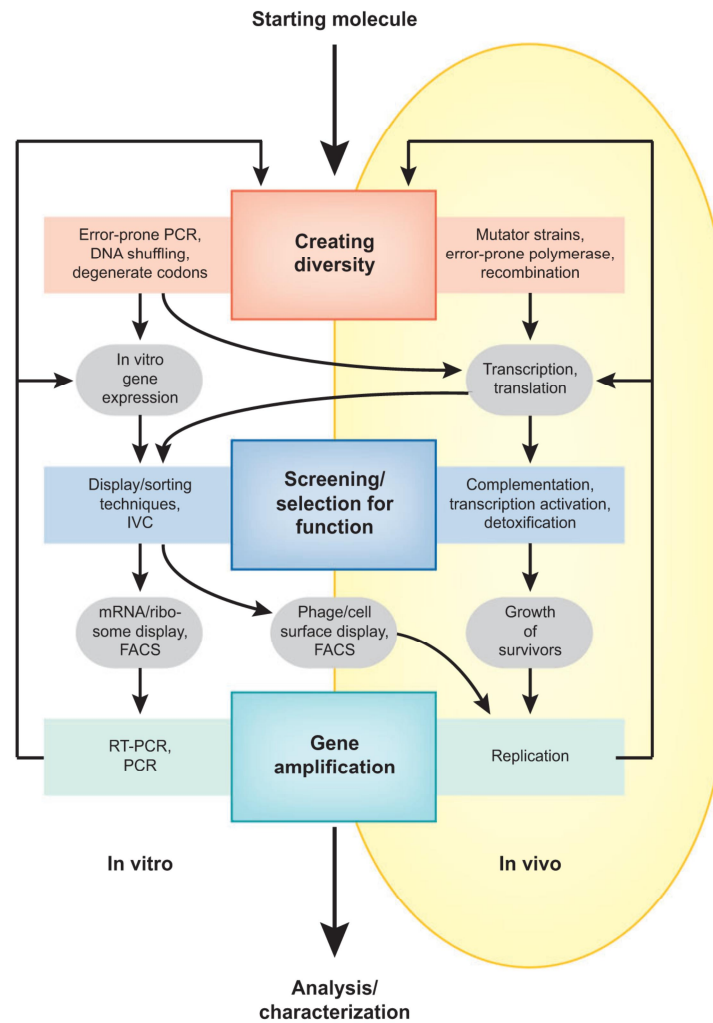
Figure 1. Schematic overview of the processes, strategies and techniques in directed molecular evolution. Figure reprinted by permission from (Jackel, Kast & Hilvert 2008).

2.1.1 Sequence space and fitness landscape

A sequence space is an abstract concept representing all the possible amino acid combinations at each position of a sequence, where each point of the space represents an individual sequence. A sequence space is usually depicted by a two-dimensional graph. In the third dimension, the fitness of each sequence in the sequence space is plotted based on its fitness against a selection pressure (Figure 2A). The goal of every directed evolution experiment is to search through the sequence space to find the highest point - the peak of the tallest mountain - in the fitness landscape (also referred to as the global maximum). A fitness landscape can

either have one single global maximum (Figure 2B, right side), or it can be rugged and have several peaks representing local fitness maxima (Figure 2B, left side). It is more difficult to find the fittest sequence in a rugged fitness landscape than in a smooth landscape (Romero, Arnold 2009).

Most of the sequence space is void of any known function, or does not even code for the desired protein (Arnold 1998a, Axe 2004), and functional sequences are usually close to each other (Axe 2004). Additionally, technical aspects, such as transformation efficiency and the amount of available mutant DNA, limit the sequence space of a mutant library. For example the sequence space for a 100-residue protein consists of 20^{100} (1.3×10^{130}) possible sequences. (As a comparison, the known universe is estimated to consist of about 10^{80} atoms). Thus, the amount of DNA needed for each sequence to be present at least once would be incomprehensibly large (Taylor et al. 2001). It is therefore important to carefully
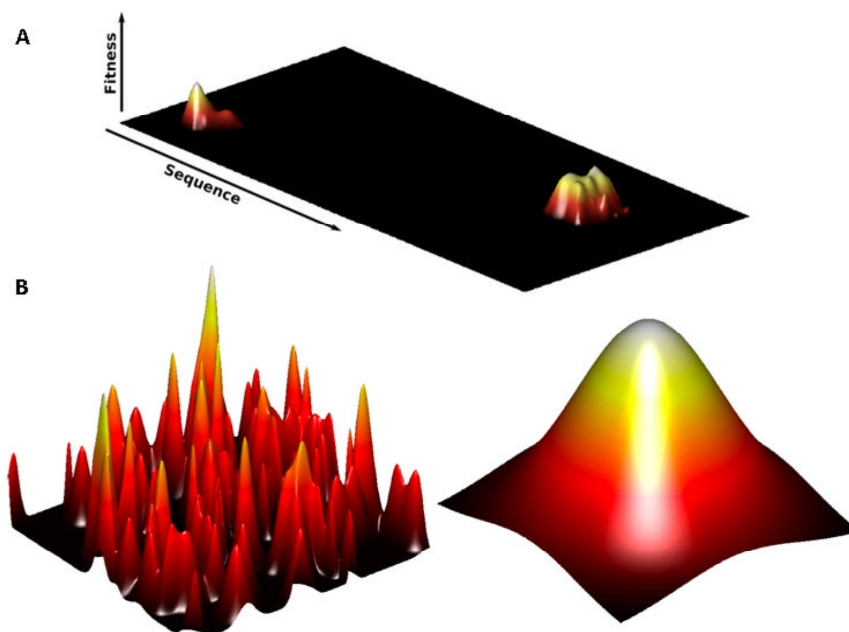
Figure 2. Sequence space and fitness landscape. (A) The plot of fitness against the sequence creates the landscape for evolution. (B) Landscapes can range from the rugged badlands landscape, which is nearly impossible to climb by mutational steps, to the Fujiyama landscape, where any beneficial mutation brings the search closer to the optimum. Figure reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology (Romero, Arnold 2009), copyright (2009).

16

choose the library design and selection method; the sparser the sequence space is populated, the more difficult and more unlikely it is to find the fittest possible sequence (Romero, Arnold 2009).

There are two main approaches for searching a sequence space. In the "greedy" approach, the researcher mutates the starting protein and subjects it to several rounds of selection with increased stringency to select for the fittest candidate from the mutant library. This fittest candidate is then mutated further and subjected to new rounds of selection. This approach works well for fitness landscapes with only one smooth peak. In a rugged fitness landscape, however, it carries the risk that in the first selection round, a candidate that lies on a local fitness maximum is selected, and from there, it is impossible to find the globally fittest available sequence in the sequence space. The "non-greedy" approach is to start the search with low stringency and select for a set of sequences that lie on several local fitness maxima (Smith, Petrenko 1997).

2.1.2 Protein families and conserved domains

A protein family is defined as a group of proteins that share a common ancestor. Even though members of a protein family can have low sequence similarities, they usually share a common three-dimensional structure and in some cases also a common function (Dayhoff et al. 1975, Murzin et al. 1995, Demuth, Hahn 2009). Members of the family either arose when a parental species divided into two ancestral species, allowing the proteins to accumulate mutations independently during divergent evolution, or they developed through gene duplication in one species; having two copies of the same gene meant that one copy could be freely mutated, while the other copy performed the original function (Dayhoff et al. 1975).

Protein families provide information on the evolutionary steps of these proteins, from a simple single cell organism to highly organized vertebrates, and can give us an idea about hidden "promiscuous" functions, which could be exploited by directed evolution (O'Brien, Herschlag 1999, Bloom et al. 2007). Because protein families already share a long common history of evolution, deleterious mutations have already been tested and discarded, making their combined sequence space more enriched with functional sequences (Crameri et al. 1998, Arnold 1998b). For this

reason protein families are an optimal starting material for directed evolution (see chapter 2.2.3).

Protein domains are independently folded parts of proteins, which display different structural folds and functions. Similar to proteins, they can be divided into families according to their structure and function (Ponting, Russell 2002). More than two thirds of all proteins are made up of two or more domains and the duplication and rearrangement of different domains within the same protein is thought to have led to new protein families with new functions and to have aided the development of complex cellular processes (Vogel et al. 2004, Chothia et al. 2003, Moore et al. 2008).

2.2 Construction of mutant libraries

A gene is mutated by changing one or more bases in its DNA sequence, which leads to a change in the amino acid sequence it encodes for. A DNA mutant library consists of a population of mutated genes. Different random mutagenesis methods are available for creating DNA mutant libraries (Figure 3). These methods yield libraries with different numbers of variants. Depending on the new function being searched for, one method can be more suitable than the other (For a detailed list of mutagenesis methods see Neylon 2004).

2.2.1 Random methods

Random methods introduce mutations randomly over the whole length of the gene (Figure 3, random methods). Caution must be taken to not introduce more than a few mutations at once per gene, since most random mutations tend to have negative effects on the protein's fold or functionality (Guo, Choe & Loeb 2004). Methods to introduce mutations by random methods include chemical deamination/methylation, the use of a DNA polymerase without a proofreading function, or the replication of the DNA in a mutator strain (Neylon 2004). For most of these
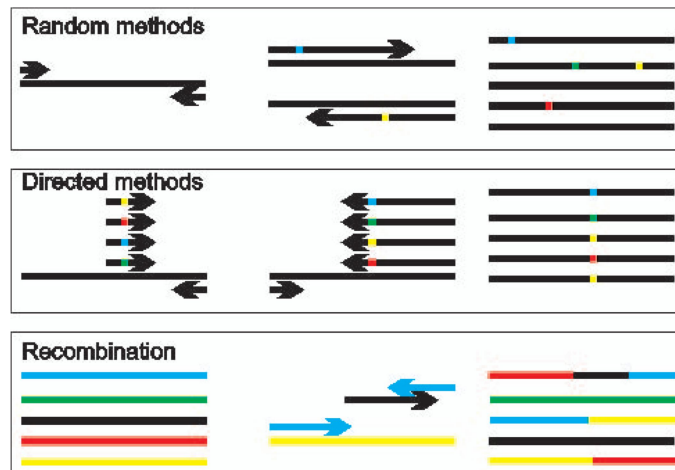
Figure 3. Overview of methods for the randomization of DNA sequences. Figure reprinted by permission from (Neylon 2004)

approaches the number of mutations introduced into a gene cannot be easily controlled. The use of a polymerase without a proofreading activity offers the most controlled way of introducing mutations, since the mutation rate of a polymerase can easily be controlled by the buffer conditions, the concentrations of dNTPs and the presence of different ions (Beckman, Mildvan & Loeb 1985, Cadwell, Joyce 1992).

One advantage random methods have over rational protein design (see chapter 2.5) is that no knowledge about the three-dimensional structure of the protein is required. However, the major disadvantage of random methods is that most of the introduced mutations are deleterious or neutral, and therefore, the mutant library contains a significant amount of non-functional or functionally neutral mutants.

2.2.2 Oligonucleotide-based methods

If the aim is to introduce mutations in a more directed manner, they are only introduced into a short stretch of the sequence, which can be as short as one randomized amino acid residue. This, however, requires knowledge of the protein's function and structure (Figure 3, directed methods). Antibody engineering, for example, usually focuses on the randomization of the region that determines complementarity i.e. the antigen binding site (Laffly et al. 2008, Kim et al. 2011). This localized randomization is accomplished by introducing short synthetic

randomized oligonucleotides where any position can be any of the four bases (Neylon 2004, Wang et al. 2006). The disadvantage of this approach is the introduction of stop codons. However, this can be avoided by limiting the randomization to the first two nucleotides in a codon. But this approach then limits the amount of possible types of amino acids that can be included (Neylon 2004). The safest option is to use synthesized trinucleotides (Virnekas et al. 1994), which guarantees that no stop codons are introduced and all possible amino acids are used. This, however, is also the most expensive option.

2.2.3 Homologous recombination

Random mutagenesis usually involves many rounds of mutation and selection. In the first round, the fittest sequences are selected and these sequences then go through a second round of mutation, in the hope that the new mutations will combine well with the previously introduced mutations to further improve the function of the protein. To facilitate this process, Stemmer (Stemmer 1994) developed a method that allows the combination of the selected advantageous mutations by a method called DNA shuffling. Here, two or more mutated sequences are cut by DNaseI into small fragments and subsequently reassembled into full length genes by a primer-less PCR. During the thermal cycling, fragments from one mutant gene base-pair with fragments from a second mutant gene and during this process combine the mutations from the two sequences (Figure 3, recombination and Figure 4, left). In contrast to applying a second round of random mutagenesis that may introduce advantageous, neutral, or deleterious mutations, by using DNA shuffling, mutations that were found to be functional are recombined in one gene and deleterious mutations are discarded, increasing the functionality of the mutant library (Stemmer 1994).

The same principle of DNA shuffling in randomly mutated genes can be applied to homologous genes that arose naturally over thousands and millions of years. Mutations found in homologous genes have already been tested by nature and are, in most cases, neutral or advantageous. Thus using homologous genes in DNA shuffling allows the recombination of more than just a handful of mutations to

20
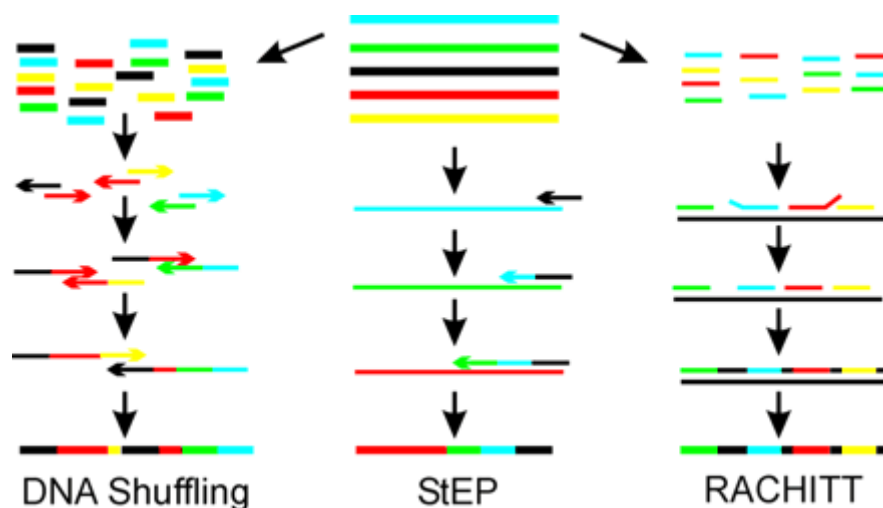
Figure 4. Homology based methods for recombining DNA sequences. Figure reprinted by permission from (Neylon 2004).

create functional genes. When natural homologous sequences are used in DNA shuffling, the method is called DNA family shuffling (Crameri et al. 1998). Importantly, DNA family shuffling samples a significantly bigger portion of the sequence space than is possible with random mutagenesis. Also, in contrast to rational protein design and directed random mutagenesis methods, no knowledge about the structural information of the protein is required. Methods that aim to refine and improve DNA shuffling include Staggered Extension Process (StEP), where the genes are shuffled by annealing to different templates during PCR (Zhao et al. 1998) (Figure 4, middle) or RAndom CHImeragenesis on Transient Templates (RACHITT), where fragmented DNA is annealed to a homologous template DNA; overhanging flaps are digested and gaps are filled by enzymatic reactions (Coco et al. 2001) (Figure 4, right). Furthermore, some studies have employed restriction enzymes instead of DNaseI for the fragmentation of the DNA (Kikuchi, Ohnishi & Harayama 1999, Engler et al. 2009). However, DNA family shuffling is still the favored method for randomly recombining homologous sequences.

2.2.4   Non-homologous recombination

The limitation of homologous recombination is the need for a certain degree of homology to accomplish base-pairing of the homologous regions during reassembly. Numerous methods have been developed to overcome this disadvantage. Methods

like ITCHY and SCRATCHY are able to recombine genes without the need of homologous regions (Ostermeier, Shim & Benkovic 1999, Lutz et al. 2001). ITCHY stands for Incremental Truncation for the Creation of HYbrid enzymes and involves an exonuclease III digestion to create random length 5'- and 3-fragments of two parental genes, followed by ligation to create chimeric genes with one cross-over (Ostermeier, Shim & Benkovic 1999). The method was later improved by cloning both genes in series into the same vector, bridged by a restriction site. The genes are amplified in the presence of the nucleotide analogue α-phosphothioate, which protects the DNA from endonuclease III digestion at its position of integration (Figure 5A). (Lutz, Ostermeier & Benkovic 2001). SHIPREC is a method similar to second generation ITCHY, however, the gene- dimer is not cloned into a vector, but digested with DNaseI and fragments of the size of wild type genes are ligated. The resulting circular DNA is then linearized by cleavage at the restriction site in between the two parental genes (Figure 5B) (Sieber, Martinez & Arnold 2001). A common feature of ITCHY and SHIPREC is that they can only create chimeric genes with a single cross-over. This disadvantage was overcome in ITCHY by applying traditional DNA shuffling using the functional chimeric genes, created by incremental truncation, after one round of selection. The method was humorously named SCRATCHY (Lutz et al. 2001).

A completely different approach is synthetic shuffling. It dismisses the use of whole wild type genes and instead designs a set of degenerated oligonucleotides that contains all the sequence diversity of a set of genes. The oligonucleotides are combined by PCR into full-length chimeric genes (Figure 5C) (Ness et al. 2002). Rather than recombining random gene fragments, exon shuffling looks at how recombination is done in nature and can recombine entire domains. The exons are amplified with chimeric oligonucleotides that define which exons are combined. The reassembly of the exons is then achieved by a primer-less PCR (Figure 5D) (Kolkman, Stemmer 2001).

22
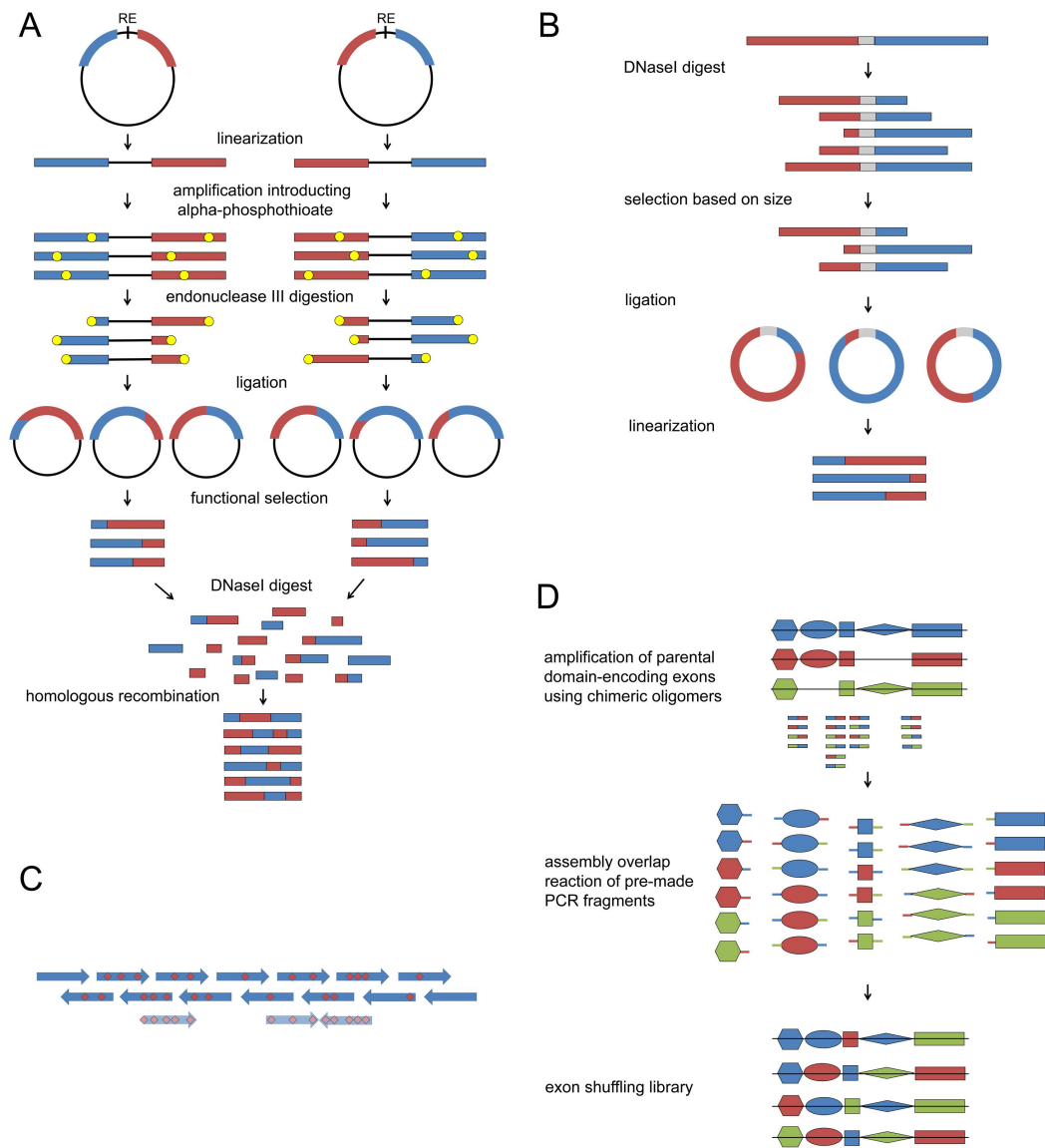
Figure 5. Methods for non-homologous recombination. (A) ITCHY and SCRATCHY. Parental genes, represented by blue and red rectangles, are cloned into a vector interrupted by a linker sequence containing a restriction site. The vector is linearized at the restriction site located in the linker sequence. Genes are amplified and α-phosphothioate (yellow circles) is incorporated. The amplified, α-phosphothioate containing genes are digested with exonuclease III. Exonuclease III activity is stopped by α-phosphothioate. The vector is ligated by blunt-end ligation and functional selection is applied to the variants. Selected genes are digested by DNaseI and chimeric genes are assembled by homologous recombination, resulting in chimeric genes with more than one cross-over (Lutz et al. 2001); (B) SHIPREC. Two parental genes are fused over a linker sequence. The fusion gene is digested by DNaseI and selection of variants is

(*Figure 5* continued) made according to the size of parental genes. The gene is circularized by ligation and linearized by restriction digest in the linker sequence. The resulting chimeric genes can now be cloned into an appropriate vector for functional selection (Sieber, Martinez & Arnold 2001); (C) Synthetic shuffling. Sequence diversity of a set of homologous genes is encoded in short degenerated oligonucleotides (blue arrows; red diamonds display degenerated positions). The oligonucleotides are used to assemble full length chimeric genes. Spiking oligonucleotides (light blue arrows) are added for additional diversity (Ness et al. 2002); (D) Exon shuffling. Exons encoding domains are amplified from a set of parental genes using chimeric oligonucleotides. The oligonucleotides add overlapping sequences specific for different parental genes. The exon fragments are assembled in a PCR reaction into full length chimeric genes, containing domains from different parental genes (Kolkman, Stemmer 2001).

2.2.5 Computational tools for library design

Mutant libraries designed by random mutagenesis or recombination are low in functionality and may be void of the desired function. More and more methods are therefore being created to bring the rational back into the design of mutant libraries by combining random mutagenesis and computational algorithms.

As discussed previously, the sequence space is only sparsely populated with sequences coding for functional proteins. One reason for this is that many mutations destabilize the protein fold, resulting in unfolded or misfolded inactive proteins. Computer algorithms have been developed to overcome this challenge. For example, protein design automation narrows down the sequence space to contain only properly folded sequences by determining the sequences that most likely form a pre-determined protein fold (Dahiyat, Mayo 1996, Hayes et al. 2002). Voigt and co-workers developed an algorithm which, based on a structural model, defines the amino acid residue positions that are more likely to tolerate substitutions and lead to an improvement in protein fitness. With the algorithm, the region of the protein that needs to be subjected to mutagenesis can be narrowed down (Voigt et al. 2001). The algorithm was developed under the observation that the more an amino acid residue interacts with other residues, the less likely it is that a substitution in it will lead to an increase in the protein's fitness. In recombination experiments, the risk of disturbing functional units is even more pronounced (e.g. interactions between residues from different parts of the polypeptide chain are disrupted by recombining

24

gene fragments). SCHEMA is a structure guided computer algorithm that defines interacting amino acids. If these interactions are disrupted due to recombination, the disruption score increases. Sequences with low disruption scores are more likely to form a properly folded chimeric protein than sequences with high disruption scores (Voigt et al. 2002). However, the disadvantage of all the discussed methods is the need for structural information on the protein.

Richard Fox created an algorithm that requires no structural information. The algorithm is fed with experimentally collected data from a directed evolution experiment and calculates a theoretical fitness landscape for the analyzed sequence space. The algorithm defines positions that can be mutated in the next round of directed evolution in order to increase the overall fitness of the tested population (Fox et al. 2003, Fox 2005).

As discussed previously, the successful generation of cross-overs by homologous recombination is limited by the need of regions with high sequence identity. The two computer algorithms eCodonOpt (Moore, Maranas 2002) and cross-over optimization for DNA shuffling (CODNS) (He, Friedman & Bailey-Kellogg 2012) increase the probability of the shuffling of parental sequences with low sequence identity by optimizing codon usage, consequently increasing the occurrence of regions of high homology.

2.3   Mutant library screening and selection

The selection method has to be designed individually for each new protein function. One major challenge in directed protein evolution is to link the sequence information found in the DNA with the properties of the protein it encodes. A vast selection of different methods has been developed for screening DNA mutant libraries and for selecting mutants (Lin, Cornish 2002, Zhao, Arnold 1997). Screening a library means probing the function of preferably all or at least a selected part of the individual variants in a library and comparing their function to the wild type function. This is a tedious approach and mostly needs to be applied for the screening of enzymatic functions. Searching for the fittest variant is easier when the desired function is the affinity of a protein towards a ligand or another protein. In this case, the whole library can be subjected to selection. Examples are phage

display (Smith, Petrenko 1997, Smith 1985), ribosome display (Hanes, Pluckthun 1997), mRNA display (Roberts, Szostak 1997), yeast display (Boder, Wittrup 1997) and bacterial display (Daugherty 2007). The following section will focus on phage display, which is the most commonly used selection method.
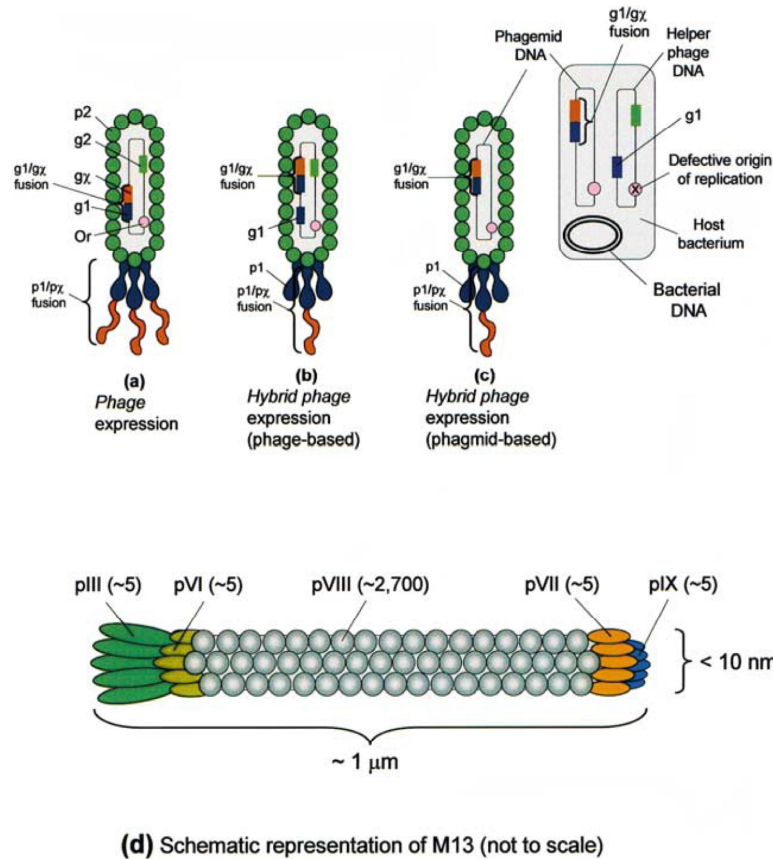
Figure 6. Phage display formats. (A) A mutated gene (gχ) is cloned in fusion with a coat protein gene (g1) directly into the phage genome. Only fusion proteins (p1/pχ) are displayed (B) The phage genome contains the g1/gχ fusion and the g1 gene. The wild type protein p1 and the fusion protein p1/pχ are displayed. (C) The fusion gene g1/gχ is available on a phagemid in addition to the wild type gene g1 in the phage genome. (D) Schematic representation of the M13 bacteriophage. The phage capsid consists of five different coat proteins (pIII, pVI-pIX). The coat proteins pIII and pVIII are mainly used to display the protein of interest. Figure reprinted by permission from (Willats 2002)

2.3.1   Phage display: principles of affinity based selection

In phage display, the DNA sequence information is linked to the protein's function by a bacteriophage that displays the mutated gene as a fusion protein on its surface. The most commonly used bacteriophages for phage display are filamentous phages like M13, fd and f1, which will be discussed here (Figure 6D). More recently developed phage display systems are based on the bacteriophage T4 (Ren et al. 1996) and the lambda phage (Sternberg, Hoess 1995).

In practice, a DNA mutant library is introduced into the genome of the phage by fusing it to one of the phage's five coat proteins, the most commonly used ones being p3 and p8. The translated protein is then displayed as a fusion protein on the surface of the phage particle and can bind to an immobilized ligand. Unbound phages are washed away and binding phages are amplified in *Echerichia coli* bacteria, and are then subjected to a second round of biopanning or sequenced and further characterized (Figure 7) (Smith, Petrenko 1997, Smith 1985, Willats 2002). Usually, three to five rounds of biopanning are carried out to select for the best
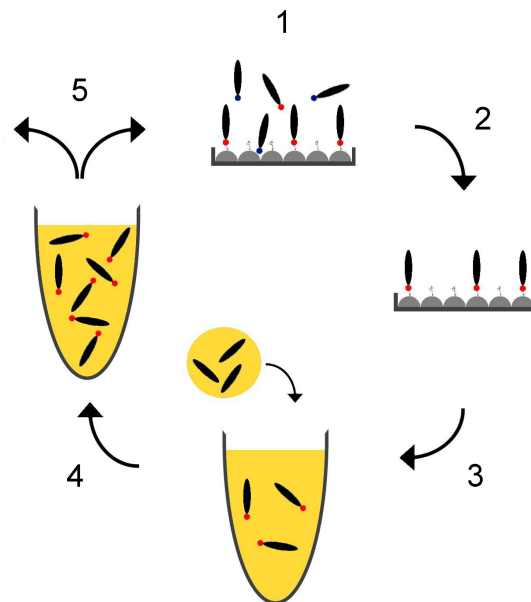
Figure 7. Phage display panning cycle. 1: The phage library is panned against an immobilized target ligand; 2: non-binders are washed away; 3: phages are eluted and added to an *E. coli* cell culture; 4: M13 helper phages are added to enrich positive binders; 5: The phages are used in a next round of panning or analyzed.

binders.

There are generally three ways to generate the phage library. The mutated gene is cloned directly as a fusion protein into the phage genome (Figure 6A); the chosen surface protein is present twice in the phage genome, once in fusion with the mutated gene and once as the wild type gene (Figure 6B); or the mutated gene in fusion with the chosen surface protein is in a separate phagemid plasmid and is inserted into the phage capsid together with the wild type surface protein encoded in the phage genome (Figure 6C). In the first approach, all the coat proteins are present as fusion proteins. This approach may be chosen if the affinity of the mutant proteins towards the ligand selected for is expected to be low. However, the fusion protein might compromise the functionality of the coat protein and therefore the viability of the phage. The other two approaches circumvent this problem by co-displaying the wild type coat protein. A low copy number of displayed proteins increases the selection pressure towards higher affinity binders (Smith, Petrenko 1997, Willats 2002).

2.3.2 Phage display for functions other than affinity

Although traditionally used to select for enhanced binding affinity, phage display can be applied to select for other functions, too. For example, it can be used to identify variants with enhanced stability by incubation at high temperatures or denaturation in guanidine hydrochloride (Jung, Honegger & Pluckthun 1999, Jung, Honegger & Pluckthun 1999). However, this approach still depends on an affinity based function. The Proside (protein stability increased by directed evolution) technology was developed to solve this problem: a phage display method is used to select for protease resistant variants independent of protein function (Sieber, Pluckthun & Schmid 1998). The surface protein pIII, commonly used as the fusion partner in phage display, consists of three domains; N1, N2 and CT. N1 and N2 are required for the phage to infect the host bacteria (Barbas III et al. 2001). In Proside, the gene of interest is cloned between CT and N2 and the fusion protein displaying phages are subjected to proteolysis. Variants with low stability against the protease lose their infectivity and disappear from the population; only the variants that resist

proteolysis are able to infect new host cells and propagate (Sieber, Pluckthun & Schmid 1998).

Phage display methods for the selection of enzyme functions are based on trapping the enzyme with suicide inhibitors (Droge et al. 2003, Droge et al. 2006) or transition state analogues (Baca et al. 1997) immobilized on a solid surface. The drawback of these methods is the difficulty to find suitable suicide inhibitors or transition state analogues. Further, selection is not based on the turnover rate of the enzymatic reaction, and thus additional screening after selection is required (Forrer, Jung & Pluckthun 1999). Two related methods that are based on turnover rate rely on the idea that the enzyme and substrate are brought into close proximity. Besides the engineered enzyme, the phage also displays a second pIII fusion protein, which depending on the method, is either a basic peptide (Pedersen et al. 1998) or the protein calmodulin (Demartis et al. 1999). The substrate of interest is then either fused to an acidic peptide that can form dimers with the displayed basic peptide, or to a calmodulin-binding peptide that binds to the displayed calmodulin. As a result, the substrate is covalently bound to the phage in close proximity to the enzyme. The product of the enzymatic reaction is then used to select for an active enzyme. The product is either bound to or released from a solid surface in the event of the enzymatic reaction, or alternatively, it is isolated using a product specific antibody (Pedersen et al. 1998, Demartis et al. 1999).

2.4   The future of directed protein evolution

Synthetic biology is an emerging field that employs directed protein evolution to manipulate whole cells to make them carry out novel functions (Khalil, Collins 2010). This can mean, for example, the engineering of whole metabolic pathways for the improved production of chemical compounds (Chatterjee, Yuan 2006). One method to accomplish this is genome shuffling, where the entire genome of a population of bacteria is shuffled and evolved, allowing the manipulation of the whole organisms (Zhang et al. 2002).

The development of high-throughput sequencing (HTS) platforms, that allow the generation of a vast amount of sequence information in a short amount of time with low costs, have facilitated the screening of mutant libraries (Metzker 2010). HTS

has been used, for example, to create a complete map of the function-structure relationship of the WW domain binding to its peptide ligand and to identify stabilizing mutations (Fowler et al. 2010, Araya, Fowler 2011), to generate a fitness landscape for the chaperone protein Hsp90 (Hietpas, Jensen & Bolon 2011), to analyze a complete phage display peptide library (Matochko et al. 2012), and to engineer the affinity of an inhibitor against H1N1 influenza hemagglutinin (Whitehead et al. 2012). The detailed sequence analysis of a mutant library in the course of a directed evolution experiment provides insights into the evolutionary path of the wilt type protein, gives a better understanding of its sequence-function relationship and helps to improve the design of an evolutionary experiment.

2.5   Protein engineering by rational design

The traditional way of engineering proteins is the introduction of mutations to specific amino acid residues at selected positions in the protein. This approach commonly requires knowledge about the structure and function of the protein. While rational protein design appears rather cumbersome in comparison with directed protein evolution, it has been used successfully in the past to alter the structural and functional properties of a number of proteins (Cedrone, Menez & Quemeneur 2000, Marshall et al. 2003, Eijsink et al. 2004). For example, the thermostability of a phytase was improved by selecting mutation sites using primary protein sequence comparison with homologous proteins (Lehmann et al. 2000). Similarly, chicken avidin was stabilized by changing a single residue based on sequence comparison with the highly stable homologous avidin related protein 4 (Hytönen et al. 2005a). The enantioselectivity of a lipase was improved by detecting unfavorable interactions between the enzyme and substrates using molecular modeling and subsequently applying site-directed mutagenesis to modify the enzyme (Rotticci et al. 2001).

   Especially when the function and structure of a protein are well known, rational protein design can be successfully used to engineer the desired, new function. The numerous mutational studies on avidin and streptavidin demonstrate the power of rational design when enough information on the protein is available (reviewed in Laitinen et al. 2006). Avidin engineering is further discussed in chapter 2.6.3.

Another example of a widely engineered protein is the green fluorescent protein, which has been structurally and functionally characterized. This enables the targeted engineering of the residues close to the chromophore to tune its emission spectra (Tsien 1998, Remington 2011).

2.6   The avidin protein family

Chicken avidin is a minor constituent of egg white and shows high affinity towards biotin (Green 1975). It was first discovered when dogs fed with raw eggs suffered from diarrhea (Bateman 1916). Later, a highly similar protein, streptavidin, was discovered in the bacterium *Streptomyces avidinii* (Chaiet, Wolf 1964). Both avidin and streptavidin are homotetrameric proteins made up of four eight stranded antiparallel β-barrels. Each monomer contains a biotin-binding site, and the sites are positioned on opposite sides of the tetramer (Figure 8) (Livnah et al. 1993, Weber et al. 1989). Biotin is a small (MW = 244.31 g/mol), soluble vitamin mainly required as a cofactor in the biosynthesis of amino acids and fatty acids and in gluconeogenesis. Biotin deficiency can lead to hair loss, dermatitis and neurological symptoms. The main sources of biotin are dietary components and the metabolites of the microflora in the large intestine. Biotin is absorbed through intestinal epithelia cells (Said et al. 1998).

Their ability to tightly, and almost irreversible bind biotin make avidin and streptavidin, often collectively called (strept)avidin, popular tools in biotechnology and life sciences. Different forms of (strept)avidin are used in applications such as affinity purification, biosensors, drug targeting, molecular diagnostics and cell labeling (Diamandis, Christopoulos 1991, Wilchek, Bayer 1989, Laitinen et al. 2007).

2.6.1   Members of the avidin protein family

Besides avidin, the chicken genome also contains a set of homologous proteins also able to tightly bind biotin. These proteins are collectively called avidin related proteins 1-7 (AVR 1-7) (Keinänen et al. 1994, Wallen, Laukkanen & Kulomaa
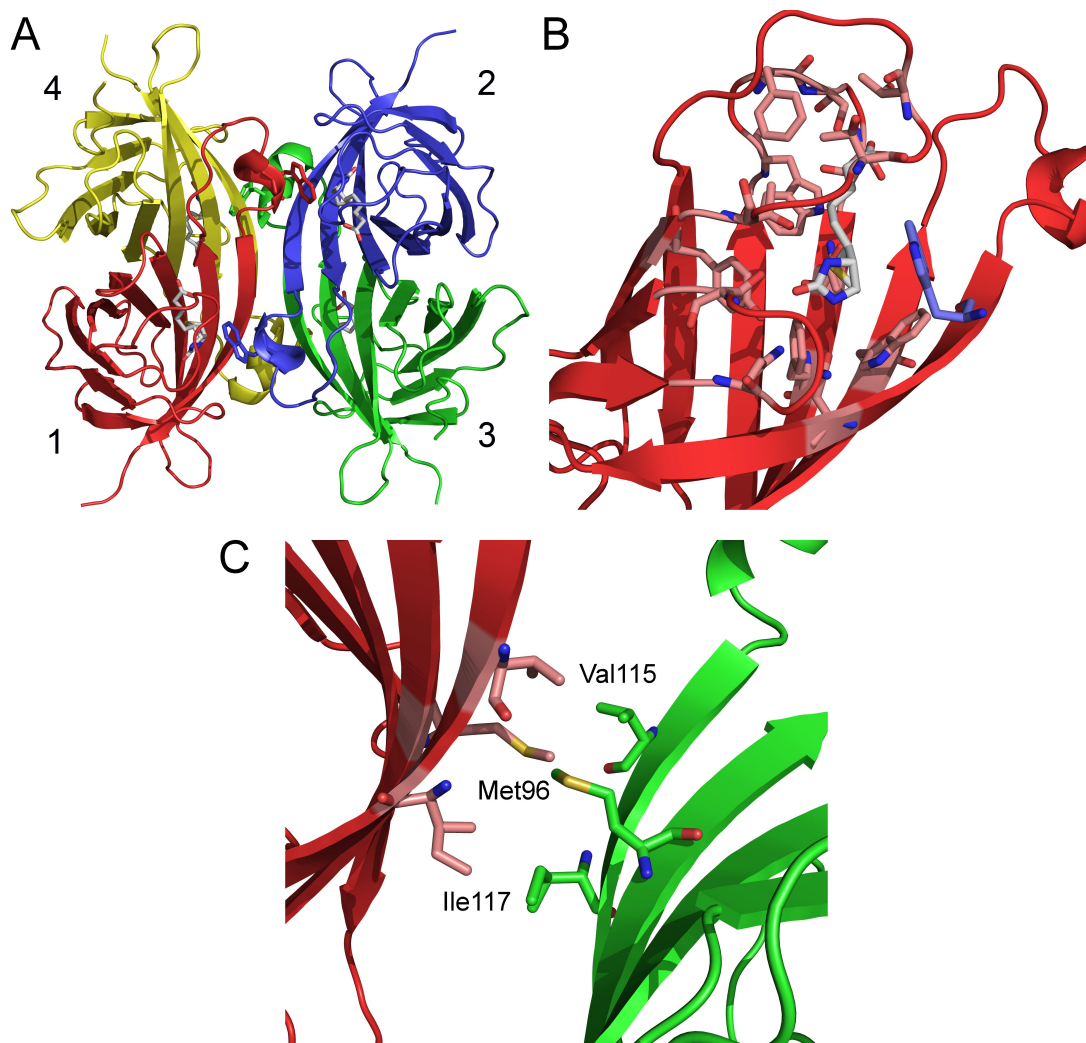
Figure 8. Structure of avidin. (A) Cartoon representation of tetrameric avidin [PDB: 1AVD] (Pugliese et al. 1993). Subunits 1-4 are numbered according to (Livnah et al. 1993) (B) Close-up view of the biotin-binding site of monomer 1. (C) Close-up view of the 1-3 subunit interface. Bound biotin and residues interacting with biotin are shown as sticks. Carbon atoms are shown in grey (biotin) or light pink and green (avidin residues), sulfur atoms in yellow, oxygen atoms in red and nitrogen atoms in blue. Trp110 from subunit 2 is shown in blue. Figure prepared using PyMOL.

1995, Hytönen et al. 2005, Hytönen et al. 2004). Not much is known about the physiological roles of AVRs. Studies in chicken have shown that their mRNAs are expressed, but it is unclear whether they are translated into proteins at all (Kunnas, Wallen & Kulomaa 1993). In addition to avidin and AVR proteins, chicken have also one or more biotin-binding proteins, which are located in the egg yolk and are

responsible for transporting biotin to the egg yolk for the developing embryo (White et al. 1976, Meslar, Camper & White 1978, Subramanian, Adiga 1995, Bush, McGahan & White 1988, Niskanen et al. 2005, Hytönen et al. 2007). Avidin like-proteins have also been characterized from other bacterial species (e.g. rhizavidin, bradavidin, shwanavidin) (Nordlund et al. 2005b, Helppolainen et al. 2007, Meir, Bayer & Livnah 2012, Helppolainen et al. 2008), frog (xenavidin) (Määttä et al. 2009), and fungi (tamavidin) (Takakura et al. 2009). These proteins are collectively referred to as the avidin protein family (Hytönen 2005).

Despite the low sequence identity between the members of the avidin protein family, all the characterized avidins are structurally similar and the residues responsible for binding biotin are highly conserved (Figure 9). With the exception of rhizavidin, shwanavidin and bradavidin II, all the characterized proteins form homotetrameric complexes. Rhizavidin and shwanavidin are dimeric proteins in the presence and absence of biotin, and bradavidin II has a highly dynamic oligomeric structure (Meir, Bayer & Livnah 2012, Meir et al. 2009, Leppiniemi et al. 2013). Homology searches have also revealed the presence of many more uncharacterized avidin-like proteins in other organisms. (unpublished data, personal communication with Vesa Hytönen and Tomi Airenne).

Differences between the avidin proteins can be seen in their functional and physicochemical properties. The proteins differ in their isoelectric points, their affinity towards biotin and other small ligands, thermal stability, and their resistance against proteases.

2.6.2 Physiological function of avidin

The physiological role of avidin has been studied in chicken. However, its exact function still remains a mystery. Avidin expression is induced by progesterone in the chicken oviduct; no avidin expression is observed in the liver or shell glands (Tuohimaa, Segal & Koide 1972). Progesterone independent avidin expression is observed during microbial or viral infection and tissue injury, and in these cases expression is also seen in the liver, plasma, intestine and muscle tissue (Kunnas, Wallen & Kulomaa 1993, Elo, Raisanen & Tuohimaa 1980, Korpela et al. 1982, Korpela et al. 1983).

```
                    *          20          *          40          *          60
bradavidin  : QS--------VN-----------WTWTNQYGSTIAITSFNSNTGAITGTYTNNAANSCD--- :  40
tamavidin2  : SD--------VQSS------LTGTWYNELNSKVELTANK--DGTLTCKYLSKVG---D--- :  39
streptavidin: -----------EAG------ITGTWYNQLGSTFIVTAGA--DGALTGTYESAVG---N--- :  36
avidin      : -----------K-----CSLTGKWTNDLGSNMTIGAVN-SRGEFTGTYITAVTATSN--- :  40
zebavidin   : qt--------vSS-----CNVTGVWRNELGSTIRVKAEG---SEVRGVYQTAVESTRGAAG :  45
AVR2        : -----------qtvARKCSLTGEWDNDLGSIMTIGAVN-DNGEFDGTYITAVADNPG--- :  45
rhizavidin  : NFKDFSSIASAS----------SSWQNQSGSTMIIQVDSF--GNVSGQYVNRAQGTGC--- :  46
                                     W  N   gS     6        g   G Y

                    *          80          *         100          *         120
bradavidin  : EGKPQGVTGWLAY--GN----TGTAISFSVNBL----G--CGSTIVWTGQLNN----ATGF :  85
tamavidin2  : VYVPYPISGRYNL--QPP-AGQGVALGWAVSWE----NSKIHSATTWSGQFFSE--SSPVI :  91
streptavidin: AESRYVLTGRYDS--APATDGSGTALGWTVAWKNNYRN--AHSATTWSGQYVGG--AEARI :  91
avidin      : EIKESPLHGTQNTINKR----TQPTFGFTVNWK----F--SESTTVFTGQCFIDRNGKEVL :  91
zebavidin   : HHRSARIIGMVSD---G----TQPTVSFSVLWE-------KGSCSAWVGQCFILDDGAQVL :  92
AVR2        : NITLSPLLGIQHK--RA----SQPTFGFTVHWN----F--SESTSVFVGQCFVDRSGKEVL :  94
rhizavidin  : QNSPYPLTGRVNG----------TFIAFSVGWNNSTEN--CNSATGWTGYAQVNGN-NTEI :  94
                 6 G             5 V 5           S 3 5 Gq

                    *         140          *         160          *         180
bradavidin  : QGLWYLSLA----EA---VAWNGISAGADTBTFSSGDKALLTKSGVDLKAGSEKLSNTK : 137
tamavidin2  : LTQWLLSSS----T-ARGDVWESTLVGNDSBTKTAPT---------------------- : 123
streptavidin: NTQWLLTSG----T-TEANAWKSTLVGHDTBTKVK----------------------- : 121
avidin      : KTMWLLRSS----VNDIGDDWKATRVGINIBTRL------------------------ : 121
zebavidin   : KTFWMLRSV----ADNLASAWGSTRMGEDIBFKTgvsn-------------------- : 126
AVR2        : KTKWLQRLA----VDDISDDWIATRVGNNDBTRQHTvee------------------- : 129
rhizavidin  : VTSWNLAYEGGSGP--------AIEQGQDTBQYVPTTE-------------------- : 124
                 t W l             w      G l F
```

Figure 9. Structure-based amino acid sequence alignment of selected avidin-like proteins. Structures used: bradavidin (unpublished results, Airenne et al.); tamavidin2, 2ZSC, chain A; streptavidin: PDB 1MK5, chain A; avidin: PDB 1VYO, chain A; zebavidin: PDB 4BJ8; AVR2: PDB 1WBI, chain A; rhizavidin, PDB 3EW2, chain C. Residues with light grey to black background display degrees of conservation with black being the most conserved. The consensus sequence is displayed below the alignment; numbers representing similarity groups (3: ST, 5: FYW, 6: LIVM). Alignment made by S. Kukkurainen using STRAP; and modified using GeneDoc (Gille 2006, Nicholas, Nicholas & and Deerfield 1997).

These studies, combined with the discovery of avidin-like proteins in bacterial species, led to a hypothesis, which proposes that avidin acts as a component of the defense mechanism and protects the developing embryo in the egg from bacterial and viral infection. Also, avidin gives a survival advantage to avidin secreting bacteria by sequestering biotin and thus making it unavailable in the surrounding environment (Kunnas, Wallen & Kulomaa 1993, Elo, Raisanen & Tuohimaa 1980, Zerega et al. 2001).

2.6.3 Rational protein mutagenesis to alter the structural and functional properties of avidin

Rational mutagenesis has been applied to (strept)avidin to learn more about the structural bases of their high affinity to biotin, but also to alter their structural and functional properties, to make them more suitable for certain biotechnological applications (Laitinen et al. 2006, Laitinen et al. 2007).

2.6.3.1 Modification of the biotin-binding pocket of (strept)avidin

The biotin-binding pocket of (strept)avidin has a perfect complementary fit for the binding of biotin. The pocket is lined with hydrophilic and hydrophobic residues, which interact with biotin. One of the hydrophobic residues in contact with biotin, W110 (W120 in streptavidin), is donated by an adjacent subunit (Figure 8B). In addition to the hydrophobic interactions, an extensive network of hydrogen bonds contributes to the tight binding of biotin (Livnah et al. 1993, Hendrickson et al. 1989). Site-specific mutagenesis of the residues involved in biotin binding leads to a reduction in the biotin affinity (Chilkoti, Tan & Stayton 1995). The most drastic reduction is seen with mutations to W110, which not only affect avidin's affinity to biotin, but also the stability of the tetramer (Sano, Cantor 1995, Laitinen et al. 1999).

Mutations in the residues involved in the hydrogen network with biotin generally lead to a reduction in affinity. Therefore, these sites have been altered in order to design (strept)avidin mutants with reversible biotin-binding affinities (Laitinen et al. 1999, Reznik et al. 1998, Morag, Bayer & Wilchek 1996b, Takakura, Sofuku & Tsunashima 2013). Other approaches to obtain reversible biotin binding have involved mutating W110 or residues at the interfaces of the subunits (Sano, Cantor 1995, Laitinen et al. 1999, Wu, Wong 2005, Qureshi et al. 2001). The disadvantage of most of the modifications leading to reversible binding is that biotin binding in general is reduced, and in the case of mutations in interface residues, the tetramer is destabilized and the protein is present in either dimeric or monomeric form.

2.6.3.2 Modifications introducing a pH-dependence to biotin binding

A handful of studies have employed pH-sensitive mutations to reverse biotin binding. A chemical modification with tetranitromethane to Y33 resulted in a protein that displayed high biotin-binding affinity at low to neutral pH, but negligible affinity at pH 10 (Morag, Bayer & Wilchek 1996b). This mutant, Nitroavidin, was suitable for reusable affinity matrices and was commercialized under the name CaptAvidin (Morag, Bayer & Wilchek 1996a). However, the chemical modification only affects approximately 70% of the biotin-binding sites, meaning that unmodified binding sites need to be blocked with biotin. This makes Nitroavidin unsuitable for applications where modified binding sites are required on both sides of the molecule (e.g. biosensor applications) (Morag, Bayer & Wilchek 1996b, Garcia-Aljaro, Munoz & Baldrich 2009). Our group genetically modified the biotin-binding site in a similar manner by mutating Y33 to histidine, allowing for pH-dependent H-bond formation. The resulting mutant behaved similar to CaptAvidin. However, it had the added benefit of being a genetically modified variant, in which all the binding sites were modified equally (Marttila et al. 2003). The Y33H mutation was combined with a dual-chain avidin mutant to create a dual-affinity avidin molecule (Hytönen et al. 2005b).

Another approach that was used to accomplish pH-dependent biotin binding was the introduction of histidine residues into the avidin subunit interfaces. The three residues making up the interface between subunits 1 and 3 (Figure 8C) and Trp110 were individually or in combination mutated to histidine, resulting in proteins whose tetrameric assembly was disrupted at low pH (Nordlund et al. 2003).

2.6.3.3 Rational design of monomeric/monovalent (strept)avidin

Much effort has been put into making monomeric (strept)avidins, because the tetrameric form of the protein causes problems with cross-linking and affects the protein's performance in certain applications. Attempts to obtain monomeric (strept)avidins have included the mutagenesis of interface residues and W110, which interacts with the biotin-binding site in the opposite subunit. But because W110 is important for tight biotin binding, most of these mutants resulted in a protein with a significantly decreased affinity towards biotin (monoavidin: 10^{-7} M

(Laitinen et al. 2003) and monomeric streptavidin: 10^{-7} M (Wu, Wong 2005)). Other approaches to reduce multivalency have included the creation of a 3:1 mix of inactive and active sptreptavidin monomers, resulting in tetrameric proteins with only one active biotin-binding site (Howarth et al. 2006) or single-chain avidin where the four subunits are encoded in one polypeptide chain, and therefore, each of the four biotin-binding sites can be individually genetically modified (Nordlund et al. 2005a).

The most stable monomeric streptavidin so far, with a biotin affinity of $K_d = 2.8$ nM, was designed by Lim et al. (Lim et al. 2013). To overcome the loss of W120, biotin-binding pocket residues in contact with biotin were exchanged with the corresponding residues in rhizavidin. Due to its dimeric quaternary structure Rhizavidin lacks a residue corresponding to W120 (Meir et al. 2009), but residues in the binding pocket compensate for this, resulting in high biotin-binding affinity. Additionally, rational design was used to improve the stability and solubility of the monomer by mutating former interface residues that were exposed to the solvent in the monomer, and shortening the flexible loops (Lim et al. 2013).

2.6.3.4   Rational design of chimeric (strept)avidin

Chimeric (strept)avidins have been developed by rational design to transfer certain properties from one protein to another. In this manner, the ability of avidin to hydrolyze biotin esters was transferred to streptavidin (Pazy et al. 2003), the high thermal stability of avidin related protein 4 (AVR4) was introduced into avidin (Hytönen et al. 2005a), and, as discussed above, the biotin affinity of monomeric streptavidin was increased by incorporating residues found in rhizavidin (see above) (Lim et al. 2013). Another type of chimeric avidin was produced by Riihimäki et al., who circularly permuted avidin with AVR4. This resulted in a pseudodimer (one subunit from avidin and one subunit from AVR4), which formed dimers (pseudotetramers) and functionally resembled wild type avidin (Riihimaki et al. 2011b).

2.6.3.5 Neutralization of avidin

The high isolectric point of avidin (pI 10.5) and its glycosylation at residue Asn17 cause high non-specific binding in many applications and are the main reasons why the bacterial, non-glycosylated streptavidin (pI = 6) is favored over avidin (Green 1975, Chaiet, Wolf 1964). Attempts have been made to create a genetically modified neutralized avidin by removing the glycosylation site and modifying basic surface residue based on sequence comparisons with streptavidin and avidin related proteins. The resulting mutants displayed reduced non-specific binding (Marttila et al. 1998, Marttila et al. 2000). Furthermore, recombinant expression of chicken avidin in *E. coli* produced non-glycosylated proteins (Hytönen et al. 2004).

2.6.4 Random protein mutagenesis to alter the ligand specificity of avidin

Only a handful of studies to date have employed directed protein evolution to engineer (strept)avidin. Most of these studies used oligonucleotide based random mutagenesis to mutate the loops in the biotin-binding site of the β-barrel to alter ligand-binding properties. For example, Levy and co-workers used this approach to lower the dissociation rate of desthiobiotin from streptavidin (Levy, Ellington 2008).

Our group employed a similar approach to alter the ligand specificity of avidin to bind a steroid. Residues in the loops between β-strands 1 and 2 and β-strands 3 and 4 were randomized and mutants with a lowered affinity towards biotin and an affinity towards testosterone were selected by phage display (Riihimaki et al. 2011a).

In a slightly different approach, Aslan and co-workers used error-prone PCR in combination with phage display against biotin to engineer a single-chain dimeric streptavidin with increased affinity towards the fluorescently labeled biotin-4-fluorescein (Aslan et al. 2005).

2.7 Studying protein function *in vivo* using model organisms

Model organisms allow the study of biological questions regarding genetics, metabolic pathways, development and diseases. Ideally, a model organism is simple to maintain, has a short generation times and large numbers of progeny and its genome can be easily manipulated. The low complexity of most model organisms allows the study of a systems and the application of the findings to the human body (Hunter 2008, Hedges 2002, Barr 2003). Popular model organisms are fission yeast (*Schizosaccharomyces pombe*) (Zhao, Lieberman 1995), fruit fly (*Drosophila melanogaster*) (Bier 2005), round worm (*Caenorhabditis elegans*) (Kaletta, Hengartner 2006), and mouse (*Mus musculus*) (Peters et al. 2007).

The recent discovery of an avidin-like protein in the model organism zebrafish (*Danio rerio*) offers the possibility to study the physiological function of avidin in detail (Sreenivasan et al. 2008). The zebrafish is a small fresh water fish that has become an important model organism in many fields of research ranging from developmental biology (Quaife, Watson & Chico 2012, Ericsson, Knight & Johanson 2013, Zhang, Patient & Liu 2013), to cancer research (Tobia, De Sena & Presta 2011, Lu et al. 2011, Liu, Leach 2011, Teittinen et al. 2012) to the study of other human diseases (Kinkel, Prince 2009, Lien et al. 2012, Berger, Currie 2012, Klee et al. 2012). The zebrafish is so popular because its simple maintenance, short generation time and the fact that zebrafish embryos are transparent and their development can thus easily be monitored under the microscope (Zon 1999, Wixon 2000).

Different tools are available to genetically engineer zebrafish (Tobia, De Sena & Presta 2011, Lu et al. 2011, Liu, Leach 2011, Teittinen et al. 2012). One simple tool that is especially important in developmental biology is the morpholino, a small, stable chemical analogue to nucleotides, which binds tightly to the translation initiation site of the gene in question and prevents its expression (Corey, Abrams 2001, Heasman 2002). This leads to a knock-down phenotype that can be used to observe the influence of the protein or RNA encoded by the silenced gene over the course of the first few days of the developing embryo (Summerton, Weller 1997, Nasevicius, Ekker 2000, Bill et al. 2009). Besides the knock-down function,

morpholinos can also be used to block pre-mRNA splicing, which can be useful in therapeutic applications (Du, Gatti 2011).

3. AIMS OF THE STUDY

Previously, our group has modified existing avidin proteins by rational mutagenesis in order to tailor them for specific purposes. This study aimed at elucidating the applicability of DNA family shuffling in the design of completely novel avidin proteins with altered structural and functional properties.

The specific aims were

I.   To study if the avidin protein family is suitable for DNA family shuffling.

II.   To structurally characterize a chimeric mutant selected from the AVD/AVR2 chimeric avidin library described in (I) and to improve the biopanning selection method.

III. To evaluate the applicability of an avidin mutant in reusable biosensor surfaces.

IV.   To characterize a novel avidin-like protein from zebrafish to extend the pool of parental proteins for DNA family shuffling and to shed light on the physiological function of avidin.

4. MATERIALS AND METHODS

Detailed descriptions of the materials and methods are presented in the original articles (I-IV).

4.1 Molecular cloning methods

4.1.1 Phagemid vector (I)

The phagemid vector (I) used to construct the DNA shuffling mutant library was based on the VTT Fab phagemid vector (pBluescript SK+ derived phagemid; VTT Research Center of Finland, Biotechnology, Espoo, Finland). The phagemid contains a Fab fragment that includes a pelB signal peptide at its N-terminus and is fused C-terminally with the pIII coat protein of M13 bacteriophage. The Fab fragment was removed by digestion with the NheI and NotI restriction enzymes (New England BioLabs, Ipswich, MA, USA). The digested linearized vector was purified from agarose gel (GFX DNA purification kit, GE Healthcare, Uppsala, Sweden) and used in a SLIC cloning reaction to create the DNA shuffling mutant library (I).

4.1.2 DNA family shuffling (I)

DNA family shuffling was done according to Stemmer (Stemmer 1994). Parental DNA was digested by DNaseI and subsequently reassembled in a primerless PCR. The reassembled DNA was purified from agarose gel and cloned by SLIC cloning (see chapter 4.1.3) into the linearized phagemid vector. To avoid including parental genes, gene specific primers that bind to one of the parental genes at its 5'-end and to the other at its 3'-end were used.

4.1.3 Sequence and ligation independent cloning (SLIC) (I)

SLIC uses homologous recombination to clone an insert into its designated vector (Li, Elledge 2007). The reassembled shuffled genes were amplified using long primers that overlap with the ends of the linearized vector. Single-stranded overhangs were created by the exonuclease activity of T4 DNA polymerase (New England Biolabs) and insertion of the insert was enabled by RecA (New England BioLabs) aided homologous recombination *in vivo*.

4.1.4 Construction of bacterial expression vectors (III, IV)

Avidin genes were amplified by a two-step stepwise elongation of sequence PCR (SES-PCR) process (Majumder 1992) to N-terminally add the bacterial secretion signal peptide OmpA from *Bordetella avium* (Hytönen et al. 2004). The PCR intermediates and products were purified from agarose gel and subsequently cloned into the pET101/D plasmid (Invitrogen, Carlsbad, CA, USA) using the TOPO® cloning protocol followed by a standard heat shock transformation of chemically competent *E. coli* Top 10 cells (Invitrogen).

Single mutations were introduced to OmpA chicken avidin pET101/D (Hytönen et al. 2004) by QuikChange mutagenesis according to the manufacturer's instructions (Stratagene, La Jolla, CA, USA). Neutralizing mutations were introduced in a two-step PCR by sequence overlap extension (Horton et al. 1989).

4.2 Phage display

4.2.1 Construction of phagemid libraries (I)

To allow the assembly of a tetrameric avidin protein on the surface of the phage, an amber stop codon was inserted between the chimeric gene and the M13 surface protein pIII (Sidhu, Weiss & Wells 2000). The amber codon is not recognized as a stop codon in *E. coli* strains containing the mutation supE44 (i.e. XL1 blue) and is instead translated to glutamine (Eggertsson, Soll 1988). Suppression of the amber codon is not completely efficient and depends highly on the bacterial growth phase

and the sequential context of the amber codon. Thus, the amber codon is read as a stop codon in 65-93% of the translated sequences (Eggertsson, Soll 1988, Singaravelan, Roshini & Munavar 2010). This allows the expression of both p3-fusion proteins and free proteins. DNA shuffled genes were cloned into a phagemid vector by SLIC cloning and subsequently transformed by electroporation into *E. coli* XL1 blue cells. Successful transformants were stored as glycerol stocks at -80 ºC.

4.2.2   Production of phages (I, II)

E. coli XL1 blue cells from glycerol stocks were grown in culture medium and infected with the VSC-M13 helper phage (Stratagene). The infected cells were diluted in culture medium and grown overnight at 28 ºC and 225 rpm. The amplified phages were purified by precipitation using 20% PEG-6000 in 2.5 M NaCl and resolved in PBS containing 1 M NaCl, 20% glycerol and 1% BSA. The phage titer was determined by plating serial dilutions onto culture plates.

4.2.3   Selection of functional mutant proteins by biopanning (I, II)

Microwell plates (MaxiSorp^{TM} Immuno 96 MicroWell^{TM} plates, Nunc, Roskilde, Denmark) were coated with biotinylated BSA, blocked with 1% milk in PBS and washed with PBS-Tween. Phages displaying the chimeric proteins were allowed to bind to the immobilized biotin and non-binders were washed away. Bound phages were eluted with hydrochloric acid and used to infect *E. coli* XL1 blue cells. The phages were amplified over night with the help of VSC-M13 helper phages (Stratagene) and precipitated the following day as described above (chapter 4.2.2). The biopanning step was repeated 2 to 4 times, increasing the amount of wash steps with every round.

In study II, a biotinylated BSA, that contains a disulfide in the linker region, was used (EZ-Link Sulfo-NHS-SS-Biotin, Thermo Scientific, Waltham, MA, USA). This allowed the elution of the bound phage from the solid surface by breaking the disulfide bond with 50 mM DTT.

4.2.4 Microwell plate assay (I, II)

In order to confirm the functionality of the phage display selected mutants, the proteins were produced in *E. coli* BL21star (DE3) or BL21-AI cell lines (Invitrogen), which do not carry the supE44 genotype. This allowed the expression and analysis of chimeric proteins without the pIII-fusion. The selected colonies were expressed in a small volume using 96-deep well plates (deep well plates, ABgene, Thermo Scientific). Cell lysates were applied to microwell plates coated with biotinylated BSA (I) or casein (II). Biotin-binding proteins were detected with biotinylated alkaline phosphatase and a phosphatase substrate.

4.2.4.1 *Avidin-biotin displacement (ABD) assay (II)*

A modified version of the microwell plate assay was used in study II. In the ABD assay the analyzed mutants were incubated on two wells, each coated with biotinylated protein. While one of the wells was incubated for 30 min with buffer containing biotin, the other was incubated with plain buffer. The amount of bound proteins was subsequently detected with biotinylated alkaline phosphatase and a phosphatase substrate. The difference in the absorbance between the two wells gave an estimate of the biotin-binding affinity; the higher the ratio between biotin treated and non-biotin treated wells, the stronger the biotin binding. Furthermore, the ratio between biotin treated and non-biotin treated wells gives an estimate of the binding affinity that is independent of the expression levels of the analyzed mutants, i.e. a high signal relates to strong biotin binding and not to the amount of functional protein present.

4.3 Protein production and purification (I-IV)

The recombinant proteins were either produced in bottle cultures (Helppolainen et al. 2007) or in a 7.5-L fermentor (Labfors 3, Infors, Bottmingen, Switzerland) using fed-batch culturing method based on pO_2-stat as described earlier (Määttä et al. 2011). The expression strain was either *E. coli* BL21star(DE3) (I) or BL21-AI (II,

III, IV) (Invitrogen). Expression was induced by the addition of IPTG (I-II) or IPTG + L-arabinose (III, IV) and carried out at 28 ºC.

Proteins were purified by affinity chromatography on either 2-iminobiotin (high affinity biotin binders, I, II, IV) or D-biotin (low affinity biotin binders, I, III) Sepharose[TM] 4 Fast Flow (Affiland S.A., Ans Liege, Belgium) as described earlier (Airenne et al. 1997). Purified proteins were eluted either at pH 4 (2-iminobiotin) or using a stepwise acetic acid gradient (D-biotin).

4.4   Protein characterization

4.4.1   Biophysical characterization

4.4.1.1   Differential scanning calorimetry (I, II, IV)

The thermal stability of the produced proteins was analyzed using an automated VP-Capillary DSC System (Microcal, Northampton, MA, USA). The thermal unfolding of the protein samples was monitored between 20 and 140 °C with a scan rate of 2 °C/min. Proteins were extensively dialyzed against the measurement buffer (50 mM NaH_2PO_4/Na_2HPO_4 with a varying concentration of NaCl [0-650 mM] and varying pH values [3-11]; for details see I- II, IV); dialysate was used in the reference cell. Samples were degased prior to the measurement. The protein concentration in the measurement cell was 15-30 μM and the ligand concentration, if applicable, was 45-90 μM. The Origin 7.0 DSC software suite (Originlab, Northampton, MA, USA) was used to analyze the data.

4.4.1.2   Size exclusion chromatography (I-IV)

The molecular mass of the protein in solution was measured by size-exclusion chromatography (SEC) using a Superdex200 10/300GL column (GE Healthcare) connected to an ÄKTA[TM] Purifier-100 equipped with a UV-900 monitor (GE Healthcare). Alternatively, a Superdex75 5/150GL column (GE Healthcare) connected to a liquid chromatography instrument (CBM-20A, Shimadzu

Corporation, Kyoto, Japan) and equipped with an autosampler (SIL-20A), UV-VIS (SPD-20A) and a fluorescence detector (RF-20Axs) as well as a Zetasizer μV SLS/DLS detector (Malvern Instruments, Worcestershire, UK) was used. The analysis was conducted by injecting 40-80 μg of protein per run. Analyses were done with flow rates of 0.25 to 0.5 ml/min and absorbance at 280 nm was used to locate the protein peaks in the chromatograms. The molecular mass calibration curve was prepared by analyzing a gel filtration standard protein mixture containing thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa) (BioRad Laboratories, Hercules, CA, USA) for Superdex200 and cytochrome C (12.4 kDa), carbonic anhydrase (29 kDa), ovalbumin (44 kDa) and BSA (66 kDa) (Sigma-Aldrich, St. Louis, MO, USA) for Superdex75. BSA was also used to calibrate the SLS detector.

4.4.1.3   Mass spectrometry (I, II, IV)

Protein identities and oligomeric states were analyzed by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS). The analysis was done with a 4.7-T or 12-T APEX-Qe instrument (Bruker Daltonics, Billercia, MA, USA). For analysis under denaturing solution conditions, the desalted protein samples were diluted with an acetonitrile/water/acetic acid (49.5:49.5:1.0, v/v) solvent and for native MS, the protein was dialyzed against 10-500 mM ammonium acetate (pH 7). A D-biotin (Sigma-Aldrich) stock solution (2 mM in water) was mixed at a desired molar ratio with the protein and incubated at room temperature for 30 min prior to the analysis.

ESI-generated ions were externally accumulated for 1 s in the hexapole ion trap before being transmitted to the ICR cell for trapping, excitation, and detection. For each spectrum, a total of 128-512 co-added 1MWord (128kWord for native-MS) time-domain transients were zero-filled once prior to a fast Fourier transformation, magnitude calculation, and external mass calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, CA, USA). The instrument was operated and the data were processed with the XMASS 6.0.2 software (Bruker Daltonics).

4.4.1.4 Dynamic light scattering (II)

The hydrodynamic radius of proteins was determined by batch dynamic light scattering using a Zetasizer Nano ZS (Malvern Instruments) at varying pH values (pH 4, 7 and 11). The analysis was done at a concentration of 1 mg/ml in the absence and presence of a 3-fold excess of D-biotin. For each sample, three measurements, consisting of 10 runs of 10 s, were done. Samples were heated in steps of 5 °C from 25 °C to 90 °C, and equilibrated for 5 min prior to data collection. The data was analyzed using the Zetasizer software v7.01 (Malvern Instruments). The peak with the highest intensity was selected for analysis based on volume mean size distribution.

4.4.1.5 Protein X-ray crystallization (II, III)

Proteins were crystallized using the vapor diffusion method, 96-well sitting drop iQ plates (TTP Laptech, Hertfordshire, UK), TTP Labtech's mosquito® liquid handling robot and a cooled crystallization incubator (RUMED® model 3201, Laatzen, Germany) set-up at 22 °C. The protein was mixed with a D-biotin solution (1 mg/ml) in 10:1 v/v ratio, respectively, before crystallization. The JCSG-plus™ Screen (Molecular Dimensions, Suffolk, UK) was used to obtain an initial hit. After buffer optimization, crystals typically formed within 1-2 weeks.

X-ray data were collected at the ESRF beam lines ID14-1 (II) and ID23-2 (III) (Grenoble, France) at 100 K from single crystals. If needed, a cryoprotectant (glycerol, 30% v/v in well solution) was added to the crystallization drop just prior to freezing in liquid nitrogen. The data was processed with XDS (Kabsch 1993) and initial phase estimates for the structure factors were obtained using the molecular replacement program Phaser (McCoy et al. 2007) within the CCP4i GUI (Collaborative Computational Project, Number 4 1994, Potterton et al. 2003)., AVR2 [PDB: 1WBI] (Hytönen et al. 2005) was used as the template structure for molecular replacement. In the case of zebavidin (III), also a tetrameric homology model, based on the AVR2 structure [PDB: 1WBI] (Hytönen et al. 2005), was created in Modeller (Sali, Blundell 1993) by Discovery Studio 3 (Accelrys Software, San Diego, CA, USA). The structures were built in Phaser, refined with Refmac5 (Murshudov, Vagin & Dodson 1997) and manually edited/rebuilt using

Coot (Emsley, Cowtan 2004). Non-protein atoms were added to the model either with the automatic procedure of Coot and ARP/wARP (Lamzin, Wilson 1993, Perrakis, Morris & Lamzin 1999, Langer et al. 2008), or manually in Coot.

The final structure was validated using the inbuilt tools of Coot (Emsley, Cowtan 2004), and the MolProbity (Davis et al. 2007) program of the Phenix software suite (Adams et al. 2002), before deposition to the Protein Data Bank (Berman et al. 2000, Berman et al. 2002) with the PDB entry codes 4BJ8 (zebavidin) and 4BCS (A/A2-1).

4.4.2   Isoelectric focusing

Isoelectric focusing was used to analyze the isoelectric point of proteins. Proteins were dialyzed against H_2O and 15 µg of protein was diluted in 2% pharmalyte 3-10 for IEF (GE Healthcare) containing 3.75 mg of the DeStreak reagent (GE Healthcare) in a volume of 250 µl (sample solution). Immobiline DryStrips pH 3-11 NL, 13 cm (GE Healthcare) were rehydrated using the sample solution at 20 °C for 6 h. The samples were run in an Ettan IPGphor3 IEF unit (GE Healthcare) for 5.3 h at 20 V, for 45 min at 150 V, for 45 min at 300 V, for 45 min at 600 V, then in a gradient to 5000 V in 1 h, to 8000 V in 1 h and at 8000 V for 6.1 h. Following focusing, the DryStrips were fixed in 20% TCA for 10 min and washed with destaining solution (0.1% $CuSO_4$, 10% acetic acid, 30% methanol) for 2 min. The DryStrips were stained with staining solution (0.25% coomassie blue, 30% methanol, 0.1% $CuSO_4$, 10% acetic acid) for 1 h and destained overnight in destaining solution. DryStrips were impregnated in 5% glycerol, 10% acetic acid for 10 min and air-dried. The isoelectric point was determined by comparison with a standard IEF protein mix 3.6-9.3 (Sigma-Aldrich).

4.4.3   Determination of ligand binding interactions

4.4.3.1   *Isothermal titration calorimetry (I, III)*

The binding affinity of avidin proteins to D-biotin was analyzed using a high-sensitivity VP-ITC titration calorimetry instrument (Microcal). Extensively dialyzed

proteins were titrated with 10 to 15 µl injections of D-biotin, dissolved in dialysate. Measurements were performed at 40 °C. Additionally, a competitive titration method as described by Sigurskjold (Sigurskjold 2000) was used for the more accurate determination of thermodynamic binding parameters. First, desthiobiotin was titrated in 15 µl aliquots in to the protein solution, and D-biotin was then titrated in 15 µl aliquots into the mixture of protein and desthiobiotin. Binding constant (K_a) and the enthalpy of binding (ΔH) were calculated from the measured heats using Origin 7.0 (Originlab). K_a and ΔH of the desthiobiotin titration were used to analyze the competitive titration reaction utilizing the "competitive binding" tool in Origin 7.0. Gibbs free energy of binding (ΔG) and entropy (ΔS) were calculated from K_a and ΔH using the equations (2) and (3) where R is the gas constant and T the temperature:

$$\Delta G = -RT \ln K_a \qquad\qquad (2)$$

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

4.4.3.2 Fluorescence spectroscopy (I-IV)

The dissociation rate of the biotinylated fluorescence probe ArcDia™ BF560 (ArcDia, Turku, Finland) was determined using a QuantaMaster™ Spectrofluorometer (Photon Technology International, Lawrenceville, NJ, USA) as described in Hytönen et al. (Hytönen et al. 2004). The changes in the fluorescence of 50 nM BF560 were measured upon the addition of 100 nM avidin protein. Dissociation of BF560 from the avidin protein was initiated by the addition of a 100-molar excess of free biotin and followed over the time span of 1 h. Measurements were performed at 25 and 50 °C.

The data was analyzed based on a one-phase dissociation model using the equation $k_{diss} \cdot t = \ln(B/B_0) + $ constant were B_0 is the maximal measured binding and B the binding measured over time.

50

4.4.3.3 Radioactive [³H]biotin assay (IV)

The dissociation of biotin from the analyzed avidin proteins was studied with a radioactive biotin dissociation assay modified from Klumb et al. (Klumb, Chu & Stayton 1998). The protein was incubated at a monomer concentration of 50 nM with 10 nM radioactive biotin ([8,9-³H]biotin, PerkinElmer, Waltham, MA, USA) for 20 min. Measurements were performed at room temperature. The unbound [³H]biotin was separated from the protein-ligand complex by centrifugal ultrafiltration through a 30,000 MW cut-off filter (Vivaspin 500 centrifugal concentrators, Sigma-Aldrich). Dissociation of [³H]biotin was initiated by the addition of excess cold biotin (50 μM final concentration) and dissociated [³H]biotin was removed by ultrafiltration at different time points. The filtrate was analyzed in a Wallac 1410 liquid scintillation counter (Wallac Oy, Turku, Finland). Triplicates of each sample were measured at each time point.

    The fraction of bound radioactive biotin at each time point was calculated using the equation (1).

$$-k_{diss}t = \ln\left[\frac{(x_t - x)}{(x_t - x_0)}\right] = \ln(fraction\ bound) \qquad (1)$$

where x_t is the total amount of radioactive biotin before the addition of protein, x is the free biotin at each time point and x_0 is the amount of free ligand in the presence of protein immediately prior to the addition of cold biotin. The dissociation rate constant (k_{diss}) was determined from the slope of the linear fit to the data points of ln(fraction bound) versus time (Klumb, Chu & Stayton 1998).

4.4.3.4 Surface plasmon resonance (I, IV)

All biosensor experiments were executed using a Biacore X biosensor (GE Healthcare). The binding of chimeric avidin proteins to 2-iminobiotin and ssDNA was analyzed as previously described (Määttä et al. 2009). A CM5-chip coated with 2-iminobiotin (Määttä et al. 2009) was used for the 2-iminobiotin analysis. Protein concentrations varying between 0.25 μM and 6 μM were used to determine the binding curves. For the ssDNA binding analysis, amino groups were introduced on

the surface of a CM5-chip by incubation with EDC/NHS followed by ethylenediamine. The amino groups were coupled with thiol groups using a NHS-PEO$_2$-maleimide linker (Thermo Scientific). Oligonucleotides (5'-SHGTCAGCCACTTTCTGGC-3', Eurogentec/Oligold, Osaka, Japan) were then coupled to the thiol groups on the surface. Unreacted groups were inactivated using 50 mM cysteine. The reference cell was prepared accordingly, but without the injection of oligonucleotides. Binding curves for ssDNA binding were measured using protein concentrations varying from 50 µM to 100 nM. The involvement of a biotin-binding site in DNA binding reaction was assessed by injecting a 1:3 mixture of protein and biotin.

To analyze the performance of AVD M96H, gold surfaces coated with a mixed self-assembly monolayer (SAM) of alkanethiol derivatives was used. The SAMs consisted of a synthesized OH-terminated penta(ethylene glycol) chain and a biotin-terminated poly(ethylene glycol) chain (n = 6) matrix components in a 80/20 (mol/mol) ratio. PBS 7.3 or HBS was used as a measurement buffer. Samples (100 µl) were injected with a 20 µl/min flow rate.

4.4.4   Experiments with zebrafish (III)

4.4.4.1   Zebrafish maintenance

In all experiments, wild-type AB zebrafish were used. Fish maintenance was performed according to standard protocols (Parikka et al. 2012, Rounioja et al. 2012). The adult fish were kept in a flow-through system with a light/dark cycle of 14 h/10 h. The water was filtered with filters and activated carbon. Sterilization was done using UV-light. The conductivity (800 µS) of the water was adjusted with sea salt (Instant Ocean, Blacksburg, VA, USA) and pH was adjusted to 7.6 with NaHCO$_3$. Water was partially (10%) exchanged daily. The fish tanks were made of FDA-approved, good grade autoclavable polycarbonate USP class VI. The fish were fed twice a day with SDS 400 food (Special Diets Services, Essex, UK). Embryos were grown in E3-H$_2$O (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl$_2$, 0.33 mM MgSO$_4$, $1 \cdot 10^{-5}$ g/l methylene blue) at 28 °C. Fish were euthanized with 4 mg/ml Tricaine (ethyl 3-aminobenzoate methanesulfonate salt) at pH 7; the pH was

adjusted using a Tris-buffer. All of the zebrafish experiments were in accordance with the Finnish Laboratory Animal Welfare Act 62/2006, the Laboratory Animal Welfare Ordinance 36/2006 and had been authorized (authorization LSLH-2007-7254/Ym-23) by the National Animal Experiment Board (Finland).

4.4.4.2   Gene expression analysis with qRT-PCR

Relative zebavidin expression was analyzed using quantitative real-time PCR (qRT-PCR). Tissue samples from adult wild type AB zebrafish (gonads, gills, kidney, tail fin, eyes and brain, as well as, whole developing embryos [0-7 days post fertilization {dpf}]) were analyzed. Total RNA was extracted with the RNeasy RNA purification kit (Qiagen, Hilden, Germany) followed by reverse transcription using the iScript™ Select cDNA synthesis kit (BioRad Laboratories). The primers for qRT-PCR (forward: 5´-CGAATGCAAAGGTGAGCTCC-3´ and reverse: 5´-ATAGCACGGAGAAAGAGACG-3´) were ordered from Oligomer (Helsinki, Finland). The cDNA was analyzed by qRT-PCR and the Maxima SYBR Green qPCR master mix (Fermentas, Burlington, Canada) and a CFX96 qPCR machine (BioRad Laboratories Inc.). Expression of zebavidin was normalized to the expression of the gene for elongation factor 1-alpha (*EF1a*, ENSDARG00000020850) expression (Tang et al. 2007). Data analysis was done using the Bio-Rad CFX Manager software (BioRad Laboratories), Microsoft Office Excel 2010, and GraphPad Prism 5.0 (GraphPad Software, La Jolla, CA, USA). Additionally, the qRT-PCR products were subjected to a melting curve analysis followed by electrophoresis in a 1.5% agarose gel (Bioline, London, UK).

4.4.4.3   Isolation of zebavidin from zebrafish oocytes

Oocytes of five individuals (20- 150 mg per individual) were washed with PBS and the supernatant was removed by centrifugation at 5,000 g and 4 °C for 5 min. The supernatant was incubated with 20 µl of D-biotin Sepharose™ 4 Fast flow (Affiland) at 4 °C for 1 h. The washed oocytes were incubated on ice in 0.5 M sucrose, 0.01 mg/ml lysozyme, 1 mM EDTA, 200 mM Tris, pH 7.4 for 30 min. After addition of 2 mM EDTA, 150 mM NaCl, 1% TritonX-100, 50 mM Tris, pH 8,

the oocytes were sonicated at 25% amplitude two times for 15 s (1 s on, 1 s off). The cell lysate was clarified by centrifugation at 13,000 g, 4 °C for 20 min. PBS + 1 M NaCl (800 µl) and 20 of µl D-biotin Sepharose™ were added to the cleared lysate and incubated at 4 °C for 1 h. The Sepharose was collected by centrifugation at 2,500 g and 4 °C for 5 min. The Sepharose was washed with 1 ml of PBS + 1 M NaCl. Samples from each step were analysed by SDS-PAGE and subsequent Coomassie Brilliant Blue staining. A molecular weight marker (Page Ruler™ Prestained Protein Ladder, Fermentas) and bacterially expressed zebavidin were used for size determination. Four selected bands with molecular weights corresponding to zebavidin (13 kDa) were excised from the gel and dried in acetonitrile, followed by in-gel digestion according to the standard protocol. Samples were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) using the LTQ Orbitrap Velos Pro mass spectrometer at the Proteomics Facility (Turku, Finland). The obtained data was searched against the NCBInr database (release 2013_02) using Mascot 2.4.0.

4.4.4.4 Morpholino studies

The zebavidin translation initiation site was determined from the genomic DNA of three zebavidin strains (+AB5, +AB7 and +AB8) by sequencing. A morpholino, binding to the translation initiation site (5'-GCCATATTAAAGAACTCATCTTGGC-3'), was ordered from GeneTools, (Philomath, USA). The morpholino was diluted to 130 µM in 200 mM KCl containing 0.2% of a rhodamine dextran tracer from a stock solution. The morpholino solution (1 nl) was injected into the yolks of 150 one-cell stage zebrafish eggs using a microinjector. Ninety eggs were injected with a random control morpholino (5'-CCTCCTACCTCAGTTACAATTTATA-3'). Embryos were grown in E3-H_2O at 28 °C. The success of the injection was determined using a fluorescent microscope. The developing embryos were analysed under a phase contrast light microscope after 24 and 48 h.

5. REVIEW OF THE RESULTS

5.1   DNA shuffling within the avidin protein family (I, II)

5.1.1 Construction of chimeric DNA mutant libraries

Genes encoding avidin (Green 1975), AVR2 (Hytönen et al. 2005), BBP-A (Hytönen et al. 2007) and rhizavidin (Helppolainen et al. 2007) were used as the parental genes in the DNA family shuffling (I). Amplified parental genes were mixed in pairs and digested with DNaseI. Subsequently, chimeric genes were reassembled from the fragments by homologous recombination. To avoid contaminating the library with parental proteins, the reassembled chimeric genes were amplified using primers specific for one gene at its 5'-end and primers specific for the other gene at its 3'-end (I, Supplementary Figure 1). Using this approach, the reassembly of full length chimeric genes containing sequence of rhizavidin was not possible. The shuffled genes were introduced into a phagemid vector by SLIC cloning, N-terminally in fusion with the pelB secretion signal peptide and C-terminally with the bacteriophage pIII coat protein. An amber stop codon was introduced between the chimeric gene and *pIII* gene allowing the expression of free chimeric avidin monomers and monomers in fusion with the pIII coat protein in amber suppressing *E. coli* cells, resulting in the display of a tetrameric protein on the phage surface.

The quality of the initial DNA mutant libraries was analyzed by sequencing a small sample of the libraries (I, Table 1). In the AVD/AVR2 library all 26 analyzed sequences were shuffled. In the AVD/BBP-A library 26 out of 28 sequences and in the AVR2/BBP-A library 22 out of 29 sequences were shuffled.

5.1.2 Characterization of selected chimeric mutants

The phages selected after three rounds of biopanning with increased stringency were analyzed by microwell plate assays and sequencing, which revealed enriched variants with higher expression levels in comparison with the parental proteins (I, Supplementary Table 1 and II, Table 1). Additionally, the ABD assay (II) enabled estimation of the mutants' dissociation rate from biotin.

In study I, the most enriched mutants, which gave high signals in the microplate assay, were expressed in *E. coli*, purified by 2-iminobiotin affinity chromatography, and characterized by size-exclusion chromatography (SEC), surface plasmon resonance analysis (SPR), differential scanning calorimetry (DSC), and isothermal titration calorimetry (ITC). All the analyzed mutants displayed lowered affinity towards 2-iminobiotin and biotin compared with the parental proteins (I, Table 2 and Figure 4). The thermal stabilities of the chimeras were in the range of the parental proteins (I, Table 3 and Figure 3). A/B-2, a chimeric form of avidin and AVR2, was the most stable mutant in the absence of biotin, and showed thermal unfolding at a slightly higher temperature (95.8 °C), even when compared with the highly stable AVR2 (93.6 °C).

The most interesting mutant was A/A2-1, a chimeric mutant containing avidin and AVR2 sequences. Analysis of biotin binding by ITC and DSC (I, Table 2 and Figure 4) showed a slightly lowered, but still strong affinity towards 2-iminobiotin and biotin. The thermal stability of the A/A2-1 protein did not increase significantly upon binding of biotin (ΔT_m: 2.3 °C), which differs from what is usually true for avidins that bind biotin tightly (chicken avidin, ΔT_m: 42.6 °C). Additionally, the elution volume of the mutant in SEC in the absence of biotin indicated a smaller oligomeric size compared with the parental proteins (I, Table 2).

Improved biopanning selection using the AVD/AVR2 chimeric library resulted in the selection of a chimera with an amino acid sequence highly similar to A/A2-1, named A/A2-B (II, Figure 1). The two proteins differed only at six amino acid positions. A/A2-B displayed a lowered dissociation rate compared with A/A2-1 and a higher thermal stability, which even exceeded the thermal stability of AVR2 in the biotin bound state (II, Figure 2 and Table 2). Further, A/A2-B also displayed high stability at high and low pH (pH 3 and pH 11), whereas the thermal stabilities of

A/A2-1 and the parental proteins were significantly lowered at extreme pH conditions (II, Table 2).

A detailed study of the oligomeric structure of A/A2-1 and A/A2-B by native-MS and SEC led to the conclusion that both mutants are stable tetramers in the absence and presence of biotin. The unusual elution pattern of A/A2-1, observed earlier (I), was found to be due to non-specific binding with the column matrix (II, Supporting Figure S1, Supporting Figure S2 and Supporting Table S1).

The crystal structure of A/A2-1 bound to biotin was solved by X-ray crystallography. It confirmed that the chimeric protein folds into the common three-dimensional structure seen for AVD and AVR2 (II, Figure 4) (Livnah et al. 1993, Hytönen et al. 2005).

5.2 Rational design of a reversible biotin-binding avidin mutant

5.2.1 Design of a reusable biosensor (III)

The chicken avidin mutant AVD M96H was found to be easily dissociated into its subunits at low pH and showed decreased stability in the presence of SDS (Nordlund et al. 2003). The combination of low pH and SDS was later found to lead to reversible biotin binding characteristics. It was therefore the prime candidate for the design of resusable biosensors based on avidin-biotin technology. For optimal performance, a flat biotinylated surface was synthesized using OH-terminated and biotin-terminated matrix alkanethiols at an 80/20 ratio (III, Figure 4). The biotin-terminated alkanethiols were slightly longer than the OH-terminated alkanethiols. This allowed for optimal spacing between individual biotin molecules for maximum binding performance. AVD M96H showed stable binding to the flat biotinylated surface. Treatment with either acetic acid (pH 2.2) or 0.5% SDS alone removed only a small amount of the bound protein. However, a combination of the two chemicals led to the complete removal of the bound protein (III, Figure 5A).

The AVD M96H sensor surface successfully bound a biotinylated lysozyme and the interaction of lysozyme and an anti-lysozyme antibody could be studied (III,
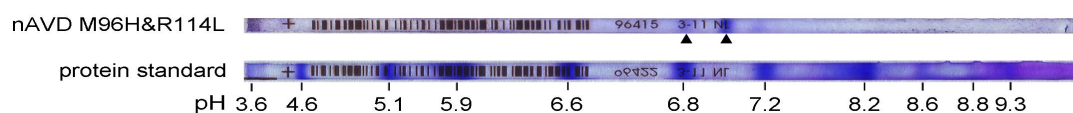
nAVD M96H&R114L	+		96415	3-11 NL
protein standard	+		96415	3-11 NL
pH 3.6	4.6	5.1	5.9	6.6	6.8	7.2	8.2	8.6	8.8 9.3

Figure 10.   Isolectric focusing of neutralized mutant nAVD M96H&R114L. The isoelectric point of nAVD M96&R114L was reduced to pI 7. The position of nAVD M96H&R114L is indicated with triangles. The analyzed proteins displayed multiple bands.

Figure 5B). However, the mutant's high isoelectric point (pI = 9.6) led to non-specific interactions with analyte proteins with a pI below 7 (III, Figure 5A and Figure 6A). Biotin surface optimization and chemical acetylation of the protein reduced the non-specific binding (III, Figure 7) and therefore the pI of AVD M96H was subsequently lowered by genetic engineering.

5.2.2 Design of a neutralized AVD M96H (nAVD M96H&R114L, unpublished data)

The neutralizing mutations K9E, R124H and R124H were chosen based on the protein structures of chicken avidin [PDB: 2AVI] and AVR2 [PDB: 1WBI] (Livnah et al. 1993, Hytönen et al. 2005), consulting also the structure of AVR4 (PDB 1Y53; Eisenberg-Domovich et al.). The lowered isoelectric points were confirmed by isoelectric focusing (Figure 10). The additional mutation R114L reduced the dissociation rate of the conjugated biotins BF560-biotin and biotin-5-fluorescein (Table 1). Biophysical characterization of nAVD M96H&R114L confirmed that the neutralizing mutations did not affect the properties of the protein (Table 2). Furthermore, non-specific binding in the biosensor experiments was reduced (Figure 11).

Table 1. Dissociation rate constants of [³H]biotin, BF560-biotin, and biotin-5-fluorescein from avidin mutants. The dissociation of labelled biotin after the addition of free biotin was measured over the course of two days ([³H]biotin) or one hour (BF560-biotin and biotin-5-fluorescein). All measurements were done at 50 °C.

	[³H]biotin	BF560-biotin		biotin-5-fluorescein	
			release		release
	k_{diss}	k_{diss}	after 1h	k_{diss}	after 1h
	(10^{-6} s^{-1})	(10^{-6} s^{-1})	(%)	(10^{-6} s^{-1})	(%)
chicken avidin	2.3	394	59.3	339.0	62.4
Streptavidin	N.D.	7.48	2.4	16.9	7.9
AVD R114L	2.40	7.30	2.1	92.2	27.1
AVD R114F	2.02	7.24	2.3	94.8	26.5
AVD M96H&R114L	N.D.	9.70	3.1	131.3	35.9
nAVD M96H&R114L	N.D.	9.23	2.7	85.2	26.7

N.D.: not determined

Table 2. Biophysical characterization of nAVD M96H&R114L. The molecular mass of AVD M96H&R114L and its neutralized version were determined by SEC and compared with chicken avidin. The molecular mass values were calculated from the elution volume using a set of protein standards (see Chapter 4.4.1.2). Thermal transition mid-point (T_m) was determined by DSC using 15 µM protein in the presence and absence of 45 µM biotin. ΔT_m is the difference between T_m(+BTN) and T_m(-BTN).

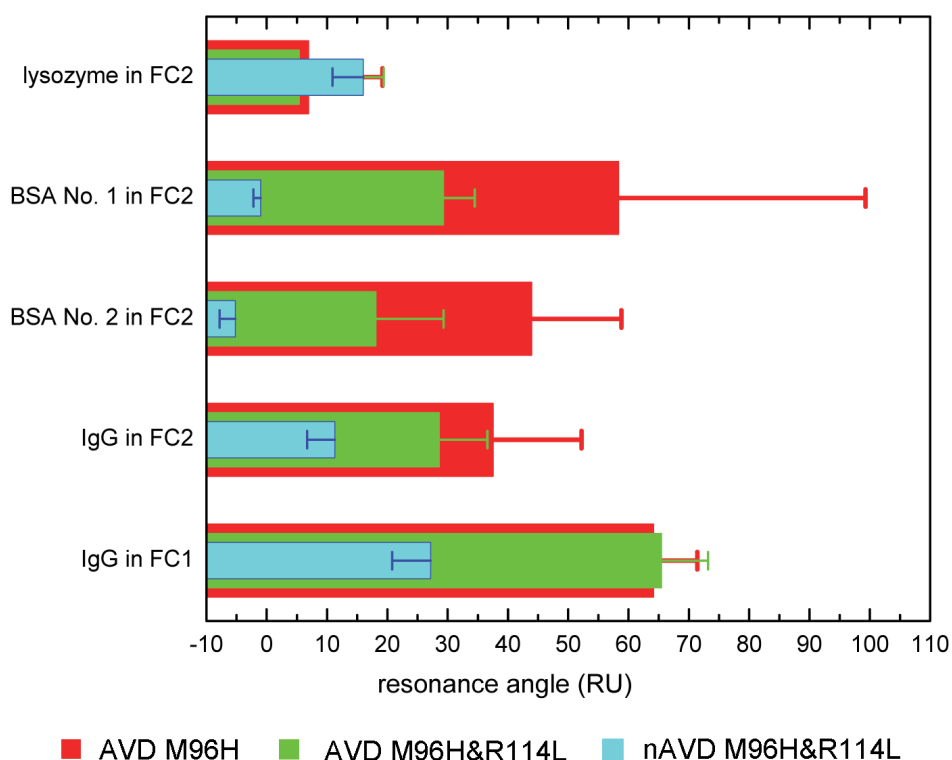|	SEC				DSC		
	elution volume		molecular mass		T_m		ΔT_m
	(ml)		(kDa)		(°C)		(°C)
	- BTN	+ BTN	- BTN	+ BTN	- BTN	+ BTN	
chicken avidin	14.8	14.6	40	44	78.0	119.6	41.6
Streptavidin	N.D.	N.D.	N.D.	N.D.	78.4	112.0	33.6
AVD R114L	N.D.	N.D.	N.D.	N.D.	74.7	117.9	43.2
AVD R114F	N.D.	N.D.	N.D.	N.D.	58.7	115.3	56.6
AVD M96H&R114L	14.8	14.5	41	47	53.6	114.9	61.3
nAVD M96H&R114L	14.1	14.2	52	58	52.0	114.2	62.2

N.D.: not determined

Figure 11. Performance of nAVD M96H&R114L in Biacore. FC1: flow cell 1, FC2: flow cell 2. The flow cells are functionalized with an avidin mutant and injected with 100 μl lysozyme, BSA No.1, BSA No.2, IgG (FC2) or only IgG (FC1). Unpublished data measured by Hermann Gruber and co-workers (Johannes Kepler University, Linz, Austria).

5.3 Characterization of zebavidin (IV)

The avidin-like protein from zebrafish, zebavidin, was expressed in *E. coli* using the bacterial secretion signal peptide OmpA (Hytönen et al. 2004, Gentry-Weeks et al. 1992). Extensive characterization of the protein revealed that zebavidin is a member of the avidin protein family. Structural analysis by X-ray crystallography showed the typical tetrameric eight-stranded β-barrel structure (IV, Figure 6), however, structural differences could be seen in the L3,4-loop and L5,6-loop, which are important for biotin binding, and in the subunit interfaces (IV, Figure 7). These differences have an effect on the structural and functional properties of the protein. Zebavidin's thermal stability was one of the lowest observed for a natural avidin protein (IV, Figure 4) and its tetrameric assembly was destabilized in the absence of

biotin and at a low ionic strength (IV, Figure 3, Supporting Figure S3). Even though the residues important for biotin binding are conserved in zebavidin (IV, Figure 1), structural differences affect its affinity towards biotin. Competitive ITC analysis suggested a K_d-value that was several orders of magnitudes (10^6) higher compared with avidin (IV, Figure 5, Table 1) (Green 1975).

Initial studies on zebavidin expression in zebrafish by qRT-PCR revealed high expression levels in the gonads of adult fish (IV, Figure 2A&B). However, according to the morpholino studies, this expression appeared not to be crucial for the development of the embryo; injection of a zebavidin expression-blocking morpholino did not lead to any phenotypic changes in up to two days old embryos (IV, Figure 2).

6. DISCUSSION

6.1   Advantages and disadvantages of DNA family shuffling within the avidin protein family

DNA family shuffling allowed the creation of diverse chimeric mutant libraries consisting of avidin-like proteins from chicken (I). Although the biopanning experiment was only designed to confirm the presence of functional biotin-binding proteins, we were also able to select for proteins that showed improved expression and thermal stability (I, II). The selected chimeric proteins A/A2-1 and A/A2-B displayed characteristics distinct from their parental proteins, which would have been difficult to achieve by rational protein design (II). This demonstrates the feasibility of creating improved avidin protein tools with random mutagenesis.

However, not every DNA shuffling method is suitable for recombining avidin genes. In addition to the DNA shuffling method developed by Stemmer (Stemmer 1994), we used StEP (see chapter 2.2.3., Zhao et al. 1998) to recombine avidin genes (I). StEP yielded only a minor amount of crossovers, which was most probably due to the length of the avidin cDNA ($\approx$ 460 bp). Earlier studies, which have successfully used StEP, have involved genes longer than 1000 bp (Callison, Hilt & Jackwood 2005, Dion et al. 2001, Zhao et al. 1998).

Variability was not only introduced by recombination, but was also due to point mutations that appeared during the cloning and phage propagation steps (I). The most prominent mutations occurred at positions that are important for biotin binding. For example, the most common point mutation was S16Y, present in the characterized mutants A/B-2 and A/A2-3. In fact, in A/B-2 it was the only residue that differed from wild type avidin. The S16Y mutation increased the thermal stability of avidin in the absence of biotin by 19.5 ºC, however, its T_m-value in the presence of biotin was reduced by 11.8 ºC (I, Table 3).

The design of the biopanning method proved to be suboptimal. The acidic elution was most probably not strong enough to allow the elution of high biotin-binding

proteins, and the proteins selected displayed weaker biotin-binding affinity (I). Thus, careful planning of the biopanning parameters is crucial for the selection of proteins with the target properties. This was shown by the selection of A/A2-B, where the elution method was made independent of the biotin binding (II). For future experiments, the biopanning parameters need to be adjusted carefully to allow the selection of mutant proteins with the desired properties.

6.2   Rational versus random protein design

Although highly promising, random protein design also has its disadvantages, as is discussed above. Furthermore, structural analysis of A/A2-1 (II) revealed that the basis for the proteins' altered properties has already been studied by rational design. For example, the main source of the increased dissociation rate of A/A2-1 was the mutation D12N. Structural studies on avidin and mutagenesis studies on streptavidin have previously demonstrated the importance of D12 for high biotin affinity (Livnah et al. 1993, Reznik et al. 1998). Additionally, the high thermal stability of A/A2-1 and A/A2-B (II) can be mainly attributed to the sequence stretch between residues 38- 58. The same sequence stretch has previously been used to improve the thermal stability of avidin (Hytönen et al. 2005a). However, some of the observed properties of A/A2-1 and A/A2-B were difficult to trace to specific changes in the sequence or structural regions of the protein. For example, the reason for the decreased biotin dissociation rate in A/A2-B over AVR2 was difficult to pin-point, and therefore, introducing this property by rational design would be challenging (II).

AVD M96H and nAVD M96H&R114L are good examples of the power of rational protein design, when enough is known about the structural and functional properties of a protein (III).We showed that it is fairly straightforward to rationally engineer mutant proteins with a desired function, without compromising other properties. Even though nAVD M96H&R114L contains five point mutations, the careful selection of the mutations, based on the sequence, structure and function of the avidin family proteins, resulted in a protein with a novel function, as well as the original stability of avidin. (III).

6.3 Avidin as a tool for reusable biotin surfaces

The properties of the chicken avidin mutant AVD M96H (Nordlund et al. 2003) proved to be optimal for use in reusable biosensors; the mutant displays stable biotin binding that is abolished by the disassembly of the tetramer at low pH and in the presence of SDS (III). A flat, biotinylated surface with bound AVD M96H was shown to be a superior reusable biotin-binding surface compared with a design using streptavidin bound to a biotinylated phospholipid monolayer, which can be removed using *n*-octylglucoside (III). The removal of the phospholipid-streptavidin-biotinylated protein layer was incomplete due to the formation of patches, where the same biotinylated protein was bound by several streptavidin molecules at once. In contrast, the removal of the AVD M96H-biotinylated protein layer from the biotinylated SAM was complete.

The biotinylated surface, further functionalized with AVD M96H, could be regenerated repeatedly without compromising its function (III). Solutions used in the regeneration of biosensors are usually harsh and even stable proteins are affected by repeated regeneration cycles. In our approach, fresh AVD M96H is bound to the surface before every measurement and thus ensures the functionality of the surface. The reusable biosensor allows the investigation of different analyte proteins without the need to change the sensor surface. The principle can be applied to any kind of biosensor based system.

6.4 The future of avidin protein engineering

The diversity between the selected parental proteins was relatively low. Employing a set of more diverse proteins could result in broader diversity in functional and structural properties, making it easier to create properties that cannot be achieved by rational protein design. Therefore, the next steps in avidin protein engineering will be to extent the DNA family shuffling to a group of more distantly related avidin-like proteins such as streptavidin, rhizavidin, zebavidin and tamavidin 2. These proteins share DNA sequence identities of 39-56%. As has already been shown here for rhizavidin, these proteins share too little homology to be successfully shuffled using the conventional DNA family shuffling approach. One way to improve the

sequence identities is to harmonize the codon usage of these proteins from diverse organisms with the help of CODNS (cross-over optimization for DNA shuffling) (He, Friedman & Bailey-Kellogg 2012). In collaboration with Chris Bailey-Kellogg and Lu He at Dartmouth College, the CODNS algorithm was extended to allow codon optimization in more than two homologous sequences. The optimized avidin sequences have been chemically synthesized and will serve as parental templates in DNA shuffling.

The work carried out in this thesis has shown that the most crucial step in directed protein evolution is to choose the correct parameter in the selection step of the mutant proteins (II). Therefore, careful consideration is needed when planning the phage display experiments. For example, the method could be optimized to select for stable monomeric avidins with a strong affinity towards biotin. This could be achieved by removing the amber stop codon, which would only allow the expression of a chimeric protein-pIII fusion protein. As a result, the chimeric protein would be displayed in monomeric form on the surface of the phages. Other possibilities include even more stable, aggregation resistant proteins (Jespers et al. 2004). This could be achieved by pre-incubation under extreme conditions (temperature, denaturants, harsh pH) before the biopanning step (Jespers et al. 2004, Famm et al. 2008). Eventually, highly stable chimeric mutants could be used to improve unstable, aggregation prone avidin mutants such as the steroid-binding avidins (Riihimaki et al. 2011a) to make them more suitable for large-scale production and diagnostic use.

6.5   Zebavidin and the physiological function of avidins

Based on our study, zebavidin from zebrafish could be categorized as a member of the avidin protein family. However, its biotin-binding affinity and thermal stability were among the lowest seen so far in avidins (IV). These differing properties can be explained by the three-dimensional structure of zebavidin. Nevertheless, the significance of these differences on the physiological function remains concealed. The difference in thermal stability could be explained by the differences in the body temperatures of chicken and zebrafish. While the chicken maintains a constant body temperature of approximately 40 °C, the body temperature of the zebrafish depends

on the water temperature, which is approximately 25 °C. Thus, zebavidin might simply not need to be as stable at high temperatures as chicken avidin.

The impact of the lowered biotin-binding affinity on the physiological function of zebavidin is not clear. The biotin scavenging ability is highly reduced in zebavidin, which displays a biotin-binding affinity approximately a million fold lower compared with that of chicken avidin. This raises the question of the physiological importance of zebavidin. Our studies did not show any biological effect caused by the morpholinos targeted to block zebavidin expression (IV). However, further studies are needed to confirm the efficiency of the morpholino. Low zebavidin expression levels measured in developing zebrafish embryos suggested that zebavidin may not play an important role in zebrafish development and most probably zebavidin is deposited to the eggs by the mother. It is also important to notice, that due to the importance of biotin for cellular processes, avidin may be very toxic. This might be the reason for its tightly controlled expression in chicken (Chan, Means & O'Malley 1973, Elo, Kulomaa & Tuohimaa 1979, Matulova et al. 2012). It has also been found that in developing bird embryos, a high avidin concentration, combined with a low amount of biotin in the diet, can be disastrous, as has been observed in studies performed in turkeys (White, Whitehead & Armstrong 1987, Robel 1987, Robel 1991). Taking into account the differences in the structures of fish and bird eggs, zebavidin's low affinity for biotin may be a compromise between toxicity and scavenging.

Alternatively to a scavenging function, zebavidin could have a role similar to the one observed for the biotin-binding proteins in chicken. These members of the avidin protein family are located in the egg yolk and display lower biotin-binding affinities in comparison with avidin. They have been suggested to act as transport proteins for the deposition of biotin into the egg yolk for the developing embryo (White et al. 1976, Niskanen et al. 2005, Hytönen et al. 2007, White 1985). However, these biotin-binding proteins are expressed in the liver, not in gonads as are chicken avidin and zebavidin (White et al. 1976, White, Whitehead 1987).

To obtain a complete picture of the physiological function of zebavidin and its effect on developing and adult fish, extensive additional experiments in zebrafish are required. One important question would be to determine the importance of zebavidin for the protection of eggs against environmental factors such as parasites,

similarly to that observed in case of tamavidin (Bleuler-Martinez et al. 2012). Zebavidin knockout zebrafish would be an attractive tool for such experiments.

7. SUMMARY AND CONCLUSION

Avidin is an important protein in biotechnology, yet it is sometimes difficult to predict the effect a mutation has on the structure and function of this protein. Therefore, the use of directed molecular evolution methods is likely to be beneficial for the engineering of avidin. In this thesis, the engineering of the avidin protein was improved through random mutagenesis and directed protein evolution. Applying DNA shuffling to members of the avidin family in combination with phage display proved to be a feasible approach for creating functional avidin proteins with altered structural and functional properties. The parameters of the selection methods proved to be crucial for the successful selection of functional mutants, and were effectively modified to select for a mutant with improved biotin binding and stability.

Despite the success of random mutagenesis approaches, rational protein design is still a powerful method for engineering proteins as is shown in the design of AVD M96H and nAVD M96H&R114L. A change in a single amino acid is sufficient to alter the biotin-binding properties in the desired fashion. This study showed the successful application of AVD M96H in the design of reusable biosensor surfaces. The undesired non-specific binding observed with AVD M96H was successfully reduced by introducing neutralizing mutations by rational design.

The characterization of zebavidin from zebrafish confirmed the hypothesis that the predicted avidin-like gene encoded a biotin-binding protein with a three dimensional structure highly similar to the other avidin protein family members. Quantitative analysis of zebavidin expression in zebrafish confirmed the previous findings in chicken. However, further studies are required to completely understand the physiological function of zebavidin as well as other avidins.

ACKNOWLEDGEMENTS

REFERENCES

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. & Terwilliger, T.C. 2002, "PHENIX: building new software for automated crystallographic structure determination", *Acta crystallographica.Section D, Biological crystallography,* vol. 58, no. 11, pp. 1948-1954.

Airenne, K.J., Oker-Blom, C., Marjomaki, V.S., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 1997, "Production of biologically active recombinant avidin in baculovirus-infected insect cells", *Protein expression and purification,* vol. 9, no. 1, pp. 100-108.

Araya, C.L. & Fowler, D.M. 2011, "Deep mutational scanning: assessing protein function on a massive scale", *Trends in biotechnology,* vol. 29, no. 9, pp. 435-442.

Arnold, F.H. 1998a, "Design by Directed Evolution", *Accounts of Chemical Research,* vol. 31, no. 3, pp. 125-131.

Arnold, F.H. 1998b, "When blind is better: protein design by evolution", *Nature biotechnology,* vol. 16, no. 7, pp. 617-618.

Aslan, F.M., Yu, Y., Mohr, S.C. & Cantor, C.R. 2005, "Engineered single-chain dimeric streptavidins with an unexpected strong preference for biotin-4-fluorescein", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 102, no. 24, pp. 8507-8512.

Axe, D.D. 2004, "Estimating the prevalence of protein sequences adopting functional enzyme folds", *Journal of Molecular Biology,* vol. 341, no. 5, pp. 1295-1315.

Baca, M., Scanlan, T.S., Stephenson, R.C. & Wells, J.A. 1997, "Phage display of a catalytic antibody to optimize affinity for transition-state analog binding", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 19, pp. 10063-10068.

Barbas III, C.F., Burton, D.R., Jamie K. Scott, Jamie K. & Silverman, G.J. 2001, *Phage Display: A Laboratory Manual, .*

Barr, M.M. 2003, "Super models", *Physiological genomics,* vol. 13, no. 1, pp. 15-24.

Bateman, W.G. 1916, "The Digestibility and Utilization of Egg Proteins", *J.Biol.Chem.,* vol. 26, pp. 263-291.

Beckman, R.A., Mildvan, A.S. & Loeb, L.A. 1985, "On the fidelity of DNA replication: manganese mutagenesis in vitro", *Biochemistry,* vol. 24, no. 21, pp. 5810-5817.

Berger, J. & Currie, P.D. 2012, "Zebrafish models flex their muscles to shed light on muscular dystrophies", *Disease models & mechanisms,* vol. 5, no. 6, pp. 726-732.

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Bourne, P.E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J.D. & Zardecki, C. 2002, "The Protein Data Bank", *Acta crystallographica.Section D, Biological crystallography,* vol. 58, no. 1, pp. 899-907.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. 2000, "The Protein Data Bank", *Nucleic acids research,* vol. 28, no. 1, pp. 235-242.

Bershtein, S., Goldin, K. & Tawfik, D.S. 2008, "Intense neutral drifts yield robust and evolvable consensus proteins", *Journal of Molecular Biology,* vol. 379, no. 5, pp. 1029-1044.

Bier, E. 2005, "Drosophila, the golden bug, emerges as a tool for human genetics", *Nature reviews.Genetics,* vol. 6, no. 1, pp. 9-23.

Bill, B.R., Petzold, A.M., Clark, K.J., Schimmenti, L.A. & Ekker, S.C. 2009, "A primer for morpholino use in zebrafish", *Zebrafish,* vol. 6, no. 1, pp. 69-77.

Bleuler-Martinez, S., Schmieder, S., Aebi, M. & Kunzler, M. 2012, "Biotin-binding proteins in the defense of mushrooms against predators and parasites", *Applied and Environmental Microbiology,* vol. 78, no. 23, pp. 8485-8487.

Bloom, J.D., Romero, P.A., Lu, Z. & Arnold, F.H. 2007, "Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution", *Biology direct,* vol. 2, no. 17.

Boder, E.T. & Wittrup, K.D. 1997, "Yeast surface display for screening combinatorial polypeptide libraries", *Nature biotechnology,* vol. 15, no. 6, pp. 553-557.

Bush, L., McGahan, T.J. & White, H.B.,3rd 1988, "Purification and characterization of biotin-binding protein II from chicken oocytes", *The Biochemical journal,* vol. 256, no. 3, pp. 797-805.

Cadwell, R.C. & Joyce, G.F. 1992, "Randomization of genes by PCR mutagenesis", *PCR methods and applications,* vol. 2, no. 1, pp. 28-33.

Callison, S., Hilt, D. & Jackwood, M. 2005, "Using DNA shuffling to create novel infectious bronchitis virus s1 genes: implications for s1 gene recombination", *Virus genes,* vol. 31, no. 1, pp. 5-11.

Cedrone, F., Menez, A. & Quemeneur, E. 2000, "Tailoring new enzyme functions by rational redesign", *Current opinion in structural biology,* vol. 10, no. 4, pp. 405-410.

Chaiet, L. & Wolf, F.J. 1964, "The Properties of Streptavidin, a Biotin-Binding Protein Produced by Streptomycetes", *Archives of Biochemistry and Biophysics,* vol. 106, pp. 1-5.

Chan, L., Means, A.R. & O'Malley, B.W. 1973, "Rates of induction of specific translatable messenger RNAs for ovalbumin and avidin by steroid hormones", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 70, no. 6, pp. 1870-1874.

Chang, C.C., Chen, T.T., Cox, B.W., Dawes, G.N., Stemmer, W.P., Punnonen, J. & Patten, P.A. 1999, "Evolution of a cytokine using DNA family shuffling", *Nature biotechnology,* vol. 17, no. 8, pp. 793-797.

Chatterjee, R. & Yuan, L. 2006, "Directed evolution of metabolic pathways", *Trends in biotechnology,* vol. 24, no. 1, pp. 28-38.

Chilkoti, A., Tan, P.H. & Stayton, P.S. 1995, "Site-directed mutagenesis studies of the high-affinity streptavidin-biotin complex: contributions of tryptophan residues 79, 108, and 120", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 92, no. 5, pp. 1754-1758.

Chothia, C., Gough, J., Vogel, C. & Teichmann, S.A. 2003, "Evolution of the protein repertoire", *Science (New York, N.Y.),* vol. 300, no. 5626, pp. 1701-1703.

Cochran, J.R., Kim, Y.S., Lippow, S.M., Rao, B. & Wittrup, K.D. 2006, "Improved mutants from directed evolution are biased to orthologous substitutions", *Protein engineering, design & selection : PEDS,* vol. 19, no. 6, pp. 245-253.

Coco, W.M., Levinson, W.E., Crist, M.J., Hektor, H.J., Darzins, A., Pienkos, P.T., Squires, C.H. & Monticello, D.J. 2001, "DNA shuffling method for generating highly recombined genes and evolved enzymes", *Nature biotechnology,* vol. 19, no. 4, pp. 354-359.

Collaborative Computational Project, Number 4 1994, "The CCP4 suite: programs for protein crystallography", *Acta crystallographica.Section D, Biological crystallography,* vol. 50, no. 5, pp. 760-763.

Corey, D.R. & Abrams, J.M. 2001, "Morpholino antisense oligonucleotides: tools for investigating vertebrate development", *Genome biology,* vol. 2, no. 5, pp. REVIEWS1015.

Crameri, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. 1998, "DNA shuffling of a family of genes from diverse species accelerates directed evolution", *Nature,* vol. 391, no. 6664, pp. 288-291.

Dahiyat, B.I. & Mayo, S.L. 1996, "Protein design automation", *Protein science : a publication of the Protein Society,* vol. 5, no. 5, pp. 895-903.

Darwin C. 1859, "On the origin of species or the preservation of favoured races in the struggle for life", *John Murray, United Kingdom.*

Daugherty, P.S. 2007, "Protein engineering with bacterial display", *Current opinion in structural biology,* vol. 17, no. 4, pp. 474-480.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B.,3rd, Snoeyink, J., Richardson, J.S. & Richardson, D.C. 2007, "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids", *Nucleic acids research,* vol. 35, no. Web Server issue, pp. W375-383.

Dawkins R. 1985, "The blind watchmaker: why the evidence of evolution reveals a universe without design", *W.W. Norton & Company, New York, London.*

Dayhoff, M.O., McLaughlin, P.J., Barker, W.C. & Hunt, L.T. 1975, "Evolution of Sequences within Protein Superfamilies", *Naturwissenschaften,* vol. 62, pp. 154-161.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L., Neri, P., Winter, G. & Neri, D. 1999, "A strategy for the isolation of catalytic activities from repertoires of enzymes displayed on phage", *Journal of Molecular Biology,* vol. 286, no. 2, pp. 617-633.

Demuth, J.P. & Hahn, M.W. 2009, "The life and death of gene families", *BioEssays : news and reviews in molecular, cellular and developmental biology,* vol. 31, no. 1, pp. 29-39.

Diamandis, E.P. & Christopoulos, T.K. 1991, "The biotin-(strept)avidin system: principles and applications in biotechnology", *Clinical chemistry,* vol. 37, no. 5, pp. 625-636.

Dion, M., Nisole, A., Spangenberg, P., Andre, C., Glottin-Fleury, A., Mattes, R., Tellier, C. & Rabiller, C. 2001, "Modulation of the regioselectivity of a Bacillus alpha-galactosidase by directed evolution", *Glycoconjugate journal,* vol. 18, no. 3, pp. 215-223.

Droge, M.J., Boersma, Y.L., van Pouderoyen, G., Vrenken, T.E., Ruggeberg, C.J., Reetz, M.T., Dijkstra, B.W. & Quax, W.J. 2006, "Directed evolution of Bacillus subtilis lipase A by use of enantiomeric phosphonate inhibitors: crystal structures and phage display selection", *Chembiochem : a European journal of chemical biology,* vol. 7, no. 1, pp. 149-157.

Droge, M.J., Ruggeberg, C.J., van der Sloot, A.M., Schimmel, J., Dijkstra, D.S., Verhaert, R.M., Reetz, M.T. & Quax, W.J. 2003, "Binding of phage displayed

Bacillus subtilis lipase A to a phosphonate suicide inhibitor", *Journal of Biotechnology,* vol. 101, no. 1, pp. 19-28.

Du, L. & Gatti, R.A. 2011, "Potential therapeutic applications of antisense morpholino oligonucleotides in modulation of splicing in primary immunodeficiency diseases", *Journal of immunological methods,* vol. 365, no. 1-2, pp. 1-7.

Eggertsson, G. & Soll, D. 1988, "Transfer ribonucleic acid-mediated suppression of termination codons in Escherichia coli", *Microbiological reviews,* vol. 52, no. 3, pp. 354-374.

Eijsink, V.G., Bjork, A., Gaseidnes, S., Sirevag, R., Synstad, B., van den Burg, B. & Vriend, G. 2004, "Rational engineering of enzyme stability", *Journal of Biotechnology,* vol. 113, no. 1-3, pp. 105-120.

Elo, H.A., Kulomaa, M.S. & Tuohimaa, P.J. 1979, "Avidin induction by tissue injury and inflammation in male and female chickens", *Comparative biochemistry and physiology.B, Comparative biochemistry,* vol. 62, no. 3, pp. 237-240.

Elo, H.A., Raisanen, S. & Tuohimaa, P.J. 1980, "Induction of an antimicrobial biotin-binding egg white protein (avidin) in chick tissues in septic Escherichia coli infection", *Experientia,* vol. 36, no. 3, pp. 312-313.

Emsley, P. & Cowtan, K. 2004, "Coot: model-building tools for molecular graphics", *Acta crystallographica.Section D, Biological crystallography,* vol. 60, no. 12, pp. 2126-2132.

Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. 2009, "Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes", *PLOS ONE,* vol. 4, no. 5, pp. e5553.

Ericsson, R., Knight, R. & Johanson, Z. 2013, "Evolution and development of the vertebrate neck", *Journal of anatomy,* vol. 222, no. 1, pp. 67-78.

Famm, K., Hansen, L., Christ, D. & Winter, G. 2008, "Thermodynamically stable aggregation-resistant antibody domains through directed evolution", *Journal of Molecular Biology,* vol. 376, no. 4, pp. 926-931.

Fasan, R., Meharenna, Y.T., Snow, C.D., Poulos, T.L. & Arnold, F.H. 2008, "Evolutionary history of a specialized p450 propane monooxygenase", *Journal of Molecular Biology,* vol. 383, no. 5, pp. 1069-1080.

Forrer, P., Jung, S. & Pluckthun, A. 1999, "Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins", *Current opinion in structural biology,* vol. 9, no. 4, pp. 514-520.

Fowler, D.M., Araya, C.L., Fleishman, S.J., Kellogg, E.H., Stephany, J.J., Baker, D. & Fields, S. 2010, "High-resolution mapping of protein sequence-function relationships", *Nature methods,* vol. 7, no. 9, pp. 741-746.

Fox, R. 2005, "Directed molecular evolution by machine learning and the influence of nonlinear interactions", *Journal of theoretical biology,* vol. 234, no. 2, pp. 187-199.

Fox, R., Roy, A., Govindarajan, S., Minshull, J., Gustafsson, C., Jones, J.T. & Emig, R. 2003, "Optimizing the search algorithm for protein engineering by directed evolution", *Protein engineering,* vol. 16, no. 8, pp. 589-597.

Garcia-Aljaro, C., Munoz, F.X. & Baldrich, E. 2009, "Captavidin: a new regenerable biocomponent for biosensing?", *The Analyst,* vol. 134, no. 11, pp. 2338-2343.

Gentry-Weeks, C.R., Hultsch, A.L., Kelly, S.M., Keith, J.M. & Curtiss, R.,3rd 1992, "Cloning and sequencing of a gene encoding a 21-kilodalton outer membrane

protein from Bordetella avium and expression of the gene in Salmonella typhimurium", *Journal of Bacteriology,* vol. 174, no. 23, pp. 7729-7742.

Gille, C. 2006, "Structural interpretation of mutations and SNPs using STRAP-NT", *Protein science: a publication of the Protein Society,* vol. 15, no. 1, pp. 208-210.

Green, N.M. 1975, "Avidin", *Advances in Protein Chemistry,* vol. 29, pp. 85-133.

Guo, H.H., Choe, J. & Loeb, L.A. 2004, "Protein tolerance to random amino acid change", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 101, no. 25, pp. 9205-9210.

Hanes, J. & Pluckthun, A. 1997, "In vitro selection and evolution of functional proteins by using ribosome display", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 10, pp. 4937-4942.

Hayes, R.J., Bentzien, J., Ary, M.L., Hwang, M.Y., Jacinto, J.M., Vielmetter, J., Kundu, A. & Dahiyat, B.I. 2002, "Combining computational and experimental screening for rapid optimization of protein properties", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, no. 25, pp. 15926-15931.

He, L., Friedman, A.M. & Bailey-Kellogg, C. 2012, "Algorithms for optimizing cross-overs in DNA shuffling", *BMC bioinformatics,* vol. 13, no. 3, pp. S3.

Heasman, J. 2002, "Morpholino oligos: making sense of antisense?", *Developmental biology,* vol. 243, no. 2, pp. 209-214.

Hedges, S.B. 2002, "The origin and evolution of model organisms", *Nature reviews.Genetics,* vol. 3, no. 11, pp. 838-849.

Helppolainen, S.H., Määttä, J.A., Halling, K.K., Slotte, J.P., Hytönen, V.P., Jänis, J., Vainiotalo, P., Kulomaa, M.S. & Nordlund, H.R. 2008, "Bradavidin II from Bradyrhizobium japonicum: a new avidin-like biotin-binding protein", *Biochimica et biophysica acta,* vol. 1784, no. 7-8, pp. 1002-1010.

Helppolainen, S.H., Nurminen, K.P., Maatta, J.A., Halling, K.K., Slotte, J.P., Huhtala, T., Liimatainen, T., Yla-Herttuala, S., Airenne, K.J., Narvanen, A., Janis, J., Vainiotalo, P., Valjakka, J., Kulomaa, M.S. & Nordlund, H.R. 2007, "Rhizavidin from Rhizobium etli: the first natural dimer in the avidin protein family", *The Biochemical journal,* vol. 405, no. 3, pp. 397-405.

Hendrickson, W.A., Pahler, A., Smith, J.L., Satow, Y., Merritt, E.A. & Phizackerley, R.P. 1989, "Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 86, no. 7, pp. 2190-2194.

Hietpas, R.T., Jensen, J.D. & Bolon, D.N. 2011, "Experimental illumination of a fitness landscape", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 108, no. 19, pp. 7896-7901.

Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. & Pease, L.R. 1989, "Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension", *Gene,* vol. 77, no. 1, pp. 61-68.

Howarth, M., Chinnapen, D.J., Gerrow, K., Dorrestein, P.C., Grandy, M.R., Kelleher, N.L., El-Husseini, A. & Ting, A.Y. 2006, "A monovalent streptavidin with a single femtomolar biotin-binding site", *Nature methods,* vol. 3, no. 4, pp. 267-273.

Hunter, P. 2008, "The paradox of model organisms. The use of model organisms in research will continue despite their shortcomings", *EMBO reports,* vol. 9, no. 8, pp. 717-720.

Hytönen, V.P. 2005, "The avidin protein family: properties of family members and engineering of novel biotin-binding protein tools", *Doctoral Dissertation, Department of Biological and Environmental Science, University of Jyväskylä*, no. 153.

Hytönen, V.P., Laitinen, O.H., Airenne, T.T., Kidron, H., Meltola, N.J., Porkka, E.J., Hörhä, J., Paldanius, T., Määttä, J.A., Nordlund, H.R., Johnson, M.S., Salminen, T.A., Airenne, K.J., Yla-Herttuala, S. & Kulomaa, M.S. 2004, "Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli", *The Biochemical journal,* vol. 384, no. 2, pp. 385-390.

Hytönen, V.P., Määttä, J.A., Kidron, H., Halling, K.K., Hörhä, J., Kulomaa, T., Nyholm, T.K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S. & Airenne, T.T. 2005, "Avidin related protein 2 shows unique structural and functional features among the avidin protein family", *BMC biotechnology,* vol. 5, no. 28.

Hytönen, V.P., Määttä, J.A., Niskanen, E.A., Huuskonen, J., Helttunen, K.J., Halling, K.K., Nordlund, H.R., Rissanen, K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S., Laitinen, O.H. & Airenne, T.T. 2007, "Structure and characterization of a novel chicken biotin-binding protein A (BBP-A)", *BMC structural biology,* vol. 7, no. 8.

Hytönen, V.P., Määttä, J.A., Nyholm, T.K., Livnah, O., Eisenberg-Domovich, Y., Hyre, D., Nordlund, H.R., Horha, J., Niskanen, E.A., Paldanius, T., Kulomaa, T., Porkka, E.J., Stayton, P.S., Laitinen, O.H. & Kulomaa, M.S. 2005a, "Design and construction of highly stable, protease-resistant chimeric avidins", *The Journal of biological chemistry,* vol. 280, no. 11, pp. 10228-10233.

Hytönen, V.P., Nordlund, H.R., Hörhä, J., Nyholm, T.K., Hyre, D.E., Kulomaa, T., Porkka, E.J., Marttila, A.T., Stayton, P.S., Laitinen, O.H. & Kulomaa, M.S. 2005b, "Dual-affinity avidin molecules", *Proteins,* vol. 61, no. 3, pp. 597-607.

Hytönen, V.P., Nyholm, T.K., Pentikäinen, O.T., Vaarno, J., Porkka, E.J., Nordlund, H.R., Johnson, M.S., Slotte, J.P., Laitinen, O.H. & Kulomaa, M.S. 2004, "Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin", *The Journal of biological chemistry,* vol. 279, no. 10, pp. 9337-9343.

Jackel, C., Kast, P. & Hilvert, D. 2008, "Protein design by directed evolution", *Annual review of biophysics,* vol. 37, pp. 153-173.

Jespers, L., Schon, O., Famm, K. & Winter, G. 2004, "Aggregation-resistant domain antibodies selected on phage by heat denaturation", *Nature biotechnology,* vol. 22, no. 9, pp. 1161-1165.

Jung, S., Honegger, A. & Pluckthun, A. 1999, "Selection for improved protein stability by phage display", *Journal of Molecular Biology,* vol. 294, no. 1, pp. 163-180.

Kabsch, W. 1993, "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants", *Journal of applied crystallography,* vol. 26, pp. 795-800.

Kaletta, T. & Hengartner, M.O. 2006, "Finding function in novel targets: C. elegans as a model organism", *Nature reviews.Drug discovery,* vol. 5, no. 5, pp. 387-398.

Keinänen, R.A., Wallén, M.J., Kristo, P.A., Laukkanen, M.O., Toimela, T.A., Helenius, M.A. & Kulomaa, M.S. 1994, "Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5", *European journal of biochemistry / FEBS,* vol. 220, no. 2, pp. 615-621.

Khalil, A.S. & Collins, J.J. 2010, "Synthetic biology: applications come of age", *Nature reviews.Genetics,* vol. 11, no. 5, pp. 367-379.

Kikuchi, M., Ohnishi, K. & Harayama, S. 1999, "Novel family shuffling methods for the in vitro evolution of enzymes", *Gene,* vol. 236, no. 1, pp. 159-167.

Kim, H.S., Lo, S.C., Wear, D.J., Stojadinovic, A., Weina, P.J. & Izadjoo, M.J. 2011, "Improvement of anti-Burkholderia mouse monoclonal antibody from various phage-displayed single-chain antibody libraries", *Journal of immunological methods,* vol. 372, no. 1-2, pp. 146-161.

Kinkel, M.D. & Prince, V.E. 2009, "On the diabetic menu: zebrafish as a model for pancreas development and function", *BioEssays : news and reviews in molecular, cellular and developmental biology,* vol. 31, no. 2, pp. 139-152.

Kirk, O., Borchert, T.V. & Fuglsang, C.C. 2002, "Industrial enzyme applications", *Current opinion in biotechnology,* vol. 13, no. 4, pp. 345-351.

Klee, E.W., Schneider, H., Clark, K.J., Cousin, M.A., Ebbert, J.O., Hooten, W.M., Karpyak, V.M., Warner, D.O. & Ekker, S.C. 2012, "Zebrafish: a model for the study of addiction genetics", *Human genetics,* vol. 131, no. 6, pp. 977-1008.

Klumb, L.A., Chu, V. & Stayton, P.S. 1998, "Energetic roles of hydrogen bonds at the ureido oxygen binding pocket in the streptavidin-biotin complex", *Biochemistry,* vol. 37, no. 21, pp. 7657-7663.

Kolkman, J.A. & Stemmer, W.P. 2001, "Directed evolution of proteins by exon shuffling", *Nature biotechnology,* vol. 19, no. 5, pp. 423-428.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. 1983, "Avidin is induced in chicken embryo fibroblasts by viral transformation and cell damage", *The EMBO journal,* vol. 2, no. 10, pp. 1715-1719.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. 1982, "Induction of avidin in chickens infected with the acute leukemia virus OK 10", *International journal of cancer.Journal international du cancer,* vol. 30, no. 4, pp. 461-464.

Kunnas, T.A., Wallen, M.J. & Kulomaa, M.S. 1993, "Induction of chicken avidin and related mRNAs after bacterial infection", *Biochimica et biophysica acta,* vol. 1216, no. 3, pp. 441-445.

Laffly, E., Pelat, T., Cedrone, F., Blesa, S., Bedouelle, H. & Thullier, P. 2008, "Improvement of an antibody neutralizing the anthrax toxin by simultaneous mutagenesis of its six hypervariable loops", *Journal of Molecular Biology,* vol. 378, no. 5, pp. 1094-1103.

Laitinen, O.H., Airenne, K.J., Marttila, A.T., Kulik, T., Porkka, E., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 1999, "Mutation of a critical tryptophan to lysine in avidin or streptavidin may explain why sea urchin fibropellin adopts an avidin-like domain", *FEBS letters,* vol. 461, no. 1-2, pp. 52-58.

Laitinen, O.H., Hytönen, V.P., Nordlund, H.R. & Kulomaa, M.S. 2006, "Genetically engineered avidins and streptavidins", *Cellular and molecular life sciences: CMLS,* vol. 63, no. 24, pp. 2992-3017.

Laitinen, O.H., Nordlund, H.R., Hytönen, V.P. & Kulomaa, M.S. 2007, "Brave new (strept)avidins in biotechnology", *Trends in biotechnology,* vol. 25, no. 6, pp. 269-277.

Laitinen, O.H., Nordlund, H.R., Hytönen, V.P., Uotila, S.T., Marttila, A.T., Savolainen, J., Airenne, K.J., Livnah, O., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 2003, "Rational design of an active avidin monomer", *The Journal of biological chemistry,* vol. 278, no. 6, pp. 4010-4014.

Lamzin, V.S. & Wilson, K.S. 1993, "Automated refinement of protein models", *Acta crystallographica.Section D, Biological crystallography,* vol. 49, no. 1, pp. 129-147.

Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. 2008, "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7", *Nature protocols,* vol. 3, no. 7, pp. 1171-1179.

Lassner, M.W. & McElroy, D. 2002, "Directed molecular evolution: bridging the gap between genomics leads and commercial products", *Omics : a journal of integrative biology,* vol. 6, no. 2, pp. 153-162.

Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L. & van Loon, A.P. 2000, "From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase", *Protein engineering,* vol. 13, no. 1, pp. 49-57.

Leppiniemi, J., Meir, A., Kahkonen, N., Kukkurainen, S., Maatta, J.A., Ojanen, M., Janis, J., Kulomaa, M.S., Livnah, O. & Hytonen, V.P. 2013, "The highly dynamic oligomeric structure of bradavidin II is unique among avidin proteins", *Protein Science,* vol. 22, no. 7, pp. 980-994.

Levy, M. & Ellington, A.D. 2008, "Directed evolution of streptavidin variants using in vitro compartmentalization", *Chemistry & biology,* vol. 15, no. 9, pp. 979-989.

Li, M.Z. & Elledge, S.J. 2007, "Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC", *Nature methods,* vol. 4, no. 3, pp. 251-256.

Lien, C.L., Harrison, M.R., Tuan, T.L. & Starnes, V.A. 2012, "Heart repair and regeneration: recent insights from zebrafish studies", *Wound repair and regeneration : official publication of the Wound Healing Society [and] the European Tissue Repair Society,* vol. 20, no. 5, pp. 638-646.

Lim, K.H., Huang, H., Pralle, A. & Park, S. 2013, "Stable, high-affinity streptavidin monomer for protein labeling and monovalent biotin detection", *Biotechnology and bioengineering,* vol. 110, no. 1, pp. 57-67.

Lin, H. & Cornish, V.W. 2002, "Screening and selection methods for large-scale analysis of protein function", *Angewandte Chemie (International ed.in English),* vol. 41, no. 23, pp. 4402-4425.

Liu, S. & Leach, S.D. 2011, "Zebrafish models for cancer", *Annual review of pathology,* vol. 6, pp. 71-93.

Livnah, O., Bayer, E.A., Wilchek, M. & Sussman, J.L. 1993, "Three-dimensional structures of avidin and the avidin-biotin complex", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 90, no. 11, pp. 5076-5080.

Lu, J.W., Hsia, Y., Tu, H.C., Hsiao, Y.C., Yang, W.Y., Wang, H.D. & Yuh, C.H. 2011, "Liver development and cancer formation in zebrafish", *Birth defects research.Part C, Embryo today : reviews,* vol. 93, no. 2, pp. 157-172.

Lutz, S., Ostermeier, M. & Benkovic, S.J. 2001, "Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides", *Nucleic acids research,* vol. 29, no. 4, pp. e16

Lutz, S., Ostermeier, M., Moore, G.L., Maranas, C.D. & Benkovic, S.J. 2001, "Creating multiple-crossover DNA libraries independent of sequence identity", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, no. 20, pp. 11248-11253.

Määttä, J.A., Eisenberg-Domovich, Y., Nordlund, H.R., Hayouka, R., Kulomaa, M.S., Livnah, O. & Hytönen, V.P. 2011, "Chimeric avidin shows stability against harsh chemical conditions-biochemical analysis and 3D structure", *Biotechnology and bioengineering,* vol. 108, no. 3, pp. 481-490.

Määttä, J.A., Helppolainen, S.H., Hytönen, V.P., Johnson, M.S., Kulomaa, M.S., Airenne, T.T. & Nordlund, H.R. 2009, "Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis", *BMC structural biology,* vol. 9, no. 63.

Majumder, K. 1992, "Ligation-free gene synthesis by PCR: synthesis and mutagenesis at multiple loci of a chimeric gene encoding OmpA signal peptide and hirudin", *Gene,* vol. 110, no. 1, pp. 89-94.

Marshall, S.A., Lazar, G.A., Chirino, A.J. & Desjarlais, J.R. 2003, "Rational design and engineering of therapeutic proteins", *Drug discovery today,* vol. 8, no. 5, pp. 212-221.

Marttila, A.T., Airenne, K.J., Laitinen, O.H., Kulik, T., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 1998, "Engineering of chicken avidin: a progressive series of reduced charge mutants", *FEBS letters,* vol. 441, no. 2, pp. 313-317.

Marttila, A.T., Hytonen, V.P., Laitinen, O.H., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 2003, "Mutation of the important Tyr-33 residue of chicken avidin: functional and structural consequences", *The Biochemical journal,* vol. 369, no. 2, pp. 249-254.

Marttila, A.T., Laitinen, O.H., Airenne, K.J., Kulik, T., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 2000, "Recombinant NeutraLite avidin: a non-glycosylated, acidic mutant of chicken avidin that exhibits high affinity for biotin and low non-specific binding properties", *FEBS letters,* vol. 467, no. 1, pp. 31-36.

Matochko, W.L., Chu, K., Jin, B., Lee, S.W., Whitesides, G.M. & Derda, R. 2012, "Deep sequencing analysis of phage libraries using Illumina platform", *Methods (San Diego, Calif.),* vol. 58, no. 1, pp. 47-55.

Matulova, M., Rajova, J., Vlasatikova, L., Volf, J., Stepanova, H., Havlickova, H., Sisak, F. & Rychlik, I. 2012, "Characterization of chicken spleen transcriptome after infection with Salmonella enterica serovar Enteritidis", *PLOS ONE,* vol. 7, no. 10, pp. e48101.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. & Read, R.J. 2007, "Phaser crystallographic software", *Journal of applied crystallography,* vol. 40, no. 4, pp. 658-674.

Meir, A., Bayer, E.A. & Livnah, O. 2012, "Structural adaptation of a thermostable biotin-binding protein in a psychrophilic environment", *The Journal of biological chemistry,* vol. 287, no. 22, pp. 17951-17962.

Meir, A., Helppolainen, S.H., Podoly, E., Nordlund, H.R., Hytonen, V.P., Maatta, J.A., Wilchek, M., Bayer, E.A., Kulomaa, M.S. & Livnah, O. 2009, "Crystal structure of rhizavidin: insights into the enigmatic high-affinity interaction of an innate biotin-binding protein dimer", *Journal of Molecular Biology,* vol. 386, no. 2, pp. 379-390.

Meslar, H.W., Camper, S.A. & White, H.B.,3rd 1978, "Biotin-binding protein from egg yolk. A protein distinct from egg white avidin", *The Journal of biological chemistry,* vol. 253, no. 19, pp. 6979-6982.

Metzker, M.L. 2010, "Sequencing technologies - the next generation", *Nature reviews.Genetics,* vol. 11, no. 1, pp. 31-46.

Moore, A.D., Bjorklund, A.K., Ekman, D., Bornberg-Bauer, E. & Elofsson, A. 2008, "Arrangements in the modular evolution of proteins", *Trends in biochemical sciences,* vol. 33, no. 9, pp. 444-451.

Moore, G.L. & Maranas, C.D. 2002, "eCodonOpt: a systematic computational framework for optimizing codon usage in directed evolution experiments", *Nucleic acids research,* vol. 30, no. 11, pp. 2407-2416.

Morag, E., Bayer, E.A. & Wilchek, M. 1996a, "Immobilized nitro-avidin and nitro-streptavidin as reusable affinity matrices for application in avidin-biotin technology", *Analytical Biochemistry,* vol. 243, no. 2, pp. 257-263.

Morag, E., Bayer, E.A. & Wilchek, M. 1996b, "Reversibility of biotin binding by selective modification of tyrosine in avidin", *The Biochemical journal,* vol. 316, no. 1, pp. 193-199.

Murshudov, G.N., Vagin, A.A. & Dodson, E.J. 1997, "Refinement of macromolecular structures by the maximum-likelihood method", *Acta crystallographica.Section D, Biological crystallography,* vol. 53, no. 3, pp. 240-255.

Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. 1995, "SCOP: a structural classification of proteins database for the investigation of sequences and structures", *Journal of Molecular Biology,* vol. 247, no. 4, pp. 536-540.

Nasevicius, A. & Ekker, S.C. 2000, "Effective targeted gene 'knockdown' in zebrafish", *Nature genetics,* vol. 26, no. 2, pp. 216-220.

Ness, J.E., Kim, S., Gottman, A., Pak, R., Krebber, A., Borchert, T.V., Govindarajan, S., Mundorff, E.C. & Minshull, J. 2002, "Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently", *Nature biotechnology,* vol. 20, no. 12, pp. 1251-1255.

Ness, J.E., Welch, M., Giver, L., Bueno, M., Cherry, J.R., Borchert, T.V., Stemmer, W.P. & Minshull, J. 1999, "DNA shuffling of subgenomic sequences of subtilisin", *Nature biotechnology,* vol. 17, no. 9, pp. 893-896.

Neylon, C. 2004, "Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution", *Nucleic acids research,* vol. 32, no. 4, pp. 1448-1459.

Nicholas, K.B., Nicholas, H.B.J. & and Deerfield, D.W.I. 1997, *GeneDoc: Analysis and Visualization of Genetic Variation*, EMBNEW.NEWS 4:14.

Niskanen, E.A., Hytönen, V.P., Grapputo, A., Nordlund, H.R., Kulomaa, M.S. & Laitinen, O.H. 2005, "Chicken genome analysis reveals novel genes encoding biotin-binding proteins related to avidin family", *BMC genomics,* vol. 6, no. 41.

Nordlund, H.R., Hytönen, V.P., Horha, J., Määttä, J.A., White, D.J., Halling, K., Porkka, E.J., Slotte, J.P., Laitinen, O.H. & Kulomaa, M.S. 2005a, "Tetravalent single-chain avidin: from subunits to protein domains via circularly permuted avidins", *The Biochemical journal,* vol. 392, no. 3, pp. 485-491.

Nordlund, H.R., Hytönen, V.P., Laitinen, O.H. & Kulomaa, M.S. 2005b, "Novel avidin-like protein from a root nodule symbiotic bacterium, Bradyrhizobium japonicum", *The Journal of biological chemistry,* vol. 280, no. 14, pp. 13250-13255.

Nordlund, H.R., Hytönen, V.P., Laitinen, O.H., Uotila, S.T., Niskanen, E.A., Savolainen, J., Porkka, E. & Kulomaa, M.S. 2003, "Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding", *FEBS letters,* vol. 555, no. 3, pp. 449-454.

O'Brien, P.J. & Herschlag, D. 1999, "Catalytic promiscuity and the evolution of new enzymatic activities", *Chemistry & biology,* vol. 6, no. 4, pp. R91-R105.

Ostermeier, M., Shim, J.H. & Benkovic, S.J. 1999, "A combinatorial approach to hybrid enzymes independent of DNA homology", *Nature biotechnology,* vol. 17, no. 12, pp. 1205-1209.

Parikka, M., Hammaren, M.M., Harjula, S.K., Halfpenny, N.J., Oksanen, K.E., Lahtinen, M.J., Pajula, E.T., Iivanainen, A., Pesu, M. & Ramet, M. 2012, "Mycobacterium marinum causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish", *PLOS pathogens,* vol. 8, no. 9, pp. e1002944.

Park, H.S., Nam, S.H., Lee, J.K., Yoon, C.N., Mannervik, B., Benkovic, S.J. & Kim, H.S. 2006, "Design and evolution of new catalytic activity with an existing protein scaffold", *Science (New York, N.Y.),* vol. 311, no. 5760, pp. 535-538.

Pazy, Y., Raboy, B., Matto, M., Bayer, E.A., Wilchek, M. & Livnah, O. 2003, "Structure-based rational design of streptavidin mutants with pseudo-catalytic activity", *The Journal of biological chemistry,* vol. 278, no. 9, pp. 7131-7134.

Pedersen, H., Holder, S., Sutherlin, D.P., Schwitter, U., King, D.S. & Schultz, P.G. 1998, "A method for directed evolution and functional cloning of enzymes", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 95, no. 18, pp. 10523-10528.

Perrakis, A., Morris, R. & Lamzin, V.S. 1999, "Automated protein model building combined with iterative structure refinement", *Nature structural biology,* vol. 6, no. 5, pp. 458-463.

Peters, L.L., Robledo, R.F., Bult, C.J., Churchill, G.A., Paigen, B.J. & Svenson, K.L. 2007, "The mouse as a model for human biology: a resource guide for complex trait analysis", *Nature reviews.Genetics,* vol. 8, no. 1, pp. 58-69.

Ponting, C.P. & Russell, R.R. 2002, "The natural history of protein domains", *Annual Review of Biophysics and Biomolecular Structure,* vol. 31, pp. 45-71.

Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. 2003, "A graphical user interface to the CCP4 program suite", *Acta crystallographica.Section D, Biological crystallography,* vol. 59, no. 7, pp. 1131-1137.

Pugliese, L., Coda, A., Malcovati, M. & Bolognesi, M. 1993, "Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 A resolution", *Journal of Molecular Biology,* vol. 231, no. 3, pp. 698-710.

Quaife, N.M., Watson, O. & Chico, T.J. 2012, "Zebrafish: an emerging model of vascular development and remodelling", *Current opinion in pharmacology,* vol. 12, no. 5, pp. 608-614.

Qureshi, M.H., Yeung, J.C., Wu, S.C. & Wong, S.L. 2001, "Development and characterization of a series of soluble tetrameric and monomeric streptavidin muteins with differential biotin-binding affinities", *The Journal of biological chemistry,* vol. 276, no. 49, pp. 46422-46428.

Remington, S.J. 2011, "Green fluorescent protein: a perspective", *Protein science : a publication of the Protein Society,* vol. 20, no. 9, pp. 1509-1519.

Ren, Z.J., Lewis, G.K., Wingfield, P.T., Locke, E.G., Steven, A.C. & Black, L.W. 1996, "Phage display of intact domains at high copy number: a system based on SOC, the small outer capsid protein of bacteriophage T4", *Protein science : a publication of the Protein Society,* vol. 5, no. 9, pp. 1833-1843.

Reznik, G.O., Vajda, S., Sano, T. & Cantor, C.R. 1998, "A streptavidin mutant with altered ligand-binding specificity", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 95, no. 23, pp. 13525-13530.

Riihimaki, T.A., Hiltunen, S., Rangl, M., Nordlund, H.R., Maatta, J.A., Ebner, A., Hinterdorfer, P., Kulomaa, M.S., Takkinen, K. & Hytonen, V.P. 2011a, "Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein", *BMC biotechnology,* vol. 11, no. 64.

Riihimaki, T.A., Kukkurainen, S., Varjonen, S., Horha, J., Nyholm, T.K., Kulomaa, M.S. & Hytonen, V.P. 2011b, "Construction of chimeric dual-chain avidin by tandem fusion of the related avidins", *PLOS ONE,* vol. 6, no. 5, pp. e20535.

Robel, E.J. 1991, "The value of supplemental biotin for increasing hatchability of turkey eggs", *Poultry science,* vol. 70, no. 8, pp. 1716-1722.

Robel, E.J. 1987, "Comparison of avidin and biotin levels in eggs from turkey hens with high and low hatchability records", *Comparative biochemistry and physiology.B, Comparative biochemistry,* vol. 86, no. 2, pp. 265-267.

Roberts, R.W. & Szostak, J.W. 1997, "RNA-peptide fusions for the in vitro selection of peptides and proteins", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 23, pp. 12297-12302.

Romero, P.A. & Arnold, F.H. 2009, "Exploring protein fitness landscapes by directed evolution", *Nature reviews.Molecular cell biology,* vol. 10, no. 12, pp. 866-876.

Rotticci, D., Rotticci-Mulder, J.C., Denman, S., Norin, T. & Hult, K. 2001, "Improved enantioselectivity of a lipase by rational protein engineering", *Chembiochem : a European journal of chemical biology,* vol. 2, no. 10, pp. 766-770.

Rounioja, S., Saralahti, A., Rantala, L., Parikka, M., Henriques-Normark, B., Silvennoinen, O. & Ramet, M. 2012, "Defense of zebrafish embryos against Streptococcus pneumoniae infection is dependent on the phagocytic activity of leukocytes", *Developmental and comparative immunology,* vol. 36, no. 2, pp. 342-348.

Said, H.M., Ortiz, A., McCloud, E., Dyer, D., Moyer, M.P. & Rubin, S. 1998, "Biotin uptake by human colonic epithelial NCM460 cells: a carrier-mediated process shared with pantothenic acid", *The American Journal of Physiology,* vol. 275, no. 5, pp. C1365-C1371.

Sali, A. & Blundell, T.L. 1993, "Comparative protein modelling by satisfaction of spatial restraints", *Journal of Molecular Biology,* vol. 234, no. 3, pp. 779-815.

Sano, T. & Cantor, C.R. 1995, "Intersubunit contacts made by tryptophan 120 with biotin are essential for both strong biotin binding and biotin-induced tighter subunit association of streptavidin", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 92, no. 8, pp. 3180-3184.

Shao, Z. & Arnold, F.H. 1996, "Engineering new functions and altering existing functions", *Current opinion in structural biology,* vol. 6, no. 4, pp. 513-518.

Shaw, A., Bott, R. & Day, A.G. 1999, "Protein engineering of alpha-amylase for low pH performance", *Current opinion in biotechnology,* vol. 10, no. 4, pp. 349-352.

Sidhu, S.S., Weiss, G.A. & Wells, J.A. 2000, "High copy display of large proteins on phage for functional selections", *Journal of Molecular Biology,* vol. 296, no. 2, pp. 487-495.

Sieber, V., Martinez, C.A. & Arnold, F.H. 2001, "Libraries of hybrid proteins from distantly related sequences", *Nature biotechnology,* vol. 19, no. 5, pp. 456-460.

Sieber, V., Pluckthun, A. & Schmid, F.X. 1998, "Selecting proteins with improved stability by a phage-based method", *Nature biotechnology,* vol. 16, no. 10, pp. 955-960.

Sigurskjold, B.W. 2000, "Exact analysis of competition ligand binding by displacement isothermal titration calorimetry", *Analytical Biochemistry,* vol. 277, no. 2, pp. 260-266.

Singaravelan, B., Roshini, B.R. & Munavar, M.H. 2010, "Evidence that the supE44 mutation of Escherichia coli is an amber suppressor allele of glnX and that it also suppresses ochre and opal nonsense mutations", *Journal of Bacteriology,* vol. 192, no. 22, pp. 6039-6044.

Smith, G.P. 1985, "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface", *Science (New York, N.Y.),* vol. 228, no. 4705, pp. 1315-1317.

Smith, G.P. & Petrenko, V.A. 1997, "Phage Display", *Chemical reviews,* vol. 97, no. 2, pp. 391-410.

Sreenivasan, R., Cai, M., Bartfai, R., Wang, X., Christoffels, A. & Orban, L. 2008, "Transcriptomic analyses reveal novel genes with sexually dimorphic expression in the zebrafish gonad and brain", *PLOS ONE,* vol. 3, no. 3, pp. e1791.

Stemmer, W.P. 1994, "DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 91, no. 22, pp. 10747-10751.

Sternberg, N. & Hoess, R.H. 1995, "Display of peptides and proteins on the surface of bacteriophage lambda", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 92, no. 5, pp. 1609-1613.

Subramanian, N. & Adiga, P.R. 1995, "Simultaneous purification of biotin-binding proteins-I and -II from chicken egg yolk and their characterization", *The Biochemical journal,* vol. 308, no. 2, pp. 573-577.

Summerton, J. & Weller, D. 1997, "Morpholino antisense oligomers: design, preparation, and properties", *Antisense & Nucleic Acid Drug Development,* vol. 7, no. 3, pp. 187-195.

Takakura, Y., Sofuku, K. & Tsunashima, M. 2013, "Tamavidin 2-REV: An engineered tamavidin with reversible biotin-binding capability", *Journal of Biotechnology,* vol. 164, no. 1, pp. 19-25.

Takakura, Y., Tsunashima, M., Suzuki, J., Usami, S., Kakuta, Y., Okino, N., Ito, M. & Yamamoto, T. 2009, "Tamavidins-novel avidin-like biotin-binding proteins from the Tamogitake mushroom", *The FEBS journal,* vol. 276, no. 5, pp. 1383-1397.

Tang, R., Dodd, A., Lai, D., McNabb, W.C. & Love, D.R. 2007, "Validation of zebrafish (Danio rerio) reference genes for quantitative real-time RT-PCR normalization", *Acta biochimica et biophysica Sinica,* vol. 39, no. 5, pp. 384-390.

Taylor, S.V., Walter, K.U., Kast, P. & Hilvert, D. 2001, "Searching sequence space for protein catalysts", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, no. 19, pp. 10596-10601.

Teittinen, K.J., Gronroos, T., Parikka, M., Ramet, M. & Lohi, O. 2012, "The zebrafish as a tool in leukemia research", *Leukemia research,* vol. 36, no. 9, pp. 1082-1088.

Tobia, C., De Sena, G. & Presta, M. 2011, "Zebrafish embryo, a tool to study tumor angiogenesis", *The International journal of developmental biology,* vol. 55, no. 4-5, pp. 505-509.

Tsien, R.Y. 1998, "The green fluorescent protein", *Annual Review of Biochemistry,* vol. 67, pp. 509-544.

Tuohimaa, P., Segal, S.J. & Koide, S.S. 1972, "Induction of avidin synthesis by RNA obtained from chick oviduct", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 69, no. 10, pp. 2814-2817.

Virnekas, B., Ge, L., Pluckthun, A., Schneider, K.C., Wellnhofer, G. & Moroney, S.E. 1994, "Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis", *Nucleic acids research,* vol. 22, no. 25, pp. 5600-5607.

Vogel, C., Bashton, M., Kerrison, N.D., Chothia, C. & Teichmann, S.A. 2004, "Structure, function and evolution of multidomain proteins", *Current opinion in structural biology,* vol. 14, no. 2, pp. 208-216.

Voigt, C.A., Martinez, C., Wang, Z.G., Mayo, S.L. & Arnold, F.H. 2002, "Protein building blocks preserved by recombination", *Nature structural biology,* vol. 9, no. 7, pp. 553-558.

Voigt, C.A., Mayo, S.L., Arnold, F.H. & Wang, Z.G. 2001, "Computational method to reduce the search space for directed protein evolution", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, no. 7, pp. 3778-3783.

Wallen, M.J., Laukkanen, M.O. & Kulomaa, M.S. 1995, "Cloning and sequencing of the chicken egg-white avidin-encoding gene and its relationship with the avidin-related genes Avr1-Avr5", *Gene,* vol. 161, no. 2, pp. 205-209.

Wang, T.W., Zhu, H., Ma, X.Y., Zhang, T., Ma, Y.S. & Wei, D.Z. 2006, "Mutant library construction in directed molecular evolution: casting a wider net", *Molecular biotechnology,* vol. 34, no. 1, pp. 55-68.

Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. & Salemme, F.R. 1989, "Structural origins of high-affinity biotin binding to streptavidin", *Science (New York, N.Y.),* vol. 243, no. 4887, pp. 85-88.

White, H.B.,3rd 1985, "Biotin-binding proteins and biotin transport to oocytes", *Annals of the New York Academy of Sciences,* vol. 447, pp. 202-211.

White, H.B.,3rd, Dennison, B.A., Della Fera, M.A., Whitney, C.J., McGuire, J.C., Meslar, H.W. & Sammelwitz, P.H. 1976, "Biotin-binding protein from chicken egg yolk. Assay and relationship to egg-white avidin", *The Biochemical journal,* vol. 157, no. 2, pp. 395-400.

White, H.B.,3rd & Whitehead, C.C. 1987, "Role of avidin and other biotin-binding proteins in the deposition and distribution of biotin in chicken eggs. Discovery of a new biotin-binding protein", *The Biochemical journal,* vol. 241, no. 3, pp. 677-684.

White, H.B.,3rd, Whitehead, C.C. & Armstrong, J. 1987, "Relationship of biotin deposition in turkey eggs to dietary biotin and biotin-binding proteins", *Poultry science,* vol. 66, no. 7, pp. 1236-1241.

Whitehead, T.A., Chevalier, A., Song, Y., Dreyfus, C., Fleishman, S.J., De Mattos, C., Myers, C.A., Kamisetty, H., Blair, P., Wilson, I.A. & Baker, D. 2012, "Optimization of affinity, specificity and function of designed influenza inhibitors using deep sequencing", *Nature biotechnology,* vol. 30, no. 6, pp. 543-548.

Wilchek, M. & Bayer, E.A. 1989, "Avidin-biotin technology ten years on: has it lived up to its expectations?", *Trends in biochemical sciences,* vol. 14, no. 10, pp. 408-412.

Willats, W.G. 2002, "Phage display: practicalities and prospects", *Plant Molecular Biology,* vol. 50, no. 6, pp. 837-854.

Wixon, J. 2000, "Featured organism: Danio rerio, the zebrafish", *Yeast (Chichester, England),* vol. 17, no. 3, pp. 225-231.

Wu, S.C. & Wong, S.L. 2005, "Engineering soluble monomeric streptavidin with reversible biotin-binding capability", *The Journal of biological chemistry,* vol. 280, no. 24, pp. 23225-23231.

Zerega, B., Camardella, L., Cermelli, S., Sala, R., Cancedda, R. & Descalzi Cancedda, F. 2001, "Avidin expression during chick chondrocyte and myoblast development in vitro and in vivo: regulation of cell proliferation", *Journal of cell science,* vol. 114, no. 8, pp. 1473-1482.

Zhang, C., Patient, R. & Liu, F. 2013, "Hematopoietic stem cell development and regulatory signaling in zebrafish", *Biochimica et biophysica acta,* vol. 1830, no. 2, pp. 2370-2374.

Zhang, Y.X., Perry, K., Vinci, V.A., Powell, K., Stemmer, W.P. & del Cardayre, S.B. 2002, "Genome shuffling leads to rapid phenotypic improvement in bacteria", *Nature,* vol. 415, no. 6872, pp. 644-646.

Zhao, H. & Arnold, F.H. 1997, "Combinatorial protein design: strategies for screening protein libraries", *Current opinion in structural biology,* vol. 7, no. 4, pp. 480-485.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. 1998, "Molecular evolution by staggered extension process (StEP) in vitro recombination", *Nature biotechnology,* vol. 16, no. 3, pp. 258-261.

Zhao, Y. & Lieberman, H.B. 1995, "Schizosaccharomyces pombe: a model for molecular studies of eukaryotic genes", *DNA and cell biology,* vol. 14, no. 5, pp. 359-371.

Zon, L.I. 1999, "Zebrafish: a new model for human disease", *Genome research,* vol. 9, no. 2, pp. 99-100.

DNA family shuffling within the chicken avidin protein family – A shortcut to more powerful protein tools

Barbara Niederhauser[a,b], Joonas Siivonen[a], Juha A. Määttä[a,b], Janne Jänis[c], Markku S. Kulomaa[a,b], Vesa P. Hytönen[a,b,*]

[a] *Institute of Biomedical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland*
[b] *BioMediTech, Tampere, Finland*
[c] *Department of Chemistry, University of Eastern Finland, Joensuu, Finland*

ABSTRACT

Avidins represent an interesting group of proteins showing high structural similarity and ligand-binding properties but low similarity in primary structure. In this study, we show that it is possible to create functional chimeric proteins from the avidin protein family when applying DNA family shuffling to the genes of the avidin protein family: avidin, avidin related gene 2 and biotin-binding protein A. The novel chimeric proteins were selected by phage display biopanning against biotin, and the selected enriched proteins were characterized, displaying diverse features distinct from the parental genes, including binding to cysteine.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Many of today's biotechnological applications rely on avidin, a small tetrameric protein commonly found in chicken egg whites, and its ability to bind the small molecule biotin with high affinity (Diamandis and Christopoulos, 1991). A number of additional proteins that are homologous to avidin can be found in the chicken (avidin related proteins (AVR1-7) and biotin-binding protein A (BBP-A)), as well as a variety of other vertebrates (xenavidin, zebavidin) and bacteria (streptavidin, rhizavidin, bradavidin), to name a few (Chaiet and Wolf, 1964; Helppolainen et al., 2007, 2008; Hytonen et al., 2007; Keinanen et al., 1994; Maatta et al., 2009). Although all members of the avidin gene family show a high affinity to biotin, their physicochemical properties differ, including pI, thermal stability, oligomerization, glycosylation or immunoreactivity.

Because of the structural similarities, one may assume that the properties can be transferred between avidin members. Indeed,

using rational design, the corresponding residues found in AVRs and streptavidin were applied to avidin, and as a result, its pI was decreased from 10.5 to 4.7 without destroying the high biotin-binding affinity or harming its stability (Marttila et al., 1998). In another study, Livnah and co-workers were able to transfer the catalytic property found in avidin to streptavidin (Eisenberg-Domovich et al., 2004). Exchanging parts of avidin with parts of AVR4 resulted in a highly thermostable and protease resistant protein (Hytonen et al., 2005a), which demonstrated not only that properties can be transferred between the avidin family proteins but also that it is possible to generate more powerful proteins by combining parts of different proteins. However, studies focusing on the residues that directly participate in biotin binding have revealed that manipulating the binding site might sometimes generate unpredicted effects. For example, Stayton's group has demonstrated that residues S45 and D128 strongly co-operate in biotin binding, making it demanding to foresee the effects of mutagenesis beforehand (Hyre et al., 2006). These studies have encouraged us to further explore the possibility to develop enhanced avidins by recombination of avidin family members.

In recent years, random mutagenesis and directed evolution methods have been developed and have been shown to be potential choices for rational directed mutagenesis, especially when the structural determinants behind molecular activity are difficult to determine. Stemmer (1994) developed a method that made it possible to create chimeric proteins from homologous genes, which

---

he named DNA (family) shuffling. Briefly, two or more genes are digested into small fragments by DNase I and reassembled in a primerless PCR reaction. This *in vitro* evolutionary approach to change the properties of proteins has a substantial advantage over the usual methods for rational mutagenesis because DNA family shuffling utilizes mutations that have already been proven to be functional in nature. The power of DNA shuffling was demonstrated, for example, by Chang et al. (1999). These authors were able to create chimeric proteins by shuffling genes from the human interferon-α gene family, which showed higher activity in mouse cells than the respective murine interferon-α.

In this study, we created functional chimeric proteins from the avidin family genes AVD, AVR2 and BBP-A using DNA family shuffling. Chimeric mutant libraries where panned against immobilized biotinylated BSA using phage display. The enrichment of chimeric sequences could be observed, which were characterized by their biophysical properties. We found that the chimeric proteins were still functional biotin binders that displayed differences in their properties as compared with their parental genes. Importantly, we were able to improve the biotin-binding affinity of AVR2, while preserving the high thermal stability. A novel ligand-receptor pair was also found: one of the mutants exhibited a moderate affinity for cysteine.

2. Materials and methods

2.1. Templates

Avidin: pFASTBAC1-AVD (Airenne et al., 1997), AVR2: pFASTBAC1-AVR2 (Laitinen et al., 2002), BBP-A: pFASTBAC1-BBP-A (Hytonen et al., 2007), Rhizavidin: pET101/D-rhizavidin (truncated form called rhizavidin-core) (Helppolainen et al., 2007).

2.2. Primers

(1) Avd_NheI_5:5′-TATTGCTAGCTGCACAACCAGCAATGGCAGCCAGAAAGTGCTCGCTGAC-3′;
(2) Avd_NotI_3b:5′-TTTGCGGCCGCCTCCTTCTGTGTGCGCTGGCGAGTGAAG-3′;
(3) Avd_NheI_SLIC_5:5′-TACGGCAGCCGCTGGATTGTTATTGCTAGCTGCACAACCAGCAATGGCAGCCAGAAAGTGCTCTCTGAC-3′;
(4) Avd_NotI_SLIC_3b:5′-GATATTCACAAACGAATGGTGCGGCCGCCTCCTTCTGTGTGCGCAGGC-3′;
(5) Avd_NotI_SLIC_3_amber:5′-GATATTCACAAACGAATGGTGCGGCCGCCTACTTCTGTGTGCGCAGGC-3′;
(6) BBP-A_5′KpnI: 5′-AAAGGTACCAGGAAGTGCGAGC-3′;
(7) BBP-A_3′HindIII: 5′-ATTTAAGCTTACTTGACACGGGTG-3′;
(8) BBP-A_NheI_SLIC_5:5′-TACGGCAGCCGCTGGATTGTTATTGCTAGCTGCACAACCAGCAATGGCATCCAGGAAGTGCGAGC-3′;
(9) BBP-A_NotI_SLIC_3_amber:5′-GATATTCACAAACGAATGGTGCGGCCGCCTAGACACGGGTGAAGACATTGGTGC-3′;
(10) Rhavd_5′:5′-TTCGATGCAAGCAACTTCAAGGATTT-3′;
(11) Rhavd_3′:5′-CGCATCCTTCAAGAGGCTTTTGTTC-3′;
(12) Rhavd_SLIC_NheI_5:5′-TACGGCAGCCGCTGGATTGTTATTGCTAGCTGCACAACCAGCAATGGCATTCGATGCAAGCAACTTCAA-3′;
(13) Rhavd_SLIC_NotI_3:5′-GATATTCACAAACGAATGGTGCGGCCGCATCCTTCAAGAGGCTTTTGTTCTCAGTCGTCGGCA-3′

2.3. Staggered extension process

The staggered extension process was performed as described in (Zhao et al., 1998). Briefly, the parental genes AVD and AVR2 were amplified from pFASTBAC1-AVD and pFASTBAC1-AVR2 using primers Avd_NheI_5 and Avd_NotI_3b. The PCR reaction (50 μl) contained 0.2 mM dNTP mix (Fermentas International Inc., Thermo Fisher Scientific Inc., Waltham, MA, USA), 30 pmol Avd_NheI_SLIC_5 primer, 30 pmol Avd_NotI_SLIC_3 primer, 0.15 pmol of each plasmid containing a parental gene, and 1.25 U Pfu DNA polymerase (Fermentas) in 1× Pfu buffer containing 2 mM MgCl$_2$ (Fermentas). The reaction mixture was applied to 80 cycles of 30 s at 95 °C and 5–10 s at 55 °C or 50 °C using a MJ Research PTC-200 thermocycler. The PCR products were analyzed on a 1% agarose gel (Top Vision LE GQ Agarose, Fermentas). The ∼500 bp bands were gel purified using

a GFX DNA purification kit (GE Healthcare). A total of 10 ng of purified PCR product was amplified with primers Avd_NheI_SLIC_5 and Avd_NotI_SLIC_3b as described above with 25 cycles of 30 s at 95 °C, 60 s at 50 °C and 60 s at 72 °C. The PCR products were gel purified as described above.

2.4. DNA family shuffling

DNA family shuffling was performed as previously described (Stemmer, 1994) with slight modifications. Briefly, parental DNA was amplified in a 100 μl reaction mixture containing 10 ng plasmid, 60 pmol of each gene specific primer (Avd_NheI_5 and Avd_NotI_3b, BBP-A_5′KpnI and BBP-A_3′HindIII or Rhizavd_5′ and Rhizavd_3′), 0.2 mM dNTP mix (Fermentas) and 2.5 U Pfu DNA polymerase (Fermentas) in 1× Pfu buffer containing 2 mM MgCl$_2$ (Fermentas). The PCR reaction was applied to 25 cycles of 30 s at 95 °C, 60 s at 50 °C and 60 s at 72 °C using a MJ Research PTC-200 thermocycler. The PCR products were analyzed on 1% agarose gel (Top Vision LE GQ Agarose, Fermentas). Bands of ∼500 bp were gel purified as described above. The purified DNA was eluted in 25 μl of water. The parental DNA was mixed in pairs using 1 μg of each gene for subsequent digestion with 0.15 U DNase I (NEB). Before adding the DNase I, the reaction mixture was equilibrated for 5 min at 15 °C. The DNA was digested for 2 min at 15 °C followed by heat inactivation at 90 °C for 10 min. A total of 10 μl of the digested DNA was combined with a 10 μl PCR premix (0.4 mM dNTP mix (Fermentas), 2× Pfu buffer (Fermentas) containing 4 mM MgCl$_2$ and 1.25 U Pfu DNA polymerase (Fermentas)). The reaction mixture was applied to the same PCR conditions as described above, except 40 reaction cycles were used. The reassembled DNA (∼500 bp) was gel purified as described above. The purified reassembled DNA (15 ng) was amplified with reaction conditions as described above. For each shuffled pair, two amplification reactions were made. In one reaction, a 5′-primer specific for one gene and a 3′-primer with specificity for the other gene were used to prevent the amplification of parental genes. In the other reaction the alternative primer combination was used (see Supplementary Fig. 1 for illustration).

2.5. Preparation of phagemid DNA libraries

The DNA products were cloned into a phagemid vector using the SLIC method (Li and Elledge, 2007). The use of the phagemid vector was based on the VTT Fab phagemid vector (VTT Technical Research Center of Finland, Espoo, Finland). The phage

libraries were constructed with an amber codon located between the chimeric gene and the M13 phage surface protein pIII to allow the expression of chimeric gene-pIII fusion proteins and free monomeric chimeric proteins, resulting in the display of tetrameric chimeric proteins on the surface of the phages (Sidhu et al., 2000) using XL1 blue cells (a supE44 containing *E. coli* strain). Moreover, the amber codon facilitated the production of free proteins from isolated phagemids using the BL21star (DE3) cell line (Invitrogen, Carlsbad, CA, USA), which does not contain the supE44 genotype. Because of a mistake in primer design, the last amino acid of avidin and AVR2, glutamic acid 128 (in avidin), was missing. However, this residue was located outside of the structural core of the protein, and previous studies have shown that the C-terminal residues can be manipulated without consequences (Hytonen et al., 2004a,b; Marttila et al., 1998).

A total of 1 µg of the phagemid vector and inserts were digested with NheI and NotI restriction enzymes (New England BioLabs Inc., Ipswich, MA, USA). The digested DNA was gel purified as described above. Approximately 250–500 ng of digested DNA was treated with 0.25 U of T4 DNA polymerase (New England Biolabs) in a 20 µl reaction for 30 min at room temperature (RT). The reaction was terminated by the addition of 2 µl of 10 mM dCTP, and the tube was placed on ice. The vector and insert were annealed in 1:1 ratio using 20 ng RecA (New England BioLabs Inc.). The reaction was incubated for 30 min at 37 °C, and 5 µl of the annealed product was transformed into electrocompetent XL1 blue cells. The transformed cells were plated onto LB$_{amp}$ plates (containing 100 µg/ml ampicillin). The positive colonies were selected, dissolved in 2 ml glycerol and stored at −80 °C.

2.6. Production of phages

A total of 10 ml of SB$_{tet+amp}$ medium (containing 10 µg/ml tetracycline and 50 µg/ml ampicillin) was inoculated with 100 µl library glycerol stock. The cells were grown overnight at 37 °C and 225 rpm. The 200-µl overnight culture was diluted in 10 ml of fresh SB$_{tet+amp}$ medium and was grown at 37 °C and 225 rpm to an OD$_{600}$ of 0.5. The cells were infected with 1 ml of VSC-M13 helper phage (Stratagene La Jolla, CA, USA; 10^{11} pfu/ml) for 30 min at 37 °C. One microliter and 10 µl cell cultures were plated onto LB$_{amp}$ plates and incubated overnight at 37 °C. The cells were diluted in 90 ml of SB$_{tet+amp}$ medium and were grown for 2 h at 37 °C and 225 rpm. Kanamycin was added to a final concentration of 70 µg/ml, and the cells were grown overnight at 28 °C and 225 rpm. The cells were centrifuged for 15 min at 4000 × g. A total of 25 ml of 20% PEG-6000 in 2.5 M NaCl was added to the supernatant, and the phages were precipitated for 30 min at 4 °C. The phage precipitate was isolated by centrifugation at 13,200 × g and 4 °C for 20 min. The supernatant was removed, and the precipitate was resolved in 2 ml PBS and centrifuged for 10 min at 16,000 × g and 4 °C. The supernatant was divided into two fresh Eppendorf tubes. The phages were precipitated by adding 250 µl of 20% PEG-6000 in 2.5 M NaCl and incubated for 30 min on ice. The precipitate was collected by centrifugation at 16,000 × g and 4 °C for 10 min. The phage precipitate was resolved in either 1 ml PBS or 1 ml of 50 mM Tris containing 1 M NaCl, 20% glycerol and 1% BSA. To determine the phage titer, 2 µl of phage dilutions (10^{-7}, 10^{-9}, 10^{-11}) were mixed with 100 µl XL1 blue cells at OD$_{600}$ of 0.5 and incubated for 15 min at 37 °C followed by plating onto LB$_{amp}$ plates.

2.7. Selection of functional chimeras by biopanning

MaxiSorp™ Immuno 96 MicroWell™ plates (Nunc A/S, Roskilde, Denmark) were coated with 2 µg of BSA and 2 µg of biotinylated BSA overnight at 4 °C. The wells were washed 3 times with 300 µl PBS and subsequently blocked with 100 µl of 1% milk in PBS for 1 h at RT. The wells were washed 3 times with 300 µl PBS. First, 100 µl of phage solution was incubated in the BSA coated wells for 1 h at RT. The phages were then transferred to wells coated with biotinylated BSA and incubated for 60 min at RT. The wells were washed 3 times with PBS-Tween (0.05%) (PBS-T) and 3 times with PBS, respectively. The phages were eluted with 100 µl of 100 mM HCl containing 17 µg of biotin followed by vigorous shaking for 10 min at RT (Barbas et al., 2001). The eluted phages were transferred to 1.5 ml-Eppendorf tubes and neutralized by adding 5 µl of 2 M Trizma® base (pH range 8–9, Sigma–Aldrich, St. Louis, MO, USA). The eluted phages were added to 10 ml XL1 blue cells at OD$_{600}$ = 0.5 and incubated for 30 min at 37 °C. A total volume of 7 ml of SB medium (containing 20 µg/ml ampicillin and 10 µg/ml tetracycline) was added. The transformed cells (10 and 100 µl, respectively) were plated onto LB$_{amp}$ plates. The cells were incubated for 30 min at 37 °C and 225 rpm. Ampicillin was added to a final concentration of 50 µg/ml, and the cells were incubated for 1 h at 37 °C and 225 rpm. The cells were superinfected with 1 ml VSC-M13 helper phage (Stratagene; 10^{11} pfu/ml) for 30 min at 37 °C. The cells were diluted in 90 ml SB$_{tet+amp}$ medium and incubated for 2 h at 37 °C and 225 rpm. Kanamycin was added to a final concentration of 70 µg/ml, and the cells were incubated overnight at 28 °C and 225 rpm. The following day, the phages were precipitated as described above. The biopanning step was repeated 2 times, and the washes were increased up to 10 times after phage binding.

2.8. Master plate

A master plate was produced for the analysis of the same colony by different methods. Selected individual colonies from the 3rd panning round were dissolved in 0.5 ml SB$_{amp+tet}$ medium in 96-deepwell plates. The cells were grown overnight at 37 °C and 700 rpm. A total of 100 µl of glycerol were added to the cells for storage at −80 °C.

2.9. Sequencing

Cells (10 µl) from the master plate were diluted in 0.5 ml SB$_{amp+tet}$ medium in 96-deepwell plates. The cells were grown for 4 h at 37 °C and 700 rpm and subsequently infected with 1 µl VSC-M13 helper phage (Stratagene; 10^{11} pfu/ml). The cells were grown overnight at 28 °C. The cells were centrifuged for 15 min at 1500 × g and 4 °C. Approximately 1 µl of supernatant was added to 9 µl of the BigDye Terminator v3.1 master mix (Applied Biosystems, Carlsbad, CA, USA) and applied to one cycle of 1 min at 96 °C, followed by 29 cycles of 30 s at 96 °C, 30 s at 50 °C and 4 min at 60 °C. The resulting PCR products were ethanol precipitated and analyzed by sequencing on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) according to the protocols recommended by the manufacturer (ABI PRISM BigDye Terminator Cycle Sequencing Kit v.1.1, Applied Biosystems).

2.10. Microplate assay for activity determination

Cells (10 µl) from master plate were diluted in 0.5 ml SB$_{amp+tet}$ medium in 96-deepwell plates. At OD$_{600}$ = 0.5, the cells were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown overnight at 28 °C and 700 rpm. The cells were centrifuged for 15 min at 4000 × g, the supernatant was removed, and the cells were lysed by resuspending in 100 µl 20% sucrose, 2 mM EDTA, 30 mM tris pH 8 containing 50 ng/µl lysozyme and then incubating for 30 min on ice. An additional cell lysis was achieved by 3 freeze–thaw cycles at −80 °C and 37 °C. The cell lysate was centrifuged for 15 min at 4000 × g. The supernatant was added to a

round-bottom 96-well microplate and 2.5 µl of 2 M Tris, 20 µl of 5 M NaCl and 20 µl of glycerol was added. The cell lysate was incubated for 1 h at RT in wells coated and blocked with biotinylated BSA as described above. The wells were washed 3 times with 300 µl PBS-Tween. A volume of 100 µl of the biotinylated alkaline phosphatase secondary antibody, diluted 1:5000 with 1% milk in PBS-T, was added and incubated for 1 h at RT. The wells were washed 3 times with PBS-T and 3 times with PBS. The phosphatase substrate (1 mg/ml in DEA buffer) was added and the absorbance was measured after 30 min at 405 nm.

2.11. Protein expression

The proteins were expressed in transformed *E. coli* BL21star (DE3) cells (Invitrogen) either in bottle cultures or in a pilot scale fermentor. When using the phagemid pelB amber construct, the expression of the protein in BL21star (DE3) cells ensured that the protein was not expressed in fusion with the pIII protein. For the bottle cultures, the cells were cultivated in Luria-Bertani medium supplemented with 100-µg/ml ampicillin at 28 °C and 175 rpm. When the cell density reached an absorbance A_{600} of 0.2–0.3, the protein expression was induced by the addition of 1 mM IPTG and 0.2% L-arabinose. Cultivation was continued overnight (16 h) at 28 °C and 175 rpm. The cells were collected by centrifugation at $1500 \times g$ and 4 °C for 15 min. The proteins were produced in a fermentor essentially as described previously (Maatta et al., 2011). Briefly, single colonies were grown overnight in 5 ml of fermentation medium containing 100 µg/ml ampicillin at 27 °C and 200 rpm. The cell culture was diluted in 500 ml of fermenting medium (see Maatta et al., 2011) at the previously mentioned parameters and was used the following day to start a 4.5 L fermentation in a Labfors Infors 3 fermentor (Infors HT, Bottmingen, Switzerland) at 28 °C. The fermenting medium contained the antifoam agent Struktol J 647 (Schill+Seilacher, Hamburg, Germany). The culture was induced at A_{600} 10–20 with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.2% (w/v) L(+)-arabinose and temperature was simultaneously decreased to 25 °C. The pO_2 was maintained at 20% (on average, oscillation was allowed) by controlling the agitation speed (200–1150 rpm) and airflow. The feed was controlled by the pO_2 status, applying feed when the oxygen level rose above 40%. Fermentation was terminated at 24 h post-induction.

2.12. Protein purification

The cells from a 500 ml cell culture were resuspended in 100 ml of 50 mM sodium carbonate, 1 M NaCl, pH 11, and were homogenized three times at 40,000 psi using an EmulsiFlex-C3 homogenizer (Avestin Inc., Ottawa, Canada). The cell lysate was clarified by centrifugation at $15,000 \times g$ and 4 °C for 30 min. The proteins were purified by affinity chromatography on a 3 ml Econo-Column® (BioRad Laboratories Inc., Hercules, CA, USA) filled with 2-iminobiotin Sepharose™ 4 Fast Flow (Affiland S.A., Ans Liege, Belgium) and connected to ÄKTA™ purifier-100 (GE healthcare/Amersham Biosciences AB, Uppsala, Sweden) using 1 ml/min flow rate. The proteins were eluted with 50 mM sodium acetate, pH 4.

2.13. Analytical gel filtration

The molecular size of the protein in solution was measured by size-exclusion chromatography using Superdex200 10/300GL column (GE Healthcare/Tricorn, Amersham Biosciences AB, Uppsala, Sweden) connected to ÄKTA™ Purifier-100 equipped with UV-900 monitor (GE Healthcare/Amersham Biosciences AB, Uppsala, Sweden). The analysis was conducted using a 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, 650 mM NaCl (pH 7.0) mobile phase (Nordlund et al., 2003a), and 40–80 µg of protein was injected per run. All analyses were done with flow rate 0.5 ml/min and absorbances at 280 and 205 nm were used to locate the protein peaks in the chromatograms. The molecular weight calibration curve was prepared by analyzing a gel filtration standard protein mixture containing thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa) (BioRad Laboratories Inc., Hercules, CA, USA).

2.14. Mass spectrometry

The electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS) analysis was performed by using a 4.7-T APEX-Qe instrument (Bruker Daltonics, Billercia, MA, USA), described in detail elsewhere (Helppolainen et al., 2007). Briefly, the protein samples were desalted with the use of PD-10 columns (GE Healthcare), diluted with acetonitrile/water/acetic acid (49.5:49.5:1.0, v/v) solvent to a final concentration of ∼2 µM and directly electrosprayed. The ions were externally accumulated for 2 s in the hexapole ion trap before being transmitted to the ICR cell for trapping, excitation and detection. A total of 500 co-added 512-k Word time-domain transients were Gaussian-multiplied and fast Fourier transformed prior to magnitude calculation and external mass calibration. The final mass spectra were charge-deconvoluted with a standard deconvolution macro implemented with the XMASS 6.0.2 software.

2.15. Analysis of biotin binding by fluorescence spectrometry

The dissociation rate constant (k_{diss}) of dye-conjugated biotin was determined by fluorescence spectrometry using the biotin-labeled fluorescent probe ArcDia™ BF560 (ArcDia, Turku, Finland BF560) as previously described (Hytonen et al., 2004a,b). In principle, the changes in fluorescence intensity of the 50 nM dye in a pH 7 buffer (50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, 650 mM NaCl, 0.1 mg/ml BSA) were measured after the addition of 100 nM biotin-binding protein. The dissociation of this complex was observed by the addition of 100-fold molar excess of free biotin. The assay was performed at 50 °C using a QuantaMaster™ Spectrofluorometer (Photon Technology International, Inc., Lawrenceville, NJ, USA).

2.16. Biacore measurements

DNA and 2-iminobiotin binding were analyzed using the optical biosensor Biacore X (GE Healthcare/Biacore Ab, Sweden) as previously described (Maatta et al., 2009). Briefly, a CM5-chip was coated with 2-iminobiotin as described (Maatta et al., 2009), and the binding curves were measured for protein concentrations varying between 6 µM and 0.25 µM. The 2-iminobiotin Biacore analysis was carried out in 50 mM sodium carbonate, pH 11, containing 1 M NaCl. To prepare a chip containing ssDNA, the chip was first treated with EDC/NHS followed by ethylenediamine to introduce amino groups on the surface. Subsequently, a 50 mM NHS-PEO$_2$-maleimide linker (Thermo Scientific, Wilmington, DE, USA) in sodium borate buffer pH 8.5 was used to couple the amino groups with thiol groups. A 1 mM solution of oligonucleotides (5′-SH-GTCAGCCACTTTCTGGC-3′, Eurogentec/Oligold) was injected onto the surface. The unreacted groups were inactivated with 50 mM cysteine in 50 mM sodium acetate, 1 M sodium chloride, pH 4. The reference cell was treated similarly, but the injection of oligonucleotide was skipped. Binding curves for DNA binding were measured for protein concentrations varying between 50 µM and 100 nM in 50 mM NaH$_2$PO$_4$/Na$_2$HPO$_4$, 100 mM NaCl, and pH 7. The surface was washed with a 300 s delay after injection with 1 M sodium chloride, 50 mM sodium hydroxide and, if needed,

Table 1
Quality of mutant libraries generated by DNA shuffling.

Library	Library size ($\times 10^4$ cfu)	No. analyzed sequences	Shuffled sequences	Unique sequences	Average number of crossovers
AVD/AVR2	2.2	26	26	25	3.2
AVD/BBP-A	2.9	28	26	12	1.4
AVR2/BBP-A	3.1	29	22	16	1.2

regenerated with 0.5% SDS. The DNA binding results were only analyzed qualitatively because of the high signal (>4500 RUs). The biotin binding site involvement in DNA binding was assessed by injections of 7 μM of protein in the presence of 21 μM d-biotin. The results were analyzed by calculating the maximum response value for both injections. The fraction of signal remaining after blocking with biotin was then calculated using Microsoft Excel.

2.17. DSC measurements

The thermal stability of the studied proteins in the presence and absence of ligands was analyzed using an automated VP-Capillary DSC System (GE Healthcare/Microcal Inc., Northhampton, MA, USA). Thermograms were recorded between 20 and 130 °C with a heating rate of 120 °C/h. The proteins were dialyzed in 50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, pH 7 or 50 mM sodium carbonate, pH 11. The protein concentration in the cell was 7 μM and the ligand concentration was 21 μM. The results were analyzed using Origin 7.0 DSC software suite, using a Gaussian fit to determine the T_m-values (GE Healthcare/MicroCal Inc., Northampton, MA, USA).

2.18. ITC measurements

The thermodynamic parameters of biotin binding were measured with a high-sensitivity VP-ITC instrument (Microcal Inc., Northampton, MA). The proteins were dialyzed against 50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, pH 7, and the last dialysate was used to dissolve and dilute the ligands and protein samples. The protein concentrations were determined spectrophotometrically, and the ligand concentrations were determined by microbalancing. All samples were degassed and pre-incubated to avoid long equilibration periods. The protein samples were loaded onto a cell, and the ligand samples were collected into the syringe. The titration was carried out with 10 μl titrations at four-minute intervals to allow the signal to return back to baseline. The data was analyzed using the Origin 7.0 ITC software suite (GE Healthcare/MicroCal Inc., Northampton, MA, USA).

2.19. Protein sequence analysis and alignment

The chimeric sequences were analyzed using DNAMAN 4.11 (Lynnon Corporation, Quebec, Canada) and MEGA 4 (Tamura et al., 2007). The sequence alignment was created using ClustalW (Larkin et al., 2007) and edited using GeneDoc (Nicholas et al., 1997) and Microsoft Office Word 2003 (Microsoft).

3. Results

3.1. Construction and quality of mutant libraries

The biotin binding proteins avidin, AVR2, BBP-A and rhizavidin were used as starting materials for the library construction using DNA family shuffling (Stemmer, 1994) and staggered extension process (StEP) (Zhao et al., 1998). Their DNA sequence homology relative to each other ranged from 38 to 80%, with the AVD/AVR2 pair having the highest homology and the AVD/rhizavidin pair having the lowest homology; their protein identities ranged from

21 to 68%. The proteins differed in isoelectric points, and they have different biotin-binding affinities (For avidin: pI 9.69; k_{diss}: 0.8×10^{-6} s^{-1}; the dissociation rate constants are calculated at 40 °C from the experimental data using the global fit model described in (Hyre et al., 2000) and the results are published in original articles (Hytonen et al., 2005, 2007). For rhizavidin, the value represents the measured value at this temperature (Helppolainen et al., 2007); AVR2: pI 4.92, k_{diss}: 6.9×10^{-4} s^{-1}; BBP-A: pI 9.75, k_{diss}: 5.5×10^{-5} s^{-1} and rhizavidin: pI 4.36 k_{diss}: 2.2×10^{-3} s^{-1} (Helppolainen et al., 2007; Hytonen et al., 2005, 2007). Initially, we wanted to construct six different mutant libraries using avidin, AVR2, BBP-A and rhizavidin; each gene paired with the other three genes. The experiments to shuffle avidin and AVR2 using the StEP method resulted in only 5% chimeric genes with chimeras containing only one crossover, and a major parental background. In contrast to the StEP method, DNA shuffling resulted in chimeras with multiple crossovers and a low parental background; however, it was not possible to construct libraries of any of the genes in combination with rhizavidin using either of these methods. This observation might be due to low similarity in DNA sequence between the rhizavidin and chicken avidin family genes (38–40%).

For the reduction of parental background genes, the original DNA shuffling protocol from Stemmer was modified. The reassembled genes were amplified using a gene specific primer of one gene at the 5′-end and gene specific primers of the other gene at the 3′-end (see Supplementary Fig. 1). This approach was inspired by DNA shuffling method used by Ikeuchi et al. (2003) where they use "screw" primers to amplify the parental genes, which add a unique sequence at the beginning and end of the gene. However, instead of adding unique sequences to the parental genes, we exploited the fact that the beginning and end of the avidin family genes differ in such an extent that different primers needed to be employed to amplify the different genes.

To confirm the quality of the chimeric libraries, 26–29 randomly selected clones of each library were analyzed by sequencing. The sequence analysis was performed at amino acid level. The AVD/AVR2 library showed the largest distribution of crossovers with all sequences being unique, and no wild-type sequences were found (Table 1). The AVD/BBP-A library contained mostly sequences with one crossover. One sequence contained a point mutation. Moreover, we observed 5 sequences containing sequence from AVR2 instead of BBP-A resulting in AVD/AVR2 chimeras. The AVR2 contamination most probably occurred in the first gene amplification step or during the DNA shuffling steps. The AVR2/BBP-A library was the least diverse library with regard to shuffled mutants.

3.2. Biopanning and sequence analysis of mutants

After three rounds of phage panning against immobilized biotinylated BSA (BTN-BSA) and BSA, an enrichment of biotin-specific phages over BSA-specific phages could be observed. From each library, BTN-BSA and BSA selected clones were analyzed. Out of the 125 clones analyzed after the biopanning experiments, 33 unique genes could be found (Supplementary Table 1). More than half of these sequences contained at least one point mutation, with the mutation S16Y being the most abundant. All three libraries showed one sequence that was highly abundant (A/A2-3, A/B-2,
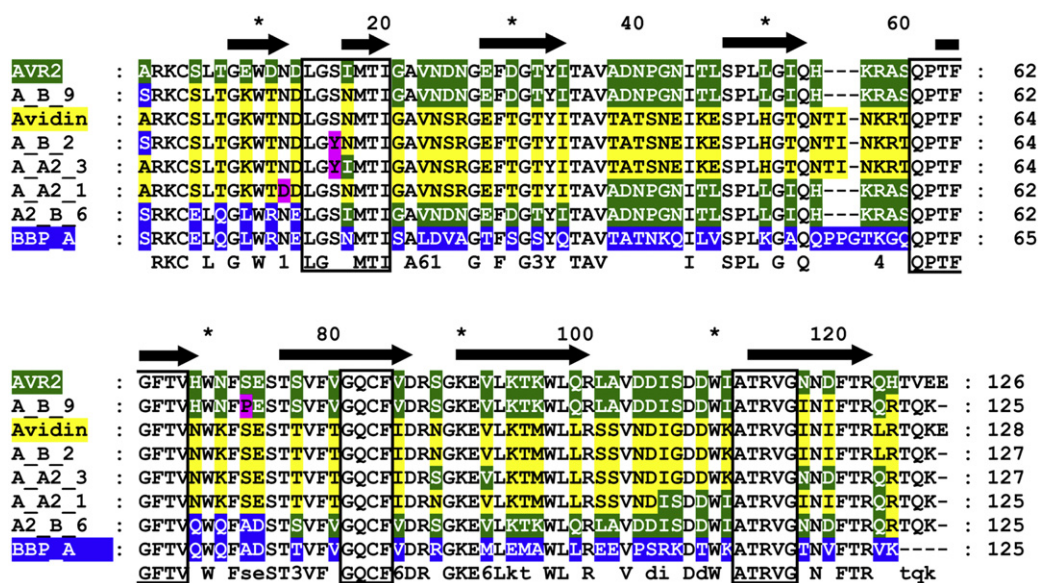
Fig. 1. Protein sequence alignment of parental proteins with selected chimeras. Residues originating from avidin are highlighted in yellow, residues from AVR2 in green and residues from BBP-A in blue. Mutations are highlighted in purple. Secondary structure according to avidin (PDB: 2AVI) is shown schematically by black arrows. The regions of frequent crossover sites are highlighted with black frames. Consensus sequence is displayed in bottom line according to GeneDoc (Nicholas et al., 1997), numbers referring to similarity groups according to Blosum 62 score table; 1: Asp and Asn, 3: Ser and Thr; 4: Lys and Arg; 6: Leu, Ile, Val, Met. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

A2/B-6, Fig. 1). Interestingly, the two mutants A/A2-3 and A/B-2 shared a high DNA sequence homology (95%). Both contained the mutation S16Y. The third clone A2/B-6 showed low DNA sequence homology (65%) with the other two, and it is a chimera of AVR2 and BBP-A.

3.3. Mutant characterization

The selected clones were analyzed in a microplate assay from the cell lysate to test their biotin binding. A total of 11 clones showed an increased response as compared with wild-type proteins. Among the most abundant sequences A/A2-3 and A/B-2 gave high signals in the microplate experiment, whereas A2/B-6 had no signal. In addition, clones A/A2-1 (found two times) and A/B-9 (found six times), showed an increased biotin-binding capacity, and they were selected for further analysis. These proteins were expressed in *E. coli* and were purified by affinity chromatography. A sequence alignment of the selected clones can be found in Fig. 1.

High-resolution ESI FT-ICR MS analysis was carried out to evaluate the quality of the expressed mutants (Supplementary Fig. 2). We selected two mutants (A/B-2 and A/B-9) for the analysis, which revealed almost perfect matches with the theoretical masses of the polypeptides. An intra-molecular disulfide bridge was present in the proteins. In addition to the expected protein forms, only a few other protein forms were detected at low abundance, most likely because of the different cleavage sites at the N-terminus (incorporation of some additional residues from the signal peptide). The majority of the proteins contained signal peptidase cleavage site between residues . . .ALA and (S/A)RK. . ., which is consistent with the prediction by SignalP 3.0.

The binding characteristics were further analyzed with the optical biosensor (Biacore X), by fluorescence spectrometry and by ITC. All five selected mutants showed a fast and complete dissociation (>90% dissociation within less than 10 s) of the fluorescent biotin conjugate ArcDia BF560 at 50 °C after the addition of free biotin (data not shown). The mutants also showed 10- to 1000-fold lower affinities and 10- to 1000-fold higher dissociation rate constants toward immobilized 2-iminobiotin in an optical biosensor

assay when compared with wild-type avidin. However, when compared with AVR2, all mutants except A/B-9 showed improved affinities to 2-iminobiotin, and the mutants A/A2-1 and A/B-2 had improved 2-iminobiotin binding properties over BBP-A (Table 2). The DNA binding was only determined qualitatively with an optical biosensor because of the high amount of binding response (discussed below). All of the mutants showed decreased binding to the DNA surface when compared with avidin, which could only be removed from the surface using harsh condition (0.5% SDS). The BBP-A showed weak binding, and A/B-9 and AVR2 did not bind DNA at all. The mutant A/A2-3 showed binding to the cysteine coated reference cell. Importantly, in all cases, DNA binding was considerably (35–95%) reduced when biotin was added prior to the measurement, and the reduction was the most significant in the case of the cysteine binding mutants A/A2-3 (95.8%), which indicates the involvement of the biotin binding site in the DNA binding (Table 2). ITC analysis revealed that, similarly to wild-type avidin, two of the mutants (A/A2-1 and A/A2-3) have a relatively high affinity for biotin (Fig. 4). Although one site binding model was fitted with an acceptable error in the K_d value in the case of A/A2-1 and A/A2-3 (below 10%), and the dependencies of the fitted model were acceptable (all below 0.3), the accuracy of determined affinities was questionable, because of the limitations of ITC to accurately quantify high-affinity interactions. Therefore, this analysis would suggest the dissociation constant $\ll 10^{-9}$ M for A/A2-1 and A/A2-3. In contrast, two of the mutants (A/B-2 and A/B-9) showed a slightly decreased biotin-binding affinity according to visual inspection of the measured data (Fig. 4). In spite of this observation, more data points on the binding isotherm and dependency values in the curve-fitting model below 0.4; errors in K_d values were over 20%, and hence, accurate K_d values could not be provided. Therefore, further studies are needed to completely understand the thermodynamics of the biotin binding process of the chimeric mutants. However, the biotin binding enthalpies were determined with high confidence and were as follows: AA2-1, −22.5 kcal/mol; AA2-3, −14.9 kcal/mol; A/B-2, −16.2 kcal/mol; A/B-9, −14.5 kcal/mol; wild-type avidin, −22.7 kcal/mol.

Table 2
Biophysical properties of selected chimeras.

	2-Iminobiotin binding (BiaCore)		DNA binding (BiaCore)		Molecular weight (analytical gel filtration, kDa)	
	$K_d \pm$ SE ($\times 10^{-8}$ M)	$k_{diss} \pm$ SE ($\times 10^{-3}$ s^{-1})	$-$BTN[a]	+BTN (%)	$-$BTN $\pm$ SD	+BTN $\pm$ SD
AVD	2.3 ± 0.1	0.54 ± 0.01	+++	11.8	46.2 ± 6.2	46.1 ± 7.1
AVR2	n.a.[b]	n.a.[b]	−	n/a	50.1 ± 0.3	55.5 ± 3.8
BBP-A	1110.0 ± 134.7	215.0 ± 5.2	+	44.9	34.9 ± 1.5	34.4 ± 1.6
A/A2-1	27.9 ± 1.7	31.9 ± 0.7	++	65.4	27.1 ± 1.6	38.2 ± 2.2
A/A2-3	668.0 ± 16.7	107.0 ± 2.7	++	4.2	39.8 ± 0.3	40.2 ± 2.1
A/B-2	7.5 ± 0.3	6.8 ± 0.1	++	40.0	41.8 ± 1.6	41.8 ± 3.4
A/B-9	n.a.[b]	n.a.[b]	−	n/a	34.7 ± 1.6	36.0 ± 0.3

Standard errors (SE) are reported for binding parameters determined by BiaCore. Standard deviations (SD) are reported for molecular weights determined using analytical gel filtration calculated for two or three independent measurements.

[a] +, positive change in signal during injection; ++, signal did not return back to zero after wash +++, surface required additional wash with 0.5% SDS for regeneration; −, no positive change in signal during injection

[b] No measurable affinity, $K_d > 10^{-4}$ M.

The oligomeric state of the mutants in solution was determined by analytical gel filtration (Table 2 and Fig. 2 and Supplementary Fig. 3). All selected mutants, except A/A2-1, showed elution chromatograms similar to avidin in the absence and presence of biotin at pH 7, suggesting tetrameric proteins (Supplementary Fig. 3). Note that the determined molecular weights are smaller than the theoretical size of the tetramer (Table 2). This behavior has been observed already in previous studies (Hytonen et al., 2005, 2007). The A/A2-1 mutant also eluted as a tetramer in the presence of biotin, but in the absence of biotin, its elution volume suggested a dimeric or trimeric state (Fig. 2A and B). Additional studies at pH 4 showed similar values, with the protein being in a monomeric state in the absence of biotin (Fig. 2C and D).

The thermal stability was determined by differential scanning calorimetry (DSC) (Table 3 and Fig. 3). Due to relatively high protein consumption by this method, the measurements could only be done once. Nevertheless, to evaluate the reliability of the analysis, the measurements were repeated for mutant A/A2-1 resulting in a standard deviation for T_m of 1 °C in the absence of biotin and 0.4 °C in the presence of biotin. These results suggest good reproducibility of the determined values. The mutants showed a relatively broad range of transition melting temperatures in the absence of ligand ranging from 70.2 °C to 95.8 °C. The A/B-2 mutant showed an even higher temperature transition midpoint (T_m) (95.8 °C) than that of AVR2 (93.6 °C). However, the chimeric proteins were not as stable as avidin in the presence of biotin, and they resembled AVR2 (Table 3). Specifically, the A/A2-1 mutant showed only a minor
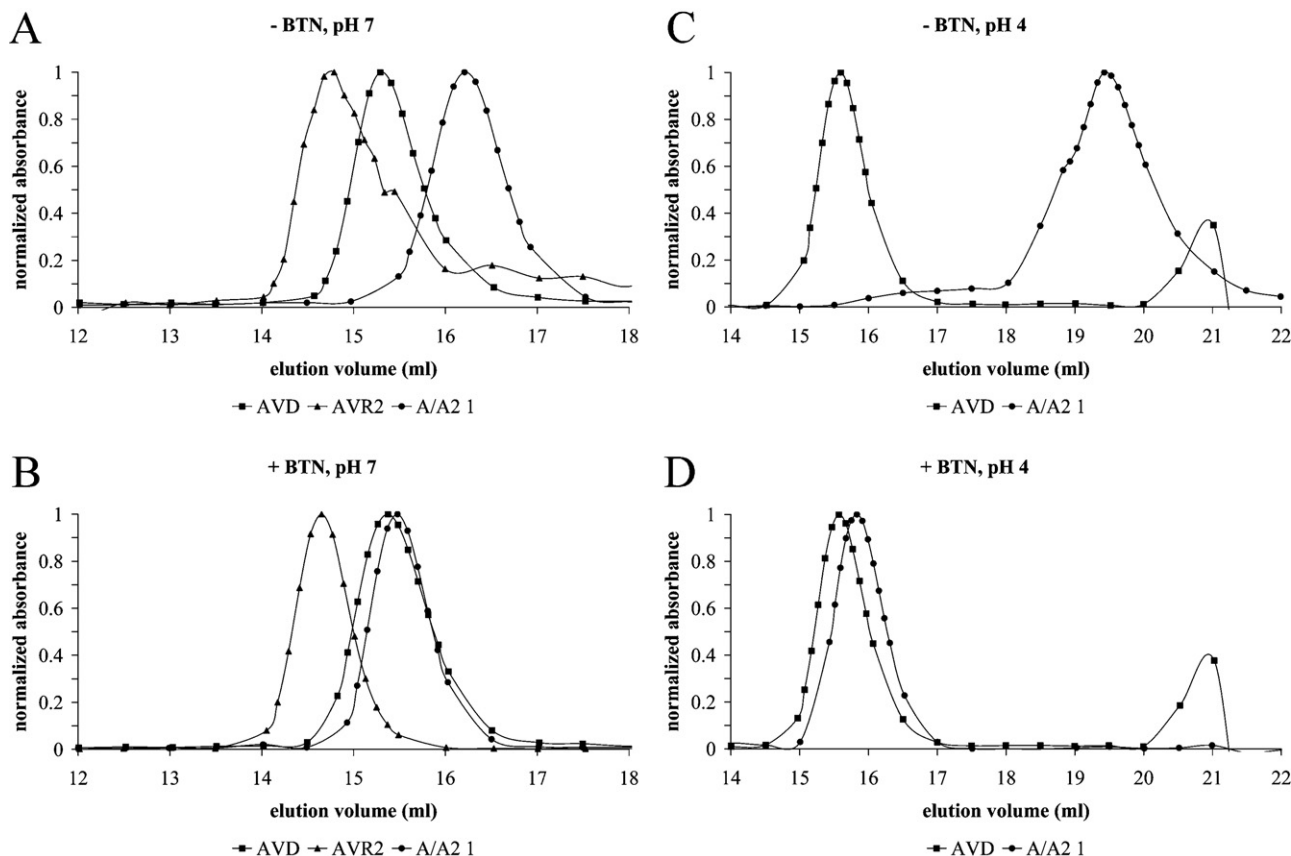

Fig. 2. Gel filtration analysis of A/A2-1. (A) Elution diagram at pH 7 in the absence of biotin. (B) Elution diagram at pH 7 in the presence of biotin. (C) Elution diagram at pH 4 in the absence of biotin. (D) Elution diagram at pH 4 in the presence of biotin.
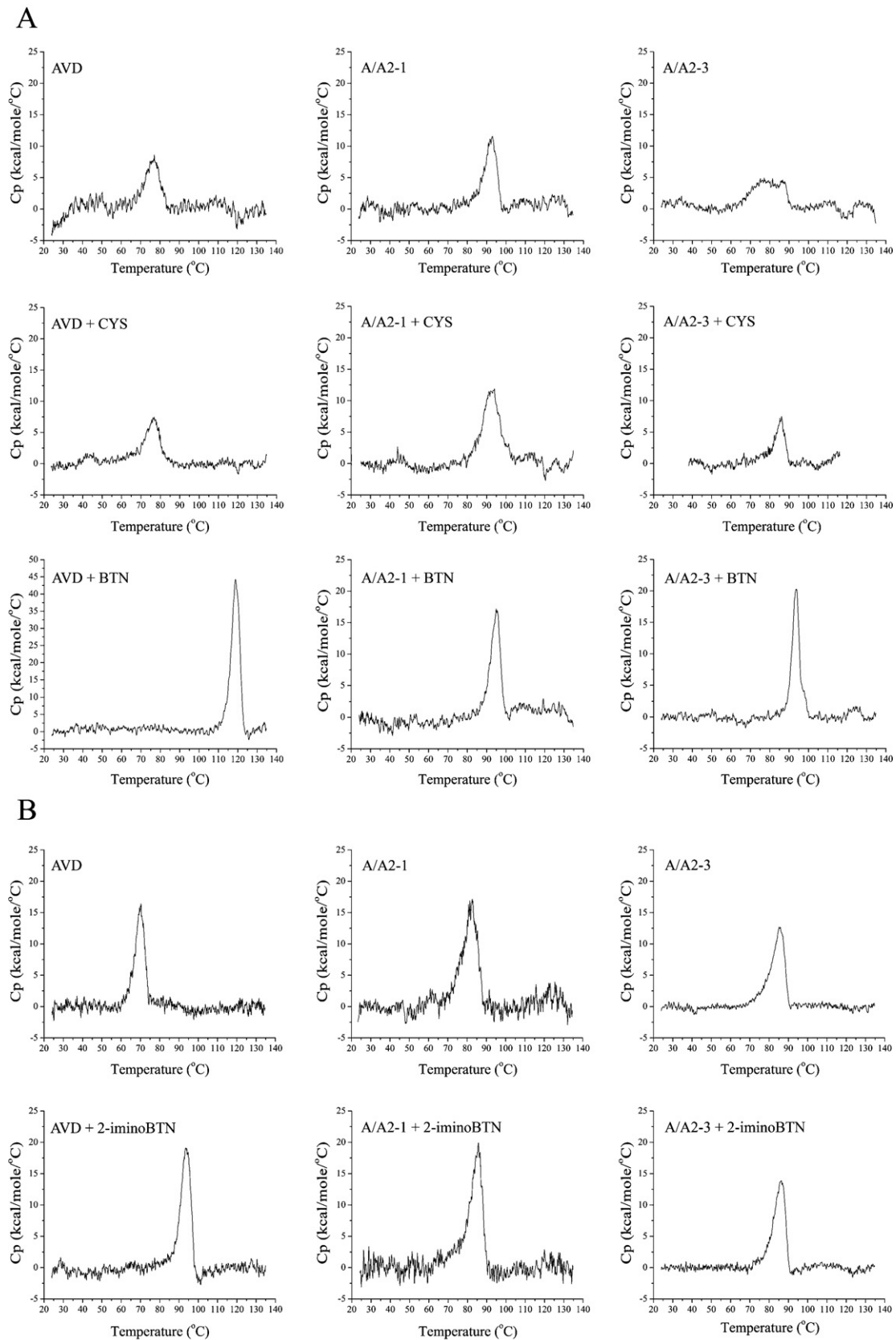
Fig. 3. Differential scanning calorimetry scans of selected mutants. (A) Thermograms of avidin (left panel) at pH 7 in the absence of ligand (top), in the presence of cysteine (center) and biotin (bottom); thermograms of A/A2-1 (middle panel) and A/A2-3 (left panel) are displayed accordingly. Please note that the *Y*-axis scale of AVD+BTN differs from others. (B) Thermograms of avidin (left panel) at pH 11 in the absence (top) and presence (bottom) of 2-iminobiotin; thermograms of A/A2-1 (middle panel) and A/A2-3 (left panel) are displayed accordingly.
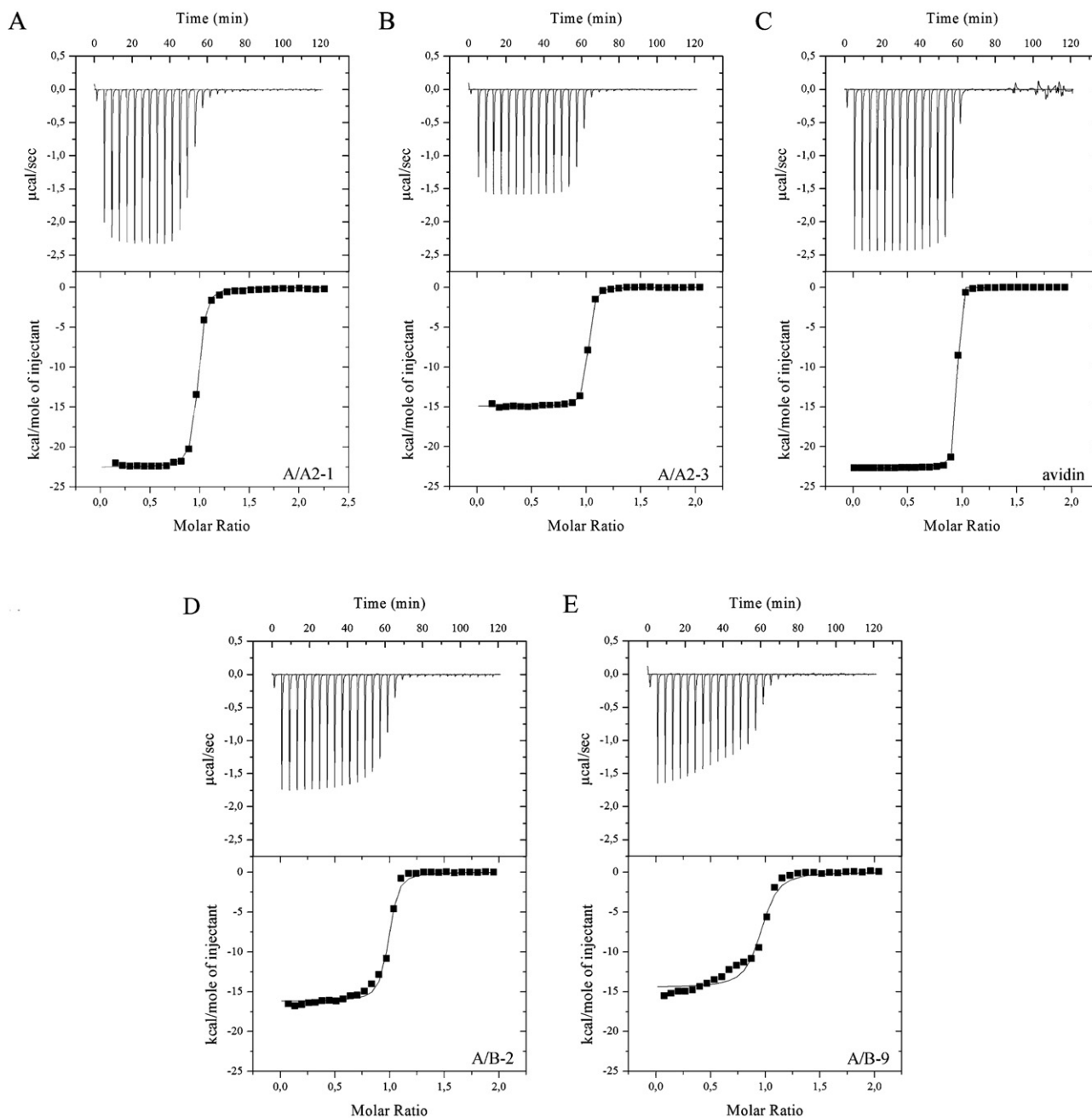
Fig. 4. Isothermal titration calorimetry binding thermograms measured for (A) A/A2-1 (B) A/A2-3 (C) avidin (D) A/B-2 (E) A/B-9; top panel: raw ITC data; bottom panel: binding isotherm derived from integrated heats.

Table 3
Thermal stability of selected chimeras analyzed by DSC.

Ligand	Biotin (pH 7)			Cysteine (pH 7)		2-Iminobiotin (pH 11)			ΔpH (pH 7–pH 11)
	T_m		ΔT_m	T_m	ΔT_m	T_m		ΔT_m	ΔT_m
	−	+		+		−	+		−
AVD	76.3	118.9	42.6	76.3	0.0	69.9	93.6	23.7	6.4
AVR2[a]	93.6	111.0	17.4	94.4	0.8	80.2	81.0	0.8	13.4
BBP-A[a]	ND[a]	102.2	ND	ND[b]	ND	ND[b]	ND	ND	ND
A/A2-1	92.4	94.7	2.3	92.3	−0.1	82.6	85.2	2.6	9.8
A/A2-3	76.5	93.4	16.9	85.2	8.6	85.2	85.9	0.7	−8.7
A/B-2	95.8	107.1	11.3	94.5	−1.3	63.7	67.9	4.2	32.1
A/B-9	70.2	84.2	14.0	69.4	−0.8	68.0	68.2	0.2	2.2

[a] BBP-A did show multiple peaks in the absence of biotin and the T_m has not been reported.
[b] Not determined due to low expression levels obtained due to pelB signal peptide.

shift (ΔT_m = 2.3 °C) in thermal stability when 21 μM of biotin was included in the sample.

The cysteine binding by A/A2-3 was confirmed by DSC. The mutant showed an increase in thermal stability (ΔT_m = 8.6 °C) in the presence of cysteine, which was not be observed in the parental or the other chimeric proteins. The thermal stability was also determined at pH 11 in presence and absence of 2-iminobiotin (Fig. 3B). Surprisingly, most mutants except for A/B-2 and A/B-9, showed a higher thermal stability at pH 11 as compared with avidin and AVR2, but the shift in thermal stability upon binding 2-iminobiotin was only between 0.2 and 4.2 °C, indicating a low binding affinity.

4. Discussion

DNA libraries constructed using the staggered extension process resulted in only small amount of chimeras and a large parental background. According to several references (Callison et al., 2005; Dion et al., 2001; Zhao et al., 1998), StEP appears to be a more suitable method for longer genes, and relatively short avidin-encoding DNAs (460 bp) might be difficult to shuffle using the StEP method. To our knowledge, all of the genes that have been successfully shuffled in previous studies had a length of at least 1 kb and were thus double the length of the avidin cDNAs.

Using the DNA shuffling method according to Stemmer, it was possible to construct the mutant libraries AVD/AVR2, AVD/BBP-A and AVR2/BBP-A with sufficient diversity in sequences and amount of crossovers, yet diversity decreased noticeably with the decreased sequence identities of the parental genes. It was not possible to produce chimeras from rhizavidin in combination with any of the other genes used, which might indicate that the DNA sequence homology between rhizavidin and the other genes (23–28%) is too low to be successfully employed in DNA family shuffling. The analysis of all the crossover sites that occurred in the genes showed that crossovers mostly happened in the four regions of high sequence identity (Fig. 1). These regions are blocks of 12–24 identical nucleotides occurring in all three genes. The absence of these regions between rhizavidin and the other avidin genes seems to be the reason for the unsuccessful production of chimeras using this method. On the basis of this knowledge, synthetic DNA using more divergent avidin genes could be produced resembling "avidin consensus" at these positions, which could allow shuffling of less similar avidin-related genes with each other (such as rhizavidin or streptavidin).

Biopanning resulted in an enrichment of the sequences. Not all of the sequences were shuffled, but they differed from the wild-type proteins by point mutations, which could have an influence on their expression level and functional properties, such as ligand binding. Surprisingly, almost all of the point mutations occurred in residues that were involved in the biotin binding in avidin, namely, Asn12, Ser16, Tyr33, Thr35, Ser75 and the interface tryptophan Trp110 (see Supplementary Table 1). The most abundant mutation observed was S16Y. This residue was revealed to be important for biotin binding in previous studies (Hytonen et al., 2005b; Klumb et al., 1998), and it makes direct hydrogen bond at the ureido oxygen of biotin (Livnah et al., 1993). The appearance of mutations linked to biotin-binding residues might reflect the selection method used, which possibly enriches proteins with decreased biotin-binding affinity compared with wild-type avidin. Although wild-type sequences were present in the libraries for AVD/BBP-A and AVR2/BBP-A, no enrichment of the parental sequences could be observed after the panning experiments. This result is consistent with the low expression level observed with wild-type sequences (Supplementary Table 1).

The low biotin-binding capacity of many of the analyzed clones suggests high non-specific binding of the phages, which could not be eliminated by pre-panning of wells coated with BSA. Yet, the setup of the microplate biotin-binding assay might reflect not only the affinity to the ligand but also a combination of the protein's expression level and affinity to biotin. The almost complete absence of positive signals of the wild-type proteins in the microplate assay indicates that the chosen secretion of the signal peptide pelB is not optimal for the expression of these proteins. Nevertheless, the high signal in some of the mutants suggests that mutations can provide a solution for this problem that favors the selection of these mutants by phage display biopanning. Our results, therefore, suggest that DNA shuffling can be an efficient way for improving the expression of secreted proteins in *E. coli*. The moderate biotin-binding affinity in all of the selected mutants is most probably due to the chosen elution conditions, which only allowed the release of inferior biotin-binders, possibly leaving the most tightly bound mutants bound to the physisorbed biotin-BSA. This problem could be circumvented in the future with the use of biotin that is immobilized via a cleavable linker. The phages could then be eluted with the ligand still bound to the protein. This technique might also help to decrease the amount of nonspecifically binding mutants.

The mutation S16Y was detected remarkably often in the phage-display enriched clones. The mutation caused a decrease in the biotin binding. In addition, a significant improvement in thermal stability was observed (T_m = 95.8 °C) when compared with wild-type avidin (T_m = 76.3 °C). The most straightforward explanation for this phenomenon is that tyrosine can act as an intrinsic ligand for the protein and stabilize the subunit. Inspection of the 3D structure of avidin suggests that the Y16 side chain could interact with side chains of residues N118 and N12.

The mutant A/A2-1 showed a distinct oligomeric behavior, which was not observed in the parental proteins. The gel filtration analysis showed an unstable oligomeric state in the absence of biotin, which seems to be stabilized in the tetrameric form by the binding of biotin. Surprisingly, differential scanning calorimetry revealed only a minor stabilizing effect due to biotin (ΔT_m = 2.3 °C). Yet, isothermal titration calorimetry (Fig. 4) estimates high affinity to biotin $K_d \ll 10^{-9}$ M. The DSC analysis revealed a high thermal stability of A/A2-1 (T_m = 92.4 °C). These results suggest that oligomerization and high thermal stability were not necessarily as tightly correlated as the previous studies had implied (Hytonen et al., 2005a; Nordlund et al., 2003a,b); the thermal stabilization or destabilization of avidin was caused by applying mutagenesis to subunit interfaces in these earlier studies.

The increased affinity of avidins for biotin can be seen from steep titration curves in ITC measurements. Strong exothermic reactions indicate a tight hydrogen bond network with biotin. However, the titration curves of the A/B-2 and A/B-9 mutants were to some extent gentler. Although the binding affinities were still relatively high, they were presumably weaker than those of other mutants. The shape of the A/B-9 titration curve could also indicate co-operative binding behavior or ligand-induced oligomerization.

Considering the fact that no specifically directed panning was used in the selection of the proteins, it is remarkable how diverse the characterized mutants are. This diversity might reflect the high structural similarity between the parental proteins utilized (Hytonen et al., 2005, 2007; Livnah et al., 1993). The previous studies have shown that the use of mutagenesis facilitated the transfer of the properties of the relative proteins between avidin family members. For example, we were able to improve the thermal stability of avidin by applying the mutation I117Y from AVR4 (Hytonen et al., 2005a), and the mutation K111I transferred from AVR2 to avidin resulted in a decrease in the biotin-binding

affinity (Hytonen et al., 2005). However, even an examination of the 3D-structure of AVR2 did not uncover the reason for lower biotin-binding affinity when compared with other members of avidin family (Hytonen et al., 2005). In contrast, the current study showed that the biotin-binding affinity of AVR2 could be improved by transferring segments from avidin while preserving the high thermal stability. Therefore, the current approach shows promise for the tailoring of avidins. We were able to generate a pool of shuffled sequences and select functionally improved proteins when compared with the parental proteins in terms of biotin binding and thermal stability. In addition, we observed a novel characteristic, namely, binding to cysteine, which might have value in applications, such as the separation of metabolites.

5. Conclusion

The potential of DNA shuffling within the avidin protein family has been demonstrated by combining the sequences of the avidin, AVR2 and BBP-A proteins. Screening the chimeric gene library resulted in novel proteins with high expression levels and novel characteristics. A novel mutation, S16Y, was resulted in thermal stabilization and a decrease in the biotin binding affinity. Moreover, we detected a mutant that displayed an affinity for cysteine. The chimeric proteins showed low affinity toward DNA, which is one of the challenges associated with wild-type avidin. The developed avidin mutants might offer characteristics suitable for applications in the life sciences.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jbiotec.2011.10.014.

References

Airenne, K.J., Oker-Blom, C., Marjomaki, V.S., Bayer, E.A., Wilchek, M., Kulomaa, M.S., 1997. Production of biologically active recombinant avidin in baculovirus-infected insect cells. Protein Expression and Purification 9, 100–108.

Barbas III, C.F., Burton, D.R., Jamie, K. Scott, Jamie, K., Silverman, G.J., 2001. Phage Display: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Callison, S., Hilt, D., Jackwood, M., 2005. Using DNA shuffling to create novel infectious bronchitis virus s1 genes: implications for s1 gene recombination. Virus Genes 31, 5–11.

Chaiet, L., Wolf, F.J., 1964. The properties of streptavidin, a biotin-binding protein produced by streptomycetes. Archives of Biochemistry and Biophysics 106, 1–5.

Chang, C.C., Chen, T.T., Cox, B.W., Dawes, G.N., Stemmer, W.P., Punnonen, J., Patten, P.A., 1999. Evolution of a cytokine using DNA family shuffling. Nature Biotechnology 17, 793–797.

Diamandis, E.P., Christopoulos, T.K., 1991. The biotin–(strept)avidin system: principles and applications in biotechnology. Clinical Chemistry 37, 625–636.

Dion, M., Nisole, A., Spangenberg, P., Andre, C., Glottin-Fleury, A., Mattes, R., Tellier, C., Rabiller, C., 2001. Modulation of the regioselectivity of a bacillus alpha-galactosidase by directed evolution. Glycoconjugate Journal 18, 215–223.

Eisenberg-Domovich, Y., Pazy, Y., Nir, O., Raboy, B., Bayer, E.A., Wilchek, M., Livnah, O., 2004. Structural elements responsible for conversion of streptavidin to a pseudoenzyme. Proceedings of the National Academy of Sciences of the United States of America 101, 5916–5921.

Helppolainen, S.H., Maatta, J.A., Halling, K.K., Slotte, J.P., Hytonen, V.P., Vainiotalo, P., Kulomaa, M.S., Nordlund, H.R., 2008. Bradavidin II from Bradyrhizobium japonicum: a new avidin-like biotin-binding protein. Biochimica et Biophysica Acta 1784, 1002–1010.

Helppolainen, S.H., Nurminen, K.P., Maatta, J.A., Halling, K.K., Slotte, J.P., Huhtala, T., Liimatainen, T., Yla-Herttuala, S., Airenne, K.J., Narvanen, A., Janis, J., Vainiotalo, P., Valjakka, J., Kulomaa, M.S., Nordlund, H.R., 2007. Rhizavidin from Rhizobium etli: the first natural dimer in the avidin protein family. The Biochemical Journal 405, 397–405.

Hyre, D.E., Le Trong, I., Freitag, S., Stenkamp, R.E., Stayton, P.S., 2000. Ser45 plays an important role in managing both the equilibrium and transition state energetics of the streptavidin–biotin system. Protein Science: A Publication of the Protein Society 9, 878–885.

Hyre, D.E., Le Trong, I., Merritt, E.A., Eccleston, J.F., Green, N.M., Stenkamp, R.E., Stayton, P.S., 2006. Cooperative hydrogen bond interactions in the streptavidin–biotin system. Protein Science: A Publication of the Protein Society 15, 459–467.

Hytonen, V.P., Laitinen, O.H., Airenne, T.T., Kidron, H., Meltola, N.J., Porkka, E.J., Horha, J., Paldanius, T., Maatta, J.A., Nordlund, H.R., Johnson, M.S., Salminen, T.A., Airenne, K.J., Yla-Herttuala, S., Kulomaa, M.S., 2004a. Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli. The Biochemical Journal 384, 385–390.

Hytonen, V.P., Maatta, J.A., Kidron, H., Halling, K.K., Horha, J., Kulomaa, T., Nyholm, T.K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S., Airenne, T.T., 2005. Avidin related protein 2 shows unique structural and functional features among the avidin protein family. BMC Biotechnology 5, 28.

Hytonen, V.P., Maatta, J.A., Niskanen, E.A., Huuskonen, J., Helttunen, K.J., Halling, K.K., Nordlund, H.R., Rissanen, K., Johnson, M.S., Salminen, T.A., Kulomaa, M.S., Laitinen, O.H., Airenne, T.T., 2007. Structure and characterization of a novel chicken biotin-binding protein A (BBP-A). BMC Structural Biology 7, 8.

Hytonen, V.P., Maatta, J.A., Nyholm, T.K., Livnah, O., Eisenberg-Domovich, Y., Hyre, D., Nordlund, H.R., Horha, J., Niskanen, E.A., Paldanius, T., Kulomaa, T., Porkka, E.J., Stayton, P.S., Laitinen, O.H., Kulomaa, M.S., 2005a. Design and construction of highly stable, protease-resistant chimeric avidins. The Journal of Biological Chemistry 280, 10228–10233.

Hytonen, V.P., Nordlund, H.R., Horha, J., Nyholm, T.K., Hyre, D.E., Kulomaa, T., Porkka, E.J., Marttila, A.T., Stayton, P.S., Laitinen, O.H., Kulomaa, M.S., 2005b. Dual-affinity avidin molecules. Proteins 61, 597–607.

Hytonen, V.P., Nyholm, T.K., Pentikainen, O.T., Vaarno, J., Porkka, E.J., Nordlund, H.R., Johnson, M.S., Slotte, J.P., Laitinen, O.H., Kulomaa, M.S., 2004b. Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin. The Journal of Biological Chemistry 279, 9337–9343.

Ikeuchi, A., Kawarasaki, Y., Shinbata, T., Yamane, T., 2003. Chimeric gene library construction by a simple and highly versatile method using recombination-dependent exponential amplification. Biotechnology Progress 19, 1460–1467.

Keinanen, R.A., Wallen, M.J., Kristo, P.A., Laukkanen, M.O., Toimela, T.A., Helenius, M.A., Kulomaa, M.S., 1994. Molecular cloning and nucleotide sequence of chicken avidin-related genes 1–5. European Journal of Biochemistry/FEBS 220, 615–621.

Klumb, L.A., Chu, V., Stayton, P.S., 1998. Energetic roles of hydrogen bonds at the ureido oxygen binding pocket in the streptavidin–biotin complex. Biochemistry 37, 7657–7663.

Laitinen, O.H., Hytonen, V.P., Ahlroth, M.K., Pentikainen, O.T., Gallagher, C., Nordlund, H.R., Ovod, V., Marttila, A.T., Porkka, E., Heino, S., Johnson, M.S., Airenne, K.J., Kulomaa, M.S., 2002. Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin. The Biochemical Journal 363, 609–617.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and clustal X version 2.0. Bioinformatics (Oxford, England) 23, 2947–2948.

Li, M.Z., Elledge, S.J., 2007. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nature Methods 4, 251–256.

Livnah, O., Bayer, E.A., Wilchek, M., Sussman, J.L., 1993. Three-dimensional structures of avidin and the avidin–biotin complex. Proceedings of the National Academy of Sciences of the United States of America 90, 5076–5080.

Maatta, J.A., Eisenberg-Domovich, Y., Nordlund, H.R., Hayouka, R., Kulomaa, M.S., Livnah, O., Hytonen, V.P., 2011. Chimeric avidin shows stability against harsh chemical conditions – biochemical analysis and 3D structure. Biotechnology and Bioengineering 108, 481–490.

Maatta, J.A, Helppolainen, S.H., Hytonen, V.P., Johnson, M.S., Kulomaa, M.S., Airenne, T.T., Nordlund, H.R., 2009. Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis. BMC Structural Biology 9, 63.

Marttila, A.T., Airenne, K.J., Laitinen, O.H., Kulik, T., Bayer, E.A., Wilchek, M., Kulomaa, M.S., 1998. Engineering of chicken avidin: a progressive series of reduced charge mutants. FEBS Letters 441, 313–317.

Nicholas, K.B., Nicholas, H.B.J., Deerfield, D.W.I., 1997. GeneDoc: Analysis and Visualization of Genetic Variation.

Nordlund, H.R., Hytonen, V.P., Laitinen, O.H., Uotila, S.T., Niskanen, E.A., Savolainen, J., Porkka, E., Kulomaa, M.S., 2003a. Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. FEBS Letters 555, 449–454.

Nordlund, H.R., Laitinen, O.H., Uotila, S.T., Nyholm, T., Hytonen, V.P., Slotte, J.P., Kulomaa, M.S., 2003b. Enhancing the thermal stability of avidin. Introduction of

disulfide bridges between subunit interfaces. The Journal of Biological Chemistry 278, 2479–2483.

Sidhu, S.S., Weiss, G.A., Wells, J.A., 2000. High copy display of large proteins on phage for functional selections. Journal of Molecular Biology 296, 487–495.

Stemmer, W.P., 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proceedings of the National Academy of Sciences of the United States of America 91, 10747–10751.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A., Arnold, F.H., 1998. Molecular evolution by staggered extension process (StEP) in vitro recombination. Nature Biotechnology 16, 258–261.

A novel chimeric avidin with increased thermal stability using DNA shuffling

Barbara Taskinen[a,b], Tomi T. Airenne[c], Janne Jänis[d], Rolle Rahikainen[a,b], Mark S. Johnson[c], Markku S. Kulomaa[a,e], Vesa P. Hytönen[a,b]*

[a]Institute of Biomedical Technology, Biokatu 6, FI-33014 University of Tampere and BioMediTech, Tampere, Finland

[b]Fimlab laboratories, Pirkanmaa hospital district, Biokatu 4, FI-33101 Tampere, Finland

[c]Department of Biosciences, Biochemistry, Åbo Akademi University, Tykistökatu 6A, FI-20520 Turku, Finland

[d]Department of Chemistry, University of Eastern Finland,Yliopistokatu 7, FI-80101 Joensuu, Finland

[e]Tampere University Hospital, Teiskontie 35, FI-33520 Tampere, Finland

*To whom correspondence should be addressed: Vesa P. Hytönen, Institute of Biomedical Technology, Biokatu 6, FI-33520 Tampere, Finland. Phone: +3584 0592 7534, Fax: +3583 364 1291, Email vesa.hytonen@uta.fi

Abstract

Avidins are a family of proteins widely employed in biotechnology. Here, we report a novel chimeric avidin form, A/A2-B, which was created using a methodology combining random mutagenesis by recombination and selection by a tailored biopanning protocol (phage display). The chimeric A/A2-B mutant showed increased thermal stability as compared to a previously reported chimeric mutant, A/A2-1, and the parental proteins chicken avidin and avidin related protein 2. The increased stability was especially evident at conditions of extreme pH as characterized using differential scanning calorimetry and dynamic light scattering. In addition, the crystal structure of the A/A2-1 mutant was solved at 1.8 Å resolution, revealing that the protein fold was not affected by the shuffled sequences and that the X-ray data supported observed physicochemical properties of the mutant.

Introduction

Common to all members of the avidin protein family is their high affinity towards a small ligand, D-biotin. This is the reason why avidin from chicken (AVD) and streptavidin from the bacterium *Streptomyces avidinii* are used in a wide variety of different biotechnology applications [1, 2]. In addition to AVD and streptavidin, several other avidins have also been described including the avidin-related proteins (AVRs) and biotin-binding proteins (BBPs) from chicken [3, 4], bradavidin I and II from *Bradyrhizobium japonicum* [5-7], rhizavidin from *Rhizobium etli* [8], shwanavidin from *Shewanella denitrificans* [9], tamavidin from mushroom [10], xenavidin from frog [11], and zebavidin from zebrafish [12]. Despite their relatively low amino acid sequence identity, all avidins share a highly similar tertiary structure, an eight-stranded β-sheet barrel structure, where the biotin-binding site is located at the open end of the barrel. With the exception of rhizavidin, shwanavidin and bradavidin II, all of the structurally characterized natural avidins are stable homotetramers with four biotin-binding sites, one on each monomer. The biotin-binding sites are not equally distributed on each side of the tetramer but

2

instead are organized as pairs on two opposite faces. Even though all avidins have similar tertiary structures, the proteins differ e.g. in their quaternary structure, ligand-binding affinities towards biotin and other small ligands, thermal stabilities and isoelectric points. These differences make the avidin protein family an excellent candidate for the design of novel proteins by DNA family shuffling [13, 14].

We have previously created functional chimeric avidins using DNA shuffled sequences of AVD and avidin related protein 2 (AVR2) [15]. Here, with an improved DNA library screening method we have selected a new mutant, A/A2-B. A/A2-B shares high sequence similarity with the previously characterized chimeric mutant A/A2-1, but has a decreased dissociation rate from biotin and a higher thermal stability, exceeding that observed for the highly stable parental protein AVR2. A/A2-B is thus the first avidin developed by directed evolution exceeding the thermal stability of any naturally existing avidin. We also present the crystal structure of A/A2-1, which elucidates the observed physicochemical features at structural level, but also proves that the avidin fold is robust and suitable for extensive genetic manipulation.

Materials and Methods

Selection of A/A2-B by phage display of AVD/AVR2 mutant library

Phage library preparation and biopanning was essentially done as described previously [15]. The chimeric mutants were displayed as a fusion with the pIII coat protein on the surface of M13 bacteriophages. An amber codon between the chimeric gene and the pIII gene allowed simultaneous expression of fusion proteins and free chimeric avidin monomers resulting in the display of tetrameric avidin proteins on the phage surface [15, 16]. *E. coli* XL1 blue cells harboring the AVD/AVR2 DNA mutant library were infected with VSC-M13 helper phages (Stratagene, La Jolla, CA, USA; 10^{11} pfu/ml) and cultured over night. Phages were collected by precipitation with 20% PEG-6000 in 2.5 M NaCl [15, 17]. Biotinylated BSA and casein used to capture functional biotin-binding proteins were

prepared by reaction with EZ-Link Sulfo-NHS-SS-Biotin (Thermo Scientific, Waltham, MA, USA) in a 10-fold molar excess in 5 mM HNa_2PO_4/H_2NaPO_4, 150 mM NaCl, pH 8 for 2 h at room temperature (RT). Unreacted biotinylation reagent was removed by dialysis against PBS. The success of biotinylation was assayed using a HABA assay according to the manufacturer's instructions (Thermo Scientific). Biotinylated proteins in PBS were used to coat MaxiSorp Immuno 96 MicroWell plates (Nunc A/S, Roskilde, Denmark) overnight at 4 °C. Biotinylated casein and BSA were used in alternating panning rounds to reduce nonspecific binding to the carrier protein. Precipitated phages were diluted 1:10 on the first panning round and 1:2 on the second, third and fourth panning round. Dilutions were made using 1% BSA in PBS in the first and third panning round and 1% casein in PBS in the second and fourth panning round. Diluted phages were incubated over night at 4 °C to allow the BSA- or casein-binding phages to bind their respective ligands. Phages (100 µl) were added to wells coated with biotinylated protein (200 ng) and incubated for 90 min with agitation at 600 rpm (Talboys orbital microplate shaker 1000MP, Troemner, Thorofare, NJ, USA) at RT. Wells were washed five times with 0.05% PBS-Tween (PBST) (300 ul). Wells were incubated with PBS (300 µl) containing D-biotin (2 µg) for 15 min at 37 °C and 600 rpm in order to elute non-specifically bound phages and phages displaying proteins with a high dissociation rate. Wells were washed with PBS (300 µl) and bound phages were eluted by addition of 50 mM DTT (100 µl) during an incubation of 1 h at 37 °C. Eluted phages were used to infect XL1 blue cells and amplified as described in [15]. Master plates were prepared from randomly picked colonies after the fourth panning round. Selected colonies were inoculated into the freezing medium (100 µl of 30 g/L tryptone, 20 g/L yeast extract, 10 g/L MOPS, 0.4 mM $MgSO_4$, 36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 6.8 mM $(NH_4)_2SO_4$, 1.7 mM sodium citrate, 4.4% glycerol, pH 7) containing 1% (v/v) glucose, 10 mg/ml tetracycline and 50 µg/ml ampicillin. Samples were incubated on 2.2 ml storage plate (ABgene, Thermo Scientific, Surrey, UK) for six hours at 37 °C and 600 rpm. In order to increase the total glycerol concentration to 15%, a solution containing freezing medium and glycerol in 1:1 ratio (24 µl) was added. Plates were shaken for 15 min at 600 rpm and stored at –70 °C.

Avidin-biotin displacement assay

In order to characterize the ligand-binding properties of selected phage colonies, a microplate based protocol was used to get an estimate of the dissociation rate constant for biotin binding. A sample (7 µl) from the master plates prepared after phage panning were used to inoculate SB medium (200 µl) containing 10 µg/ml tetracycline, 100 µg/ml ampicillin and 1% glucose. Cells were grown in a 2.2 ml storage plate (ABgene, Thermo Scientific) for 6 h at 37 °C and 600 rpm (Talboys orbital microplate shaker 1000MP, Troemner) after which the protein expression was induced by adding SB medium (50 µl) containing 5 mM IPTG, 10 µl/ml tetracycline, 50 µg/ml ampicillin and 0.1% glucose. The final concentration of IPTG was 1 mM. Cells were incubated over night at 28 °C and 600 rpm. Three MaxiSorp™ 96 MicroWell™ plates (Nunc A/S) were coated with BSA (200 ng) in PBS or with a mixture of BSA (50 ng) and BSA-BTN (150 ng) in PBS (200 µl). For each clone to be analysed, one well was coated with BSA and two wells were coated with BSA-BTN. Coating was done over night at 4 °C. Cells were harvested by centrifugation for 10 min at 1,000 g and 4 °C. Pellets were frozen at –20 °C. Frozen cells were lysed by vigorous shaking in EasyLyse reagent (40 µl) (Epicentre Biotechnologies, Madison, Wisconsin, USA) for 10 min. The cell lysate was diluted with 10 mM Tris-HCl buffer, pH 7.5, 1 M NaCl (200 µl). Cell debris was removed by centrifugation at 8'000 g for 5 min at 4 °C. Coated wells were washed 3 times with PBST and blocked with 1% (w/v) milk in PBST (300 µl) for 1 h at RT and then washed again. The clarified cell lysate (70 µl) was added to each well and incubated for 1 h at RT, shaking at 500 rpm (Talboys orbital microplate shaker 1000MP, Troemner). Unbound proteins were washed off with PBST.

In order to detect the bound proteins, biotinylated alkaline phosphatase diluted 1:5000 in 1% milk in PBST (100 µl) was added to each well. Plates were incubated for 1 h at RT and 500 rpm. After washing, PBS (200 µl) was added to each well of the plate coated with 1% milk in PBS and one plate coated with BSA-BTN. PBS + 10 µg/ml biotin (200 µl) was added to each well of the second plate coated with BSA-BTN and incubated for 30 min at RT and 500 rpm. Wells were washed 3 times with

PBST and 3 times with PBS. "Phosphatase substrate" (100 µl of 1 mg/ml, Sigma-Aldrich, St. Louis, MO, USA) in DEA buffer (1M diethanol amine, 0.5 mM $MgCl_2$, pH 9.8) was added to each well and the absorbance at 405 nm was measured at 10-minute intervals for one hour using a microplate reader 680XR (Bio-Rad Laboratories Inc., Hercules, CA, USA). The signals were compared between wells treated with free biotin and the untreated wells in order to analyze the rate of biotin displacement instead of just assaying the functionality of the mutants by detection of bound biotinylated alkaline phosphatase. The ratio of the signals gives an estimate for biotin dissociation rate, which is independent of the amount of functional protein expressed.

Protein production and purification

The A/A2-B mutant was produced in an *E. coli* bottle culture using BL21-AI cells, whereas the A/A2-1 mutant as well as the parental proteins AVD and AVR2 were produced using a pilot-scale fermentor using *E. coli* BL21-AI fed-batch culture as described in [15, 18]. Protein purification by 2-iminobiotin affinity chromatography was performed as described previously [15]. The purified proteins were analyzed with denaturing mass spectrometry to verify their purity and sequence (data not shown).

Dissociation rate determination by fluorescence spectrometry

The dissociation rate constant (k_{diss}) of fluorescently labelled biotin was determined by fluorescence spectrometry using the biotin-labelled fluorescent probe ArcDia^{TM} BF560 (ArcDia, Turku, Finland) as described in [19]. In principle, the changes in the fluorescence intensity of a 50 nM dye in 50 mM sodium phosphate, 650 mM NaCl, 0.1 mg/ml BSA, pH 7 was measured after the addition of biotin-binding protein to a final concentration of 100 nM. The dissociation of this complex was observed by addition of a 100-fold molar excess of free biotin. The assay was performed at 25 °C using a QuantaMaster^{TM} Spectrofluorometer (Photon Technology International, Inc., Lawrenceville, NJ, USA)

equipped with circulating water bath thermostat. The fluorescence probe was excited at 560 nm and emission was measured at 578 nm.

Size exclusion chromatography with light scattering analysis

Proteins were analysed using a liquid chromatography instrument (CBM-20A, Shimadzu Corporation, Kyoto, Japan) equipped with autosampler (SIL-20A), UV-VIS (SPD-20A) and fluorescence detector (RF-20Axs) as well as Zetasizer µV light scattering detector (Malvern Instruments Ltd, Worcestershire, UK) for molecular weight (static light scattering) and hydrodynamic size (dynamic light scattering) determination. The instrument was controlled using Lab Solutions Version 5.51 (Shimadzu Corporation) and OmniSEC 4.7 (Malvern Instruments Ltd.). Samples (70 µg in 40-100 µl) were injected on a Superdex75 5/150GL column (GE healthcare, Uppsala, Sweden) equilibrated with Na_2HPO_4/NaH_2PO_4, 650 mM NaCl, pH 7. Runs were executed with a flow rate of 0.25 ml/min at 12 °C. Molecular weight determination was either done by calculating a standard curve of molecular weight markers (cytochrome C, 12.4 kDa; carbonic anhydrase, 29 kDa; ovalbumin, 44 kDa; BSA, 66 kDa, Sigma-Aldrich) or was based on the light-scattering intensity of the eluting protein, for which BSA was used for instrument calibration using the OmniSEC software (Malvern Instruments Ltd.).

Mass spectrometry analysis

All mass spectrometric (MS) analyses were performed on a 12-T APEX-Qe™ Fourier transform ion cyclotron resonance (FT-ICR) instrument (Bruker Daltonics, Billerica, MA, USA), interfaced to an electrospray ionisation (ESI) source. The protein sample's buffer was exchanged into 10 mM ammonium acetate (pH 6.9) with the use of PD-10 columns (GE Healthcare). The resulting fractions, which eluted in a volume of 3-4 ml, were pooled and concentrated to approximately 250 µl using Millipore Ultrafree-0.5 Biomax-5 (5-kDa cut-off) centrifugal filter devices (Millipore, Billerica, MA, USA). The concentrations of the protein stock solutions were determined by measuring the UV-absorbance

at 280 nm and by using a sequence-derived extinction coefficient calculated using General Protein/Mass Analysis for Windows (GPMAW, Lighthouse Data, Odense, Denmark). In order to analyze the protein under denaturing solution conditions, the stock solution was further diluted with an acetonitrile/water/acetic acid (49.5:49.5:1.0, v/v/v, pH 3.2) solvent. Alternatively, the sample was diluted with 10–500 mM ammonium acetate (pH ~7) to perform native-MS analysis. For biotin-binding experiments, D-biotin (Sigma-Aldrich) stock solution (2 mM in water) was mixed at a desired molar ratio with the protein and incubated at RT for 30 min prior to the analysis. For each spectrum, a total of 128-512 co-added 1MWord (128kWord for native-MS) time-domain transients were zero-filled once prior to a fast Fourier transformation, magnitude calculation and external mass calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, CA, USA). The instrument was operated and the data were processed with the use of XMASS 6.0.2 software.

Dynamic light scattering

The hydrodynamic radius of proteins in varying conditions was determined by batch dynamic light scattering using Zetasizer Nano ZS (Malvern Instruments Ltd.). Proteins were dialysed either into pH 7 buffer (50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, pH 7), pH 3 buffer (50 mM sodium citrate, 100 mM NaCl, pH 3), or pH 11 buffer (50 mM sodium carbonate, 100 mM NaCl, pH11). Proteins were analysed at a concentration of 1 mg/ml in the absence and presence of a 3-fold excess of D-biotin. Samples were heated in 5 °C steps from 25 °C to 90 °C, and equilibrated for 5 min prior to data collection. Three measurements were made with 10 runs of 10 s per measurement. Data was analysed using Zetasizer software v7.01 (Malvern Instruments Ltd). Analysis was based on the volume distribution, and the volume mean size was plotted against the temperature in order to obtain the transition temperature for thermal unfolding.

Differential scanning calorimetry

The thermal stability of the studied proteins in the presence and absence of ligands was analysed using an automated VP-Capillary DSC System (Microcal Inc., Northampton, MA, USA). Thermograms were recorded between 20 and 140 °C with a heating rate of 2.0 °C/min. Proteins were dialysed into a pH 3, pH 7 or pH 11 buffer. Samples were degassed prior to the measurement. The protein concentration in the cell was 30 μM, and the ligand concentration was 105 μM. Results were analysed using the Origin 7.0 DSC software suite (Microcal Inc.).

X-ray structure determination of A/A2-1

The vapour diffusion method, 96-well sitting drop iQ plates (TTP Laptech) and TTP Labtech's mosquito® liquid handling robot were used to crystallize the A/A2-1 mutant (7.4 mg/ml; 50 mM sodium acetate, pH 4), at 22 °C. The protein was mixed with a D-biotin solution (1 mg/ml; 5 mM Tris, pH 8.8, 8 mM CHES, pH 9.5) in a 10:1 v/v ratio before crystallization. The JCSG-plus[TM] Screen (Molecular Dimensions Ltd., Suffolk, UK) gave an initial hit and, after optimization, the A/A2-1 mutant crystals formed typically in 1-2 weeks in drops of protein-biotin solution (300 nl) and well solution (300 nl; 0.09 M phosphate/citrate, pH 4.2; 36% v/v PEG 300); well solution (55 μl) was used at the bottom of the well.

X-ray data were collected at the ESRF beam line ID14-1 (Grenoble, France) at 100 K from a single crystal, which was frozen in liquid nitrogen; no cryoprotectant was added. The data were processed with XDS [20] (see Table 3 for statistics) and the phase problem was solved using the molecular replacement program Phaser [21] within the CCP4i GUI [22, 23]. In the molecular replacement, the A chain of AVR2 [PDB: 1WBI] [24] was used as the search ensemble; two chains were searched for. The structure from Phaser was first rebuilt using the automatic procedure of ARP/wARP [25-27] and then, in several cycles, manually edited/rebuilt within Coot [28] and refined

with Refmac5 [29]. Non-protein atoms were added to the model either with the automatic procedure of Coot and ARP/wARP, or manually in Coot. For structure determination statistics, see Table 3.

The final structure of the A/A2-1 mutant was validated using the inbuilt tools of Coot [28], and using MolProbity [30] of the Phenix software suite [31], before depositing the coordinates and structure factors to the Protein Data Bank [32, 33] with PDB entry code 4BCS.

Miscellaneous methods

The sequence alignment was created using ClustalW [34] and edited using GeneDoc [35] and Microsoft Office Word 2010 (Microsoft). The structural superimpositions were made using the "align" command of PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.2, Schrödinger, LLC) and the chain A of the A/A2-1 mutant as the reference structure. PyMOL and the visualization and modelling package Bodil [36] was used for visual analyses of the structures. PyMOL was used to create all the figures relating to structural representations. GIMP 2.6.9 and CorelDRAW X5 were used to edit the figures. Theoretical hydrodynamic radius from the obtained crystal structure of the A/A2-1 mutant was estimated using HYDROPRO Version 10 [37].

Results

Selection of a chimeric mutant with improved biotin binding

Biopanning of a AVD/AVR2 DNA shuffling library against a biotin-coated surface was previously done by phage display, resulting in enrichment of a mutant protein, A/A2-1, with an increased dissociation rate towards biotin, reflecting less-tight biotin binding in comparison to either AVD or AVR2 [15]. Here, we modified the panning method by 1) including an additional washing step with free biotin prior to

elution, 2) adding a disulphide bond in the linker between biotin and the coating protein, which allows cleavage of the linker, and 3) developing a powerful microplate assay for functional screening of the clones. This led to successful selection of the A/A2-B mutant with a decreased dissociation rate, i.e. D-biotin binds more tightly to this mutant than to A/A2-1.

The initial analysis of the freely expressed protein was made using an avidin-biotin displacement assay (ABD-assay). The assay compares two biotinylated wells that were incubated with cell lysate containing the expressed protein. While one well is incubated with buffer, the other well is incubated with buffer containing biotin. Bound biotin binding proteins are subsequently detected by biotinylated alkaline phosphatase. The ratio between the resulting signals of the two wells gives an indication of the biotin dissociation rate. A mutant, named A/A2-B, showed a good binding response and had a high response ratio for the biotin-treated versus untreated wells, thus indicating a slow dissociation from the biotin-functionalized surface (Table 1). The observed response ratio was higher (meaning a lower dissociation rate) than for the previously analyzed mutant A/A2-1 [15], yet still lower as compared to AVD. Although the dissociation rates of A/A2-1 and A/A2-B were clearly different, the two mutants only differ at six residue positions (Figure 1).

The observed lowered dissociation rate of mutant A/A2-B as compared to A/A2-1 was further confirmed by following the dissociation of the biotin labeled with the fluorescent dye BF560 (Figure 2). At 25 °C, AVR2 and A/A2-1 showed very rapid dissociation of BF560-biotin; within seconds after the addition of free biotin, all of the probe was released from the protein, indicating a $k_{diss} > 0.1$ s^{-1}. In comparison, AVD showed a slow dissociation rate of $0.13 \cdot 10^{-4}$ s^{-1}; even one hour after the addition of free biotin less than 20% of the probe was released from the protein. Moreover, A/A2-B showed a dissociation rate between that of AVD and AVR2, and clearly lower than A/A2-1. Yet, the dissociation rate ($3.08 \cdot 10^{-4}$ s^{-1}) of A/A2-B measured was about 25-fold faster than that observed for AVD, and 76% of the probe was released from A/A2-B one hour after the addition of free biotin.

11

The mutants A/A2-1 and A/A2-B were produced in *E. coli*, purified using 2-iminobiotin affinity chromatography and identified by denaturing MS. The integrity of the oligomeric states was analyzed by native MS as well as by size-exclusion chromatography with on-line static and dynamic light scattering detector (SEC-LS/DLS) and batch DLS. These analyses showed that both proteins formed stable tetramers in the presence and absence of biotin (Supporting Figures S1 and S2). The experimental molecular masses were in agreement with the theoretical molecular masses of the homotetrameric proteins. In addition, the hydrodynamic radius for the tetramer, calculated from the crystal structure of A/A2-1 [PDB: 4BCS] using the HYDROPRO software (Table S1) [37] was consistent with the experimental data (Table S1). Furthermore, the SEC analysis indicated strong interaction with the column for both mutants, seen by a low UV response and elution in multiple peaks that contain protein species of identical molecular weight (Supporting Figure S2A&C). This explains the previously observed high elution volume in SEC analysis and reveals why the determined molecular mass was lower than would have been predicted for the A/A2-1 tetramer [15].

High thermal stability of the chimeric mutants

High stability of proteins over a broad pH range is required for many applications. Therefore, the thermal stability of the chimeric avidin mutants was analyzed at varying pH values. In differential scanning calorimetry (DSC) analysis without biotin, A/A2-1 and A/A2-B showed increased thermal stability at all pH values studied (pH3, pH 7, pH 11) in comparison to AVD; the most clear difference was found at pH 3 (Table 2). However, AVD was the most stable protein at pH 7 when biotin was added. The pH-dependence of thermal stability, with or without biotin, was apparent for AVD and A/A2-1, whereas A/A2-B seemed to be less sensitive to changes in the pH. For example, the transition midpoint (T_m) of AVD at pH 7 dropped from 79.9 °C to 51.2 °C at pH 3 and to 67.9 °C at pH 11, whereas the difference between the highest (at pH 7) and lowest (at pH 11) T_m value of A/A2-B was only 7.2 °C. The addition of biotin stabilized AVD and A/A2-B notably, but the thermal stability of A/A2-1 was not substantially affected by biotin binding, as has been previously reported [15]. However, at

pH 3, the increase of thermal stability upon D-biotin binding (ΔT_m) for A/A2-1 was 23.1 °C, which is substantially higher than at neutral pH (8.5°C) and in the range observed for A/A2-B (ΔT_m = 26.8 °C). Unfortunately, it was not possible to analyze AVR2 at pH 3, because the protein precipitated upon dialysis at this pH, most probably because of the low isoelectric point (pI = 4.7) of AVR2. In general, the thermal stability of A/A2-B did not drastically change at any of the analyzed conditions and A/A2-B was the most stable protein studied (Table 2).

The thermal stability of the proteins was further analyzed by DLS and by following the hydrodynamic radius during the heating of the samples from 25 to 90 °C (Figure 3). In this analysis, particles with a hydrodynamic radius reflecting tetramers (~3 nm) were observed until the denaturing temperature was reached, indicated by the appearance of large particles due to protein aggregation. In the absence of biotin, AVD was the least stable of the analyzed proteins and aggregated depending on the pH in a temperature range between 45-60 °C in line with the results from the DSC analysis. According to the volume mean size distribution (Figure 3), AVD was present in the tetrameric form in the temperature range 65-85 °C, however, when the derived count rate was plotted against the temperature (Figure S3), it was clear that at temperatures higher than 45 °C, aggregates were the dominant species in the solution. Generally, the observed denaturing temperatures were lower than those observed using DSC due to at least the following reasons: (1) a lower temperature scan rate was used in DLS experiments, and (2) DSC reports the transition temperature of thermal unfolding, whereas the onset of thermal unfolding was recorded by DLS. For all conditions studied, both chimeric mutants showed higher thermal stabilities in comparison to AVD, with A/A2-B being substantially more stable than A/A2-1. However, at pH 7 in the absence of biotin, none of the mutants was more stable than AVR2. In the presence of biotin, protein denaturation could not be observed by DLS due to the upper temperature limit (90 ºC) of the instrument. However, the low ΔT_m observed in DSC analysis for A/A2-1 could be confirmed at pH 11, where the protein started to aggregate already at 80 °C.

Crystal structure of A/A2-1

In order to better understand the structural origin of the physicochemical properties, the A/A2-1 mutant

structure was solved in complex with biotin at 1.8 Å resolution (see Table 3 for the structure

determination statistics). Two monomers were found in the asymmetric unit and, as expected based

on the high sequence identity to its parental proteins AVD [38] and AVR2 [24], the quaternary

structure was homotetrameric (Figure 4A), each monomer having an eight-stranded β-sheet barrel

structure with the biotin-binding site located at the open end of the barrel (Figure 4B,C; [39]). When

superimposed, A/A2-1 fits better to the AVR2 [PDB: 1WBI] than the AVD structure [PDB: 1AVD], the

r.m.s.d values for the displacement of Cα atoms being 0.2 Å and 0.5 Å, respectively. Almost all of the

sequence and structural differences to AVD are in the region derived from AVR2, starting at A38 of the

L3,4-loop of A/A2-1 and ending with S58 of the L4,5-loop, whereas A/A2-1 differs from AVR2 at many

single positions at the N- and C-terminal regions originating from AVD (Figure 1).

The individual or cumulative functional effect of the many mutations found in A/A2-1 is not

trivial to predict but probably the two most interesting residues of A/A2-1 in terms of ligand binding and

stability are D12 and D39. The D12 of A/A2-1 is replaced by an asparagine in AVD, AVR2 and A/A2-

B, which has been shown to hydrogen bond to the ureido ring of biotin [39]. Based purely on the X-ray

structure presented here, the D12 of A/A2-1 also seems to be hydrogen bonded (2.7 Å) to the

carbonyl oxygen atom of the ureido ring of biotin (Figure 5A). In bulk water, protonation of the side

chain of aspartate occurs only at a lower pH (pK_a=3.9).

D39 of A/A2-1 forms a salt bridge with R112 in a way similar to that seen in the AVR2 structure

[PDB: 1WBI], whereas in AVD the salt bridge is missing (Figure 5B-D). The main-chain nitrogen atom

of D39 in A/A2-1 and AVR2, as well as the nitrogen atom of the equivalent residue A39 in AVD, is

hydrogen bonded to biotin; the salt bridge is hence indirectly linked to BTN via D39. Moreover, L97

and S99 are in the close vicinity of the salt bridge in A/A2-1 and are equivalent to Q97 and L99 in

AVR2 (Figure 1, 5D). These differences could explain the slightly different geometry of the salt bridge of A/A2-1 and AVR2. Furthermore, residues T75 and L97 of A/A2-1 and S75 and Q97 of AVR2 are in contact with biotin and may thus directly influence the biotin-binding properties of these proteins, too. T75, L97 and S99 of A/A2-1 originated from AVD (T77, L99 and S101).

Discussion

The extreme ligand-binding affinity of chicken avidin has made it an important tool in biotechnological applications. We have earlier shown that DNA family shuffling in combination with phage display is a powerful method to create and select novel avidins with altered physicochemical properties [15]. Choosing the optimal parameters during the phage display biopanning process is crucial for the successful selection of high-affinity binders, but that is not a trivial task. We were able to fine-tune the selection conditions by employing harsher washing conditions with free biotin present, by using ligand-binding independent elution conditions during phage display biopanning and by using a more refined screening method. As a result of the new selection strategy, we could select a new mutant, A/A2-B, which displayed a decreased dissociation rate as compared to a previously characterized mutant (A/A2-1) selected from the same library. The additional wash step with free biotin removes mutants with high dissociation rate towards biotin and the use of a biotinylation reagent with a cleavable linker ensures the release of even the tightest and acid-resistant binders. Furthermore, our screening of the selected mutants using the ABD-assay confirmed the strength of the biotin affinity and enabled the selection of efficiently expressing mutants. The improved ligand-binding properties of A/A2-B over A/A2-1 demonstrate the importance of choosing the optimal parameters for biopanning experiments.

There are only six amino acid differences between A/A2-1 and A/A2-B:K9E, D12N, S105G, I115N, I117D and Q121L (Figure 1), yet they cause distinct changes In the biotin-binding properties and protein stability. The difference between the biotin dissociation rates of A/A2-1 and A/A2-B can be

traced back to the mutation N12D. In AVD, the side chain of N12 hydrogen bonds to the ureido

oxygen of biotin [39] and mutation of the corresponding residue in streptavidin to alanine lowered

drastically the biotin affinity of the mutant via disabling hydrogen bonding of the mutated residue to

biotin [40]. At pH 7, D12 of A/A2-1 may not be able to act as an effective hydrogen bond donor,

because of the low pK_a value (3.9) of its side chain. A/A2-1 was crystallized at very acidic conditions,

and the X-ray structure of A/A2-1 represents a snapshot of conditions near the pK_a, where about 50%

of the D12 residues in the crystal would be expected, if exposed to solvent, to be protonated. In the

A/A2-1 structure the OD2 oxygen atom of D12 is surrounded by aromatic and nonpolar residues that

would act to reduce the effective dielectric constant (in comparison to bulk water), effectively raising

the pKa value and increasing the probability of a protonated D12. This would support the formation of

the hydrogen bond more fully at pH 4 and this is reflected in the hydrogen bond between the two

oxygen atoms (2.7 Å apart). This hypothesis is supported by our DSC data, which showed that at pH 3

biotin stabilized A/A2-1 substantially more (ΔT_m = 23.1 °C) than at pH 7 (ΔT_m = 8.5 °C). This

interaction seen between the ureido oxygen atom of biotin and the side chain oxygen atom of D12

(Figure 5A) is similar to that seen e.g. in the crystal structures of the parental proteins AVD [PDB:

2AVI] and AVR2 [PDB: 1WBI], but where an asparagine, unaffected by pH fluctuation in this regard, is

involved in the hydrogen bond.

The question remains as to what factors are responsible for the lowered dissociation rate of

A/A2-B in comparison to AVR2. Two residues that are in contact with biotin may be involved: residue

75, which is threonine in A/A2-1, A/A2-B and AVD but serine in AVR2, and residue 97, which is

leucine in A/A2-1, A/A2-B and AVD, but glutamine in AVR2. The presence of a serine residue at

position 75 of AVR2 has been reported to increase the size of the biotin-binding pocket and hence

decrease the affinity to biotin (Figure 5D) [24]. The additional methyl group on T75, as seen in the

A/A2-1 structure, forms multiple hydrophobic interactions with biotin (C6 and S1 atoms) and with F77

(CD1 and CE1 atoms) that would increase the biotin affinity in comparison to S75 in AVR2. In the

A/A2-1 structure, the terminal methyl groups of L97 pack against the centrally located C2 and C8 carbon atoms on one face of biotin. Replacement by glutamine in AVR2 negates these nonpolar interactions but the amide side chain is then ideally located to hydrogen bond elsewhere: to the side chains of S75 and R112 (altering coplanarity of the salt bridge slightly with D39) and to the O12 atom of biotin.

Both mutants, A/A2-1 and A/A2-B, displayed high thermal stability in the absence of biotin, similarly to that measured for AVR2. In the presence of D-biotin, the thermal stability of A/A2-B exceeded that of AVR2, which may reflect the higher biotin-binding affinity of A/A2-B. The sequence stretch between residues 38-58, which represents the β-strand 4 and most of the adjacent L3,4- and L4,5-loops is likely to be the main source for the increased stability. The sequence region 38-58 is from AVR2 and also found in AVR4, another highly thermostable avidin-related protein found in chicken [41]. Moreover, residues 38-58 have successfully been transferred to AVD, to create an ultra-stable chimeric protein termed ChiAVD with T_m = 111.1°C in the absence of ligand [42].

The crystal structure of A/A2-1 shows that the L3,4-loop conformation of the mutant is highly similar to AVR2 and AVR4 and might at least partially explain the observed high thermal stability. The L3,4-loop is rather flexible in AVD without biotin and upon binding of biotin serves as a lid, locking the biotin-binding site [39]. In the L3,4-loop of AVR2, AVR4, A/A2-1 and A/A2-B (Figure 1), there is a proline residue, which introduces rigidity, and the loop is further stabilized by the intermonomeric salt bridge between D39 and R112 as described above for A/A2-1 (and A/A2-B) and reported earlier for AVR2 and AVR4 [24, 43]. As shown for AVR4 [43], the conformation of the proline-containing L3,4-loop does not change drastically upon biotin binding and the bound biotin is partially exposed to the solvent [43]. Moreover, A/A2-1 and A/A2-B have an isoleucine residue at position 109 similarly to AVR2, whereas in AVD and AVR4 the equivalent residue is lysine. When K109I mutation was introduced into AVD or AVR4, it reduced the affinity towards biotin [24]; the mutation K109I was shown to have a destabilizing effect on AVD and AVR4 and, conversely, changing isoleucine to lysine at this position in AVR2

increased the thermal stability [24]. It would be interesting to see what effect mutation I109K would have in A/A2-B, which already displays high thermal stability.

In conclusion, DNA shuffling of AVD and AVR2 resulted in a chimeric mutant (A/A2-B) with higher thermal stability than that observed for the parental proteins and a lower biotin dissociation rate than that measured for a previously reported mutant, A/A2-1. A/A2-B displays high stability at a broad pH range. The crystal structure of A/A2-1 revealed structural details that support the physicochemical observations. The highly stable chimeric avidin reported here may help in the design of improved, higher stability engineered avidins such as highly-stable steroid-binding avidins [44].

Authors Contributions

BT, TA, JJ and RR carried out the experiments. BT, TA, JJ, RR, MSJ and VPH analyzed the data. TA, MSJ, MSK and VPH contributed to writing the research proposal and all authors were involved in the preparation of the manuscript. All authors have seen and approved the final version.

References

1. Diamandis EP, Christopoulos TK (1991) The biotin-(strept)avidin system: principles and applications in biotechnology. Clin Chem 37: 625-636.

2. Laitinen OH, Nordlund HR, Hytönen VP, Kulomaa MS (2007) Brave new (strept)avidins in biotechnology. Trends Biotechnol 25: 269-277.

3. Keinänen RA, Wallén MJ, Kristo PA, Laukkanen MO, Toimela TA et al. (1994) Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5. Eur J Biochem 220: 615-621.

4. Hytönen VP, Määttä JA, Niskanen EA, Huuskonen J, Helttunen KJ et al. (2007) Structure and characterization of a novel chicken biotin-binding protein A (BBP-A). BMC Struct Biol 7: 8.

5. Nordlund HR, Hytönen VP, Laitinen OH, Kulomaa MS (2005) Novel avidin-like protein from a root nodule symbiotic bacterium, Bradyrhizobium japonicum. J Biol Chem 280: 13250-13255.

6. Helppolainen SH, Määttä JA, Halling KK, Slotte JP, Hytönen VP et al. (2008) Bradavidin II from Bradyrhizobium japonicum: a new avidin-like biotin-binding protein. Biochim Biophys Acta 1784: 1002-1010.

7. Leppiniemi J, Meir A, Kahkonen N, Kukkurainen S, Maatta JA et al. (2013) The highly dynamic oligomeric structure of bradavidin II is unique among avidin proteins. Protein Sci 22:980-94.

8. Helppolainen SH, Nurminen KP, Maatta JA, Halling KK, Slotte JP et al. (2007) Rhizavidin from Rhizobium etli: the first natural dimer in the avidin protein family. Biochem J 405: 397-405.

9. Meir A, Bayer EA, Livnah O (2012) Structural adaptation of a thermostable biotin-binding protein in a psychrophilic environment. J Biol Chem 287: 17951-17962.

10. Takakura Y, Tsunashima M, Suzuki J, Usami S, Kakuta Y et al. (2009) Tamavidins--novel avidin-like biotin-binding proteins from the Tamogitake mushroom. FEBS J 276: 1383-1397.

11. Määttä JA, Helppolainen SH, Hytönen VP, Johnson MS, Kulomaa MS et al. (2009) Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis. BMC Struct Biol 9: 63.

12. Taskinen B, Zmurko J, Ojanen M, Kukkurainen S, Parthiban M et al. Zebavidin- An avidin-like protein from zebrafish. PLoS One 8:e77207.

13. Stemmer WP (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Natl Acad Sci USA 91: 10747-10751.

14. Crameri A, Raillard SA, Bermudez E, Stemmer WP (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391: 288-291.

15. Niederhauser B, Siivonen J, Maatta JA, Janis J, Kulomaa MS et al. (2012) DNA family shuffling within the chicken avidin protein family - A shortcut to more powerful protein tools. J Biotechnol 157: 38-49.

16. Sidhu SS, Weiss GA, Wells JA (2000) High copy display of large proteins on phage for functional selections. J Mol Biol 296: 487-495.

17. Barbas III, CFBurton, DRJamie K. Scott, Jamie K.and Silverman, GJ , (2001) *Phage Display: A Laboratory Manual.*

18. Määttä JA, Eisenberg-Domovich Y, Nordlund HR, Hayouka R, Kulomaa MS et al. (2011) Chimeric avidin shows stability against harsh chemical conditions--biochemical analysis and 3D structure. Biotechnol Bioeng 108: 481-490.

19. Hytönen VP, Laitinen OH, Airenne TT, Kidron H, Meltola NJ et al. (2004) Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli. Biochem J 384: 385-390.

20. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. Journal of applied crystallography 26: 795-800.

21. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658-674.

22. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760-763.

23. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr 59: 1131-1137.

24. Hytönen VP, Määttä JA, Kidron H, Halling KK, Hörhä J et al. (2005) Avidin related protein 2 shows unique structural and functional features among the avidin protein family. BMC Biotechnol 5: 28.

25. Lamzin VS, Wilson KS (1993) Automated refinement of protein models. Acta Crystallogr D Biol Crystallogr 49: 129-147.

26. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol 6: 458-463.

27. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc 3: 1171-1179.

28. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-2132.

29. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240-255.

30. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35: W375-83.

31. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ et al. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58: 1948-1954.

32. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242.

33. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE et al. (2002) The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 58: 899-907.

34. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

35. Nicholas, K.B., Nicholas, H.B.J., and and Deerfield, D.W.I. (1997) GeneDoc: Analysis and Visualization of Genetic Variation. EMBNEW NEWS 4:14.

36. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Bjorklund D et al. (2004) BODIL: a molecular modeling environment for structure-function analysis and drug design. J Comput Aided Mol Des 18: 401-419.

37. Ortega A, Amoros D, Garcia de la Torre J (2011) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys J 101: 892-898.

38. Green NM (1975) Avidin. Adv Protein Chem 29: 85-133.

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

40. Reznik GO, Vajda S, Sano T, Cantor CR (1998) A streptavidin mutant with altered ligand-binding specificity. Proc Natl Acad Sci USA 95: 13525-13530.

41. Hytönen VP, Nyholm TK, Pentikäinen OT, Vaarno J, Porkka EJ et al. (2004) Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin. J Biol Chem 279: 9337-9343.

42. Hytönen VP, Määttä JA, Nyholm TK, Livnah O, Eisenberg-Domovich Y et al. (2005) Design and construction of highly stable, protease-resistant chimeric avidins. J Biol Chem 280: 10228-10233.

43. Eisenberg-Domovich Y, Hytönen VP, Wilchek M, Bayer EA, Kulomaa MS et al. (2005) High-resolution crystal structure of an avidin-related protein: insight into high-affinity biotin binding and protein stability. Acta Crystallogr D Biol Crystallogr 61: 528-538.

44. Riihimaki TA, Hiltunen S, Rangl M, Nordlund HR, Maatta JA et al. (2011) Modification of the loops in the ligand-binding site turns avidin into a steroid-binding protein. BMC Biotechnol 11: 64-6750-11-64.

Figure Legends

Figure 1. Sequence alignment of AVD, A/A2-1, A/A2-B and AVR2. A/A2-1 and A/A2-B show higher similarity to AVD [NP_990651.1] than to AVR2 [NP_001025519.1]. The two mutants differ from each other only at six amino acid positions. Amino acids originating from AVD and AVR2 are respectively indicated with yellow or green background. A point mutation in A/A2-1 is indicated with pink background. The secondary structure is based on AVD [PDB: 2AVI]; the β-strands are indicated by black arrows.

Figure 2. Release of the fluorescence-labeled biotin probe. The biotinylated fluorescence probe BF560 was released from the proteins at 25 °C and followed over one hour after addition of free biotin. Please note the immediate release of the fluorescence probe in case of AVR2 and A/A2-1, indicating high dissociation rate.

Figure 3. Hydrodynamic radius analysis. The analysis was carried out over a temperature range of 25-90 °C at three different pH-values, both in the absence and presence of biotin. The hydrodynamic radius was derived from the volume distribution. A. AVD, B. AVR2, C. A/A2-1, D. A/A2-B.

Figure 4. X-ray structure of A/A2-1. A. Cartoon representation of the homotetramer. Subunits I-IV are numbered according to Livnah et al. (1993) [39]: I, blue; II, green; III, light grey; IV, brown. B. Cartoon models of the superimposed A/A2-1 (blue) and chicken AVD (brown; [PDB:1AVD]) subunits I. C. Cartoon models of the superimposed A/A2-1 (blue) and AVR2 (brown; [PDB:1WBI]) subunits I. A-C. The bound biotin ligands are drawn as spheres; carbon atoms are shown in white, sulfur atoms in yellow, nitrogen atoms in blue and oxygen atoms in red. L1,2, the loop between β-strand 1 and 2; L2,3, the loop between β-strand 2 and 3, etc.

Figure 5. Unique features of A/A2-1. A. A weighted $2F_O$-F_C contour map (sigma level 1) showing electron density around biotin (BTN) and D12. B. Salt-bridge between D39 and R112 in the A/A2-1 structure. The side chains of T75, L97 and S99 in the close vicinity of the salt bridge are also shown

22

as stick models; oxygen atoms are shown in red and nitrogen atoms in blue. C. The salt bridge cannot form in the chicken AVD structure [PDB:1AVD] between residues A39 and R114 equivalent to residues D39 and R112 of A/A2-1. Residues T77, L99 and S101 equivalent to T75, L97 and S99 of A/A2-1 are shown. D. Salt-bridge between D39 and R112 in the AVR2 structure [PDB:1WBI]. Residues S75, Q97 and L99 equivalent to T75, L97 and S99 of A/A2-1 are shown. B-D. Cartoon models: subunit I, blue; subunit II green. Biotin (BTN) molecules are shown as spheres; coloring as in Figure 4.

Tables

Table 1. Avidin-biotin displacement assay. Absorbance measured at 404 nm from samples on differently treated microplate wells are shown.

mutant	BTN coating + BTN treatment	BTN coating	BSA coating	ratio between BTN-treated and untreated samples
A/A2-1	0.002	0.033	0.034	0.06
A/A2-B	0.156	0.198	0.000	0.79
AVD[a]	1.185	1.282	0.005	0.92

[a] 5 µg/ml of purified AVD protein from chicken egg white (Belovo) was used as a positive control.

Table 2. Thermal transition midpoints (T_m) at different pH-values as determined by DSC. Errors are derived from two independent measurements.

	T_m (°C)				ΔT_m (°C)
	- BTN		+ BTN		
AVD					
pH 3	51.2	±1.1	97.5	±0.9	46.4
pH 7	79.6	±0.4	121.0	±0.1	41.3
pH 11	67.9	±1.2	105.0	±6.2	36.0
AVR2					
pH 3	N.A.[a]		N.A.[a]		
pH 7	95.6	±0.1	115.3	±0.2	19.6
pH 11	79.8	±0.0	101.9	±0.0	22.2
A/A2-1					
pH 3	80.4	±0.6	103.5	±0.6	23.1
pH 7	90.5	±0.3	99.9	±0.2	8.5
pH 11	72.2	±3.2	83.8	±0.7	11.6
A/A2-B					
pH 3	89.3	±0.3	116.4	±0.1	27.1
pH 7	92.5	±0.1	119.3	±0.5	26.8
pH 11	85.3	±0.9	109.7	±1.9	24.4

[a]AVR2 precipitated completely upon against the pH 3 buffer

Table 3. Structure determination statistics for A/A2-1 [PDB: 4BCS].

Cell parameters

Space group	$P2_12_12$
Unit cell:	
a, b, c (Å)	81.1, 47.2, 60.6
α, β, γ (°)	90, 90, 90

Data collection[a]

Wavelength (Å)	0.93340
Beamline	ID14-1 (ESRF)
Detector	MarCCD
Resolution (Å)[b]	20-1.8 (1.9-1.8)
Unique observations[b]	22157 (3250)
I/sigma[b]	29.8 (4.7)
R_{factor} (%)[b]	4.3 (40.6)
Completeness[b]	99.7 (99.9)
Redundancy[b]	7.2 (7.2)

Refinement

R_{work} (%)[c]	18.9
R_{free} (%)[c]	22.6
Monomers (asymmetric unit)	2
Protein atoms	1893
Heterogen atoms	53
Solvent atoms	112
R.m.s.d:	
Bond lengths (Å)	0.014
Bond angles (°)	1.870

[a]The numbers in parenthesis refer to the highest resolution bin.

[b]From XDS [20].

[c]From Refmac 5 [29].

Figure 1
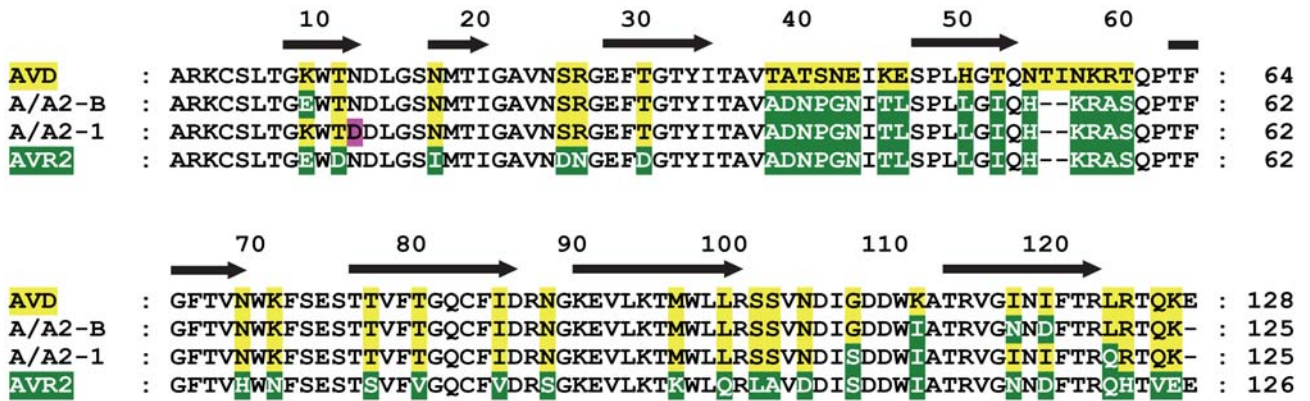
Figure 2
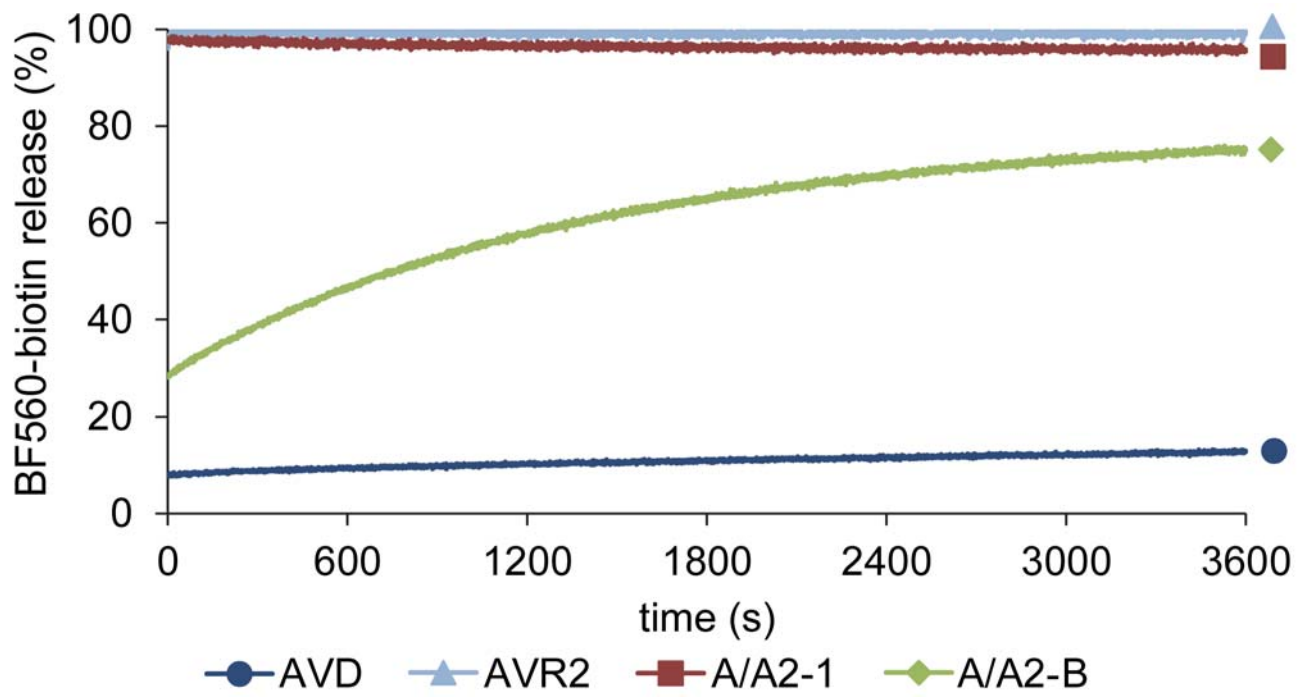
Figure 3
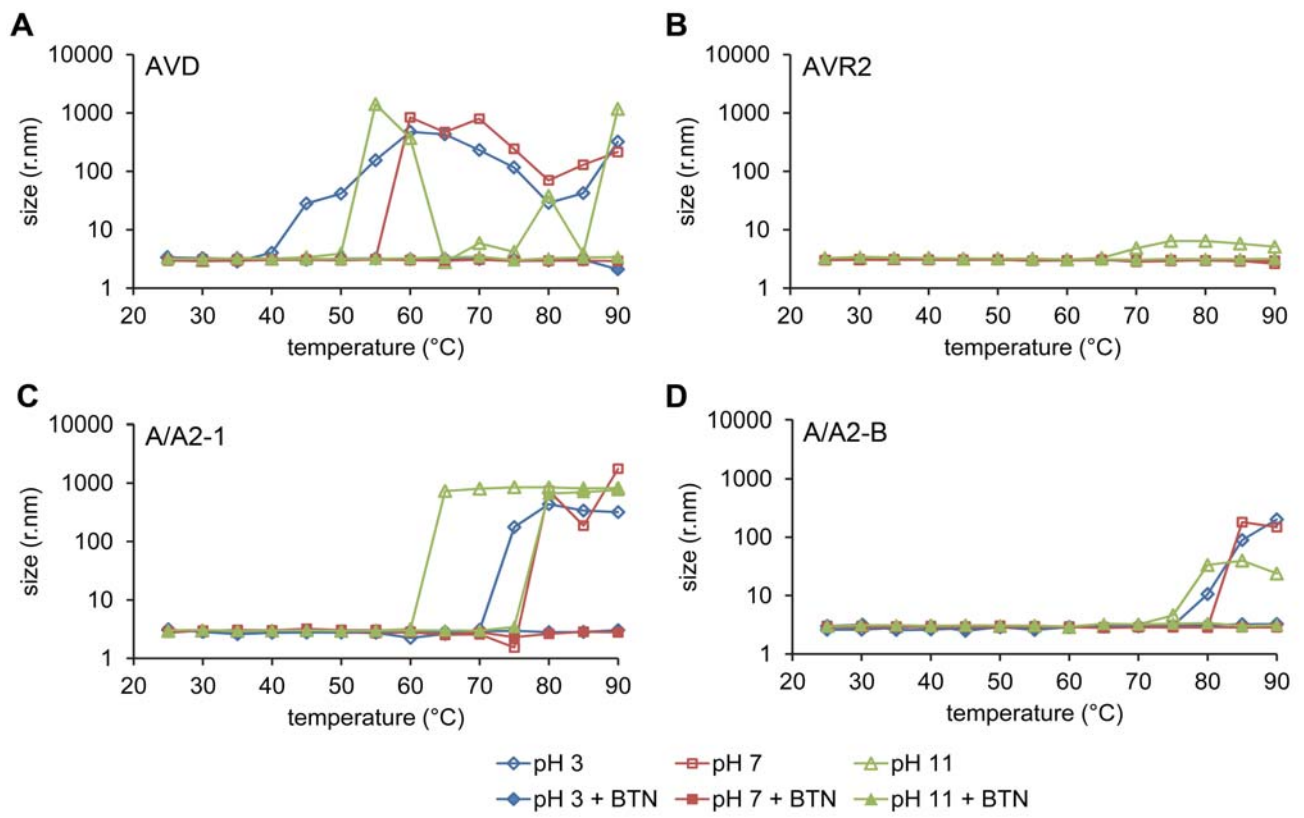
Figure 4
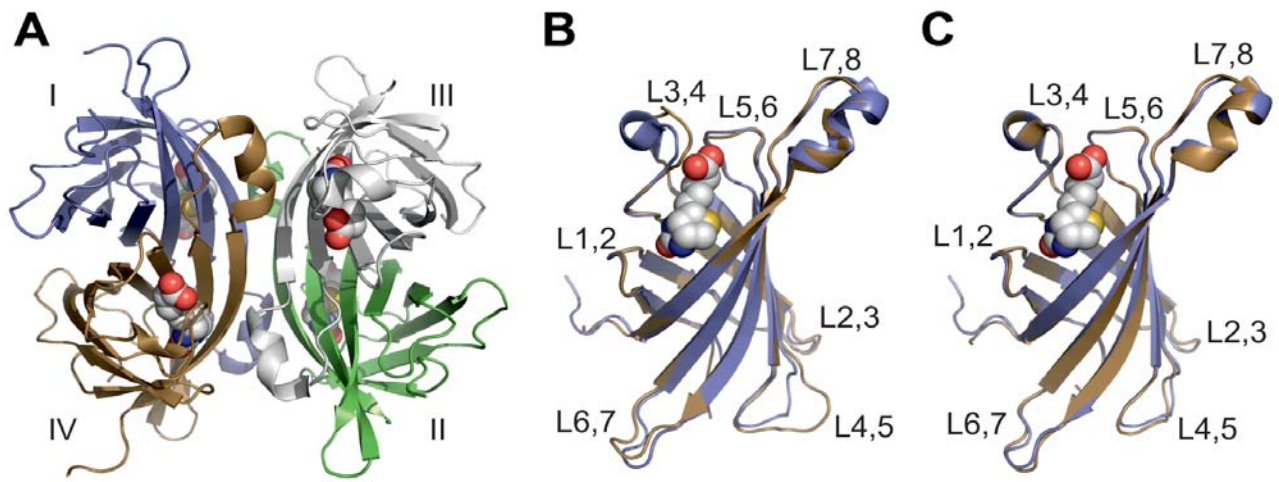
Figure 5
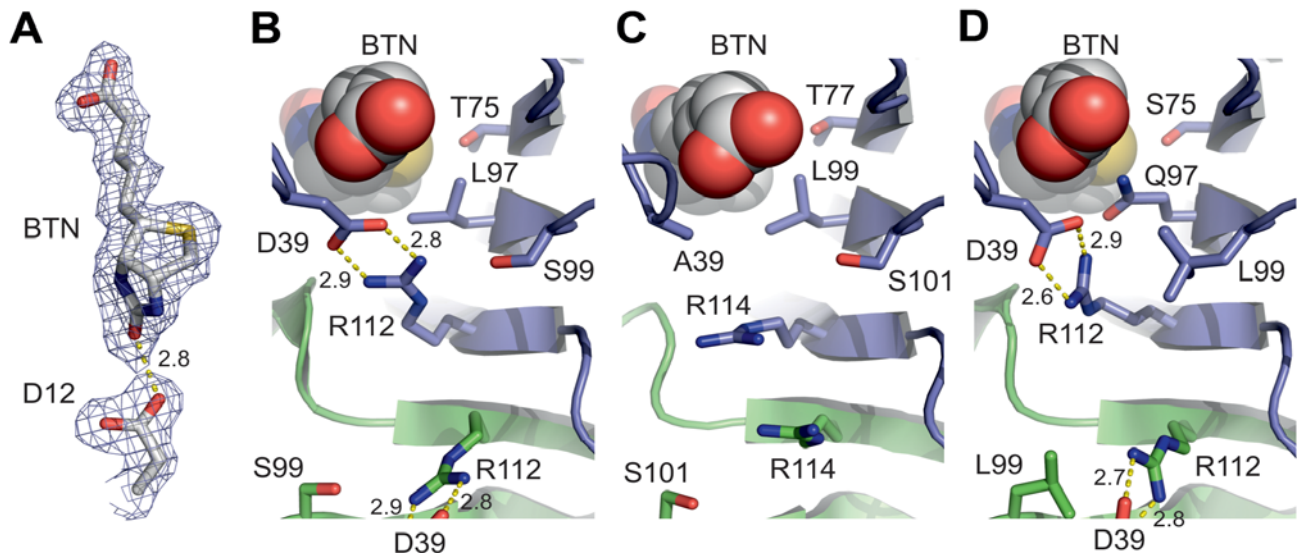
Supporting Information

Supporting Table S1. Molecular weight and hydrodynamic radius obtained by SEC-LS and DLS

analysis

protein sample	elution volume (ml) SEC-LS	molecular weight (kDa) SEC-LS	theoretical	hydrodynamic radius[a] (nm) SEC-DLS	(nm) DLS[b]
BSA	1.57	66.5	66.4	3.74	
AVR2	1.70	54.6	55.9	3.07	3.57
AVR2 BTN	1.69	56.8	56.9	3.11	3.50
AVD	1.62	50.2	57.4	3.14	3.52
AVD BTN	1.70	63.2	58.4	3.09	3.53
	2.70	73.1	58.4	N.A.	
A/A2-1	1.96	51.9	55.6	N.A.	3.39
A/A2-1 BTN	1.72	57.3	56.6	2.83	3.26
	2.61	57.2	56.6	N.A.	
A/A2-B	1.89	55.3	55.5	N.A.	3.38
	2.34	52.3	55.5	N.A.	
	2.78	50.9	55.5	N.A.	
A/A2-B BTN	1.75	56.7	56.4	2.90	3.25
	2.66	57.9	56.4	N.A.	

N.A.: DLS signal too low to obtain values

[a]The hydrodynamic radius for the tetramer of A/A2-1 was 3.145 nm, as calculated based on the PDB

coordinates [4BCS]

[b]determined with batch DLS with at protein concentration of 1 mg/ml

Supporting Figure Legends

Supporting Figure S1. Native-MS analysis of analyzed proteins. Native mass spectra of 10 µM protein in 10 mM ammonium acetate, pH 6.9. A. A/A2-1, B. A/A2-1 + BTN, C. A/A2-B, D. A/A2-B + BTN, E. AVD and F. AVD + BTN. Te corresponds to the protein tetramer and numbers denote different charge states.

Supporting Figure S2. SEC-LC analysis. Proteins were run in phosphate buffer containing 650 mM sodium chloride on a Superdex75 column at 12 °C. UV-VIS absorbance at 280 nm (UV), static light scattering (LS) and dynamic light scattering (DLS) of the eluting protein was recorded. The left Y-axis shows the scale of the UV and LS signal intensities. Molecular weight (MW) and hydrodynamic radius were calculated from the LS and DLS signal, respectively, using BSA for the calibration of the LS detector. The right Y-axis shows the scale for MW and hydrodynamic radius. A. A/A2-1, B. A/A2-1 in the presence of 3-fold molar excess of biotin, C. A/A2-B, D. A/A2-B in the presence of 3-fold molar excess of biotin.

Supporting Figure S3. Derived count rate as a function of temperature as determined by DLS. Analysis was carried out from 25 to 90 °C in three different pH-values in the absence and presence of biotin. A. AVD, B. AVR2, C. A/A2-1, D. A/A2-B.
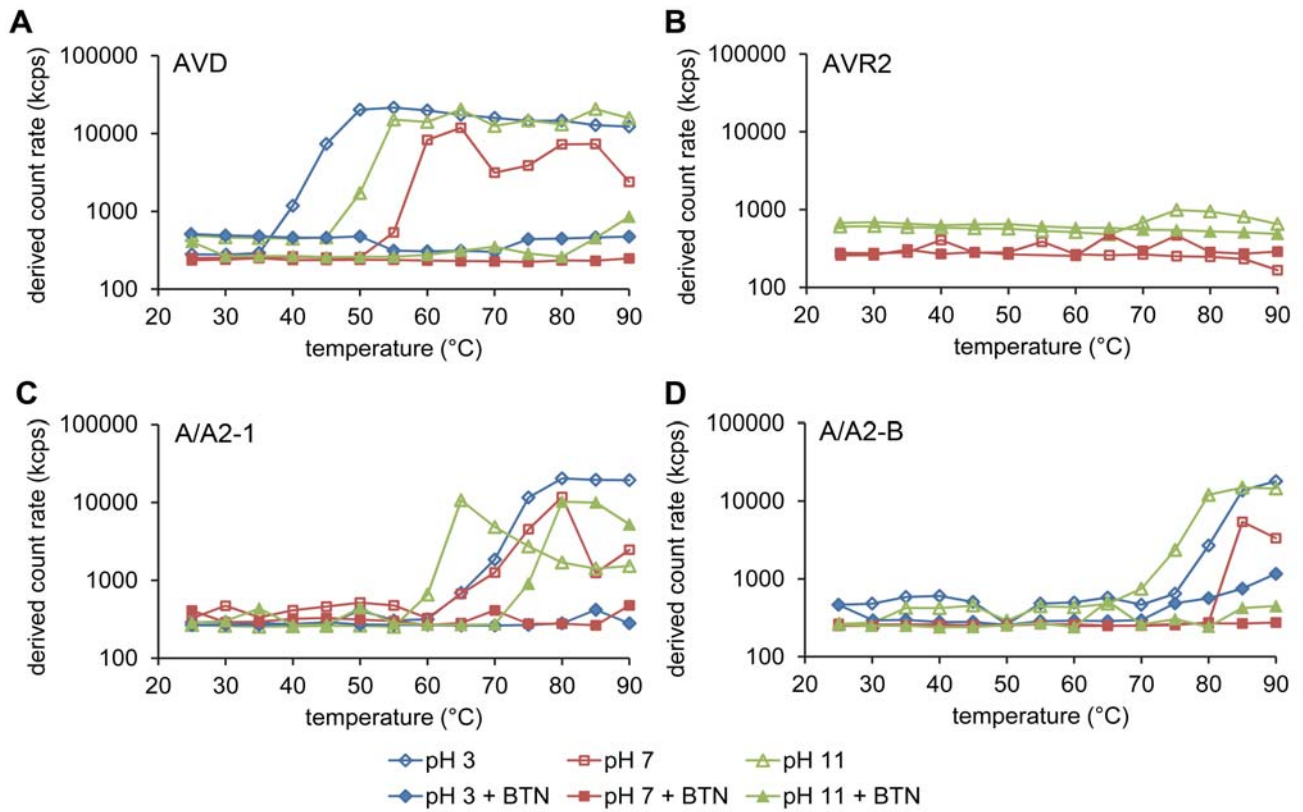
Supporting Figure S1
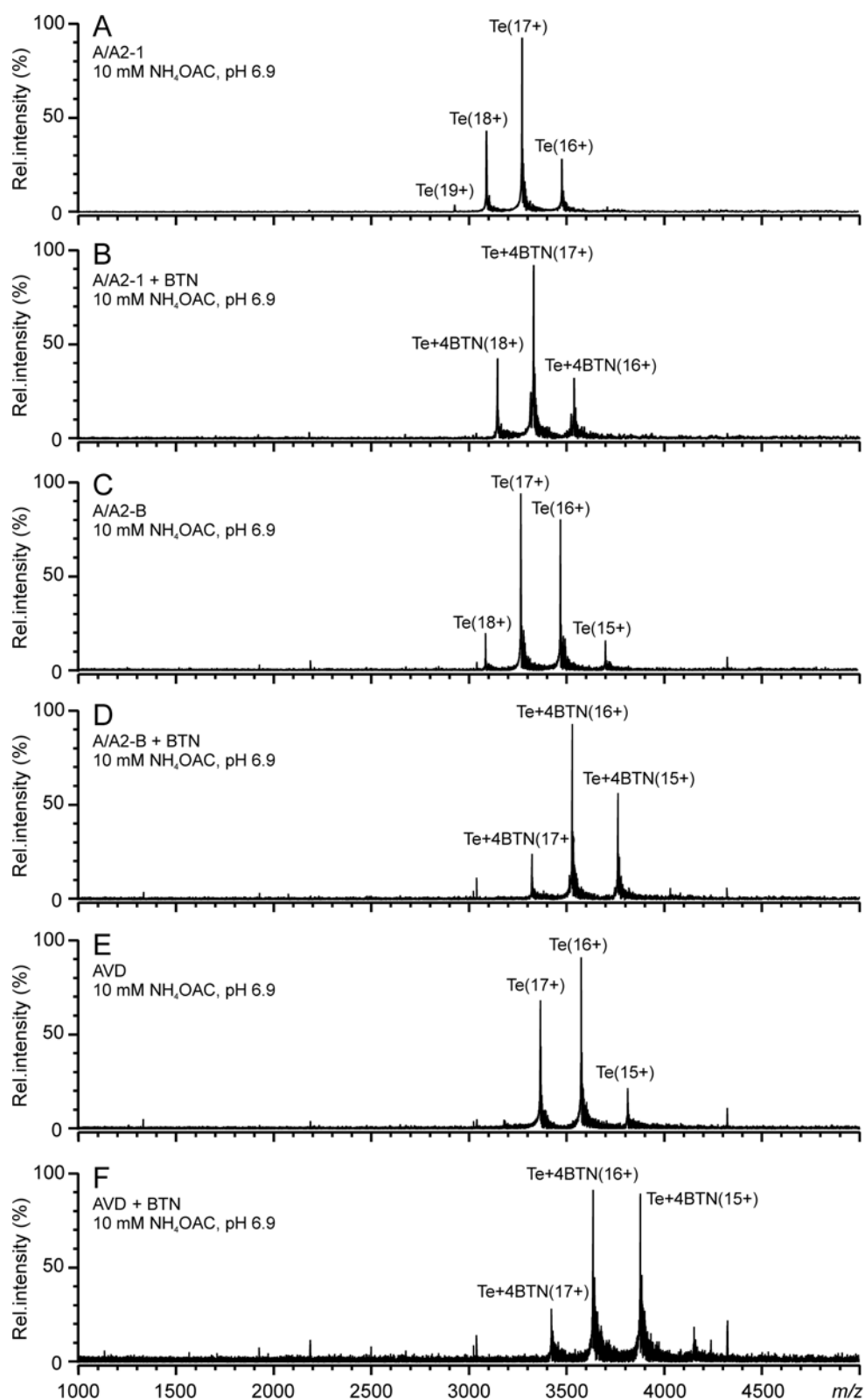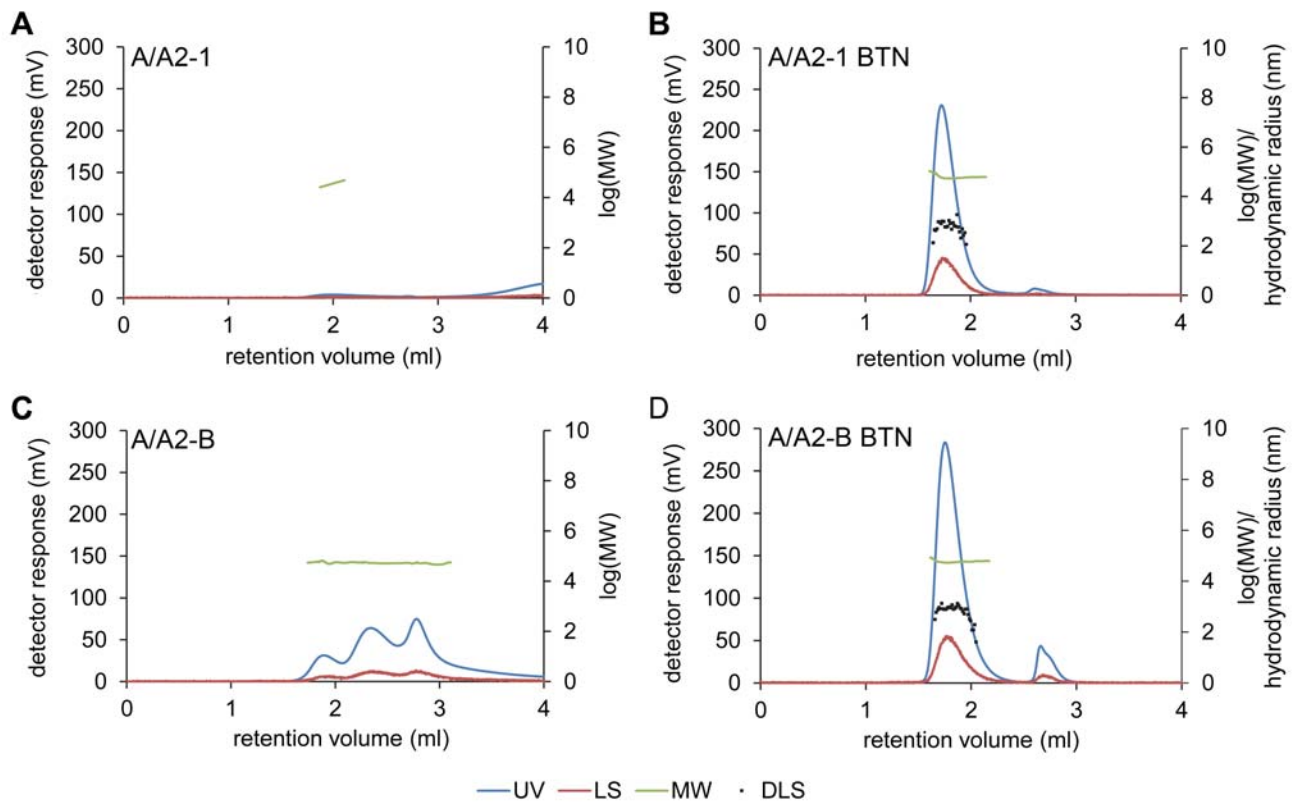
Supporting Figure S2

Supporting Figure S3

Bioconjugate Chemistry

Article

pubs.acs.org/bc

Reversible Biofunctionalization of Surfaces with a Switchable Mutant of Avidin

Philipp Pollheimer,[†] Barbara Taskinen,[‡,▲] Andreas Scherfler,[†] Sergey Gusenkov,[§] Marc Creus,[∥]
Philipp Wiesauer,[†] Dominik Zauner,[†] Wolfgang Schöfberger,[⊥,∇] Clemens Schwarzinger,[#] Andreas Ebner,[†]
Robert Tampé,[¶] Hanno Stutz,[§] Vesa P. Hytönen,[‡,▲] and Hermann J. Gruber*,[†]

[†]Institute of Biophysics, Johannes Kepler University, Gruberstr. 40, 4020 Linz, Austria

[‡]Institute of Biomedical Technology and BioMediTech, University of Tampere, Biokatu 6, FI-33014 Tampere, Finland

[▲]Fimlab Laboratories, Biokatu 4, 33520 Tampere, Finland

[§]Department of Molecular Biology, Division of Chemistry and Bioanalytics, University of Salzburg, Hellbrunner Str. 34, 5020 Salzburg, Austria

[∥]Department of Chemistry, University of Basel, Spitalstr. 51, CH-4056 Basel, Switzerland

[⊥]Institute of Inorganic Chemistry and [#]Institute of Chemical Technology of Organic Compounds, Johannes Kepler University, Altenberger Str. 69, 4040 Linz, Austria

[¶]Institute of Biochemistry, Biocenter, Johann Wolfgang Goethe-University, Max-von-Laue-Strasse 9, D-60438 Frankfurt a. M., Germany

Ⓢ *Supporting Information*

ABSTRACT: Label-free biosensors detect binding of prey molecules ("analytes") to immobile bait molecules on the sensing surface. Numerous methods are available for immobilization of bait molecules. A convenient option is binding of biotinylated bait molecules to streptavidin-functionalized surfaces, or to biotinylated surfaces via biotin—avidin—biotin bridges. The goal of this study was to find a rapid method for reversible immobilization of biotinylated bait molecules on biotinylated sensor chips. The task was to establish a biotin—avidin—biotin bridge which was easily cleaved when desired, yet perfectly stable under a wide range of measurement conditions.

The problem was solved with the avidin mutant M96H which contains extra histidine residues at the subunit—subunit interfaces. This mutant was bound to a mixed self-assembled monolayer (SAM) containing biotin residues on 20% of the oligo(ethylene glycol)-terminated SAM components. Various biotinylated bait molecules were bound on top of the immobilized avidin mutant. The biotin—avidin—biotin bridge was stable at pH ≥3, and it was insensitive to sodium dodecyl sulfate (SDS) at neutral pH. Only the combination of citric acid (2.5%, pH 2) and SDS (0.25%) caused instantaneous cleavage of the biotin—avidin—biotin bridge. As a consequence, the biotinylated bait molecules could be immobilized and removed as often as desired, the only limit being the time span for reproducible chip function when kept in buffer (2—3 weeks at 25 °C). As expected, the high isolectric pH (p*I*) of the avidin mutant caused nonspecific adsorption of proteins. This problem was solved by acetylation of avidin (to p*I* < 5), or by optimization of SAM formation and passivation with biotin-BSA and BSA.
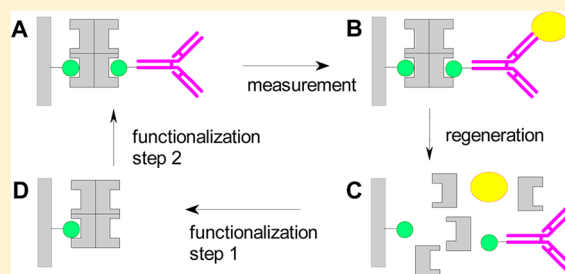
■ INTRODUCTION

Affinity biosensors are prepared by immobilization of specific bait molecules (nucleic acids, proteins, carbohydrates, etc.) on the surfaces of suitable transducers.[1] Label-free biosensors have the additional requirement that the sensor surface must be resistant to nonspecific adsorption.[1] This is achieved either by dense coverage of the sensor surface with bait molecules[2] or by coating of the sensor surface with a "protein-resistant" monolayer of short or long polymer chains[3−5] whereby the bait molecules are attached to the inert polymer chains.

Conventionally, the bait molecules are immobilized via covalent bonds, e.g., between the amino groups of proteins and aldehyde or carboxyl groups on the sensor surface. The reason

for the popularity of these methods is the high abundance of amino groups on most proteins (exemplified by 80−90 lysine residues per antibody molecule[6]). Covalent coupling of amines works well on flat chip surfaces which are densely covered with amino-reactive groups.[2,7] On protein-resistant surfaces, however, covalent coupling is retarded by steric repulsion between the protein-resistant polymer layer and the bait molecules. As a consequence, aldehydes are inapplicable on protein-resistant sensor surfaces[8] and activated carboxyls can only be used with

special precautions.[9,10] Similar problems as with proteins were encountered when trying to immobilize oligonucleotides with aminohexyl groups on chip surfaces with activated COOH groups.[11]

A widely used alternative to covalent coupling is immobilization of biotinylated bait molecules on preimmobilized (strept)-avidin as shown in Figure 1. Immobilization of (strept)avidin
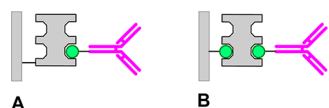

Figure 1. Two alternative strategies of chip functionalization with biotinylated bait molecules. (A) Direct attachment of (strept)avidin to the chip surface. (B) Formation of a biotin−avidin−biotin bridge.

can be achieved by covalent bonds to the chip surface (Figure 1A) or by attachment to surface-bound biotin groups (Figure 1B). The latter scheme is based on the fact that the four biotin-binding sites in (strept)avidin are grouped in two pairs on opposite sides of the streptavidin tetramer.[12] Consequently, (strept)avidin can be used as a mediator between the biotinylated sensor surface and biotinylated bait molecules.[13]

The advantages of biotinylated bait molecules are numerous: (i) Due to the robustness of streptavidin, sensor surfaces prefunctionalized with streptavidin are commercially available. (ii) Most bait molecules are commercially available with a biotin label or can be biotinylated on the time scale of an hour. (iii) No chemical skill is required for immobilization of a biotinylated bait molecule on a streptavidin surface. (iv) Due to the high affinity and fast association kinetics between streptavidin and biotin,[12,14] low concentrations (<1 μM) of bait molecules are sufficient for dense functionalization of streptavidin surfaces on a short time scale. (v) Streptavidin chips proved ideal for immobilization of single-stranded DNA (ssDNA) or double-stranded DNA (dsDNA) at the optimal density for hybridization or for interaction studies with DNA-binding proteins, respectively.[11,15] (vi) The streptavidin−biotin bond is insensitive to short pulses (30 s) of 10 mM HCl or NaOH, as frequently used to regenerate the biosensor after binding of prey molecules to the immobile bait molecules. (vii) The streptavidin−biotin bond is stable on the time scale of days at/near neutral pH; therefore, the chip can be reused for many measurements.

The high stability of the streptavidin−biotin bond has its downside also. At the end of a measurement series it would often be desirable to cleave the streptavidin−biotin bond for subsequent functionalization of the same chip with a new biotinylated bait molecule. The motivation for exchange of the biotinylated bait is as follows: (i) It eliminates human intervention for chip exchange and concomitant interruption of the work flow. (ii) It eliminates the deviation between different sensor chips, which is a problem of some product lines. (iii) It allows for extended use of valuable sensor chips which contain complex instrumentation or complex nanostructures.[16,17] (iv) The most urgent need for in situ exchange of bait molecules is encountered in those cases when the prey molecule cannot be removed from the biotinylated bait molecule without denaturing the latter. In such a case it is necessary to discard the whole sensor chip after each single measurement.

A viable solution to these problems has been found by inserting a DNA double helix between a dextran-coated chip

surface and streptavidin. For this purpose, both the chip surface and streptavidin have been derivatized by attachment of single-stranded DNA (Biotin CAPture Kit, GE Healthcare product no. 28920233, GE data file 28−9577−47 AA, available at https://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314787424814/litdoc28957747AA1_20110831132219.pdf). The two ssDNA molecules have a complementary sequence (undisclosed, possibly poly dT and poly dA)[18] which allows for stable immobilization of streptavidin by double helix formation. Subsequently, a biotinylated bait molecule is attached for a series of measurements with cognate prey molecules. The series is ended by injection of 100 mM HCl which dissociates the double helix whereupon new modified streptavidin molecules are immobilized for a new series of measurements. The drawbacks of this method are that it is not applicable to the analysis of DNA-binding proteins and that it is only available for the surface plasmon resonance (SPR) biosensors of one company.

An alternative strategy for reversible anchorage of streptavidin on surfaces relies on a biotinylated phospholipid monolayer (Figure 2A). Such monolayers are formed if a hydrophobic self-assembled monolayer (SAM) is incubated with lipid vesicles.[19,20] Lipid monolayers can be used to present specific lipid components for binding of bacterial toxins or blood-clotting factors.[21,22] After completion of a measurement
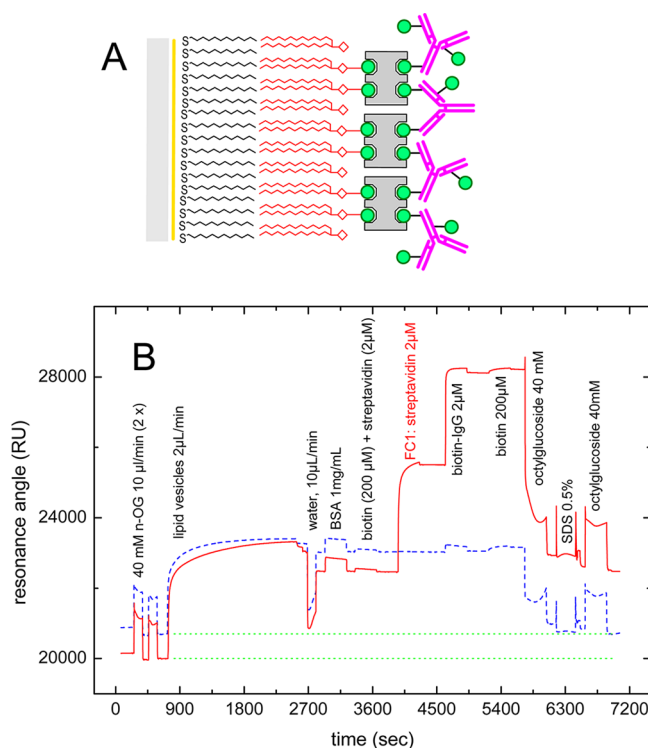

Figure 2. Test for reversible HPA chip functionalization with a phospholipid monolayer, streptavidin, and biotinylated IgG. In both flow cells, a lipid monolayer was formed by slow injection of sonicated lipid vesicles (2 mg/mL, 90% DOPC and 10% biotin-cap-DOPE). In FC1 (solid red line), streptavidin was bound to the biotin residues of the lipid. All subsequent injections were simultaneously applied to both flow cells: biotin-IgG (2 μM), free D-biotin (200 μM), octyl glucoside (40 mM), SDS (0.5%), and again octyl glucoside (40 mM). PBS 7.3 was used as running buffer and for preparation of the lipid vesicles. Unless stated otherwise, the flow was 20 μL/min.

series, the lipid monolayer is removed with 40 mM *n*-octylglucoside or 20 mM CHAPS (http://www.sprpages.nl/Index.php), thereby regenerating the hydrophobic SAM on which a new lipid monolayer can be formed.[23] Monolayers with 10−50% biotinylated lipid have been used for immobilization of streptavidin and biotinylated bait molecules.[24−26] One of these studies reports on repeated functionalization of the identical chip with different biotinylated peptides.[24] It was not clear, however, whether this method is also applicable to statistically biotinylated proteins; therefore, this important question was examined in the present study (see Results section).

Reversible immobilization of (strept)avidin on biotinylated surfaces (Figure 1B) can also be achieved by derivatives of biotin which exhibit a reduced (or pH-dependent) affinity for (strept)avidin. Three derivatives of biotin have been reported to bind (strept)avidin with lower affinity than biotin itself: iminobiotin,[27−29] desthiobiotin,[30−32] and N$_3$-ethylbiotin.[33] Iminobiotin binds well at pH 11,[27] whereas at neutral pH its rapid dissociation from (strept)avidin[34] prohibits its use for immobilization of bait molecules on sensor surfaces. Unstable binding at neutral pH is also seen with N$_3$-ethylbiotin.[33] The opposite problem is encountered with desthiobiotin. Due to its high affinity, complete dissociation of streptavidin from a desthiobiotinylated sensor chip takes ~1 day,[31] and imperfect chip regeneration is seen on a shorter time scale.[30] The same problems are expected if the bait molecules carry modified biotin residues. Moreover, a wealth of biotinylated bait molecules is commercially available, while none are available with iminobiotin, desthiobiotin, or N$_3$-ethylbiotin as labels. In conclusion, biotin derivatives are useful for affinity chromatography[27−30,32−36] but not for reversible biosensor functionalization.

The obvious alternative to modified biotin is the use of modified (strept)avidin with reduced or switchable affinity. In this respect, the design shown in Figure 1A has widely been used for affinity chromatography (or affinity capture on latex beads). The first version was "monomeric avidin".[37−39] The method has been improved by a double mutant of streptavidin which is also monomeric but exhibits less leaching of biotinylated ligands.[35] Reduced affinity for biotinylated target molecules was also achieved by nitration of a tyrosine residue in avidin ("captavidin") and this protein has largely been used for affinity chromatography.[40,41] Recently an attempt has been published in which captavidin was irreversibly attached to the sensor chip (as in Figure 1A) and the pH was switched between pH 7.4 and pH 10 to achieve specific binding to, or efficient release of biotinylated proteins from, the immobilized captavidin.[42] Unfortunately, the reproducibility of the on/off switch was insufficient for practical application in biosensing. As an alternative, streptavidin has been derivatized with "smart polymers" next to the entrance of each biotin-binding site,[43] thereby introducing a pH- and temperature-dependent switch. This would mean that all measurements have to be performed at 4 °C and the sensor must be heated to 37 °C in order to remove the biotinylated bait molecule from the sensor surface. Such temperature cycling, however, is incompatible with label-free biosensors which must be operated at a very constant temperature. Moreover, the switch is not responding with 100% efficiency. In conclusion, the use of modified (strept)avidin as depicted in Figure 1A is suitable for affinity chromatography but not for reversible biosensor functionalization.

In biosensing, the strategy in Figure 1B has a big advantage over that in Figure 1A if switchable variants of (strept)avidin are to be employed: In the case of the biotin−avidin−biotin bridge it is not a problem if the (strept)avidin molecules are irreversibly destroyed during "rigorous regeneration" of the chip surface because they are completely removed anyway (see Figure 3). With this in mind, we screened the literature for a
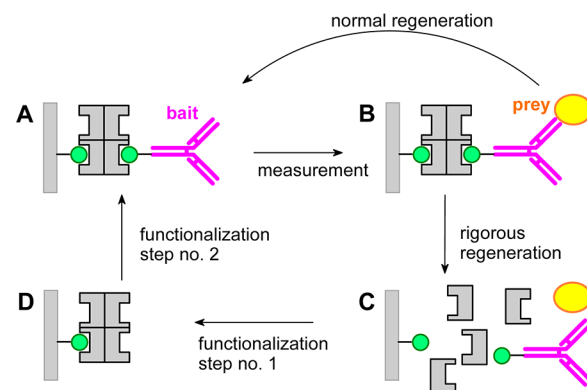

Figure 3. Reversible chip functionalization by denaturation of mutated avidin. (A) A biotinylated chip with bound avidin and biotinylated bait is ready for the measurement. (B) Prey molecules are bound during the measurement. If possible, the prey molecules are removed by "normal regeneration" after the measurement. (C) Rigorous regeneration of the sensing surface is achieved by denaturation of avidin. (D) A new cycle of chip functionalization starts with binding of mutated avidin.

good candidate among the published mutants of (strept)avidin. Proper pH dependence of protein stability and of biotin binding was the most interesting parameter. Among a list of pH-dependent avidin mutants,[44] the single mutant M96H of avidin looked most promising because it shows high stability at neutral pH, while a severe drop in pH is required to abolish biotin binding and to dissociate the native tetramer into nonfunctional monomers. As shown in the present study, avidin M96H is a valuable tool for implementation of reversible chip functionalization (Figure 3) if the nonspecific adsorption of proteins toward avidin is suppressed by proper precautions.

■ EXPERIMENTAL PROCEDURES

Materials. Captavidin was purchased from Invitrogen. Bovine serum albumin (BSA, fatty acid-free) was obtained from Roche Applied Science. BSA with covalently attached hexahistidine was prepared as described.[45] The avidin mutant M96H[44] was constructed for bacterial expression in *E. coli* by introducing mutation M96H to OmpA chicken avidin in pET101/D (Hytönen, V. P., Laitinen, O. H., Airenne, T. T., Kidron, H., Meltola, N. J., Porkka, E. J., Hörhä, J., Paldanius, T., Määttä, J. A., Nordlund, H. R., Johnson, M. S., Salminen, T. A., Airenne, K. J., Ylä-Herttuala, S., and Kulomaa, M. S. (2004) Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli. *Biochem. J. 384*, (Part 2), 385−390, Määttä, J. A., Eisenberg-Domovich, Y., Nordlund, H. R., Hayouka, R., Kulomaa, M. S., Livnah, O., and Hytönen, V. P. (2011) Chimeric avidin shows stability against harsh chemical conditions−biochemical analysis and 3D structure. *Biotechnol. Bioeng. 108*, 481−490), by QuikChange mutagenesis according to manufacturer's instructions (Stratagene, La Jolla, CA, USA). The protein was produced and purified as described

in (Määttä, J. A., Eisenberg-Domovich, Y., Nordlund, H. R., Hayouka, R., Kulomaa, M. S., Livnah, O., and Hytönen, V. P. (2011) Chimeric avidin shows stability against harsh chemical conditions—biochemical analysis and 3D structure. *Biotechnol. Bioeng. 108*, 481−490). Maltose-binding protein (MBP) with a C-terminal His$_6$ tag (MBP-His$_6$) was the kind gift of Jacob Piehler, University of Osnabrück, Germany). Anti-MBP antibody, avidin, biotin-protein G, D-biotin, lysozyme, immunoglobulin G (IgG) from goat, and streptavidin were purchased from Sigma-Aldrich. Sulfo-NHS acetate and sulfo-NHS-SS-biotin were obtained from Pierce. The IgG-binding peptide acetyl-HWRGWVC-NH$_2$ (>95% purity) was custom-synthesized by Peptide 2.0 (www.peptide2.com). The conjugate of this peptide and biotin, with a poly(ethylene glycol) linker (PEG) in between (termed HWRGWVC-PEG$_{18}$-biotin), as well as the components of the mixed self-assembled monolayer (SAM) shown in Figure 4, were synthesized as described in the
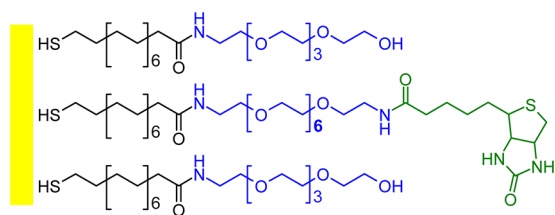
Figure 4. Composition of the mixed SAM used in this study. Two kinds of 16-mercaptohexadecanoic acid derivatives were mixed at the ratio 80/20 (mol/mol). The "matrix alkane thiol" (MAT) carried an OH-terminated penta(ethylene glycol) chain which provided for high protein resistance[51] and the "biotin-terminated alkane thiol" (BAT) contained a longer poly(ethylene glycol) chain ($n = 6$) with a terminal biotin, as required for optimal binding of streptavidin.[52]

Supporting Information. Biotin-tris-NTA was prepared as described.[46,47] Biotin-cap-NHS, biotin-IgG, and biotin-S-S-lysozyme were prepared as published.[48,49] 1,2-Dioleoylphosphatidylcholine (DOPC) and *N*-[6-(biotinamido)hexanoyl]-1,2-dioleoylphosphatidylethanolamine (biotin-cap-DOPE) were obtained from Avanti Polar Lipids. Sonicated vesicles with the desired molar ratio of these phospholipids (90/10) were prepared as described.[50]

Buffers. Phosphate-buffered saline (PBS 7.3) contained 140 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, and 1.8 mM KH$_2$PO$_4$. Buffer 11.0 was prepared from PBS 7.3 by first adjusting pH 10.0 with Na$_2$CO$_3$ (7.8 mg/mL, same refractory index as PBS 7.3) and then to pH 11.0 with NaOH (6.1 mg/mL, same refractory index as PBS 7.3). HEPES-buffered saline (HBS) contained 150 mM NaCl and 10 mM HEPES, adjusted to pH 7.4 with NaOH. Immediately before the corresponding BIAcore experiments, this buffer was supplemented with a final concentration of 10 μM NiSO$_4$ and sterile-filtered (0.22 μm) with strong aspirator suction to degas the buffer. The same procedure was applied to PBS 7.3 before BIAcore experiments, except that NiSO$_4$ was omitted.

Surface Plasmon Resonance Experiments. Unless stated otherwise in the figure legend, cleaning of bare glass chips, evaporation of chromium (3 nm) and gold (41 nm), as well as cleaning of the gold surface and coating with a mixed SAM of the alkanethiol derivatives shown in Figure 4 was performed as described,[51] using a total thiol concentration of 50 μM. However, the SAM components were different from those in the cited publication. Most SAMs were formed with

the OH-terminated matrix component and the biotin-terminated anchor component, the syntheses of which are described in the Supporting Information. SAMs of 1-octadecanethiol were formed in the same way, except that ethanol (analytical grade) was used in place of acetonitrile. This chip is equivalent to the commercial HPA chip (BIAcore). The chips were mounted on the chip holders with double-sided adhesive tape (nonpermanent) and inserted in a BIAcore X device for measurement of binding by SPR. Degassed buffer (PBS 7.3 or HBS) was constantly run over the chip surface, typically at 10−20 μL/min, as stated in the figure legends. Unless stated otherwise, all injection volumes were 100 μL.

■ RESULTS AND DISCUSSION

Reversible Immobilization of Streptavidin on Monolayers of Biotinylated Phospholipid. In a preceding study, the scheme in Figure 2A was successfully used for reversible immobilization of peptides which had a single biotin residue on their C-terminus.[24] Here the biotinylated peptide was replaced by IgG carrying ~5 biotin labels on average.[48] The results are shown in Figure 2B. First, the monolayer of gold-bound octadecyl chains was washed by two short injections of octyl glucoside. Second, a lipid monolayer was formed during slow injection of lipid vesicles containing 10% of biotinylated lipid. Third, water was applied to dissociate the adsorbed lipid vesicles. Fourth, BSA was injected to cover empty patches void of lipid, but evidently there were none. The transient rise of the resonance angle during the BSA injection was due to the high concentration of BSA in solution, but after the injection there was no net change of the resonance angle. Absence of binding was also seen when streptavidin (2 μM) was blocked with a large excess of biotin (200 μM) before the injection (FC2, blue dashed trace). In contrast, extensive binding was seen in flow cell 1 (FC1, red solid trace) when streptavidin was injected in the absence of biotin. Subsequently, biotinylated IgG was applied in both flow cells, leading to pronounced binding on top of streptavidin (FC1, red solid trace) and no binding on top of the phospholipid monolayer (FC2, blue dashed trace). The protein double layer in FC1 (as illustrated in Figure 2A) was completely resistant to free biotin (200 μM). Octyl glucoside was able to remove all of the lipid monolayer in FC2 (blue dashed trace) but only part of the material in FC1 (red solid trace). Further injections of SDS and octyl glucoside did not cause much improvement either.

The partial resistance of the protein double layer in FC1 to detergent was interpreted by formation of patches where biotinylated IgG was bound to several streptavidin molecules and the latter was simultaneously connected to two adjacent IgG molecules (as illustrated in Figure 2A). This model predicts that complete chip regeneration with detergent should be achieved if molecules with a single biotin group bind on top of streptavidin, as actually observed for biotinylated peptides.[24]

Switchable Binding of the Avidin Mutant M96H to Biotin Residues on a Sensor Chip. The above results indicate that reversible immobilization of (strept)avidin alone does not solve the problem of reversible chip functionalization with biotinylated bait molecules. At the same time it is necessary to prevent the formation of large aggregates of (strept)avidin and biotinylated bait molecules (see Figure 2), or to find an efficient way for the dispersion of such aggregates. As depicted in Figure 3, denaturation of (strept)avidin appears ideal for the latter strategy. Denaturation is accompanied by dissociation of (strept)avidin into its four subunits;[12] therefore,

clusters of (strept)avidin and biotinylated proteins (as in Figure 2) are dissociated into soluble proteins which can be washed away.

We first tested this strategy with wild-type streptavidin on a mixed SAM on gold (see Figure 4). The long 16-mercaptohexadecanoic acid linker and was chosen because it provides for long-term stability of the SAM, and the penta(ethylene glycol) chain on the matrix molecule was selected for minimization of nonspecific protein adsorption.[51] The difference in length between the OH-terminated matrix alkanethiol (MAT) and the biotin-terminated alkanethiol (BAT) was adjusted for optimal binding[52] and the molar ratio (80:20) for particularly slow dissociation of streptavidin.[52] We tried to find a denaturant which was compatible with the microfluidic flow cell of the BIAcore instrument and at the same time strong enough to remove streptavidin at 25 °C within several minutes. Unfortunately, no such reagent could be found.

The next step was a literature search for a derivative or a mutant of (strept)avidin with distinct switching behavior: Binding of biotin was desired to be strong over a wide range of measurement conditions (e.g., 4 < pH < 10). At the same time the protein was to be easily denatured with a buffer that was compatible with a microfluidic flow cell. No derivative/mutant of streptavidin was found that meets these criteria. In contrast, a number of pH-sensitive avidin mutants are known, in which hydrophobic amino acids at the subunit interfaces are replaced by histidine residues.[44] Protonation of these histidines at low pH causes subunit dissociation and denaturation.[44]

From a condensed overview of the literature data (see Table S1 in the Supporting Information) it was concluded that the avidin mutant M96H was the most promising candidate for successful implementation of the scheme in Figure 3. The first test of avidin M96H is shown in Figure 5A. After a short wash with SDS, the mixed SAM did not adsorb any BSA, indicating perfect protein resistance. The avidin mutant M96H, however, was instantaneously bound to the biotin-terminated SAM. At the end of the avidin injection, the resonance angle shows a sudden jump to a higher value. This was caused by the fact that the buffer salt concentration of the avidin sample was slightly lower than of the running buffer. The level of avidin binding (determined after the end of the avidin injection) was 2300 RU, which is similar to the "thickness" of a streptavidin monolayer.[7] The subsequent injection of citric acid (1%, pH 2.2) removed only a small fraction of bound avidin M96H. In contrast, a 1:1 mixture of 0.5% SDS and 5% citric acid (final concentrations 0.25% SDS and 2.5% citric acid, pH 2.0) caused almost instantaneous removal of avidin M96H from the biotinylated chip surface. This can be seen from the short spike at the beginning of the corresponding injection. The spike reflects two processes with opposite effects upon the resonance angle: rapid replacement of running buffer (PBS 7.3) by SDS/citric acid which has a higher bulk refractory index, and rapid dissociation of avidin M96H.

After the SDS/citric acid injection, the resonance angle was the same as before binding of avidin M96H (see green dotted line). The same observation was made in two further cycles of binding and subsequent desorption by SDS/citric acid. Interestingly, the application of SDS alone (0.25%) had a minor effect on chip-bound avidin M96H, as shown in the second cycle of avidin binding and desorption (Figure 5A).

In conclusion, a 1:1 mixture (v/v) of 0.5% SDS and 5% citric acid is the optimal reagent for rapid dissociation of avidin
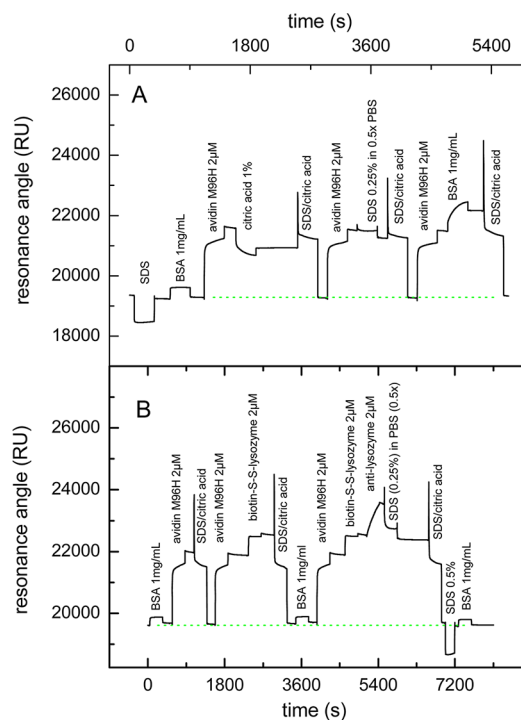

Figure 5. Test for reversible binding of the avidin mutant M96H to the mixed SAM shown in Figure 4. The running buffer was PBS 7.3 and the flow was 20 μL/min. 100 μL volumes of the samples noted in the above figure were injected, corresponding to 300 s injection times. (A) Search for the optimal reagent for desorption of avidin M96H. (B) Test for dissociation of a network consisting of avidin M96H, biotin-S-S-lysozyme, and antilysozyme, without cleavage of the disulfide bonds.

M96H from the biotin-terminated SAM shown in Figure 4. The logic next step was to examine whether this reagent can dissolve the networks between avidin and a biotinylated protein (Figure 3A and 3B) by dissociation of avidin M96H into four denatured subunits (Figure 3C). The corresponding data are shown in Figure 5B.

The first segment of Figure 5B was a mere repetition of the tests in panel A, i.e., only avidin M96H alone was bound to the chip and subsequently removed with SDS/citric acid. At the end of the SDS/citric acid injection, the baseline was the same as at the beginning, providing further support for complete removability of avidin M96H from the biotinylated SAM.

In the second segment of Figure 5B, biotinylated lysozyme was bound on top of avidin M96H. The small response was due to the small molecular mass of lysozyme (M_r = 14000) as compared to avidin M96H (M_r = 58000). Thus, the signal indicates full coverage of avidin M96H with biotinylated lysozyme. In spite of cross-linking of avidin M96H with statistically biotinylated lysozyme, the subsequent injection of SDS/citric acid could instantaneously remove all proteins from the chip surface.

In the third segment of Figure 5B, a trilayer of three cognate proteins was formed: avidin M96H, biotin-S-S-lysozyme, and anti-lysozyme. Subsequently, a 1:1 mixture of SDS (0.5%) and PBS 7.3 (final SDS concentration 0.25%) was injected which obviously removed the antibody only. In contrast, the mixture of SDS and citric acid caused complete regeneration of the chip surface. A tiny remnant was removed by an injection of pure SDS (0.5%) in water.

From Figure 5 it is concluded that avidin M96H fulfills important criteria for reversible functionalization of biotinylated chip surfaces: (i) This avidin mutant rapidly binds to the biotinylated SAM and remains firmly attached. (ii) Biotinylated proteins quickly bind on top of immobilized avidin M96H and remain firmly bound. (iii) The biotin−avidin−biotin bridge is completely resistant to SDS (0.5%) and surprisingly resistant to acid, even at pH 2.2. (iv) When desired, all biotin−avidin−biotin bridges can instantaneously be broken by the combination of SDS with citric acid, even in the case of a multiply biotinylated protein!

However, the third cycle in Figure 5A gives a first indication of the downside of avidin, as compared to streptavidin: Due to its high isoelectric point (pI ∼ 10−10.5)[12,53] it is prone to adsorb nucleic acids, as well as most proteins which are usually negatively charged at neutral pH. Here, BSA was adsorbed on top of immobilized avidin M96H. Fortunately, both proteins were completely removed by SDS/citric acid at the end of the third cycle in Figure 5A.

The problem of nonspecific adsorption was studied more closely in Figure 6A. Lysozyme has a high pI value (∼11.2);[54]

therefore, it has a net positive charge at neutral pH, just as avidin M96H. This explains why lysozyme did not adsorb on top of avidin M96H, in contrast to BSA and IgG which were significantly bound in the subsequent injections (Figure 6A, FC2, blue dashed line). Unfortunately, the problem of nonspecific protein adsorption could not be solved by "saturation of the adsorptive sites" with BSA, as is common practice in many solid phase assays such as ELISA:[55] Two successive injections of BSA and the subsequent injection of IgG caused a progressive rise of the resonance angle in each step (Figure 6A, FC2, blue dashed line). A similar overall rise of the resonance angle was also obtained when injecting IgG alone in flow cell 1 (Figure 6A, FC1, red solid line).

Taken together, the data in Figures 5 and 6A demonstrate that the avidin mutant M96H has the desired switching behavior, being tightly bound to a biotinylated sensor chip at neutral or weakly acidic pH, while being easily removed at pH 2.0 in combination with detergent. At the same time, the positive surface potential of avidin M96H at neutral pH results in pronounced nonspecific adsorption of most proteins because the majority of proteins are negatively charged at neutral pH.

Reversible Sensor Chip Functionalization with Acetylated Avidin M96H. Since the undesired nonspecific adsorption of proteins on top of immobilized avidin M96H is caused by the excess of positive charges, the obvious solution to this problem is modification of the surface charges, either by suitable chemical reagents or by genetic engineering. In principle, genetic engineering is to be preferred because it results in a homogeneous sample in which all avidin molecules have exactly the same modifications. The alternative was modification of lysine residues with acetyl groups[56] or with succinic anhydride.[57] The disadvantage of chemical modification is its statistical mechanism, which yields a range of pI values. Its advantage over genetic engineering is simplicity and a shorter time scale. We therefore decided to perform a preliminary study with chemical modification in order to find out whether genetic engineering is worth the effort. As shown below, the answer was clearly yes: the problem of nonspecific adsorption was largely reduced by chemical modification of avidin M96H while retaining its favorable switching behavior. Even better results are to be expected from genetic engineering which has the advantage of producing a homogeneous population of modified avidin molecules.

First, it was necessary to determine the optimal degree of acetylation that caused suppression of nonspecific adsorption, without impeding specific binding of biotin. Given the close similarity between wild-type avidin and avidin M96H, the test series with different degrees of acetylation was performed with the wild-type protein. As described in the Supporting Information, suppression of nonspecific adsorption was maximal when avidin was treated with a 40-fold excess of sulfo-NHS acetate under carefully controlled conditions. The same procedure was then applied to avidin M96H, yielding so-called "avidin M96H-40" which had a narrow pI distribution between 4.5 and 4.8 (see Supporting Information). After acetylation, avidin M96H-40 was tested for specific binding to biotin and for nonspecific adsorption of proteins (Figures 7−9).

The most important finding in Figure 7 is that the avidin mutant M96H-40 has fully retained its switching behavior, in spite of acetylation. After some initial tests of protein adsorption (BSA, IgG) to the bare SAM, two cycles of specific binding and subsequent desorption were performed with avidin
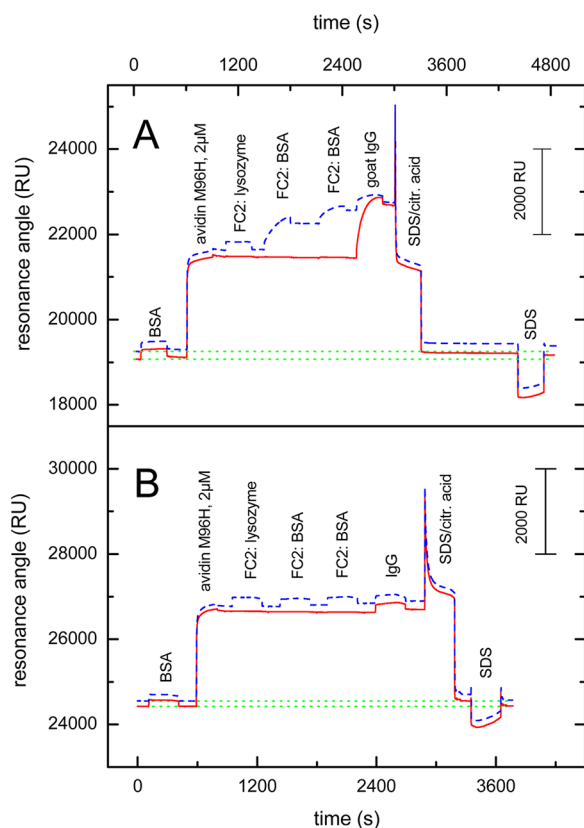

Figure 6. Test for nonspecific adsorption of lysozyme, BSA, and goat IgG (blue dashed trace, FC2) on top of the monolayer of avidin M96H. In FC1 (red solid trace), only goat IgG was applied to the layer of avidin M96H. Protein concentrations were 1 mg/mL, except for avidin M96H (2 μM). All other conditions were as in Figure 5. In panel A, the mixed SAM shown in Figure 2 had been prepared by the usual procedure.[8] In panel B, the SAM components had been mixed and subsequently reduced with zinc/acetic acid to ensure a molar ratio of 8/2 (MAT/BAT) not only in solution but also on the gold surface (for details, see text and Supporting Information). Solid traces and red color refer to flow cell 1 (FC1), dashed traces and blue color to flow cell 2 (FC2).
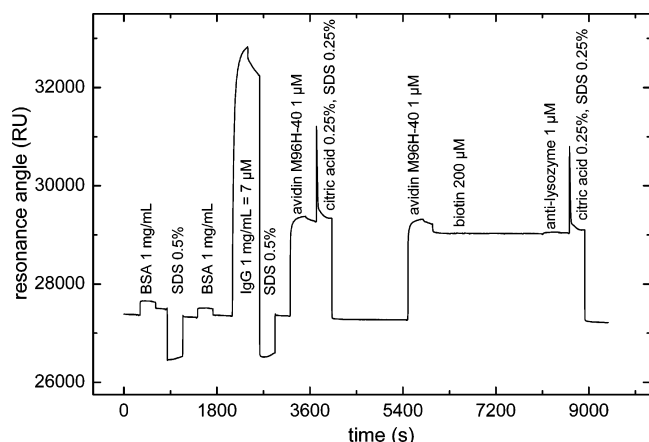
Figure 7. Reversible binding of acetylated avidin M96H-40 to the mixed SAM shown in Figure 4 and suppression of nonspecific adsorption by acetylation. The operating conditions were the same as in Figure 5
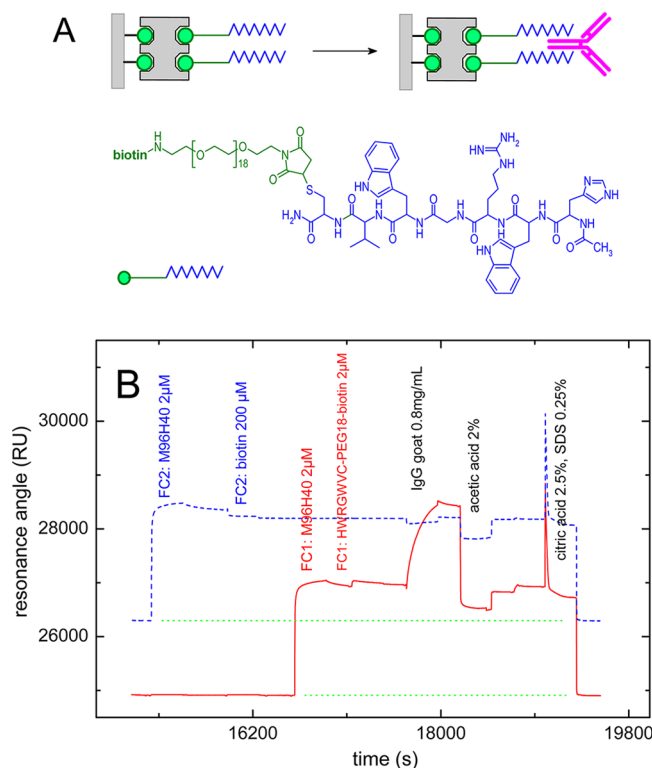

Figure 8. Functionalization of chip-bound avidin M96H-40 with an IgG-binding heptapeptide. The peptide HWRGWVC had been linked to biotin via a PEG$_{18}$ chain with 18 ethylene glycol units (for details see Supporting Information). The operating conditions were the same as in Figure 5. Solid traces and red color refer to flow cell 1 (FC1), dashed traces and blue color to flow cell 2 (FC2).

Figure 9. Application of acetylated avidin M96H-40 for immobilization of biotin-tris-NTA, thereby creating a chip surface with high affinity for His$_6$-tagged bait molecules. The latter were repeatedly bound to immobilized biotin-tris-NTA in the presence of 200 μM Ni^{2+} and released by 500 mM imidazole. The operating conditions were the same as in Figure 5, except that HBS (supplemented with 10 μM Ni^{2+}) was used as running buffer in place of PBS 7.3. Solid traces and red color refer to flow cell 1 (FC1), dashed traces and blue color to flow cell 2 (FC2).

while before the addition of free D-biotin, some dissociation was visible from the SPR trace.

Biotin-induced dissociation from the mixed SAM was not unique for acetylated avidin M96H-40. It was also seen with nonacetylated avidin M96H (6 ± 2%, data not shown). The most probable explanation for the distinction between stable and unstable binding is that firmly bound avidin molecules simultaneously bind to two adjacent biotin groups on the chip surface, while the small fraction of loosely bound avidin "sits" on a single surface-bound biotin (as stated in ref 52). In the case of acetylated avidin M96H-40, the loosely bound fraction shows spontaneous dissociation even in the absence of biotin and it is instantaneously released upon addition of free biotin. In the case of nonacetylated avidin M96H, the "loosely bound fraction" shows no spontaneous dissociation and can only be released by addition of free biotin.

Before the last injection of SDS/citric acid in Figure 7, an antibody (antifluorescein antibody, 1 μM) was injected. Fortunately, no nonspecific adsorption was seen at this concentration which is typical for biosensing experiments. This finding is particularly significant because in the same figure the bare SAM had shown a high level of nonspecific adsorption of IgG (at $t \sim 2400$ s).

Figures 8 and 9 exemplify the reversible immobilization of biotinylated bait molecules on top of acetylated avidin M96H-40. In Figure 8, the hexapeptide HWRGWV was applied for

M96H-40. In both cycles, a similar level of binding was achieved. Moreover, injection of SDS/citric acid always resulted in perfect regeneration of the chip surface by complete removal of bound avidin M96H-40.

In the second cycle, free D-biotin was applied on top of avidin M96H-40, with two effects: (i) About 10% of surface-bound avidin M96H-40 were dissociated. (ii) The remaining 90% of avidin M96H-40 were perfectly retained on the chip surface,

reversible capture of IgG molecules on a sensor surface. Among all cysteine-free hexapeptides, this peptide sequence exhibits the highest affinity for IgG molecules from various mammalian species.[58,59] The binding site for the peptide is known to be located in the Fc domain of IgG molecules.[58] Here, we added an extra cysteine (HWRGWVC) for subsequent linking to maleimide-PEG$_{18}$-biotin (see Supporting Information). The resulting conjugate HWRGWVC-PEG$_{18}$-biotin was bound on top of a monolayer of avidin M96H-40 (Figure 8, FC1, red solid trace). It was not possible to actually see the binding of the peptide-PEG-biotin conjugate in the SPR trace, for two reasons: (i) The small molecular mass of this molecule leads to a small SPR signal. (ii) As explained above, a small fraction of avidin M96H-40 dissociates when challenged with biotin (or a small, flexible derivative of biotin). These processes have opposite effects on the SPR trace. Nevertheless, it is obvious that peptide-PEG-biotin was bound because the subsequent injection of IgG (1 mg/mL) resulted in pronounced binding of IgG (Figure 8, FC1, red solid trace), followed by slow dissociation after the end of injection. No binding of IgG was seen in the control cell (Figure 8, FC2, blue dashed trace) where the monolayer of avidin M96H-40 had been blocked with D-biotin before the injection of IgG.

Interestingly, all bound IgG could be dissociated from the immobilized peptide with 2% acetic acid (pH 2.6), without much effect on the avidin M96H-40 monolayer (see solid and dashed trace in Figure 8). The latter could only be removed by a solution containing 2.5% citric acid (pH 2.0) and 0.25% SDS. These findings emphasize the robustness of immobilized M96H-40 over a wide range of measurement conditions which nicely contrasts the instantaneous response to SDS/citric acid when its removal is actually desired.

In Figure 9, the biotinylated bait molecule was biotin-tris-NTA. It is known to function as a general adaptor for His$_6$-tagged proteins.[46,47] When occupied with Ni^{2+} ions, each of the three NTA groups in biotin-tris-NTA binds two histidine residues, with the effect that a His$_6$ tag is tightly bound on the time scale of hours.[46,60] Here, we combined biotin-tris-NTA with acetylated avidin M96H-40, resulting in a new chip surface which holds significant potential for practical application (see sketch on top of Figure 9). After immobilization of avidin M96H-40 on the biotinylated SAM, biotin-tris-NTA was bound on top of avidin (red solid trace in Figure 9). The next injection was a test for nonspecific adsorption. In spite of the high concentration of BSA (1 mg/mL), nonspecific adsorption was absent, both in the measuring cell with the immobilized biotin-tris-NTA (red solid line) and in the control cell with biotin-blocked avidin M96H-40 (blue dashed line).

Then, many cycles were performed in which a His$_6$-tagged protein was specifically bound to biotin-tris-NTA in the presence of 200 μM Ni^{2+} and subsequently released with 500 mM imidazole. As can be seen from the data, imidazole caused perfect release of the His$_6$-tagged proteins, without affecting the sublayer of avidin M96H-40. In the fourth cycle, anti-MBP antibody was bound on top of MBP-His$_6$, thereby demonstrating the practical purpose of this design: The His$_6$-tagged protein serves as bait molecule and the interaction with a complementary molecule is to be monitored by SPR. While the antibody showed intense binding in the measuring cell (red solid line), no binding was seen in the control cell (blue dashed line). The only imperfection in this fourth cycle was that imidazole could not fully release the antigen–antibody complex from the chip surface, in spite of two successive injections. The

problem was solved by injecting a 1:1 mixture of the running buffer with 0.5% SDS. The resulting baseline was identical with the baseline before the first injection of a His$_6$-tagged protein. Moreover, the remaining layer of avidin M96H-40 was still able to bind His$_6$-tagged maltose-binding protein in the next cycle. These observations in Figure 9 confirm that SDS alone does not harm avidin M96H-40—just like acetic acid alone did not hurt avidin M96H-40 in Figure 8.

Suppression of Nonspecific Protein Adsorption without Acetylation. Acetylation of avidin M96H-40 allowed us to prove that nonspecific adsorption of proteins can be eliminated by lowering of the pI of the avidin mutant. The logic next step was to combine the mutation M96H with other mutations of avidin which lower the pI and abolish nonspecific adsorption.[61] This study is under way in our laboratories.

Fortuitously we discovered a simple alternative for complete suppression of nonspecific adsorption which works well with unmodified avidin M96H. The trick was to improve the quality of the biotin-terminated SAM. So far, the mixed SAMs were formed by combining stock solutions of the individual SAM components (1 mM in ethanol) at the desired molar ratio before diluting with acetonitrile and starting the incubation with the cleaned gold surface,[8,51] following an earlier example.[62]

In the course of this study we remembered an important report about the kinetics of mixed SAM formation.[63] It was shown that, in mixtures of thiols and disulfides, the rate of chemisorption to the gold surface is >10 times faster for the thiol form than for the corresponding disulfide (even when the disulfide concentration was counted twice). This made us worry about the oxidation state of MAT and BAT (Figure 4) in the 1 mM stock solutions. If disulfide formation of MAT was more extensive than of BAT, then BAT would be greatly overrepresented in the SAM (and vice versa). Deviation of the MAT/BAT ratio in the SAM from the optimal value of 80/20 was expected to have very adverse effects on binding of (strept)avidin.[52]

In order to ensure the identical oxidation state of MAT and BAT, we mixed solid MAT and BAT at the desired molar ratio and treated the mixture in THF with zinc/acetic acid which is known to ensure complete reduction of dialkyldisulfides.[64] The resulting thiol mixture was completely reoxidized to the disulfidic state, containing a statistical mixture of disulfidic MAT−MAT, MAT−BAT, and BAT−BAT adducts (see Supporting Information). The mixture was dissolved in chloroform, dried down in aliquots, and stored under argon at −25 °C where no further chemical changes are to be expected. One 0.44 mg aliquot was used to prepare the mixed SAMs on five gold chips with a simplified procedure (see Supporting Information).

The big surprise with these newly prepared SAMs was a greatly reduced level of nonspecific protein adsorption. The dramatic effect of the SAM preparation method is exemplified in Figure 6 where panel A shows pronounced nonspecific adsorption for the old SAM type and panel B near absence of protein adsorption on the new SAM type. The beneficial effect of the new SAM preparation method was consistently seen on all (seven) chips tested from two different batches of chip formation, in many cycles of functionalization with avidin M96H and subsequent removal with SDS/citric acid. In a series of experiments, the extent of BSA adsorption (from 1 mg/mL BSA) was 26 ± 23 RU and the extent of IgG adsorption (from 2 μM IgG, applied after BSA) was only 12 ± 4 RU.

Based on the above findings, an optimized protocol for the differential functionalization of sample cell and control cell was established, as depicted in Figure 10A. Biotin−BSA was used to
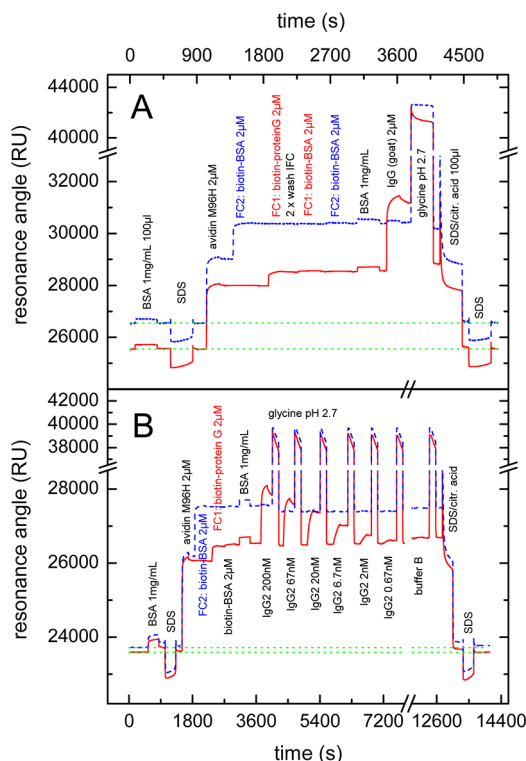

Figure 10. Selective binding of IgG to immobilized biotin−protein G after inertization of avidin M96H with biotin−BSA and free BSA. (A) Protocol for the preparation of an inert control cell (FC2, blue dashed trace) and a sample cell with high specificity for goat IgG (FC1, red solid trace). (B) Reversible association and dissociation of human IgG2κ on immobilized biotin−protein G, using 50 mM glycine (pH 2.7) for selective removal of IgG from avidin-bound biotin−protein G. The different concentrations of IgG2κ were prepared by serial dilution of a 7 μM stock solution of IgG2κ (in PBS 7.3) with buffer B (1 μM BSA in PBS 7.3). In all experiments of this figure, the SAM had been prepared by the improved procedure which is described in the Supporting Information. Solid traces and red color refer to flow cell 1 (FC1), dashed traces and blue color to flow cell 2 (FC2).

block the control cell (FC2, blue dashed line). Subsequently, biotin−protein G was immobilized in the sample cell (FC1, red solid line). The integrated flow cell (IFC) was washed (2×) and both flow cells were treated with biotin−BSA and also with BSA to further minimize nonspecific adsorption of proteins. Finally, free IgG (2 μM) was injected, resulting in extensive specific binding (2600 RU) on top of immobilized protein G (FC1, red solid line), while very little IgG (30 RU) was bound in the control cell (FC2, blue dashed line).

The time course of IgG binding and dissociation in Figure 10A looked inviting for kinetic analysis by the BIAcore evaluation software. This IgG, however, was a mixture of all IgG subtypes from goat, thus no meaningful data were to be gained from such analysis. In Figure 10B, the heterogeneity problem was solved by injecting different concentrations of human IgG2κ, always removing bound IgG2κ after a short dissociation period with 50 mM glycine (pH 2.7). IgG2κ was diluted with 1 μM BSA (in the same buffer) in order to prevent

loss of IgG2κ in dilute solutions by denaturation or adsorption at the solid/liquid or air/liquid interface.

The association/dissociation periods at different concentrations of IgG2κ were evaluated by the double referencing method,[65] as shown in Figure 11. The Langmuir model (1:1
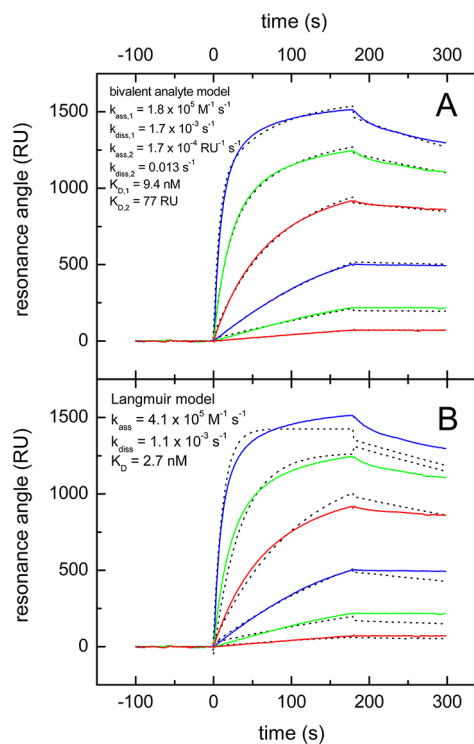

Figure 11. Kinetic analysis of reversible binding of human IgG2κ to immobilized biotin−protein G. The data in Figure 10B were analyzed according the double referencing method[65] using the program BIAevaluation 3.2 RC1. The results of global fits are shown for the ″bivalent analyte model″ (A) and for the conventional Langmuir model (B). The measured traces are shown in color (solid lines), the calculated traces in black (dotted lines).

binding between immobile protein G and IgG2κ) was clearly inadequate (Figure 11B) while the ″bivalent analyte model″ proved to be well suited to describe both association and dissociation. Given the C_2 symmetric nature of IgG, it makes sense that the two identical heavy chains in the Fc portion are bound by two adjacent biotin−protein G molecules on the avidin surface.

Figures 10 and 11 show that avidin M96H can well be used for sophisticated biosensing, in spite of its high pI value, if proper precautions are taken. One important aspect is the new method of mixed SAM preparation and the other is passivation with biotin−BSA and BSA.

Examination of Captavidin as Potential Alternative to Avidin M96H. Many years ago we tested whether switchable biotin−avidin−biotin bridges can be formed on biotin-terminated SAM with the help of captavidin. We saw pronounced leaching of captavidin from the biotinylated surface. At that time, however, it was not clear whether this was due to captavidin or to insufficient SAM quality. Having optimized the preparation of biotin-terminated SAMs, as described above, we now repeated the test with newly purchased captavidin. The data (in Figure 12) show that captavidin cannot compete with avidin M96H in case of biotin−avidin−biotin bridges (Figure 1B): (i) The layer
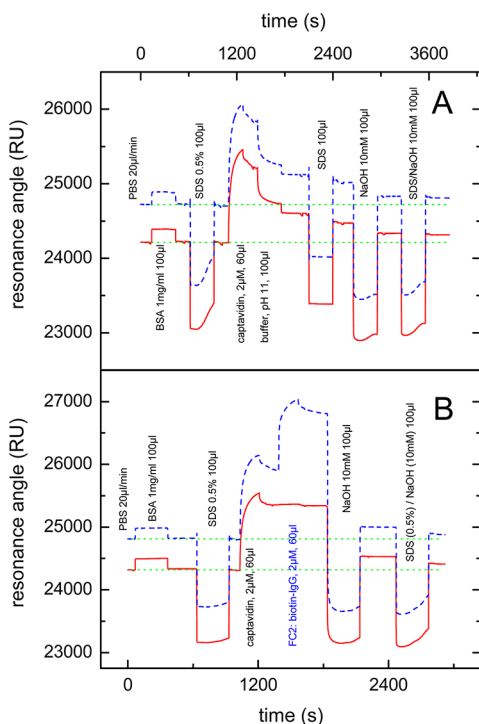
Figure 12. Test for reversible binding of captavidin the mixed SAM shown in Figure 4. (A) Binding of captavidin (2 μM) and subsequent desorption by buffer 11.0 (see Experimental Procedures). (B) Successive binding of captavidin (2 μM) and biotin-IgG (2 μM), followed by an attempt of complete chip regeneration with NaOH and NaOH/SDS. The SAM had been prepared by the improved procedure which is described in the Supporting Information. The running buffer was PBS 7.3 and the flow was 20 μL/min. Solid traces and red color refer to flow cell 1 (FC1), dashed traces and blue color to flow cell 2 (FC2).

thickness of captavidin is a little over 1000 RU, as compared to >2000 RU in the case of avidin M96H. (ii) Bound captavidin shows rapid dissociation from the biotin-terminated SAM. Progressive dissociation was also seen after binding of biotin-IgG (Figure 12B) which had been expected to stabilize captavidin on the chip by cross-linking of adjacent molecules. (iii) In spite of weak binding to the sensor surface, it was not possible to remove all bound captavidin ($\pm$biotin-IgG) on the same time scale as shown above for avidin M96H.

This comparison with captavidin demonstrates that the avidin mutant M96H is uniquely suited for reversible biofunctionalization of biotinylated surfaces. It shows very stable biotin-binding over a wide range of measurement conditions, even in SDS, and yet its function can be switched off within minutes when desired.

Hypothesis for the Synergy between SDS and Citric Acid in Dissociating Avidin M96H. It is easy to rationalize the combined effect of acid (pH 2.0) and SDS upon biotin-bound avidin M96H (see Figure 13). This avidin mutant has four additional histidines at the subunit−subunit interfaces which are accessible to protonation at low pH.[44] If protonated, these cationic sites should exhibit an enhanced affinity for negatively charged SDS molecules, thereby facilitating intrusion of hydrophobic SDS tails between the subunit interfaces. Together with the global destabilization by other protonated sites, the additionally bound SDS may be responsible for the
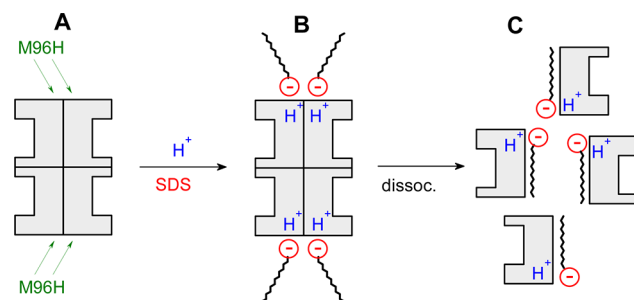

Figure 13. Hypothesis for the synergy between H[+] and SDS with respect to the denaturation of avidin M96H. Position 96 of each monomer is located at a subunit−subunit interface[44] (A). Protonation of the artificially introduced histidine residues is thought to promote binding of the negatively charged SDS molecules (B) which cause dissociation and denaturation of the subunits (C).

unusually quick denaturation and subunit dissociation from the chip surface caused by SDS at pH 2.0.

CONCLUSIONS

The goal of this study was to identify a variant of (strept)avidin which functions as a reversible bridge between surface-bound biotin residues on one hand and biotinylated bait molecules on the other hand. This goal has successfully been reached by use of the avidin mutant M96H. The latter proved to be very robust, retaining full function at neutral pH, even in the presence of SDS, or at low pH (>3) in the absence of SDS. Only the combination of SDS and low pH (2.0) was able to denature avidin M96H, causing instantaneous dissociation from a biotinylated chip surface.

The problem of nonspecific protein adsorption on top of chip-bound avidin M96H was solved by simple measures: (i) an improved method of mixed SAM formation and (ii) passivation with biotin−BSA and free BSA. The alternative method (acetylation of avidin M96H) proved equally suited to suppress nonspecific adsorption. The adverse side effect of acetylation was the appearance of a small fraction of avidin derivatives with reduced binding strength. Nevertheless, the goals of the acetylation studies were fully reached because its main purpose was to test whether lowering of the pI value was sufficient to abolish nonspecific protein adsorption on top of avidin M96H. The method of choice for this purpose is of course genetic engineering of the surface charges on avidin M96H and the corresponding study is under way.

ASSOCIATED CONTENT

Ⓢ Supporting Information
Screening of published avidin mutants for optimal pH dependence of biotin binding. Acetylation of avidin and avidin M96H and characterization of the products. Syntheses of the OH-terminated matrix alkanethiol (MAT) and of the biotin-terminated alkanethiol (BAT). Synthesis of the peptide-PEG$_{18}$-biotin conjugate with the peptide sequence HWRGWVC. Optimized version of mixed SAM formation. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*Phone: +43 (732) 2468-7597. Fax: +43 (732) 2468-7609. E-mail: hermann.gruber@jku.at.

■ ABBREVIATIONS

Avidin M96H-40, avidin mutant M96H derivatized with a 40-fold excess of sulfo-NHS acetate; BAT, biotin-terminated alkanethiol (see Figure 4); biotin-cap-DOPE, N-(6-(N-biotinyl)aminocaproyl-1,2-dioleoyl-phophatidylethanolamine; biotin-S-S-lysozyme, lysozyme derivatized with sulfo-NHS-SS-biotin; biotin-tris-NTA, see Figure 11; BSA, bovine serum albumin; DOPC, 1,2-dioleoylphosphatidylcholine; dsDNA, double-stranded DNA; HBS, HEPES-buffered saline; HEPES, N-(2-hydroxyethyl)-piperazine-N'-ethanesulfonic acid; MBP-His$_6$, maltose-binding protein with a C-terminal His$_6$ tag; HPA chip, BIAcore chip with a hydrophobic SAM of long alkyl chains; HWRGWVC-PEG$_{18}$-biotin, see Figure 8; IgG, immunoglobulin G; MAT, matrix alkanethiol (see Figure 4); MBP, maltose-binding protein; MBP-His$_6$, maltose-binding protein with a C-terminal His$_6$ tag; NHS, N-hydroxysuccinimide; NTA, nitrilotriacetate; PBS, phosphate-buffered saline; PEG, poly(ethylene glycol); pI, isoelectric point; RU, resonance units (1 RU = 0.0001° change of the resonance angle); SAM, self-assembled monolayer; SDS, sodium dodecyl sulfate; SPR, surface plasmon resonance; ssDNA, single-stranded DNA; (strept)avidin, streptavidin or avidin; sulfo-NHS acetate, N-2-sulfosuccinimidyl acetate; sulfo-NHS-SS-biotin, N-2-sulfosuccinimidyl 3-[2-(biotinamido)ethyl-dithio]-propionate sodium salt

■ REFERENCES

(1) Cooper, M. A. (2009) Sensor surfaces and receptor deposition. In *Label-free biosensors*, 1st ed. (Cooper, M. A., Ed.) pp 110–142, Cambridge University Press, Cambridge.

(2) Su, X., Chew, F. T., and Li, S. F. Y. (2000) Piezoelectric quartz crystal based label-free analysis for allergy disease. *Biosens. Bioelectron.* 15, 629–639.

(3) Prime, K. L., and Whitesides, G. M. (1993) Adsorption of proteins onto surfaces containing end-attached oligo(ethylene oxide): A model system using self-assembled monolayers. *J. Am. Chem. Soc.* 115, 10714–10721.

(4) Lofas, S., and Johnsson, B. (1990) A novel hydrogel matrix on gold surfaces in surface plasmon resonance sensors for fast and efficient covalent immobilization of ligands. *J. Chem. Soc., Chem. Commun.*, 1526–1528.

(5) Cha, T., Guo, A., Jun, Y., Pei, D., and Zhu, X.-Y. (2004) Immobilization of oriented protein molecules on poly(ethylene glycol)-coated Si(111). *Proteomics* 4, 10965–10976.

(6) Dorner, M. M., Bassett, E. W., Beiser, S. M., Kabat, E. A., and Tanenbaum, S. W. (1967) Studies on human antibodies: V. Amino acid composition of antidextrans of the same and of different specificities from several individuals. *J. Exp. Med.* 125, 823–831.

(7) Hahn, C. D., Leitner, C., Weinbrenner, T., Schlapak, R., Tinazli, A., Tampé, R., Lackner, B., Steindl, C., Hinterdorfer, P., Gruber, H. J., and Hölzl, M. (2007) Self-assembled monolayers with latent aldehydes for protein immobilization. *Bioconjugate Chem.* 18, 247–253.

(8) Hölzl, M., Tinazli, A., Leitner, C., Hahn, C. D., Lackner, B., Tampé, R., and Gruber, H. J. (2007) Protein-resistant self-assembled monolayers on gold with latent aldehyde functions. *Langmuir* 23, 5571–5577.

(9) Lahiri, J., Isaacs, L., Tien, J., and Whitesides, G. M. (1999) A strategy for the generation of surfaces presenting ligands for studies of binding based on an active ester as a common reactive intermediate: A surface plasmon resonance study. *Anal. Chem.* 71, 777–790.

(10) Lofas, S., Johnsson, B., Edstrom, A., Hansson, A., Lindquist, G., Hillgren, R.-M. M., and Stigh, L. (1995) Methods for site controlled coupling to carboxymethyldextran surfaces in surface plasmon resonance sensors. *Biosens. Bioelectron.* 10, 813–822.

(11) Sawada, J.-I., and Suzuki, F. (2000) DNA-Protein interactions. In *Real-time analysis of biomolecular interactions* (Nagata, K, and Handa, H, Eds.) pp 127–132, Springer Publishing Co., Tokyo.

(12) Green, N. M. (1990) Avidin and streptavidin. *Methods Enzymol.* 184, 51–67.

(13) Ternynck, T., and Avrameas, S. (1990) Avidin-biotin system in enzyme immunoassays. *Methods Enzymol.* 184, 469–481.

(14) Los, G. V., Darzins, A., Karassina, N., Zimprich, C., Learish, R., McDougall, M. G., Encell, L. P., Friedman-Ohana, R., Wood, M., Vidugiris, G., Zimmerman, K., Otto, P., Klaubert, D. H., and Wood, K. V. (2005) HaloTag interchangeable labeling technology for cell imaging and protein capture. *Cell Notes*, 2–6.

(15) Su, X. D., Wu, Y. J., Robelek, R., and Knoll, W. (2005) Surface plasmon resonance spectroscopy and quartz crystal microbalance study of streptavidin film structure effects on biotinylated DNA assembly and target DNA hybridization. *Langmuir* 21, 348–353.

(16) Stern, E., Klemic, J. F., Routenberg, D. A., Wyrembak, P. N., Turner-Evans, D. B., Hamilton, A. D., LaVan, D. A., Fahmy, T. M., and Reed, M. A. (2007) Label-free immunodetection with CMOS-compatible semiconducting nanowires. *Nature* 445, 519–22.

(17) Munroe, D. J., and Harris, T. J. (2010) Third-generation sequencing fireworks at Marco Island. *Nat. Biotechnol.* 28, 426–8.

(18) Yuan, Y. J, Gopinath, S. C. B., and Kumar, P. K. R. (2011) Regeneration of commercial Biacore chips to analyze biomolecular interactions. *Opt. Eng.* 50, 034402.

(19) Plant, A. L. (1993) Self-assembled phospholipid/alkanethiol biomimetic bilayers on gold. *Langmuir* 9, 2764–2767.

(20) Plant, A. L., Brigham-Burke, M., Petrella, E. C., and O'Shannessy, D. J. (1995) Phospholipid/alkanethiol bilayers for cell-surface receptor studies by surface plasmon resonance. *Anal. Biochem.* 226, 342–348.

(21) Kuziemko, G. M., Stroh, M., and Stevens, R. C. (1996) Cholera toxin binding affinity and specificity for gangliosides determined by surface plasmon resonance. *Biochemistry* 35, 6375–6384.

(22) Arai, M. (2000) Blood coagulation proteins. In *Real-time analysis of biomolecular interactions* (Nagata, K., and Handa, H., Eds.) pp 163–172, Springer, Tokyo.

(23) Hodnik, V., and Anderluh, G. (2010) Capture of intact liposomes on Biacore sensor chips for protein-membrane interaction studies. *Methods Mol. Biol.* 627, 201–11.

(24) Bergsmann, J., Derler, I., Muik, M., Frischauf, I., Fahrner, M., Pollheimer, P., Schwarzinger, C., Gruber, H. J., Groschner, K., and Romanin, C. (2011) Molecular determinants within the N-terminus of Orai3 controlling channel activation and gating. *J. Biol. Chem.* 286, 31565–31575.

(25) Mun, S., and Choi, S.-J. (2009) Optimization of the hybrid bilayer membrane method for immobilization of avidin on quartz crystal microbalance. *Biosens. Bioelectron. 24*, 2522−2527.

(26) Sagmeister, B. P., Graz, I., Schwödiauer, R., Bauer, S., and Gruber, H. J. (2009) User-friendly, biocompatible miniature flow cell for fragile high fundamental frequency quartz crystal resonators. *Biosens. Bioelectron. 24*, 2643−2648.

(27) Hofmann, K., Wood, S. W., Brinton, C. C., Montibeller, J. A., and Finn, F. M. (1980) Iminobiotin affinity columns and their application to retrieval of streptavidin. *Proc. Natl. Acad. Sci. U. S. A. 77*, 4666−4668.

(28) Orr, G. A. (1981) The use of the 2-iminobiotin-avidin interaction for the selective retrieval of labeled plasma membrane components. *J. Biol. Chem. 256*, 761−6.

(29) Gauthier, D. J., Gibbs, B. F., Rabah, N., and Lazure, C. (2004) Utilization of a new biotinylation reagent in the development of a nondiscriminatory investigative approach for the study of cell surface proteins. *Proteomics 4*, 3783−3790.

(30) Yoon, H. C., Hong, M.-Y., and Kim, H.-S. (2001) Reversible association/aissociation reaction of avidin on the dendrimer monolayer functionalized with a biotin analogue for a regenerable affinity-sensing surface. *Langmuir 17*, 1234−1239.

(31) Knoll, W., Zizlsperger, M., Liebermann, T., Arnold, S., Badia, A., Liley, M., Piscevic, D., Schmitt, F.-J., and Spinke, J. (2000) Streptavidin arrays as supramolecular architectures in surface-plasmon optical sensor formats. *Colloids Surf., A 161*, 115−137.

(32) Hirsch, J. D., Eslamizar, L., Filanoski, B. J., Malekzadeh, N., Haugland, R. P., Beechem, J. M., and Haugland, R. P. (2002) Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: Uses for protein labeling, detection, and isolation. *Anal. Biochem. 308*, 343−57.

(33) Ying, L. Q., and Branchaud, B. P. (2011) Design of a reversible biotin analog and applications in protein labeling, detection, and isolation. *Chem. Commun. 47*, 8593−5.

(34) Määttä, J. A., Helppolainen, S. H., Hytönen, V. P., Johnson, M. S., Kulomaa, M. S., Airenne, T. T., and Nordlund, H. R. (2009) Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis. *BMC Struct. Biol. 9*, 63.

(35) Qureshi, M. H., and Wong, S. L. (2002) Design, production, and characterization of a monomeric streptavidin and its application for affinity purification of biotinylated proteins. *Protein Expr. Purif. 25*, 409−15.

(36) Wu, S. C., and Wong, S. L. (2004) Development of an enzymatic method for site-specific incorporation of desthiobiotin to recombinant proteins in vitro. *Anal. Biochem. 331*, 340−348.

(37) Guchhait, R. B., Polakis, S. E., Hollis, D., Fenselau, C., and Lane, M. D. (1974) Acetyl coenzyme A carboxylase system of Escherichia coli. Site of carboxylation of biotin and enzymatic reactivity of 1′-N-(ureido)-carboxybiotin derivatives. *J. Biol. Chem. 249*, 6646−56.

(38) Kohanski, R. A., and Lane, M. D. (1990) Monovalent avidin affinity columns. *Methods Enzymol. 184*, 194−200.

(39) Friis, S., Godiksen, S., Bornholdt, J., Selzer-Plon, J., Rasmussen, H. B., Bugge, T. H., Lin, C. Y., and Vogel, L. K. (2010) Transport via the transcytotic pathway makes prostasin available as a substrate for matriptase. *J. Biol. Chem. 286*, 5793−802.

(40) Morag, E., Bayer, E. A., and Wilchek, M. (1996) Immobilized nitro-avidin and nitro-streptavidin as reusable affinity matrices for application in avidin-biotin technology. *Anal. Biochem. 243*, 257−263.

(41) Morag, E., Bayer, E. A., and Wilchek, M. (1996) Reversibility of biotin-binding by selective modification of tyrosine in avidin. *Biochem. J. 316*, 193−9.

(42) García-Aljaro, C., Xavier-Munõz, F., and Baldrich, E. (2009) Captavidin: a new regenerable biocomponent for biosensing? *Analyst 134*, 2338−2343.

(43) Ding, Z., Fong, R. B., Long, C. J., Stayton, P. S., and Hoffman, A. S. (2001) Size-dependent control of the binding of biotinylated proteins to streptavidin using a polymer shield. *Nature 411*, 59−62.

(44) Nordlund, H. R., Hytönen, V. P., Laitinen, O. H., Uotila, S. T., Niskanen, E. A., Savolainen, J., Porkka, E., and Kulomaa, M. S. (2003)

Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. *FEBS Lett. 555*, 449−54.

(45) Artelsmair, H., Kienberger, F., Tinazli, A., Schlapak, R., Zhu, R., Preiner, J., Wruss, J., Kastner, M., Saucedo-Zeni, N., Hölzl, M., Rankl, C., Baumgartner, W., Howorka, S., Blaas, D., Gruber, H. J., Tampé, R., and Hinterdorfer, P. (2008) Atomic force microscopy-derived nanoscale chip for the detection of human pathogenic viruses. *Small 4*, 847−54.

(46) Reichel, A., Schaible, D., Al Furoukh, N., Cohen, M., Schreiber, G., and Piehler, J. (2007) Noncovalent, site-specific biotinylation of histidine-tagged proteins. *Anal. Chem. 79*, 8590−600.

(47) Strunk, J. J., Gregor, I., Becker, Y., Lamken, P., Lata, S., Reichel, A., Enderlein, J., and Piehler, J. (2009) Probing protein conformations by in situ non-covalent fluorescence labeling. *Bioconjugate Chem. 20*, 41−46.

(48) Kamruzzahan, A. S. M., Ebner, A., Wildling, L., Kienberger, F., Riener, C. K., Hahn, C, D., Pollheimer, P. D., Winklehner, P., Hölzl, M., Lackner, B., Schörkl, D. M., Hinterdorfer, P., and Gruber, H. J. (2006) Antibody linking to atomic force microscope tips via disulfide bond formation. *Bioconjugate Chem. 17*, 1473−1481.

(49) Ebner, A., Wildling, L., Kamruzzahan, A. S. M., Rankl, C., Wruss, J., Hahn, C. D., Hölzl, M., Kienberger, F., Blaas, D., Hinterdorfer, P., and Gruber, H. J. (2007) A new, simple method for linking of antibodies to atomic force microscopy tips. *Bioconjugate Chem. 18*, 1176−1184.

(50) Gruber, H., Wilmsen, H., Schurga, A., Pilger, A., and Schindler, H. (1995) Measurement of intravesicular volumes by salt entrapment. *Biochim. Biophys. Acta 1240*, 266−76.

(51) Hahn, C. D., Tinazli, A., Hölzl, M., Leitner, C., Frederix, F., Lackner, B., Müller, N., Klampfl, C., Tampé, R., and Gruber, H. J. (2007) Pragmatic studies on protein-resistant self-assembled mono-layers. *Chemical Monthly 138*, 245−252.

(52) Jung, L. S., Nelson, K. E., Stayton, P. S., and Campbell, C. T. (2000) Binding and dissociation kinetics of wild-type and mutant streptavidins on mixed biotin-containing alkylthiolate monolayers. *Langmuir 16*, 9421−9432.

(53) Zocchi, A., Jobe, A. M., Neuhaus, J. M., and Ward, T. R. (2003) Expression and purification of a recombinant avidin with a lowered isoelectric point in Pichia pastoris. *Protein Expr. Purif. 32*, 167−74.

(54) Kondo, A., and Higashitani, K. (1992) Adsorption of model proteins with wide variation in molecular properties on colloidal particles. *J. Colloid Interface Sci. 150*, 344−351.

(55) Herrmann, J. E., Hendry, R. M., and Collins, M. F. (1979) Factors involved in enzyme-linked immunoassay and evaluation of the method of identification of enteroviruses. *J. Clin. Microbiol. 10*, 210−217.

(56) Rosebrough, S. F., and Hartley, D. F. (1996) Biochemical modification of streptavidin and avidin: in vitro and in vivo analysis. *J. Nucl. Med. 37*, 1380−4.

(57) Finn, F. M., and Hofmann, K. (1990) Isolation and characterization of hormone receptors. *Methods Enzymol. 184*, 244−274.

(58) Yang, H., Gurgel, P. V., and Carbonell, R. G. (2006) Hexamer peptide affinity resins that bind the Fc region of human immunoglobulin G. *J. Pept. Res. 66* (Suppl 0.1), 120−137.

(59) Yang, H., Gurgel, P. V., and Carbonell, R. G. (2009) Purification of human immunoglobulin G via Fc-specific small peptide ligand affinity chromatography. *J. Chromatogr., A 1216*, 910−8.

(60) Lata, S., and Piehler, J. (2005) Stable and functional immobilization of histidine-tagged proteins via multivalent chelator headgroups on a molecular poly(ethylene glycol) brush. *Anal. Chem. 77*, 1096−1105.

(61) Marttila, A. T., Airenne, K. J., Laitinena, O. H., Kulik, T., Bayer, E. A., Wilchek, M., and Kulomaa, M. S. (1998) Engineering of chicken avidin: a progressive series of reduced charge mutants. *FEBS Lett. 441*, 313−317.

(62) Tinazli, A., Tang, J., Valiokas, R., Picuric, S., Lata, S., Piehler, J., Liedberg, B., and Tampé, R. (2005) High-affinity chelator thiols for

switchable and oriented immobilization of histidine-tagged proteins: a generic platform for protein chip technologies. *Chem.—Eur. J. 11*, 5249−5259.

(63) Bain, C. D., Troughton, E. B., Tao, Y.-T., Evall, J., Whitesides, G. M., and Nuzzo, R. G. (1989) Formation of monolayer films by the spontaneous assembly of organic thiols from solution onto gold. *J. Am. Chem. Soc. 111*, 321−335.

(64) Svedhem, S., Hollander, C. A., Shi, J., Konradsson, P., Liedberg, B., and Svensson, S. C. T. (2001) Synthesis of a series of oligo(ethylene glycol)-terminated alkanethiol amides designed to address structure and stability of biosensing interfaces. *J. Org. Chem. 66*, 4494−4503.

(65) Rich, R. L., and Myszka, D. G. (2000) Advances in surface plasmon resonance biosensor analysis. *Curr. Opin. Biotechnol. 11*, 54−61.

PLOS ONE

Zebavidin - An Avidin-Like Protein from Zebrafish

Barbara Taskinen[1,2], Joanna Zmurko[1,3¤a], Markus Ojanen[1], Sampo Kukkurainen[1,2], Marimuthu Parthiban[4], Juha A. E. Määttä[1,2], Jenni Leppiniemi[1,5], Janne Jänis[6], Mataleena Parikka[1], Hannu Turpeinen[1], Mika Rämet[1,5], Marko Pesu[1,2], Mark S. Johnson[4], Markku S. Kulomaa[1,5], Tomi T. Airenne[4], Vesa P. Hytönen[1,2*]

1 Institute of Biomedical Technology, University of Tampere, BioMediTech, Tampere, Finland, **2** Fimlab Laboratories, Pirkanmaa Hospital District, Tampere, Finland, **3** Institute of Molecular, Cell and Systems Biology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom, **4** Department of Biosciences, Biochemistry, Åbo Akademi University, Turku, Finland, **5** Tampere University Hospital, Tampere, Finland, **6** Department of Chemistry, University of Eastern Finland, Joensuu, Finland, **7** Department of Paediatrics, Tampere University Hospital, Tampere, Finland

Abstract

The avidin protein family members are well known for their high affinity towards D-biotin and high structural stability. These properties make avidins valuable tools for a wide range of biotechnology applications. We have identified a new member of the avidin family in the zebrafish (*Danio rerio*) genome, hereafter called zebavidin. The protein is highly expressed in the gonads of both male and female zebrafish and in the gills of male fish, but our data suggest that zebavidin is not crucial for the developing embryo. Biophysical and structural characterisation of zebavidin revealed distinct properties not found in any previously characterised avidins. Gel filtration chromatography and native mass spectrometry suggest that the protein forms dimers in the absence of biotin at low ionic strength, but assembles into tetramers upon binding biotin. Ligand binding was analysed using radioactive and fluorescently labelled biotin and isothermal titration calorimetry. Moreover, the crystal structure of zebavidin in complex with biotin was solved at 2.4 Å resolution and unveiled unique ligand binding and subunit interface architectures; the atomic-level details support our physicochemical observations.

Introduction

Chicken avidin [1] and its bacterial cousin streptavidin from *Streptomyces avidinii* [2] have been studied extensively for decades. Many biotechnological applications would be unthinkable without the presence of these stable proteins that have extremely high affinity ($K_d = 10^{-14} - 10^{-16}$ M) towards biotin (D-biotin) [3].

Avidin is found in the egg white of chicken and other egg-laying species [1]. Its expression in the oviduct is induced by progesterone, but it has also been shown to be induced in most tissues by bacterial infection or physical damage [4-6]. These findings led to the suggestion that avidin acts as an anti-microbial agent, yet its physiological function is still not completely understood.

The sequencing of the zebrafish (*Danio rerio*) genome revealed a gene similar to that encoding chicken avidin [7]. Zebrafish, a tropical freshwater fish, is a model organism used traditionally to study developmental biology and more recently to investigate many other fields of biomedical research including haematological disorders and infectious diseases [8,9]. Although avidins have been found in many organisms including chickens and other birds, frog (xenavidin from *Xenopus tropicalis*, [10]), mushrooms (tamavidin from Tamogitake mushroom, [11]) and several bacterial species [2,12-14], no evidence for avidin proteins in fish species have

been published to date. In an early study, Korpela et al. observed avidin expression in various vertebrates, but their studies did not include any fish species [15].

In the current study, we present the characterisation of an avidin-like protein in zebrafish, called zebavidin. Expression of zebavidin and its effect on embryonic development in zebrafish was investigated. We used biochemical and biophysical methods to evaluate the ligand-binding properties, oligomeric state and thermal stability of the produced recombinant zebavidin. X-ray crystallography was used to analyse the structure at atomic detail. Although zebavidin has many unique physicochemical and structural features, it clearly can be classified as a new member of the avidin protein family based on its structure and function.

Materials and Methods

Zebrafish maintenance

Wild-type AB zebrafish were used in all experiments. The fish were maintained according to standard protocols [16,17]. The adult fish were kept in a flow-through system at 28 °C with a light/dark cycle of 14 h/10 h. The fish tanks were made of FDA-approved, good grade autoclavable polycarbonate USP class VI. Reverse osmosis water was used with conductivity (800 µS) and pH levels (pH 7.6) adjusted using sea salt (Instant Ocean, Blacksburg, VA, USA) and $NaHCO_3$ respectively. The water was filtered mechanically using filters, chemically using activated carbon and sterilized with UV-light. Nitrifying bacteria were added to the system to convert nitrogenous wastes to less toxic substances. Water was partially (10%) exchanged daily. The fish were fed with SDS 400 food (Special Diets Services, Essex, UK) twice a day. Embryos were grown in E3-H_2O (5mM NaCl, 0.17 mM KCl, $CaCl_2$, 0.33 mM $MgSO_4$, 1 · 10^{-5}% methylene blue) at 28 °C. Fish were euthanized with 4 mg/ml Tricaine (ethyl 3-aminobenzoate methanesulfonate salt) pH 7; the pH was adjusted using Tris-buffer. All of the zebrafish experiments were in accordance with the Finnish Laboratory Animal Welfare Act 62/2006, the Laboratory Animal Welfare Ordinance 36/2006 and have been authorized (authorization LSLH-2007-7254/Ym-23) by the National Animal Experiment Board (Finland).

Zebavidin gene expression analysis with qRT-PCR

Relative zebavidin mRNA quantification was performed with adult wild-type AB zebrafish tissue samples (gonads, gills, kidney, tail fin, eyes and brain) of both sexes as well as from whole developing wild-type AB strain embryos of different ages (0-7 days post fertilization (dpf)). Because of the difficulties of separating gametes from gonads, "gonad" samples contain tissue from both gonads and gametes. Developmental samples were pooled from 15-50 individual embryos depending on the age. Total RNA was extracted with an RNeasy RNA purification kit (Qiagen, Hilden, Germany) followed by reverse transcription using iScript™ Select cDNA synthesis kit (Bio-Rad, California, USA) according to the manufacturer's instructions. Zebavidin primers (F: 5′-CGAATGCAAAGGTGAGCTCC-3′ and R: 5′-ATAGCACGGAGAAAGAGACG-3′) for quantitative real-time

PCR (qRT-PCR) were ordered from Oligomer (Helsinki, Finland). The mRNA expression was measured using cDNA, qRT-PCR, Maxima SYBR Green qPCR master mix (Fermentas, Burlington, Canada) and a CFX96 qPCR machine (Bio-Rad Laboratories). Adult tissue samples and 0 dpf developmental samples were run as technical duplicates from 3-7 biological replicates and the 1-7 dpf developmental samples as technical triplicates from one pooled biological sample. Expression of zebavidin was normalized to elongation factor 1-alpha gene (EF1a, ENSDARG00000020850) expression [18]. Bio-Rad CFX Manager software (Bio-Rad) and Microsoft Office Excel 2010 were used in the analysis of the results. GraphPad Prism 5.0 (GraphPad Software, Inc., La Jolla, CA, USA) was used in the statistical analysis. Additionally, the qRT-PCR products were subjected to melting curve analysis followed by 1.5% agarose gel (Bioline, London, UK) electrophoresis.

Morpholino studies

The zebavidin translation initiation site was sequenced from the genomic DNA of three zebavidin strains (+AB5, +AB7 and +AB8). The Morpholino, aimed at blocking the translation of zebavidin by binding to the translation initiation site (5′-GCCATATTAAAGAACTCATCTTGGC-3′), was ordered from GeneTools, LCC, (Philomath, USA). The Morpholino stock was diluted to a final concentration of 130 µM using 200 mM KCl containing 0.2% rhodamine dextran tracer (Invitrogen, Carlsbad, CA, USA). The rhodamine dextran tracer allowed the confirmation of a successful injection of the Morpholino using fluorescence microscopy. Morpholino solution, 1nl, was injected into the yolk of 150 one-cell stage zebrafish eggs. Ninety eggs were injected with a random control Morpholino with random sequence (5′-CCTCCTACCTCAGTTACAATTTATA-3′). The control Morpholino served as a negative control. Embryos were grown in E3-H_2O at 28 °C and screened at 24 h post injection using a fluorescent microscope in order to determine the success of the injection. Developing embryos were analysed under a contrast light microscope after 24 and 48 h.

Isolation of zebavidin from zebrafish oocytes

Zebrafish oocytes were extracted from zebrafish gonads of five individuals. The oocytes (20- 150 mg per individual) were washed with 1 ml PBS and centrifuged for 5 min at 5,000 g and 4 °C. Supernatant was removed, mixed with 20 µl biotin Sepharose™ 4 Fast flow (Affiland S.A., Ans Liege, Belgium) and incubated at 4 °C for 1 hour. Washed oocytes were resuspended in 400 µl 0.5 M sucrose, 0.01 mg/ml lysozyme, 1 mM EDTA, 200 mM Tris, pH 7.4 and incubated on ice for 30 min. 400 µl 2 mM EDTA, 150 mM NaCl, 1% TritonX-100, 50 mM Tris, pH 8 was added and oocytes were sonicated at 25% amplitude two times for 15 s (alternating 1 s on, 1 s off). Sonicated oocytes were centrifuged at 13,000 g and 4 °C for 20 min. PBS + 1 M NaCl (800 µl) and 20 µl biotin Sepharose™ were added and samples were incubated at 4 °C for 1 hour. Sepharose was collected by centrifugation at 2,500 g and 4 °C for 5 min. Sepharose was washed with 1 ml PBS + 1 M NaCl. Samples from each step were taken and analysed on 15%

SDS PAGE and subsequent Coomassie Brilliant Blue staining. A molecular marker (Page Ruler™ Prestained Protein Ladder, Fermentas) and bacterial expressed zebavidin were used for the size determination. Four selected bands with a molecular weight corresponding to zebavidin (13 kDa) were excised from the gel and dried in acetonitrile. Samples were in-gel digested at the Proteomics Facility (Turku, Finland) according to the standard protocol and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) using the LTQ Orbitrap Velos Pro mass spectrometer. The obtained data was searched against the NCBInr database (release 2013_02) using Mascot 2.4.0 (Matrix Science, Boston, MA, USA).

Construction of expression vector

Zebavidin cDNA (GenBank: BC127392.1) was amplified from the cDNA obtained from the I.M.A.G.E cDNA collection (IMAGE:5413136) using the primers Zeb_pET101_5' (5'-CACCATGAGTTCTTTAATATGGCATTTGG-3') and zeb_3'_HindIII (5'-TCAAGCTTAATTTGAAACTCCAGTCTTG-3') for amplification of the whole cDNA including the natural secretion signal peptide. The sequence encoding for the OmpA secretion signal peptide from *Bordetella avium* was added to cDNA encoding the core region of zebavidin in a two-step stepwise elongation of sequence PCR (SES-PCR) process [19], essentially as described earlier for chicken avidin [20]. Primers zeb_OmpA_B.a_5' (5'-GCCGCCGTTACGGCCTCTGGTGTTGCCTCGGCTCAGACCGTGAGCTCCTGTAATGTGACC-3') and zeb_3'_HindIII were used in the first step and primers chim_to_2_5' (5'-CACCATGAACAAACCCTCCAAATTCGCTCTGGCGCTTGCCTTCGCCGCCGTTACGGCCTC-3') and zeb_3'_HindIII in the second step of SES-PCR. The PCR intermediates and products were purified from the gel with an Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare, Buckinghamshire, UK) and subsequently cloned into the pET101/D plasmid (Invitrogen) using the TOPO® cloning protocol followed by a standard heat shock transformation of chemically competent Top10 *E. coli* cells (Invitrogen). The sequence of the construct obtained by this method was confirmed using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA).

Recombinant protein production

Zebavidin was produced in a pilot scale fermentor as previously described [21]. The pET101/D vector containing the zebavidin ORF was transformed into *E. coli* BL21-AI (Invitrogen) cells. Single colonies were grown overnight in 5 ml of fermenting medium containing 100 µg/ml ampicillin and 10 µg/ml tetracycline at 27 °C with shaking at 200 rpm. The cell culture was diluted in 500 ml of fermenting medium (see 21) under the same conditions and used the following day to start a 4.5 l fermentation in a Labfors Infors 3 fermentor (Infors HT, Bottmingen, Switzerland) at 25 °C. The fermenting medium contained the antifoam agent struktol J 647 (Schill + Seilacher, Hamburg, Germany). The culture was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 0.2% (w/v) L(+)-arabinose when the OD_{600} reached a value of approximately 20. The target pO_2 of the culture was set to 20%, which was controlled by the agitation speed (200-1150 rpm) and air flow rate. The feeding of the culture was controlled by a pO_2-stat, which applies feed when the oxygen level rises above 40%. In practice, this was achieved by allowing the pO_2 to oscillate during the feeding phase. Fermentation was stopped 24 h after inoculation.

Zebavidin purification by affinity chromatography

The cell pellet was suspended in PBS buffer containing 2 mM EDTA and 1 M NaCl (for biotin affinity chromatography), or 50 mM Na-carbonate (pH 11) containing 1 M NaCl (for 2-iminobiotin affinity chromatography). The lysate was homogenised twice using an EmulsiFlex C3 homogeniser (Avestin Inc., Ottawa, Canada) using guiding pressure set to 40 psi to obtain homogenizing pressure of 15,000 psi. The crude cell lysate was clarified by centrifugation at 16,000 g and 4 °C for 60 min. The protein was purified by affinity chromatography as described by K. Airenne et al. [22] using biotin or 2-iminobiotin Sepharose™ 4 Fast flow (Affiland S.A., Ans Liege, Belgium). The protein concentrations in the eluted fractions were determined using the theoretical molar absorption coefficient of zebavidin at 280 nm (28,990 $M^{-1}cm^{-1}$) and the estimated molecular weight (13,588.70 g/mol). The A_{280} values were measured using a NanoDrop 2000 (Thermo Scientific, Wilmington, DE, USA). The molecular size and purity of the protein were analysed by SDS-PAGE (15% gel) and subsequent Coomassie Brilliant Blue staining. A molecular marker (Page Ruler™ Prestained Protein Ladder, Fermentas) was used for the size determination.

Analytical gel filtration

The molecular size of the protein in solution was measured by size-exclusion chromatography using a Superdex200 10/300GL column connected to an ÄKTA™ purifier-100 equipped with a UV-900 monitor (GE healthcare/Amersham Biosciences AB, Uppsala, Sweden). As a mobile phase, either 50 mM sodium phosphate (NaH_2PO_4/Na_2HPO_4), 2 mM EDTA, pH 7 or 10 mM ammonium acetate, 0.05% β-mercaptoethanol, pH 7 with varying concentrations of sodium chloride (0, 100 and 650 mM) was used. Approximately 40-80 µg of protein was injected per run. All analyses used a flow rate of 0.3 ml/min and were executed at 4 °C. Absorbance at 280 nm was used to detect the eluting protein. A molecular weight calibration curve was prepared for each buffer condition by using a gel filtration standard protein mixture containing thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa) (Bio-Rad).

Mass spectrometry

Mass spectrometric (MS) analyses for the characterization of structural and functional properties of recombinant zebavidin were performed on a 12-T APEX-Qe™ Fourier transform ion cyclotron resonance (FT-ICR) instrument (Bruker Daltonics, Billerica, MA, USA), interfaced to an electrospray ionisation (ESI) source. The protein sample in 2 M acetic acid was buffer exchanged into 10 mM ammonium acetate (pH 3.0) with the use of PD-10 columns (GE Healthcare, Uppsala, Sweden). The

resulting fractions, which eluted between 4 and 6 ml, were pooled and concentrated to approximately 250 µl using Millipore Ultrafree-0.5 Biomax-5 (5-kDa cut-off) centrifugal filter devices (Millipore, Billerica, MA, USA). The concentrations of the protein stock solutions were determined as described above. In order to analyse the protein under denaturing solution conditions, the stock solution was further diluted with an acetonitrile/water/acetic acid (49.5:49.5:1.0, v/v) solvent. Alternatively, the sample was diluted with 10–500 mM ammonium acetate (pH ~7) to perform native-MS analysis. ESI-generated ions were externally accumulated for 1 s in the hexapole ion trap before being transmitted to the ICR cell for trapping, excitation and detection. For each spectrum, a total of 250 co-added 1MWord (128kWord for native-MS) time-domain transients were zero-filled once prior to a fast Fourier transformation, magnitude calculation and external mass calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, CA, USA). The instrument was operated and the data were processed with the use of Bruker XMASS 6.0.2 software.

Fluorescence spectroscopy

The affinity of the zebavidin protein towards the biotin-labelled fluorescent dye ArcDia™ BF560 (ArcDia, Turku, Finland) was measured by a method based on the quenching of the dye as a result of its binding to avidin [20]. In practice, a 10 nM solution of the BF560 probe in 50 mM sodium phosphate, 650 mM NaCl and 0.1 mg/ml BSA was titrated with a known concentration of the protein at 25 °C with continuous stirring. The sample was allowed to equilibrate for 10 min after each addition of protein, and the fluorescence intensity at 578 nm was measured for 20 s following excitation at 560 nm. The protein was added to the sample until no clear decrease in fluorescence intensity was observed, indicating the point where nearly all available ligand-binding sites were occupied by biotin. The concentration of the protein where half of the available ligand-binding sites are occupied corresponds to the equilibrium dissociation constant (K_d), which was calculated by fitting a binding curve to the data by nonlinear regression using Origin 7.0 (Originlab Corporation, Northampton, MA, USA).

The dissociation rate constant (k_{diss}) of fluorescently labelled biotin was determined by fluorescence spectrometry using the biotin-labelled fluorescent probe ArcDia™ BF560 as described in [20]. In principle, the changes in the fluorescence intensity of a 50 nM dye in a pH 7 buffer (50 mM sodium phosphate, 650 mM NaCl, 0.1 mg/ml BSA) were measured after the addition of 100 nM biotin-binding protein. The dissociation of this complex was observed by addition of a 100-fold molar excess of free biotin (D-biotin, Sigma-Aldrich Co. LLC., St. Louis, MO, USA). The assay was performed at 25 °C and 50 °C using a QuantaMaster™ Spectrofluorometer (Photon Technology International, Inc., Lawrenceville, NJ, USA).

Radioactive [3H]biotin assay

The biotin-binding properties were further studied with a radioactive biotin assay modified from that described in [23]. The protein at 50 nM subunit concentration was incubated with 10 nM radioactive biotin ([8,9-3H]biotin, PerkinElmer, Waltham,

MA, USA) at room temperature (22 ± 1 °C) for 20 min. Measurements were performed in 50 mM NaH_2PO_4/Na_2HPO_4, pH 7, containing 100 mM NaCl and 10 µg/ml BSA to prevent non-specific binding. Centrifugal ultrafiltration was performed through 30,000 MW cut-off filter (Vivaspin 500 centrifugal concentrators, Sigma-Aldrich) to separate the unbound [3H]biotin from the protein-ligand complex. Excess cold biotin (D-biotin, Sigma-Aldrich) was added to a final concentration of 50 µM to replace [3H]biotin and measure its dissociation. The dissociated [3H]biotin was separated from the protein-ligand complex at different time points by ultrafiltration and the radioactivity of the filtrate was analysed in a Wallac 1410 liquid scintillation counter (Wallac Oy, Turku, Finland). Triplicates of each sample were measured at each time point.

Equation (1) was fitted to the data in order to determine the fraction of bound radioactive biotin at each time point:

$$-k_{diss}t = \ln\left[\frac{(x_t - x)}{(x_t - x_0)}\right] = \ln(fraction\_bound) \qquad (1)$$

where x_t is the total amount of radioactive biotin before the addition of protein, x is the free biotin at each time point and x_0 is the amount of free ligand in the presence of protein immediately prior to the addition of cold biotin. The dissociation rate constant (k_{diss}) was determined from the slope of the linear fit to the data points of ln(fraction bound) versus time [23].

Isothermal Titration Calorimetry (ITC)

Affinity towards biotin was determined using a high-sensitivity VP-ITC titration calorimetry instrument (Microcal Inc., Northampton, MA, USA) by using direct titration with biotin as described in [24] and by using a competitive titration method as described earlier in [25]. Measurements were made in 50 mM NaH_2PO_4/Na_2HPO_4, 100 mM NaCl, 2 mM EDTA, pH 7. Ligand solutions of biotin and desthiobiotin (Sigma-Aldrich) with a concentration of 75 µM were prepared in the same dialysis buffer as the protein. The protein solution in the cell was 4.65 µM. At first, a biotin titration experiment to zebavidin was carried out by using 15 µl aliquots. Competitive binding experiments were carried out as follows: In the first experiment, desthiobiotin was titrated in 15 µl aliquots to the zebavidin. In a second experiment, biotin was titrated in 15 µl aliquots to a mixture of 4.65 µM zebavidin and 14 µM desthiobiotin. Measurements were done at 40 °C. Origin 7.0 (Originlab Corporation) was used to derive the association binding constant (K_a) and enthalpy of binding (ΔH) from the measured biotin or desthiobiotin binding data. The observed values for desthiobiotin were then used to analyse the competitive titration reaction utilizing the "competitive binding" tool. The Gibbs free energy of binding (ΔG) and entropy (ΔS) were calculated from K_a and ΔH using equation (2) and (3) where R is the gas constant and T the temperature:

$$\Delta G = -RT\ln K_a \qquad (2)$$

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

Differential Scanning Calorimetry (DSC)

The thermal stability of the studied proteins in the presence and absence of ligands was analysed using an automated VP-Capillary DSC System (Microcal Inc.). Thermograms were recorded between 20 and 140 °C with a heating rate of 120 °C/h. Proteins were dialysed into 50 mM NaH_2PO_4/Na_2HPO_4, 2 mM EDTA, pH 7 or 10 mM ammonium acetate, 0.05% β-mercaptoethanol, pH 7. Samples with different salt concentrations were prepared by adding 5 M NaCl. Samples were degassed prior to the measurement. The protein concentration in the cell was 20-30 µM, and the ligand concentration was 60-90 µM. The results were analysed using the Origin 7.0 DSC software suite (Originlab Corporation).

Stability analysis

The thermostability and oligomeric state of zebavidin in the presence and absence of biotin was measured using the SDS-PAGE stability assay as described previously [26,27]. In brief, zebavidin was chemically acetylated, SDS-PAGE sample buffer containing SDS and 2-mercaptoethanol was added, and the sample was heated to the target temperature for 20 minutes. The oligomeric state after treatment was then assayed by SDS-PAGE.

Crystallization and X-ray structure determination

Zebavidin (1.6 mg/ml; 50 mM Tris-HCl, pH 7) was crystallized using the vapour diffusion method, 96-well sitting drop iQ plates (TTP Laptech), TTP Labtech's mosquito® liquid handling robot and a cooled temperature-controlled crystallization incubator (RUMED® model 3201) set-up for 22 °C. The protein was mixed with biotin solution (1 mg/ml; 5 mM Tris, pH 8.8, 8 mM CHES, pH 9.5) in 10:1 v/v ratio, respectively, before crystallization. The initial hit was found from the JCSG-plus™ Screen (Molecular Dimensions) and, after optimization, 300 nl of the protein-biotin solution and 150 nl of well solution (0.18 M magnesium chloride, 0.09 M Bis-Tris, pH 5.5, 23% w/v PEG 3350) were used to crystallize zebavidin. The crystals formed typically in 1-2 weeks and their X-ray diffraction properties were initially analyzed directly from drops on 96-well plates using a PX Scanner (Agilent Technologies).

X-ray data were collected at the ESRF beam line ID23-2 (Grenoble, France) at 100 K from a single crystal. As a cryoprotectant, 1 µl of glycerol (30% v/v in well solution) was added to the crystallization drop just prior to freezing in liquid nitrogen. The data were processed with XDS [28] (see Table 1 for statistics) and initial phase estimates for the structure factors were obtained using the molecular replacement program Phaser [29] within the CCP4i GUI [30,31]. For molecular replacement, a tetrameric homology model of zebavidin was created using Modeller [32] of Discovery Studio 3 (Accelrys Software Inc.) and AVR2 [PDB: 1WBI] [33] as the template structure; the zebavidin sequence was aligned with the sequences of chains A-D of the AVR2 structure using the program Malign [34] of the Bodil multi-platform software package for biomolecular visualization and modeling [35]. In Phaser, a solution could be found by searching four poly(Ala/Gly) models with trimmed loops and termini. The initial X-ray structure of zebavidin was refined with Refmac5 [36] and

Table 1. X-ray structure determination statistics for zebavidin [PDB: 4BJ8].

Cell parameters	
Space group	$P2_12_12$
Unit cell:	
a, b, c (Å)	182.2, 196.8, 52.6
α, β, γ (°)	90, 90, 90
Data collection[a]	
Wavelength (Å)	0.87260
Beamline	ID23.2 (ESRF)
Detector	MarCCD
Resolution (Å)[b]	25 - 2.4 (2.5 - 2.4)
Unique observations[b]	74935 (8359)
I/sigma[b]	14.4 (3.4)
R_{factor} (%)[b]	12.0 (60.8)
Completeness[b]	100 (99)
Redundancy[b]	8.2 (8.3)
Refinement	
R_{work} (%)[c]	19.2
R_{free} (%)[c]	26.6
Monomers (asymmetric unit)	16
Protein atoms	14505
Heterogen atoms	262
Solvent atoms	503
R.m.s.d:	
Bond lengths (Å)	0.013
Bond angles (°)	1.7

a. The numbers in parenthesis refer to the highest resolution bin.
b. From XDS [28].
c. From Refmac 5 [36].
doi: 10.1371/journal.pone.0077207.t001

manually edited/rebuilt using Coot [37]. Solvent atoms, biotin molecules and a glycerol molecule were added to the model either with the automatic procedure of Coot and ARP/wARP [38-40], or manually in Coot. For structure determination statistics, see Table 1.

The final structure of zebavidin was validated using the inbuilt tools of Coot [37], and using MolProbity [41] of the Phenix software suite [42], before deposition to the Protein Data Bank [43,44] with PDB entry code 4BJ8.

Miscellaneous methods

Primer design and sequence analysis were performed using DNAMAN 4.11 (Lynnon Corporation, Quebec, Canada). Signal peptide profiling was performed with SignalP [45,46]. N-glycosylation sites were predicted using NetNGlyc 1.0 Server [47]. The structural superimpositions were made using the "align" command of PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.2, Schrödinger, LLC) and the chain A (or tetramer ABCD) of zebavidin as the reference structure. The structure-based sequence alignment was done using Malign [34] in the visualization and modelling package Bodil [35]. The diagram showing hydrogen bonds to biotin was made using LigPlot+ v.1.4 [48]. ESPript [49] was used for visualization of
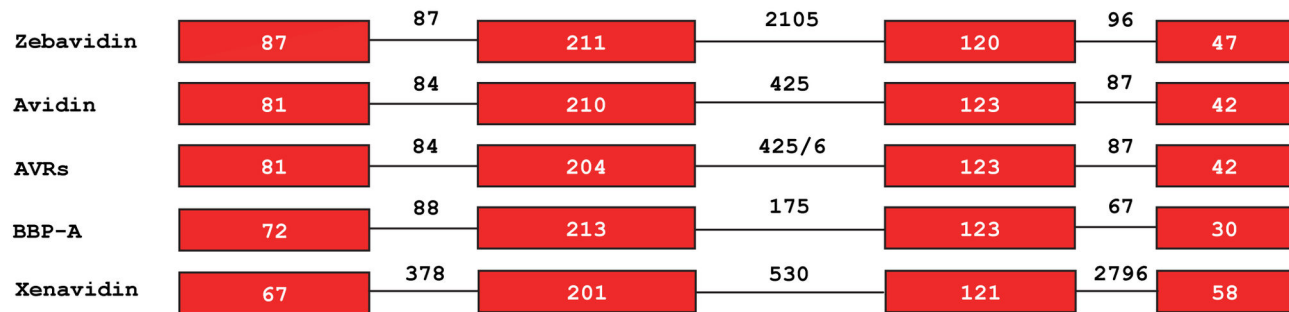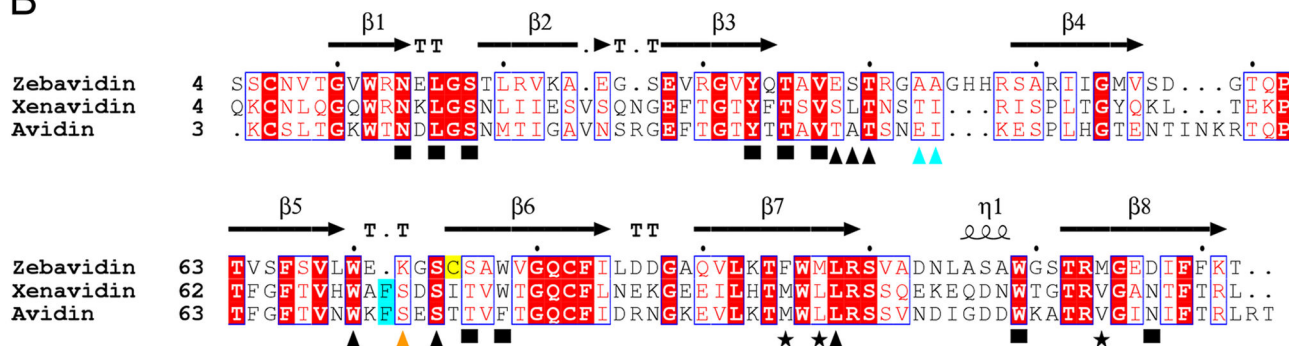
Figure 1. Gene organisation and structure-based sequence alignment of selected members of the avidin protein family. (A) Gene organization shown for selected avidins. The exons are shown as red boxes and the length of introns and exons are indicated. (B) Structure-based sequence alignment. The sequences of zebavidin [PDB: 4BJ8; reported here], xenavidin [PDB: 2UYW] and chicken avidin [PDB: 1AVD] are shown. Black squares: residues surrounding the bicyclic ring system of biotin; black triangles: residues around the valeric acid moiety; cyan triangles: zebavidin-specific Ala43 and Ala44 facing the valeric acid moiety of biotin; cyan background: the conserved Phe72/Phe71 of avidin/xenavidin missing from zebavidin; orange triangle: indicates the residue directly involved in biotin binding in avidin/xenavidin (Ser73/Ser72) but not in zebavidin (Lys72); asterisks: three zebavidin specific residues (Phe95, Met97 and Met114) found at IF1,2 and IF1,3 subunit interfaces; yellow background: the free cysteine (Cys75) of zebavidin. The conserved residues are indicated by the default coloring scheme of the ESPript program [49]. The β-strands of zebavidin are numbered. TT, β-turn and α1, 3/10-helix.
doi: 10.1371/journal.pone.0077207.g001

the sequence alignment. PyMOL was used to create all of the figures relating to structural representations (except the diagram describing the hydrogen bonds). GIMP 2.6.9, CorelDRAW X5 and Photoshop CS5.1 were used to edit the figures.

Results

Analysis of an avidin-related gene in zebrafish

The DNA sequence from the genome assembly of zebrafish revealed a gene similar to the chicken avidin gene (Gene ID: 567678, Gene Symbol LOC567678). The gene is located on chromosome 10 of the *Danio rerio* genome, and it is 2981 base pairs in length in a region from base 18'857'183 to base 18'860'163. Alignment of zebavidin cDNA with its corresponding DNA sequence revealed that the zebavidin gene consists of four exons (87 bp, 211 bp, 120 bp and 42 bp) and three introns (87 bp, 2105 bp and 96 bp). This exon-intron

structure is highly similar to that observed in all previously characterised avidins (Figure 1A). Unlike the chicken, which carries several avidin or avidin-related genes [50,51], zebrafish appears to have only one avidin-related gene. A protein sequence alignment of three eukaryotic avidins, zebavidin, xenavidin [10] and chicken avidin [52], shows that zebavidin shares many of the highly conserved residues involved in biotin binding but has also unique features (Figure 1B; see below).

Zebavidin expression in zebrafish

In order to analyze the spatiotemporal expression of zebavidin, we quantified the expression of zebavidin with qRT-PCR in adult zebrafish and zebrafish embryos. The expression of zebavidin was normalized to the elongation factor 1-alpha gene (*EF1a*) housekeeping gene expression, and the relative zebavidin mRNA quantification was used to assess zebavidin mRNA amounts in male gonads (n=3), female gonads (n=5), male gills (n=3), female gills (n=3), kidney (n=6), tail fin (n=5),

eyes (n=7) and brain (n=7). Zebavidin expression levels were also quantified in developing fish embryos of 0-7 days post fertilization (dpf, Figure 2 A&B).

In adult zebrafish (see Figure 2A), the highest zebavidin expression was observed in female gonads. Expression of the zebavidin gene in ovaries was already observed to be higher as compared to the standard in a cDNA microarray [7]. Expression was also evident in male gonads and male gills. Interestingly, zebavidin mRNA was present only in very low amounts in female gills. Low concentrations of zebavidin mRNA were also present in kidney, tail fin, eyes and brain. Supporting these findings, the zebavidin protein was isolated from oocytes and its surrounding tissue and its identity was confirmed by MS analysis (Table S1 and Figure S1).

In the developmental samples (see Figure 2B), the relative zebavidin expression was observed to be the highest immediately after spawning at the 0 dpf timepoint. This expression was over a thousand fold higher than at the next timepoint of 1 dpf (P < 0.0001, two-sample t-test). This, together with the high expression levels in the maternal gonads, implies that zebavidin has been expressed in the mother prior to spawning and that the high level of zebavidin mRNA in the 0 dpf embryo is of maternal origin. This is further supported by the results from the Morpholino study: injection of a zebavidin-specific Morpholino, designed to block zebavidin protein production by binding to the translation initiation site caused no detectable effects during the 2-day follow-up (Figure 2C). However, the efficiency of the Morpholino was not confirmed by any other experiments.

Expression of recombinant zebavidin in *E. coli*

The signal peptide prediction program SignalP [46,53] suggested two probable N-terminal signal peptide cleavage sites located between amino acid residues 21 and 22 and residues 29 and 30. According to this analysis with SignalP, a secretion signal was found that is potentially functional in gram-negative bacteria too. Therefore, the entire open reading frame of zebavidin cDNA was first constructed into the pET101/D expression vector. However, no active protein was expressed using this construct. Therefore, the first 30 amino acids of the open reading frame were replaced with an N-terminal secretion signal peptide OmpA from *Bordetella avium* [20,54]. The N-terminal truncation was based on the results from the SignalP analysis and on sequence comparison of known avidins. The residues analogous to those cleaved away from zebavidin are not known to have any important role in other avidins. Purification of *E. coli* expressed zebavidin was attempted with two different affinity chromatography columns; purification with 2-iminobiotin Sepharose™ yielded an insignificant amount of the protein, whereas zebavidin bound well to the biotin column and could be eluted using a stepwise gradient of acetic acid (0-4 M) with typical yields of 175 µg of recombinant zebavidin per gram of cell pellet (wet weight).

Identification of the *E. coli* expressed and purified protein

The identity of the purified protein was confirmed by ESI FT-ICR mass spectrometry analysis under denaturing solution conditions (Figure S2). The most abundant isotopic mass for the zebavidin monomer M (averaged over the detected protein ion charge states 8+ to 13+) was 13,588.80 ± 0.03 Da, which is in a good agreement with the theoretical mass (13,588.70 Da) calculated from the amino acid sequence. This analysis suggests the presence of one intramolecular disulphide bond, one free cysteine per monomer and no intermolecular disulphide bonds between subunits, all of which is in agreement with the X-ray structure.

Dynamic oligomeric state of zebavidin

The oligomeric state of zebavidin was analysed by native mass spectrometry (native-MS) analysis. Mass spectra obtained under native conditions and at low ionic strength (10 mM ammonium acetate, pH 7) at a protein (monomer) concentration of 10 µM without biotin (Figure 3A) revealed the presence of both protein monomers (charge states 7+ and 8+) and dimers (charge states 9+ to 11+) in solution. The determined mass of the dimer (27,177.34 ± 0.01) proves its noncovalent character without any ligands attached. In addition, very small peaks at ~*m/z* 3500-4000 were also detected corresponding to the protein tetramer. In contrast, in the presence of biotin only the protein tetramer (54353.3 ± 0.4 Da; charge states 14+ to 16+) was detected (Figure 3B); addition of a 3-fold molar excess of biotin increased the tetramer mass by ~980 Da consistent with the saturative binding of four biotin molecules. We also tested whether the protein tetramers could be stabilized at a higher ionic strength. Indeed, a mass spectrum acquired in 500 mM ammonium acetate, in the absence and presence of biotin (Figure 3C-D), revealed only the presence of protein tetramer. No protein monomers or dimers were detected under these conditions. The dynamic oligomeric state of zebavidin was further confirmed by analytical gel filtration in two different buffer systems and three different NaCl concentrations (0 mM, 100 mM, 650 mM). In a sodium phosphate buffer (50 mM NaH_2PO_4/Na_2HPO_4, 2 mM EDTA, pH 7) the protein eluted in a single sharp peak, independent of the salt concentration or the presence of biotin (Figure S3). The determined molecular masses (38.9-47.3 kDa) (see also Table S2) were comparable to that of bacterially expressed chicken avidin, which displayed a size of 39.5 and 38.2 kDa in the absence and presence of biotin, respectively. The determined masses were thus smaller than the theoretical masses of the tetramers. This behaviour has also been observed for the other bacterially expressed avidins, including biotin-binding protein A (BBP-A) [27] and avidin-related protein 4 (AVR4) [33], all of which have an apparent molecular mass of approximately 40 kDa in gel filtration analysis. This behaviour is believed to be caused by column interaction due to a high isoelectric point of the proteins (pI > 7.5). This is supported by AVR2, an avidin-related protein with an isoelectric point of 4.7, which displays the molecular mass of 55.1 kDa in gel filtration analysis, which is close to the value as determined by MS analysis (57.2 kDa) [33]. For conditions without sodium chloride, the elution peak was delayed by 2 ml. This delay was found not to be due to a difference in the size of the protein, but rather due to the differences in the performance of the column; low ionic strength
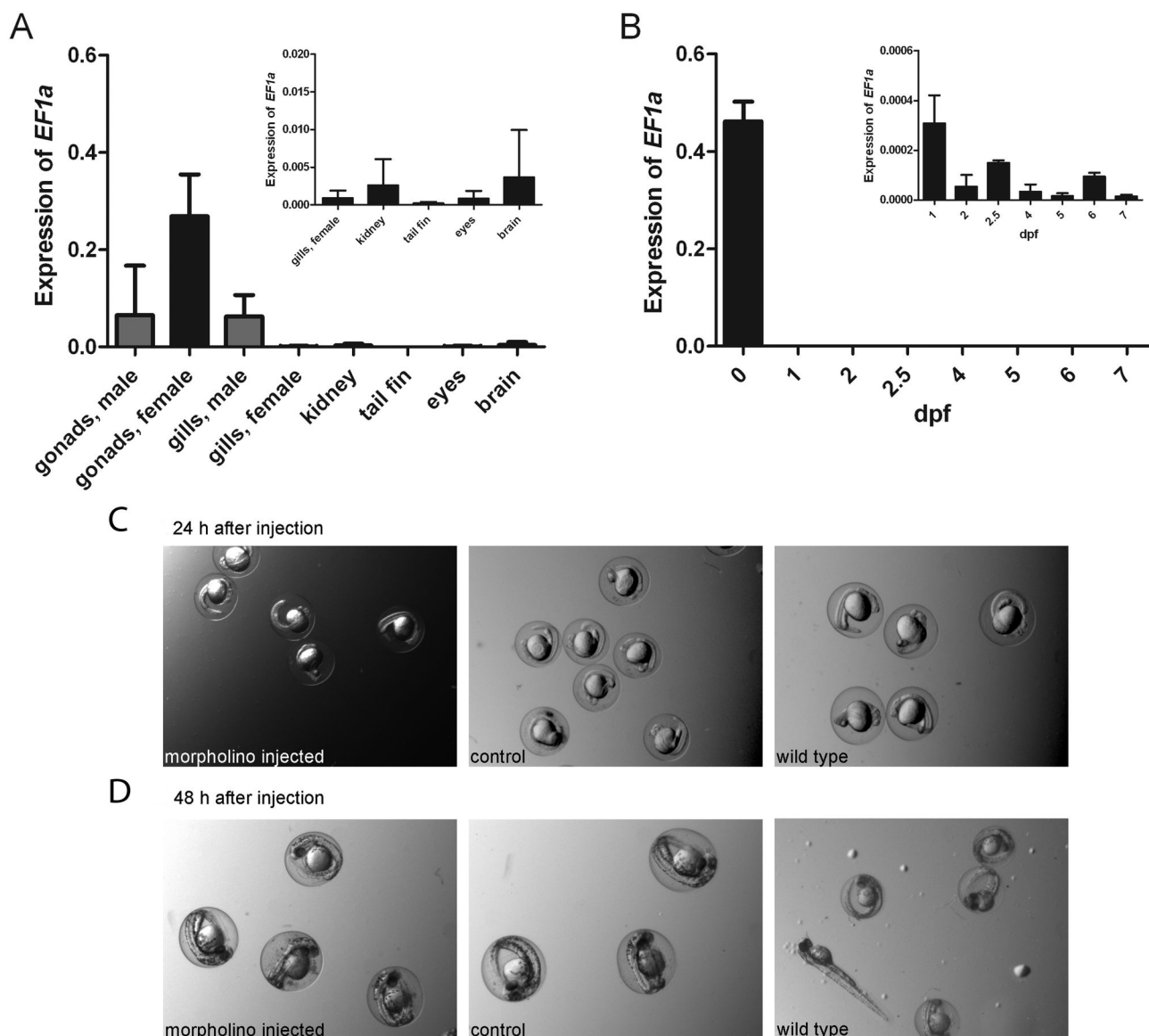
Figure 2. Expression of zebavidin in zebrafish. Relative zebavidin mRNA quantification was performed with qRT-PCR (A) in adult zebrafish tissues and (B) in developing zebrafish (0-7 days post fertilization (dpf)). Expression analysis of each tissue in (A) are shown as the mean (error bar = SD) of the biological replicates. The two-sample t-test was used to compare the expression between female gonads and female gills (P = 0.002), male gonads and male gills (nonsignificant, NS), female gonads and male gonads (P = 0.023) and female gills and male gills (NS). Expression analysis of the developmental samples in (B) are shown as the mean (SD) of the biological replicates on 0 dpf and as the mean (SD) of the technical replicates on 1-7 dpf. All of the developmental samples were pooled from 15 to 50 individuals depending on the age of embryos. The two-sample t-test was used to compare the expression of 0 dpf and 1 dpf samples to later time points. A statistically significant difference in expression was found between the 0 dpf sample and all of the other samples (P < 0.0001 in all comparisons) and also between the 1 dpf sample and the samples of 2, 4, 5, 6 and 7 dpf (P < 0.05 in all comparisons). In both (A) and (B) the expression of zebavidin is normalized to *EF1a* expression. The insets in the upper right corners represent enlargements of parts of the actual figures. (C) Embryos, injected with a Morpholino blocking the translation initiation site or with a random control Morpholino, 24 h after injection in comparison with WT embryos. (D) Embryos 48 h after injection.

doi: 10.1371/journal.pone.0077207.g002

of the mobile phase also affected the elution of protein standard accordingly (Figure S4). Similar elution volumes were observed in ammonium acetate buffer (10 mM ammonium acetate, 0.05% β-mercaptoethanol, pH7) at salt concentrations
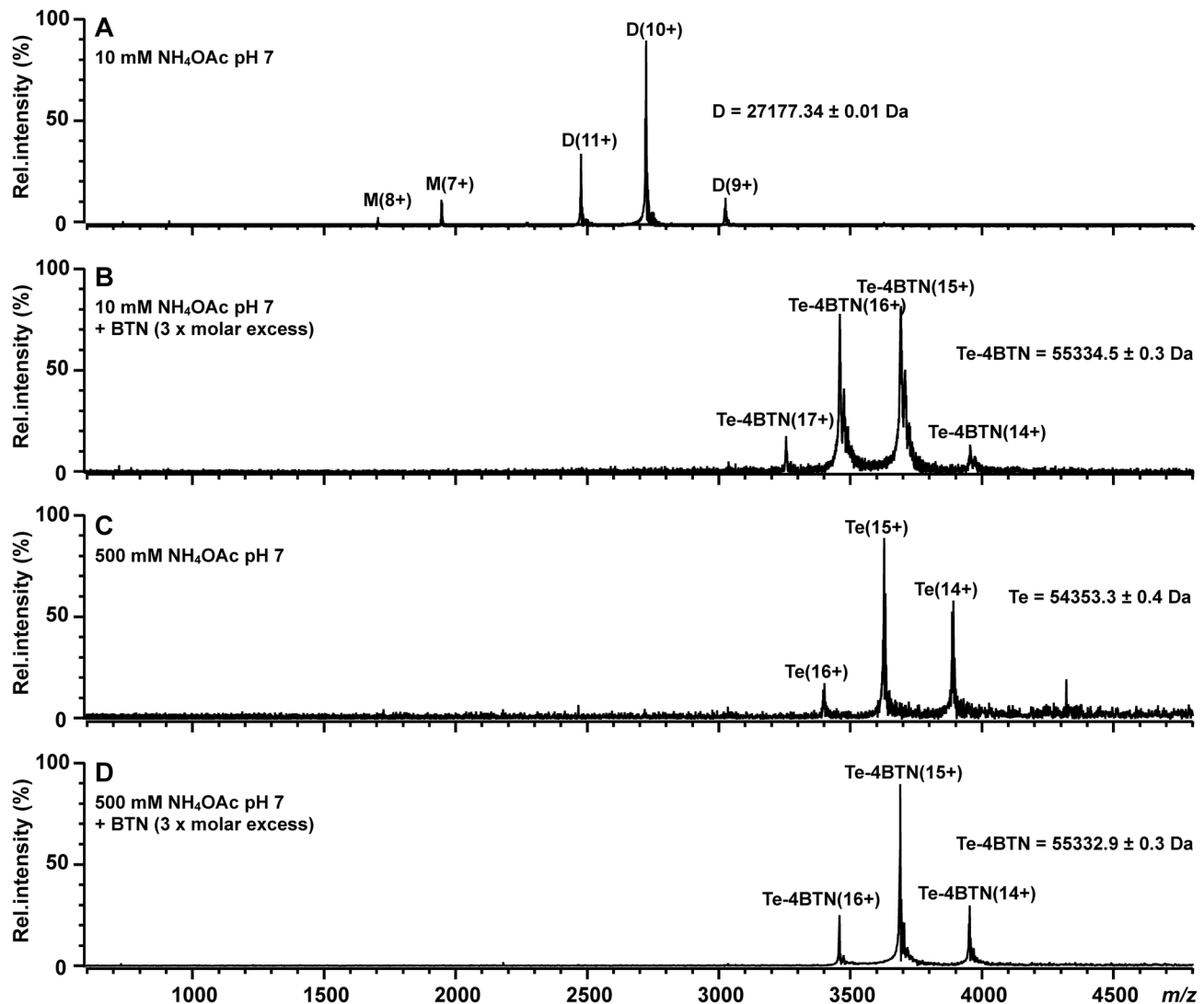
Figure 3. Native-MS analysis of zebavidin. Native mass spectra of 10 µM zebavidin in four native solution conditions: ammonium acetate buffer 10 mM (A), 10 mM + 30 µM biotin (BTN) (B), 500 mM (C) and 500 mM + 30 µM biotin (D). Letters M, D and Te correspond to the protein monomer, dimer and tetramer, respectively, and numbers denote different charge states. Molecular masses of the different protein forms (averaged over all detected charge states) are also presented.

doi: 10.1371/journal.pone.0077207.g003

of 100 and 650 mM (Figure S3 A&B). In the absence of salt, it was not possible to obtain an elution peak in the absence of biotin (Figure S3A). In the presence of biotin the peak was not sharp, displaying a maximum at 17.6 ml and a shoulder towards higher elution volumes (Figure S3B).

We also used an SDS-PAGE-based analysis to determine the oligomeric assembly of zebavidin at different temperatures. In the absence of biotin, the protein migrated as a monomer, even at ambient temperatures. In the presence of biotin, at ambient temperatures, the protein did not migrate as a perfect band, but it instead migrated in a smeared band indicating that the oligomer disassembled within the timescale of the analysis. The smeared band was visible up to a temperature between 60

°C and 70 °C, after which only monomeric species were observed (Figure S5). In comparison, chicken avidin exists as a clear tetramer up to 60 °C in the absence of biotin and up to 90 °C in the presence of biotin in this assay [55].

Thermal stability

The thermal stability of zebavidin was analysed by DSC. The thermal transition midpoint (T_m) in the absence and presence of biotin were 67.8 °C and 80.0 °C, respectively (Figure 4A&B). These values are considerably lower than those measured previously for chicken avidin (83.5 °C and 117.0 °C (+biotin)) [56] and for bacterial streptavidin (75.0 °C and 112.0 °C (+biotin)) [57]. The moderate increase in T_m (12.2 °C) in the

presence of biotin also suggests a lower binding affinity than that observed for avidin, which is dramatically stabilised by the addition of biotin (ΔT_m = 41.8 °C). In comparison, ΔT_m = 21.2 °C have been measured for AVR2 in a DSC analysis after the addition of biotin [33]. According to our knowledge, zebavidin has the lowest thermal stability of the "natural" avidins characterised to date. Another avidin with a substantially reduced thermal stability is BBP-A. This protein exhibited two peaks in the DSC analysis in the absence of biotin at 51 °C and 68 °C, whereas biotin increased the T_m value to 103.4 °C [27].

In order to analyse the influence of ionic strength on the thermal stability, T_m values were determined at different sodium chloride concentrations in either sodium phosphate or ammonium acetate buffer. With both buffer systems, clear trends could be observed with increasing salt concentrations (Figure 4C, see also Table S3 and Figure S6). In the absence of biotin, the thermal stability increased slightly with increasing salt concentration in both buffer systems. Surprisingly, in the presence of biotin the thermal stability decreased with increasing salt concentrations in ammonium acetate buffer, whereas the thermal stability stayed constant in the sodium phosphate buffer.

Zebavidin biotin binding

In fluorescence spectroscopy, affinity for a fluorescent biotin conjugate was determined by titration of the protein into the dye. The equilibrium dissociation constant (K_d) between zebavidin and the fluorescent biotin conjugate ArcDia BF560 was $1.4 \cdot 10^{-7}$ M. In comparison, the affinity of avidin to the probe was too high to be determined by this method (Figure 5A). Furthermore, zebavidin showed rapid and complete dissociation of the fluorescent biotin conjugate ArcDia BF560 at 25 and 50 °C after the addition of free biotin (>90% of the complex dissociated within less than ten seconds, suggesting a k_{diss} value in the range of 0.1 - 1 s^{-1}, data not shown). It was therefore not possible to determine the dissociation rate constant accurately using this method. In contrast, the dissociation rate constant of chicken avidin to the fluorescent biotin conjugate has been determined to be $2.3 \cdot 10^{-5}$ s^{-1} at 25 °C [20].

In the radioactive biotin ([³H]biotin) dissociation study, a dissociation rate constant k_{diss} of $6.5 \cdot 10^{-3}$ s^{-1} was measured for zebavidin. This is ~100,000-fold higher than the dissociation rate of [³H]biotin determined for chicken avidin by extrapolation from measurements performed at higher temperatures using a global fit model ($5.0 \cdot 10^{-8}$ s^{-1}) [33]. An indication of the reduced affinity of zebavidin for [³H]biotin relative to the affinity of other avidins was also found when the assay was initialised: the amount of free ligand immediately prior to the addition of cold biotin was 28% in the zebavidin assay and ~3% in the avidin assay. In addition, the majority of [³H]biotin was released from zebavidin within five minutes after the addition of a 1000-molar excess of cold biotin, and the release reached 100% within six hours (Figure 5B). For avidin, the dissociation of [³H]biotin was comparably lower, with only 10% of the probe being released 24 hours after addition of a 1000-molar excess of cold biotin.

The dissociation rate of conjugated biotins from avidins is faster because the conjugate prevents interactions between the
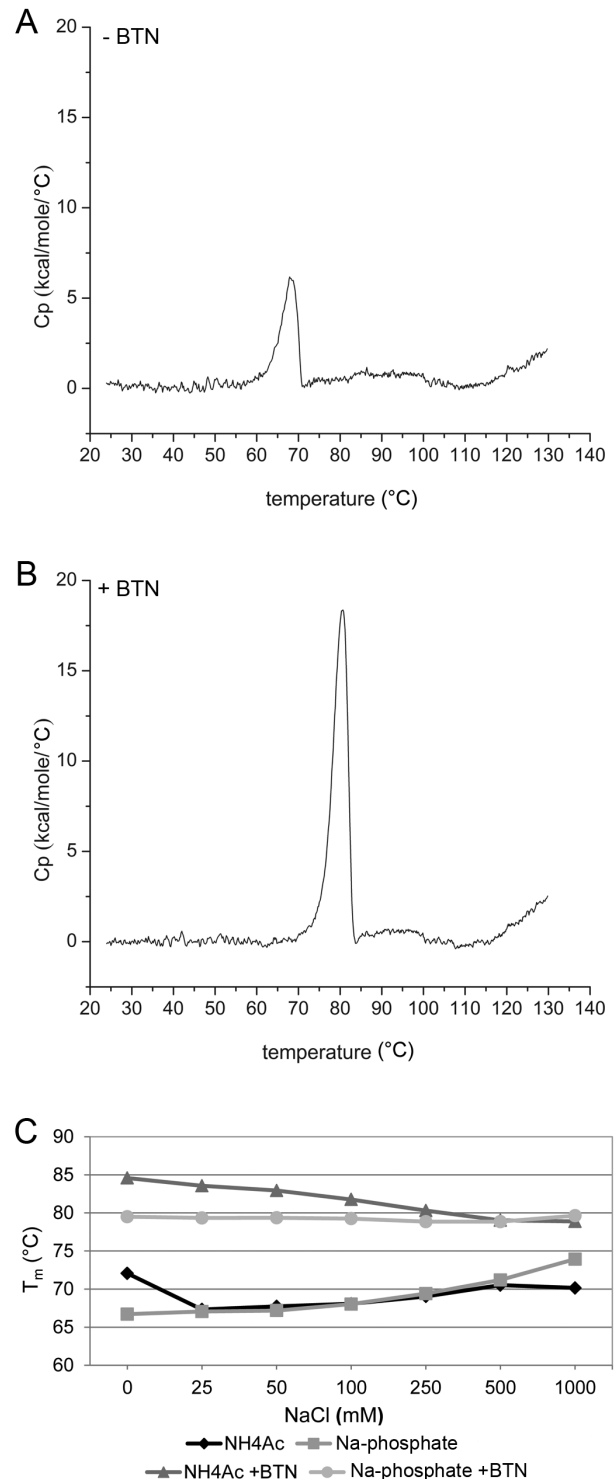
Figure 4. Thermal stability of zebavidin. DSC thermogram of 30 µM zebavidin in the absence (A) and in the presence (B) of 90 µM biotin (BTN). Protein was analysed in 50 mM Na_2HPO_4/NaH_2PO_4, 100 mM NaCl, 2 mM EDTA, pH 7. (C) Thermal transition midpoints of zebavidin as a function of the sodium chloride concentration in ammonium acetate (NH_4Ac) or sodium phosphate buffer (Na-phosphate).

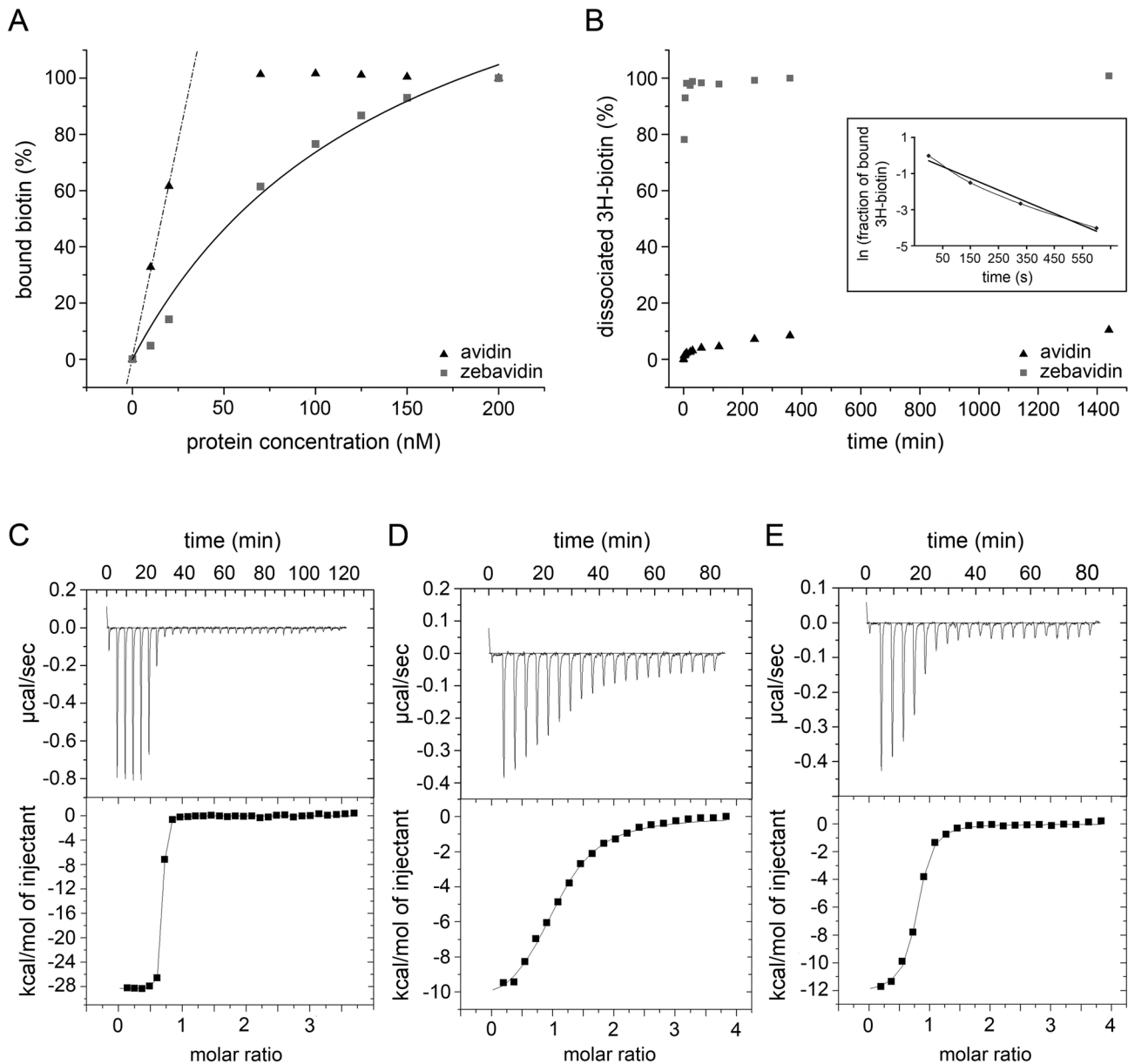doi: 10.1371/journal.pone.0077207.g004

Figure 5. Biotin binding of zebavidin. (A) Determination of the binding affinity to fluorescently labelled biotin. The amount of bound ligand was determined by following the quenching of the fluorescence of the probe. A non-linear regression has been used to fit a binding curve to the zebavidin data. Avidin showed a very high affinity to the probe, and the line has been fitted to the first data points in order to visualise the tight binding at lower protein concentrations. (B) The dissociation of radio-labelled [³H]-biotin from the protein-ligand complex after addition of cold biotin measured at different time points. The inset shows the fit for the first data points of zebavidin used for the determination of the dissociation rate constant. (C) ITC binding thermogram of ligand titration to zebavidin. (D) Desthiobiotin titration to zebavidin. (E) Biotin titration to zebavidin-desthiobiotin. (C-E) Top panel: raw ITC data. (C-E) Bottom panel: binding isotherm derived from integrated heats.

doi: 10.1371/journal.pone.0077207.g005

residues of the L3,4-loop and biotin. The structural details behind this phenomenon have been discussed in previous publications [58,59].

Isothermal titration calorimetry analysis of biotin binding to zebavidin showed an affinity constant in the nanomolar range, which lies at the sensitivity limit of the instrument and therefore resulting in rather high error values of over 35% (Figure 5C and Table 2). Because of the low reliability of ITC for quantification of extremely tight binding, a competitive binding experiment was performed using desthiobiotin. This analysis gave a similar

Table 2. Biotin-binding constants and enthalpies determined by ITC at 40 °C.

protein	ligand	K_a	K_d	ΔH	ΔS	ΔG
		$(\times 10^7 \text{ M}^{-1})$	$(\times 10^{-9} \text{ M})$	(kcal/mol)	(cal/molK)	(kcal/mol)
zebavidin	biotin	19.5/42.3	5.13/2.36	-26.3/-28.3	-46.1/-51.0	-11.9/-12.4
zebavidin	desthiobiotin	0.15/0.14	660/700	-11.1/-11.3	-7.1/-8.0	-8.9/-8.8
zebavidin- desthiobiotin	biotin	20.8/22.6	4.8/4.4	-22.9/-22.6	-38.1/-38.3	-11.9/-12.0

Values from two independent measurements are displayed.
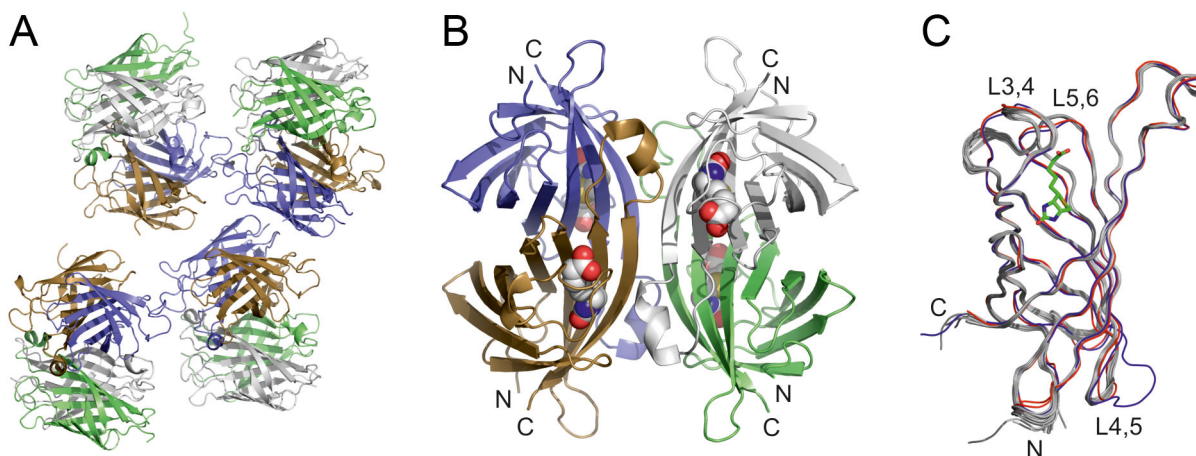
doi: 10.1371/journal.pone.0077207.t002

Figure 6. X-ray structure of zebavidin. A cartoon model of the contents of the asymmetric unit (A) and a tetramer (B) are coloured as follows: subunit I, blue; subunit II, green; subunit III light grey; and subunit IV, brown. (C) Ribbon representation of superimposed chains A-P of zebavidin (light grey), chain A of chicken avidin (blue) [PDB:1AVD] and chain A of xenavidin (red) [PDB:2UYW]. The biotin molecules of zebavidin are shown as space-filling models (B) or as a stick model (C). The N-termini (N) and C-termini (C) are indicated (B-C), as well as the L3,4-loop, L4,5-loop and L5,6-loop (C).

doi: 10.1371/journal.pone.0077207.g006

affinity constant, but with substantially decreased deviation in between two measurements (Figure 5D&E and Table 2).

X-ray structure of zebavidin

Zebavidin crystallized in the space group $P2_12_12$ and, surprisingly, contained 16 monomers (four homotetramers) in the asymmetric unit (Figure 6A; see Table 1 for structure determination statistics). The homotetramer of zebavidin (Figure 6B) resembles that of chicken avidin and is a typical "dimer of dimers". Each of the 16 monomers in the asymmetric unit is structurally highly similar to each other (Figure 6C) and the Cα atoms of the monomers superimpose with a root mean square deviation (rmsd) of less than 0.3 Å; the tetramers superimpose with the same accuracy. When compared to other eukaryotic avidin structures, chicken avidin [PDB: 1AVD, 2AVI] [52,60] and xenavidin [PDB: 2UYW] [10], the rmsd (Cα-atom positions) is 0.7 Å and 0.6 Å, respectively. The sequence identity between zebavidin and chicken avidin and xenavidin is less than 40%. The major structural differences are found at the biotin-binding loop, the L3,4-loop, which is three or more residues longer in zebavidin than in the other known avidin structures; the sequence and conformation of the L3,4-loop is

also unique and especially affects the conformation of the adjacent L5,6-loop, which, in turn, is one amino acid shorter than in other avidin structures (Figure 6C, Figure 7A-D). When compared to its closest homolog, chicken avidin, there are clear differences in the L4,5-loop of zebavidin, too.

The unique L3,4-loop and L5,6-loop affect the mode of biotin binding in zebavidin (Figure 7A), even though most biotin-binding residues are conserved in zebavidin when compared to other avidins. Highly conserved residues include Asn14, Leu16, Ser18, Tyr33, Thr35, Val37, Ser76, Trp78, Trp96, Trp109 (from another subunit), and Asp117 that surround the bicyclic ring moiety (ureido/tetrahydroimidizalone ring fused with a tetrahydrothiophene ring) of biotin; they also have very similar conformations as the equivalent residues in the structures of chicken avidin [PDB: 1AVD, 2AVI] and xenavidin [PDB: 2UYW]. Even though Asp117 is replaced in avidin/ xenavidin by Asn118/Asn117 and Ser76 by Thr77/Thr76, and Trp78 by Phe79 in avidin, these substitutions are common within the avidins. In contrast, Glu38, Ser39, Thr40, Ala43, Ala44, Trp70, Ser74 and Leu98, which surround the valeric acid moiety of biotin, mostly have unique interactions (Figure 1B; Figure 7A-D). For example, the side-chain oxygen atom of
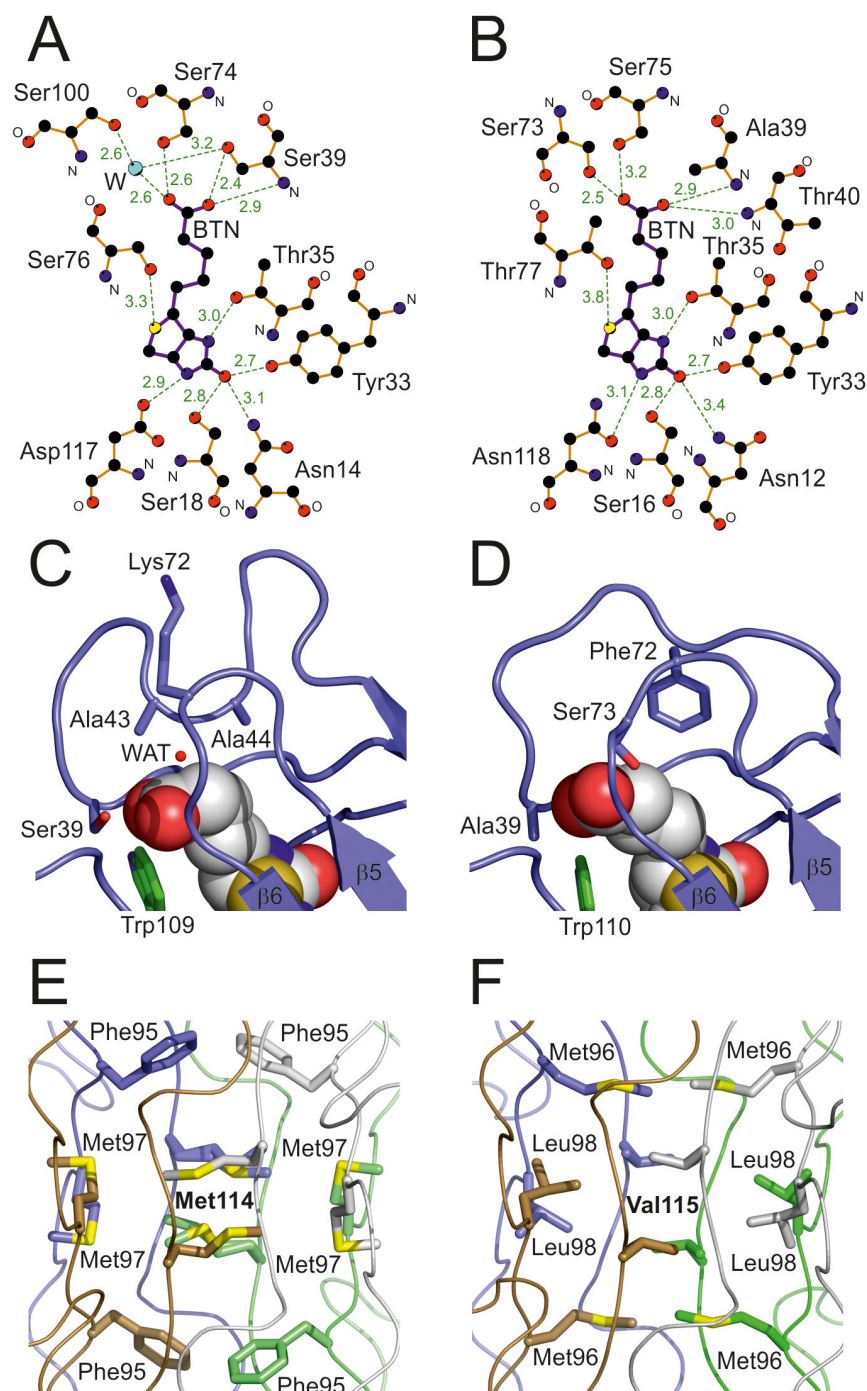
Figure 7. Unique structural properties of zebavidin. (A-B) A diagram showing hydrogen bonds to biotin (BTN) in zebavidin (A) and chicken avidin [PDB:1AVD] (B). Nitrogen atoms are coloured blue, oxygen atoms red, carbon atoms black and sulphur atoms yellow. A structural water molecule (W) is shown in cyan. The hydrogen bonds are drawn as dashed green lines and their distances shown in ångströms. The main-chain nitrogen (N) and oxygen (O) atoms are labelled. (C-D) Side chains of selected residues lining the valeric acid moiety of biotin in (C) zebavidin and (D) chicken avidin [PDB:1AVD] are shown as blue stick models, the oxygen and nitrogen atoms are coloured red and dark blue, respectively. The side chain of a tryptophan residue from a neighbouring subunit is coloured green. A structural water molecule (WAT) is indicated by a small red sphere (C). Biotin molecules are shown as space-filling models and the β5 and β6 strands are labelled. (E-F) The tetramer core of zebavidin (E) and chicken avidin (F) is shown as ribbon models. The subunits I-IV are coloured as in Figure 6A&B. Side-chain atoms of selected residues are shown as stick models; sulphur atoms are indicated by yellow colour.

Ser39 of zebavidin may form a hydrogen bond with one oxygen atom of the valeric acid moiety of biotin (or, depending on the chain, with a side-chain nitrogen atom of Arg113); in avidin/xenavidin a hydrophobic residue Ala39/Leu41 is found. The Cα-atoms of Ala43 and Ala44 (L3,4-loop) point towards the carbon atoms of the valeric acid moiety, Ala44 occupying the space that is occupied by the side chain of Phe72/Phe71 of the L5,6-loop in avidin/xenavidin (Figure 7C&D). The L5,6-loop is shorter than in other avidins and a residue equivalent to Phe72 in avidin is missing, affecting the positioning and conformation of Lys72, which is not in direct contact with biotin even though the equivalent residue Ser73/Ser72 in avidin/xenavidin is hydrogen bonded to biotin. However, in zebavidin a structural water molecule is found at a position matching that of the side-chain oxygen atom of Ser73 of avidin suggesting that this water molecule, and at least two other structural water molecules near the valeric acid moiety, may be important for biotin binding (Figure 7A&C).

The subunit interfaces of zebavidin have many interactions unique within avidin proteins, the most characteristic ones being located at the core of all subunits (Figure 7E&F). The electron density maps indicates high thermal motion and the presence of alternative rotamers for several out of the sixteen Met114 residues found in the cores of four tetramers of the asymmetric unit. The equivalent residue in avidin/xenavidin is Val115/Val114. Moreover, Phe95 (weak electron density for many of the side chain atoms) and Met97 are located near the "Met114-core". Together with Met114 and Glu116, they are the key residues at the subunit 1-2 and 1-3 interfaces (IF1,2 and IF1,3 interfaces; subunit numbering according to Livnah et al, 1993 {{17 Livnah,O. 1993;}}); these residues are also unique to zebavidin. The residues equivalent to Phe95, Met97 and Glu116 in zebavidin are Met96/Met95, Leu98/Leu97 and Ile117/Ala116 in avidin/xenavidin. To our knowledge, Phe95, Met97, Met114 and Glu116 are not found in any other characterized members of the avidin protein family, suggesting that the architecture of the IF1,2 and IF1,3 interfaces of zebavidin are at least partially responsible for the observed physicochemical properties (see above), such as low thermal stability.

In addition to shared interactions with the IF1,2 and IF1,3 interfaces, the large IF1,4 interface of zebavidin is defined with many additional non-conserved interactions, too; including the salt bridge formed by residues Glu28 and Arg30, and numerous hydrophobic and weak interactions in which residues Ile53, Met55, Ser57, Thr63, Ser65, Ser67, Leu69, Ala77, Val79, Lys93, Arg99, Leu105 and Thr112 are involved. Probably the most unique feature of the IF1,4 interface is the presence of a free cysteine residue, Cys75. This cysteine may form weak hydrogen bonds with the side chain oxygen atoms of Thr63 and Ser65, or with the main chain oxygen atom of Val64, but it is not close to any other cysteine residues. None of the known other avidin structures has a cysteine at an equivalent position.

Discussion

The availability of genetic tools in zebrafish makes it an attractive model for studying the function of avidin. Although much effort has been put into investigating the biological role of avidin [61], a comprehensive understanding is still missing. The characterisation of zebavidin reported here provides new insights towards a more complete picture of the biological importance of avidins.

Besides avidin, which is present in egg white, chicken appears to have a set of different biotin-binding proteins localised in the egg yolk; these proteins have been reported to be responsible for the deposition of biotin into the egg yolk for the developing embryo [27,51,62,63]. The biotin-binding proteins are expressed in liver and show lower thermal stability and lower biotin-binding affinity in comparison to avidin [62,64]. In contrast to chicken, zebrafish appears to carry only one avidin-like gene in its genome, which brings up the question whether zebavidin fulfils the role of avidin or the biotin-binding protein. Like biotin-binding protein A [27], zebavidin has low thermal stability and lower biotin-binding affinity. However, the highest expression of zebavidin was observed in the gonads of female fish and no expression could be observed in the liver. This suggests that zebavidin could have a functional role similar to avidin, possibly functioning as an anti-microbial agent. It is possible that zebrafish is missing a biotin-binding protein that is responsible for the transport of biotin to the oocyte and that the transport of the vitamin happens through a different mechanism. However, not much is known so far about the transport and uptake of vitamins by the developing oocyte [65]. Our results indicate that zebavidin is not necessarily needed during the development of the embryo and might not be expressed during embryogenesis (Figure 2).

Differences in the biotin-binding properties between zebavidin and the chicken avidin were observed already during the purification stage - it was not possible to purify zebavidin with 2-iminobiotin affinity chromatography. Avidin-like proteins with low or no affinity towards 2-iminobiotin have also been found previously, such as AVR2 having negligible affinity for 2-iminobiotin [33]. In contrast to avidin, which sticks almost irreversibly to the biotin-coated affinity resin, both zebavidin and AVR2 can be eluted from the resin with a relatively mild treatment. Moreover, the biotin complexes of zebavidin and AVR2 have similar dissociation rate constants at 22 °C: k_{diss} = $6.5 \cdot 10^{-3}$ s^{-1} for zebavidin, and k_{diss} = $1.5 \cdot 10^{-4}$ s^{-1} for AVR2, which are both substantially higher than that determined for chicken avidin (k_{diss} = $5.0 \cdot 10^{-8}$ s^{-1}) [33]. Structural differences in the L3,4- and L5,6-loop as well as the differences in the residues in contact with biotin, especially hydrogen bonding to the valeric acid moiety of biotin, are likely to have a cumulative impact on the observed lower biotin-binding affinity of zebavidin in comparison to other avidins. The observed oligomeric instability (see below) may contribute to the decrease in the biotin-binding affinity, too.

Most of the avidins analysed to date form tetramers in solution. Our native-MS and gel filtration analysis indicated that zebavidin forms stable tetramers in solutions having an appropriate ionic strength, but dissociates into monomers and

dimers in the low ionic strength. The addition of biotin also stabilised the tetrameric assembly and only the tetrameric form was detected in the presence of biotin. Such transient oligomerization has not been described for other natural avidins. For example, rhizavidin was found to be exclusively dimeric, both in the absence and presence of biotin [13] and also at a lower ionic strength (e.g. in 10 mM ammonium acetate buffer). Bradavidin II has been found to form oligomers as a function of protein concentration, but no clearly defined oligomeric states were observed and biotin had no effect on the oligomeric state [66]. However, transient oligomerisation was observed for mutant avidins. Monomeric streptavidin mutants containing either mutation Ser45Ala or Asp128Ala were stable monomers in the absence of biotin, but in the presence of biotin, they partially formed tetramers [67]. Similarly, a monomeric recombinant avidin mutant with mutations Asn54Ala, Asn69Ala, Met96Ala, Val115Ala and Ile117Ala was tetrameric in the presence of biotin [68]. At the structural level, the Met114 residues forming the core of all tetramers in the crystal structure of zebavidin, together with the IF1,2 interface residue Phe95 and the IF1,3 interface residue Met97, may be responsible for the transient oligomeric behaviour that we observed experimentally. The methionine side-chain is long, unbranched and inherently very flexible. The four Met114 and four Phe95 residues at the core of the tetramer showed high thermal motion, which may lead to instability of the zebavidin oligomers. The free and relatively polar Cys75 in nonpolar environment at the IF1,4 interface of zebavidin, together with several other unique interactions found at the same interface, may also weaken the quaternary structure, enabling transient oligomeric states under different buffer conditions.

The observed thermal stability of zebavidin was lower than that observed for other avidins. Like the transient oligomeric stability, the low thermal stability can also be explained by the unique subunit architecture of zebavidin, as described above. In addition, the dimer-dimer interface contains two hydrogen-bonded pairs of negatively charged glutamic acid residues (Glu116), which we propose would contribute to the stability of the zebavidin tetramers. Even though this site is one of the key interactions sites at the IF1,3 subunit interface of all tetrameric avidins characterized to date, it is not well conserved. Interestingly, mutation of the corresponding site in streptavidin from histidine to a negatively charged aspartic acid (His127Asp) disrupted the dimer-dimer interface [69]. Other previously reported mutagenesis studies have also shown that the subunit interfaces are important for the stability of the avidin. For example, chicken avidin was stabilised by the Ile117Tyr mutation, which resulted in a more compact packing of the side chains at the IF1,3 interface [70]. Analogously, introducing a disulphide bond between subunits 1 and 3 by the mutation Ile117Cys led to increased thermal stability, too [56]. In contrast, the IF1,2 interface mutant Met96His substantially destabilised the avidin tetramer [71].

The low thermal stability of zebavidin might reflect the temperature of the natural living environment of zebrafish. Whereas chicken has a body temperature of approximately 40 °C, zebrafish lives at a relatively constant temperature of approximately 25 °C. The eggs of chicken and zebrafish may experience even more deviations in terms of the temperature. Therefore, one could speculate that zebrafish avidin does not need to be as thermostable as chicken avidin in order to resist the changes in its external environment.

Conclusion

Zebavidin is an avidin-like protein from zebrafish and highly expressed in the gonads of adult fish but our preliminary results suggest that it is not critical for the development of the embryo. Our results do not exclude a potential role for zebavidin as an anti-microbial agent as suggested for chicken avidin [5,6,72], and additional experiments with e.g. bacterial infected embryos and adult zebrafish would be needed to fully elucidate the physiological role of zebavidin. Structural and functional characterisation confirmed the hypothesis that zebavidin is a new member of the avidin protein family and binds tightly to biotin, even though the biotin-binding affinity of zebavidin is substantially lower than that measured for avidin or streptavidin. Moreover, the oligomeric state of zebavidin seems to be unstable at least in conditions with low ionic strength, which could be explained by the unique subunit interface architecture of the protein.

Supporting Information

Figure S1. Isolation of zebavidin from zebavidin oocytes. SDS-PAGE of fractions taken during protein isolation with biotin sepharose. Oocytes from five different individuals were analysed (1-5). R1: biotin sepharose incubated with wash fraction before homogenisation of oocytes; R2: biotin sepharose incubated with clarified homogenised oocytes after one wash; W3: wash fraction of R2; FT1: non- bound proteins from R1; FT2: non- bound proteins from R2. Protein bands A-D, which were analysed by LC-MS/MS, are indicated by a black rectangle. Selected bands A and B came from samples taken during the isolation process from the same sample of oocytes. Selected bands C and D resulted from another sample of oocytes. For band A and D, biotin sepharose was incubated with the wash fraction before sonication of the oocytes and represent biotin-binding protein that is located in the oviduct of the mother. For band B, biotin sepharose was incubated with the ruptured oocytes and represent biotin-binding protein that is located inside the oocyte. Band C was taken from the supernatant fraction after incubation of wash fraction with biotin sepharose and represents non biotin-binding proteins. Zebavidin has two potential glycosylation sites at position 22 and 35. The reason for the observed unsharp zebavidin bands might therefore partly be due to heterogeneous glycosylation.
(TIF)

Figure S2. Mass spectrometry analysis of zebavidin. ESI FT-ICR mass spectra of 10 µM zebavidin in denaturing solution conditions (acetonitrile/water/acetic acid 49.5:49.5:1, v/v, pH 3.2).
(TIF)

Figure S3. Analytical gel filtration elution diagrams of zebavidin in different buffer systems and salt concentrations. Elution chromatograms in 10 mM ammonium acetate, pH 7 (NH$_4$Ac) with 0, 100 or 650 mM NaCl in the absence (**A**) and presence (**B**) of biotin (BTN) and in 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7 (Na-phosphate) with different salt concentrations in the absence (**C**) and presence (**D**) of biotin.
(TIF)

Figure S4. Analytical gel filtration elution diagrams of standard proteins in different conditions. (**A**) In 10 mM ammonium acetate, pH 7 with 0, 100, 650 mM NaCl (NH$_4$Ac). (**B**) In 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, pH 7 with 0, 100, 650 mM NaCl (Na-phosphate).
(TIF)

Figure S5. Oligomeric state of zebavidin in dependence of temperature. SDS-PAGE of chemically acetylated zebavidin incubated at variant temperatures in the absence and presence of biotin. M: molecular weight marker (kDa).
(TIF)

Figure S6. Thermal stability of zebavidin in dependence of sodium chloride concentration. (**A**) in 10 mM ammonium acetate buffer (NH$_4$Ac) in the absence of biotin (-BTN). (**B**) in the presence of biotin (+BTN). (**C**) in 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$ buffer (Na-phosphate) in the absence of biotin. (**D**) in the presence of biotin. Zebavidin concentration of 20 μM and biotin concentration of 60μM was used in all measurements.
(TIF)

Table S1. Peptides from LC-MS/MS analysis matching zebavidin sequence.

(DOCX)

Table S2. Elution volumes and molecular weights (MW) of zebavidin in different conditions obtained by analytical gel filtration.
(DOCX)

Table S3. Thermal stability of zebavidin as a function of sodium chloride concentration obtained by DSC.
(DOCX)

Author Contributions

Conceived and designed the experiments: BT JM JJ HT M. Parikka M. Pesu TA VH. Performed the experiments: BT JZ MO SK M. Parthiban JL JJ M. Parikka HT TA. Analyzed the data: BT JZ MO SK M. Parthiban JM JL JJ M. Parikka HT TA VH. Contributed reagents/materials/analysis tools: MR M. Pesu MJ MK VH. Wrote the manuscript: BT MO SK JM JJ M. Parikka HT M. Pesu MJ MK TA VH.

References

1. Green NM (1975) Avidin. Adv Protein Chem 29: 85-133. doi:10.1016/S0065-3233(08)60411-8. PubMed: 237414.
2. Chaiet L, Wolf FJ (1964) The Properties of Streptavidin, a Biotin-Binding Protein Produced by Streptomycetes. Arch Biochem Biophys 106: 1-5. doi:10.1016/0003-9861(64)90150-X. PubMed: 14217155.
3. Diamandis EP, Christopoulos TK (1991) The biotin-(strept)avidin system: principles and applications in biotechnology. Clin Chem 37: 625-636. PubMed: 2032315.
4. O'Malley BW (1967) In vitro hormonal induction of a specific protein (avidin) in chick oviduct. Biochemistry 6: 2546-2551. doi:10.1021/bi00860a036. PubMed: 6058123.
5. Elo HA, Jänne O, Tuohimaa PJ (1980) Avidin induction in the chick oviduct by progesterone and non-hormonal treatments. J Steroid Biochem 12: 279-281. doi:10.1016/0022-4731(80)90279-4. PubMed: 7421213.
6. Elo HA, Räisänen S, Tuohimaa PJ (1980) Induction of an antimicrobial biotin-binding egg white protein (avidin) in chick tissues in septic Escherichia coli infection. Experientia 36: 312-313. doi:10.1007/BF01952296. PubMed: 6989624.
7. Sreenivasan R, Cai M, Bartfai R, Wang X, Christoffels A et al. (2008) Transcriptomic analyses reveal novel genes with sexually dimorphic expression in the zebrafish gonad and brain. PLOS ONE 3: e1791. doi:10.1371/journal.pone.0001791. PubMed: 18335061.
8. Wixon J (2000) Featured organism: Danio rerio, the zebrafish. Yeast 17: 225-231. doi:10.1002/1097-0061(20000930)17:3. PubMed: 11025533.
9. Lohi O, Parikka M, Rämet M (2013) The zebrafish as a model for paediatric diseases. Acta Paediatr 102: 104-110. doi:10.1111/j.1651-2227.2012.02835.x. PubMed: 22924984.
10. Määttä JA, Helppolainen SH, Hytönen VP, Johnson MS, Kulomaa MS et al. (2009) Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis. BMC Struct Biol 9: 63. doi:10.1186/1472-6807-9-63. PubMed: 19788720.
11. Takakura Y, Tsunashima M, Suzuki J, Usami S, Kakuta Y et al. (2009) Tamavidins-novel avidin-like biotin-binding proteins from the Tamogitake mushroom. FEBS J 276: 1383-1397. doi:10.1111/j.1742-4658.2009.06879.x. PubMed: 19187241.
12. Nordlund HR, Hytönen VP, Laitinen OH, Kulomaa MS (2005) Novel avidin-like protein from a root nodule symbiotic bacterium, Bradyrhizobium japonicum. J Biol Chem 280: 13250-13255. doi:10.1074/jbc.M414336200. PubMed: 15695809.
13. Helppolainen SH, Nurminen KP, Määttä JA, Halling KK, Slotte JP et al. (2007) Rhizavidin from Rhizobium etli: the first natural dimer in the avidin protein family. Biochem J 405: 397-405. doi:10.1042/BJ20070076. PubMed: 17447892.
14. Helppolainen SH, Määttä JA, Halling KK, Slotte JP, Hytönen VP et al. (2008) Bradavidin II from Bradyrhizobium japonicum: a new avidin-like biotin-binding protein. Biochim Biophys Acta 1784: 1002-1010. doi:10.1016/j.bbapap.2008.04.010. PubMed: 18486632.
15. Korpela JK, Kulomaa MS, Elo HA, Tuohimaa PJ (1981) Biotin-binding proteins in eggs of oviparous vertebrates. Experientia 37: 1065-1066. doi:10.1007/BF02085010. PubMed: 7308390.
16. Parikka M, Hammarén MM, Harjula SK, Halfpenny NJ, Oksanen KE et al. (2012) Mycobacterium marinum causes a latent infection that can be reactivated by gamma irradiation in adult zebrafish. PLOS Pathog 8: e1002944. PubMed: 23028333.
17. Rounioja S, Saralahti A, Rantala L, Parikka M, Henriques-Normark B et al. (2012) Defense of zebrafish embryos against Streptococcus

pneumoniae infection is dependent on the phagocytic activity of leukocytes. Dev Comp Immunol 36: 342-348. doi:10.1016/j.dci. 2011.05.008. PubMed: 21658407.

18. Tang R, Dodd A, Lai D, McNabb WC, Love DR (2007) Validation of zebrafish (Danio rerio) reference genes for quantitative real-time RT-PCR normalization. Acta Biochim Biophys Sin (Shanghai) 39: 384-390. doi:10.1111/j.1745-7270.2007.00283.x. PubMed: 17492136.

19. Majumder K (1992) Ligation-free gene synthesis by PCR: synthesis and mutagenesis at multiple loci of a chimeric gene encoding OmpA signal peptide and hirudin. Gene 110: 89-94. doi: 10.1016/0378-1119(92)90448-X. PubMed: 1544581.

20. Hytönen VP, Laitinen OH, Airenne TT, Kidron H, Meltola NJ et al. (2004) Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli. Biochem J 384: 385-390. doi: 10.1042/BJ20041114. PubMed: 15324300.

21. Määttä JA, Eisenberg-Domovich Y, Nordlund HR, Hayouka R, Kulomaa MS et al. (2011) Chimeric avidin shows stability against harsh chemical conditions-biochemical analysis and 3D structure. Biotechnol Bioeng 108: 481-490. doi:10.1002/bit.22962. PubMed: 20939005.

22. Airenne KJ, Oker-Blom C, Marjomäki VS, Bayer EA, Wilchek M et al. (1997) Production of biologically active recombinant avidin in baculovirus-infected insect cells. Protein Expr Purif 9: 100-108. doi: 10.1006/prep.1996.0660. PubMed: 9116491.

23. Klumb LA, Chu V, Stayton PS (1998) Energetic roles of hydrogen bonds at the ureido oxygen binding pocket in the streptavidin-biotin complex. Biochemistry 37: 7657-7663. doi:10.1021/bi9803123. PubMed: 9601024.

24. Määttä JA, Airenne TT, Nordlund HR, Jänis J, Paldanius TA et al. (2008) Rational modification of ligand-binding preference of avidin by circular permutation and mutagenesis. Chembiochem 9: 1124-1135. doi:10.1002/cbic.200700671. PubMed: 18381715.

25. Sigurskjold BW (2000) Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. Anal Biochem 277: 260-266. doi:10.1006/abio.1999.4402. PubMed: 10625516.

26. Bayer EA, Ehrlich-Rogozinski S, Wilchek M (1996) Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. Electrophoresis 17: 1319-1324. doi:10.1002/elps. 1150170708. PubMed: 8874057.

27. Hytönen VP, Määttä JA, Niskanen EA, Huuskonen J, Helttunen KJ et al. (2007) Structure and characterization of a novel chicken biotin-binding protein A (BBP-A). BMC Struct Biol 7: 8. doi: 10.1186/1472-6807-7-8. PubMed: 17343730.

28. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Crystallogr 26: 795-800. doi:10.1107/S0021889893005588.

29. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658-674. doi:10.1107/S0021889807021206. PubMed: 19461840.

30. Collaborative Computational Project, Number 4 (1994); the CCP4. (1994) suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760-763. doi:10.1107/S0907444994003112. PubMed: 15299374.

31. Potterton E, Briggs P, Turkenburg M, Dodson E (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr 59: 1131-1137. doi:10.1107/S0907444903008126. PubMed: 12832755.

32. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234: 779-815. doi:10.1006/jmbi.1993.1626. PubMed: 8254673.

33. Hytönen VP, Määttä JA, Kidron H, Halling KK, Hörhä J et al. (2005) Avidin related protein 2 shows unique structural and functional features among the avidin protein family. BMC Biotechnol 5: 28. doi: 10.1186/1472-6750-5-28. PubMed: 16212654.

34. Johnson MS, Overington JP (1993) A structural basis for sequence comparisons. An evaluation of scoring methodologies. J Mol Biol 233: 716-738. doi:10.1006/jmbi.1993.1548. PubMed: 8411177.

35. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Björklund D et al. (2004) BODIL: a molecular modeling environment for structure-function analysis and drug design. J Comput Aid Mol Des 18: 401-419. doi:10.1007/s10822-004-3752-4. PubMed: 15663001.

36. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240-255. doi:10.1107/S0907444996012255. PubMed: 15299926.

37. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126-2132. doi:10.1107/S0907444904019158. PubMed: 15572765.

38. Lamzin VS, Wilson KS (1993) Automated refinement of protein models. Acta Crystallogr D Biol Crystallogr 49: 129-147. doi:10.1107/S0108767378096270. PubMed: 15299554.

39. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol 6: 458-463. doi:10.1038/8263. PubMed: 10331874.

40. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc 3: 1171-1179. doi:10.1038/nprot.2008.91. PubMed: 18600222.

41. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35: W375-W383. doi:10.1093/nar/gkm216. PubMed: 17452350.

42. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ et al. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58: 1948-1954. doi:10.1107/S0907444902016657. PubMed: 12393927.

43. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242. doi: 10.1093/nar/28.1.235. PubMed: 10592235.

44. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE et al. (2002) The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 58: 899-907. doi:10.1107/S0907444902003451. PubMed: 12037327.

45. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1-6. doi:10.1093/protein/10.1.1. PubMed: 9051728.

46. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP. J Mol Biol 3.0: 340: 783-795

47. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S (2004) Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4: 1633-1649. doi: 10.1002/pmic.200300771. PubMed: 15174133.

48. Laskowski RA, Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 51: 2778-2786. doi:10.1021/ci200227u. PubMed: 21919503.

49. Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15: 305-308. doi:10.1093/bioinformatics/15.4.305. PubMed: 10320398.

50. Keinänen RA, Wallén MJ, Kristo PA, Laukkanen MO, Toimela TA et al. (1994) Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5. Eur J Biochem 220: 615-621. doi:10.1111/j.1432-1033.1994.tb18663.x. PubMed: 8125122.

51. Niskanen EA, Hytönen VP, Grapputo A, Nordlund HR, Kulomaa MS et al. (2005) Chicken genome analysis reveals novel genes encoding biotin-binding proteins related to avidin family. BMC Genomics 6: 41. doi:10.1186/1471-2164-6-41. PubMed: 15777476.

52. Livnah O, Bayer EA, Wilchek M, Sussman JL (1993) Three-dimensional structures of avidin and the avidin-biotin complex. Proc Natl Acad Sci U S A 90: 5076-5080. doi:10.1073/pnas.90.11.5076. PubMed: 8506353.

53. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int J Neural Syst 8: 581-599. doi:10.1142/S0129065797000537. PubMed: 10065837.

54. Gentry-Weeks CR, Hultsch AL, Kelly SM, Keith JM, Curtiss R 3rd (1992) Cloning and sequencing of a gene encoding a 21-kilodalton outer membrane protein from Bordetella avium and expression of the gene in Salmonella typhimurium. J Bacteriol 174: 7729-7742. PubMed: 1447140.

55. Laitinen OH, Airenne KJ, Marttila AT, Kulik T, Porkka E et al. (1999) Mutation of a critical tryptophan to lysine in avidin or streptavidin may explain why sea urchin fibropellin adopts an avidin-like domain. FEBS Lett 461: 52-58. doi:10.1016/S0014-5793(99)01423-4. PubMed: 10561495.

56. Nordlund HR, Laitinen OH, Uotila ST, Nyholm T, Hytönen VP et al. (2003) Enhancing the thermal stability of avidin. Introduction of disulfide bridges between subunit interfaces. J Biol Chem 278: 2479-2483. doi: 10.1074/jbc.M210721200. PubMed: 12446662.

57. González M, Argaraña CE, Fidelio GD (1999) Extremely high thermal stability of streptavidin and avidin upon biotin binding. Biomol Eng 16: 67-72. doi:10.1016/S1050-3862(99)00041-8. PubMed: 10796986.

58. Pazy Y, Raboy B, Matto M, Bayer EA, Wilchek M et al. (2003) Structure-based rational design of streptavidin mutants with pseudo-catalytic activity. J Biol Chem 278: 7131-7134. doi:10.1074/jbc.M209983200. PubMed: 12493758.

59. Hayouka R, Eisenberg-Domovich Y, Hytönen VP, Määttä JA, Nordlund HR et al. (2008) Critical importance of loop conformation to avidin-enhanced hydrolysis of an active biotin ester. Acta Crystallogr D Biol Crystallogr 64: 302-308. doi:10.1107/S0907444907067844. PubMed: 18323625.

60. Pugliese L, Coda A, Malcovati M, Bolognesi M (1993) Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 A resolution. J Mol Biol 231: 698-710. doi:10.1006/jmbi.1993.1321. PubMed: 8515446.

61. Tuohimaa P, Joensuu T, Isola J, Keinänen R, Kunnas T et al. (1989) Development of progestin-specific response in the chicken oviduct. Int J Dev Biol 33: 125-134. PubMed: 2485692.

62. White HB 3rd, Dennison BA, Della Fera MA, Whitney CJ, McGuire JC et al. (1976) Biotin-binding protein from chicken egg yolk. Assay and relationship to egg-white avidin. Biochem J 157: 395-400. PubMed: 962874.

63. White HB 3rd (1985) Biotin-binding proteins and biotin transport to oocytes. Ann N Y Acad Sci 447: 202-211. doi:10.1111/j.1749-6632.1985.tb18438.x. PubMed: 3860173.

64. White HB 3rd, Whitehead CC (1987) Role of avidin and other biotin-binding proteins in the deposition and distribution of biotin in chicken eggs. Discovery of a new biotin-binding protein. Biochem J 241: 677-684. PubMed: 3593216.

65. Lubzens E, Young G, Bobe J, Cerdà J (2010) Oogenesis in teleosts: how eggs are formed. Gen Comp Endocrinol 165: 367-389. doi:10.1016/j.ygcen.2009.05.022. PubMed: 19505465.

66. Leppiniemi J, Meir A, Kähkönen N, Kukkurainen S, Määttä JA et al. (2013) The highly dynamic oligomeric structure of bradavidin II is unique among avidin proteins. Protein Sci 22: 980-994. doi:10.1002/pro.2281. PubMed: 23661323.

67. Qureshi MH, Yeung JC, Wu SC, Wong SL (2001) Development and characterization of a series of soluble tetrameric and monomeric streptavidin muteins with differential biotin binding affinities. J Biol Chem 276: 46422-46428. doi:10.1074/jbc.M107398200. PubMed: 11584006.

68. Laitinen OH, Marttila AT, Airenne KJ, Kulik T, Livnah O et al. (2001) Biotin induces tetramerization of a recombinant monomeric avidin. A model for protein-protein interactions. J Biol Chem 276: 8219-8224. doi:10.1074/jbc.M007930200. PubMed: 11076945.

69. Sano T, Vajda S, Smith CL, Cantor CR (1997) Engineering subunit association of multisubunit proteins: a dimeric streptavidin. Proc Natl Acad Sci U S A 94: 6153-6158. doi:10.1073/pnas.94.12.6153. PubMed: 9177186.

70. Hytönen VP, Määttä JA, Nyholm TK, Livnah O, Eisenberg-Domovich Y et al. (2005) Design and construction of highly stable, protease-resistant chimeric avidins. J Biol Chem 280: 10228-10233. doi:10.1074/jbc.M414196200. PubMed: 15649900.

71. Nordlund HR, Hytönen VP, Laitinen OH, Uotila ST, Niskanen EA et al. (2003) Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. FEBS Lett 555: 449-454. doi:10.1016/S0014-5793(03)01302-4. PubMed: 14675754.

72. Nordback I, Kulomaa M, Joensuu T, Tuohimaa P (1982) Characterization of progesterone-independent avidin production of chicken tissues in culture. Comp Biochem Physiol Comp Physiol 71: 389-393. doi:10.1016/0300-9629(82)90423-6. PubMed: 6120783.