



JUHA MÄÄTTÄ

Structural and Functional Characterization
of Engineered Avidin Proteins



ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty of Medicine of the University of Tampere,
for public discussion in the 5th floor lecture hall of Finn-Medi 1,
Biokatu 6, Tampere, on June 11th, 2010, at 12 o'clock.

UNIVERSITY OF TAMPERE

ACADEMIC DISSERTATION

University of Tampere, Institute of Medical Technology
The National Graduate School in Informational and Structural Biology (ISB)
University of Jyväskylä, Department of Biological and Environmental Science
Finland

Supervised by

Professor Markku Kulomaa
University of Tampere
Finland
Docent Vesa Hytönen
University of Tampere
Finland

Reviewed by

Professor Reijo Lahti
University of Turku
Finland
Professor Antti Poso
University of Eastern Finland
Finland

Distribution

Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 40 190 9800

Fax +358 3 3551 7685

taju@uta.fi

www.uta.fi/taju

<http://granum.uta.fi>

Cover design by

Juha Siro

Acta Universitatis Tamperensis 1529

ISBN 978-951-44-8103-1 (print)

ISSN-L 1455-1616

ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 969

ISBN 978-951-44-8104-8 (pdf)

ISSN 1456-954X

<http://acta.uta.fi>

Table of Contents

Table of Contents	3
List of original communications	6
My contributions to the thesis articles	7
Abbreviations	8
YHTEENVETO	10
ABSTRACT	12
1. Introduction	14
2. Review of the literature	16
2.1 Covalent bonds and non-covalent interactions	16
2.2 Introduction to non-covalent interactions in protein-ligand binding	17
2.3 Thermodynamics of ligand binding	18
2.3.1 Enthalpy, entropy - Gibbs energy of binding	19
2.4 Water, hydration and hydrophobic effect	22
2.4.1 Hydrogen bonding	22
2.4.1.1 Furcation of hydrogen bonds	23
2.4.2 Hydrophobic effect, desolvation of water molecules from the surface of the macromolecules	24
2.4.3 Bridging water in binding interfaces	27
2.5 Ligand binding kinetics	27
2.5.1 Rate of reaction, kinetics	29
2.5.1.1 Factors influencing the rate of the reaction	32
2.5.1.2 Methods to determine equilibrium constants	33
2.6 Structural features in protein-ligand binding	36
2.6.1 Mobile domains and flexible loop regions of proteins	37
2.6.2 Co-operation between subunits	37
2.7 Avidin – a high affinity and highly stable ligand-binding protein	38
2.7.1 Structure of the ligand-binding pocket of avidin – cooperative hydrogen bonds	39

2.7.1.1	The effects of hydrated water on ligand binding in avidin.....	42
2.7.2	Ligands of avidin	43
2.7.2.1	Biotin and biotin analogues.....	44
2.7.2.2	Peptide ligands of (strept)avidin	45
2.7.2.3	Azocompounds.....	45
2.7.3	Stability of the avidin protein	48
3.	Aims of the study	50
4.	Materials and Methods.....	51
4.1	Construction of expression vectors for avidins	51
4.1.1	Production of the chimeric avidin mutant in insect cells.....	51
4.1.2	<i>E. coli</i> expression vectors	52
4.2	Protein expression and purification.....	53
4.2.1	Fermentation.....	54
4.3	Characterization of ligand binding.....	55
4.3.1	Isothermal titration calorimetry	55
4.3.2	Fluorescence emission spectroscopy assay	56
4.3.3	Radiobiotin dissociation assay.....	56
4.3.4	Surface plasmon resonance (SPR).....	57
4.4	Thermostability and solvent stability	57
4.4.1	Differential scanning calorimetry	57
4.4.2	SDS-PAGE based transition temperature analysis.....	57
4.4.3	Microplate assay	58
4.4.4	Solvent stability assay with resin linked proteins and ³ H- labelled biotin	58
4.5	Structural analysis by X-ray crystallography	59
4.5.1	Circular permuted avidin Loop3-4 and Biotin Binding Protein A (II,III).....	59
4.5.2	Chimeric avidin (I117Y) (IV).....	59
4.6	Mass spectrometry.....	59
4.7	Fast Protein Liquid Chromatography Gel Filtration	60
4.8	Computer programs.....	60
5.	Review of the Results.....	61
5.1	Chimeric avidin mutants (I, IV)	61
5.1.1	Expression vectors, fermentation and purification of recombinant proteins	61

5.1.2	Ligand-binding analyses.....	62
5.1.3	Thermal and solvent stability of avidins.....	64
5.1.3.1	Solvent stability assay with resin linked proteins and tritium labelled biotin	65
5.1.4	Structural analysis by X-ray crystallography	66
5.2	Structure, stability and ligand binding mode of BBP-A (II)	67
5.3	Rational modification of avidin by circular permutation and mutagenesis (III).....	68
6.	Discussion	71
6.1.1	Water molecules in binding events.....	71
6.2	Development of a stable chimeric avidin (I117Y).....	73
6.3	Structure-function relationships of avidin like proteins	74
7.	Summary and conclusion	78
	Acknowledgements	80
	References	82

List of original communications

This text is based on the following original communications that are referred to in text by their Roman numerals (I-IV)

- I Hytönen VP, **Määttä JA**, Nyholm TK, Livnah O, Eisenberg-Domovich Y, Hyre D, Nordlund HR, Hörhä J, Niskanen EA, Paldanius T, Kulomaa T, Porkka EJ, Stayton PS, Laitinen OH, Kulomaa MS (2005): Design and construction of highly stable, protease-resistant chimeric avidins. *J Biol Chem* 280(11):10228-33.
- II Hytönen VP, **Määttä JA**, Niskanen EA, Huuskonen J, Helttunen KJ, Halling KK, Nordlund HR, Rissanen K, Johnson MS, Salminen TA, Kulomaa MS, Laitinen OH, Airene TT (2007): Structure and characterization of novel chicken biotin-binding protein A (BBP-A). *BMC Struct Biol* 7:8.
- III **Määttä JA**, Airene TT, Nordlund HR, Jänis J, Paldanius TA, Vainiotalo P, Johnson MS, Kulomaa MS, Hytönen VP (2008): Rational modification of ligand-binding preference of avidin by circular permutation and mutagenesis. *ChemBioChem* 9(7):1124-35. *Erratum in ChemBioChem* 9(8):1181.
- IV **Määttä JA**, Eisenberg-Domovich Y, Nordlund HR, Hayouka R, Kulomaa MS, Livnah O, Hytönen VP (2010): Chimeric avidin shows good stability in harsh chemical conditions – Biochemical analysis and 3D structure. Submitted

My contributions to the thesis articles

- Article I Dr. Vesa Hytönen was mainly responsible for the planning, practical work and writing up of this study. I was responsible for a significant part of the practical work. The calorimetric analyses were done by Dr. Thomas Nyholm. Tuomas Kulomaa performed the radiobiotin dissociation analyses under the supervision of Dr. David Hyre. Professor Oded Livnah and Yael Eisenberg-Domovich participated in the structural analysis of the proteins.
- Article II Dr. Vesa Hytönen was mainly responsible for the planning and writing up of this study. I was mainly responsible for the calorimetric analyses and describing the calorimetric experiments in the article. I also participated in other parts of the practical work together with Dr. Hytönen. The differential scanning calorimetry analyses were done by Dr. Katrin Halling. The organic synthesis was done by Kaisa Helttunen under the supervision of Dr. Juhani Huuskonen. Dr. Tomi Airene was responsible for the crystallographic structural determination of the proteins.
- Article III I was mainly responsible for the planning, practical work and writing of this study. Dr. Janne Jänis carried out the mass spectrometric analyses. Dr. Tomi Airene carried out the structural analysis of the proteins. Dr. Airene and Dr. Vesa Hytönen participated in the writing of this article.
- Article IV I was mainly responsible for the planning, practical work and writing of this study. Professor Oded Livnah and Yael Eisenberg-Domovich participated in the structural analysis of the proteins. Dr. Vesa Hytönen and Professor Oded Livnah contributed to the writing of the article.

Abbreviations

3D	three-dimensional
A	amplitude
<i>A</i>	area
ASA	accessible surface area
<i>AVR</i>	avidin related gene
AVR	<i>AVR</i> encoded protein
AVD	avidin
BBP	biotin binding protein
BSA	bovine serum albumin
ChiAVD	chimeric avidin mutant
C_p	heat capacity
ΔC_pL	change in heat capacity upon ligand binding
DOTA	1,4,7,10-tetraazacyclododecane-1,4,7,10-tetra acetic acid
DSC	differential scanning calorimetry
dcAVD	dual chain avidin
H	enthalpy
HABA	2-(4'-hydroxyazobenzene) benzoic acid
ITC	isothermal titration calorimetry
K_a	association constant
k_{ass}	association rate constant
k_B	Boltzmann's constant
K_b	association constant
K_d	dissociation constant
k_{ass}	association rate constant
k_{diss}	dissociation rate constant
Da	dalton
LB	Lysogeny broth
log P	octanol- water partition coefficient
M	molar concentration
MHC	major histocompatibility complex
PCR	polymerase chain reaction
PDB	protein data bank

pI	isoelectric point
R	universal gas constant
S	entropy
SI	Système International d'Unités
SDS-PAGE	sodium-dodecyl sulphate polyacrylamide gel electrophoresis
Sf9 or Sf21	Spodoptera frugiperda cell lines
T _m	transition midpoint of heat denaturation
T _r	transition midpoint of oligomeric disassembly
UV	ultraviolet
wt	wild type
Å	ångström (0.1 nm)

YHTEENVETO

Avidiini löydettiin ensimmäistä kertaa kanasta, mutta myöhemmin sitä on todettu esiintyvän yleisesti lähes kaikissa munivien eläinlajien munissa. Avidiinin uskotaan estävän vieraiden organismien kasvua, koska se sitoo niiden kasvuun tärkeitä biotiinia, ja siten toimivan munassa puolustusproteiinina. Avidiinin kolmiulotteisen rakenteen ratkaiseminen paljasti, miksi se sitoo sen luonnollista ligandia biotiinia hyvin tiukasti. Avidiinin rakenteesta kävi ilmi, että sen ja biotiinin välille muodostuu useita vetysidoksia, sekä että ligandin sitomispaikan muoto on hyvin ideaalinen biotiinille. Vahvan sitoutumisen vuoksi avidiinin ja sen bakteerista löydetyn vastineen, streptavidiinin, käyttö monissa biotieteiden sovellutuksissa on ollut runsasta. Avidiinin tai streptavidiinin käyttöä biotieteiden sovellutuksissa kutsutaan yleisesti (strept)avidini-biotiini teknologiaksi. Molempien proteiinien käyttökelpoisuutta on edistänyt niiden kestävyys.

Tässä tutkimuksessa kehitettiin uusia avidiinin muotoja ja tutkittiin kehitettyjen uusien proteiinien ominaisuuksia. Avidiinin ja samaan geeniperheeseen kuuluvien avidiinin kaltaisten geenien koodaamien proteiinien (AVR) lämmönkestävyyksissä on eroja. Avidiinia lämpökestävämmän AVR4/5 proteiinin lämmönkestävyydelle tärkeitä alueita pyrittiin kartoittamaan kohdennetun mutageneesin avulla. Samalla haluttiin tutkia, onko mahdollista siirtää AVR4/5:n kriittisiä alueita avidiiniin ja siten parantaa sen lämmönkestävyyttä. Tuloksena syntyi kimeerinen avidiini, ChiAVD(I117Y), jonka lämmönkestävyys on ylivertainen aikaisempiin avidiiniproteiineihin nähden. Jatkotutkimuksena selvitettiin korreloiko lämmönkestävyys orgaanisissa liuottimissa selviytymisen kanssa. Havaitimme, että vaikka kimeeristä avidiinia käsiteltäisiin erilaisilla orgaanisilla liuottimilla, se säilyttää useimmiten sitomiskykynsä. Siten kimeerista avidiinia on mahdollista hyödyntää aivan uudentyyppisissä avidiini-biotiini teknologian sovellutuksissa.

Tutkimuksen toisessa osatyössä tutkittiin avidiinin β -juosteiden 3 ja 4 välisen silmukka-alueen (L3,4) vaikutusta biotiinin sekä 4'-hydroksiatzobentseeni-2-karboksyylin hapon (HABA) sitoutumiseen. Avidiinin, jonka silmukka-alue L3,4 poistettiin sirkulaarisella permutaatiolla, sitomisominaisuuksia verrattiin

isotermaalisella titraatio kalorimetrialla (ITC) villityypin avidiiniin. Silmukka-alueen poiston (L3,4) ja yhden pistemutaation (Asp118Met) avulla pystyimme heikentämään biotiinin sitoutumista ja parantamaan HABA:n affiniteettia huomattavasti. Muokatun avidiinin (cpAVD4→3(N118M)) biotiinin ja HABA:n välinen suhteellinen sitomiskyky muuttui villityypin avidiiniin nähden yli miljardikertaisesti (10^9).

Tämän työn kolmannessa osatyössä pyrittiin selvittämään kananmunan keltuaisessa esiintyvän biotiinia sitovan proteiini A:n (BBP-A) ominaisuuksia. Proteiinin röntgensädekrystallograffisesti määritetty rakenne ja ITC:llä määritetyt sitomisominaisuudet paljastivat, että BBP-A sitoo tiukasti paitsi biotiinia, myös d-biotiini-d-sulfoksidia (BSO). Proteiini muistuttaa rakenteeltaan avidiinia, mutta on lämmönkestävyydeltään heikompi.

Kokonaisuudessaan tämä monitieteellinen tutkimus selventää avidiinin ja sen kaltaisten proteiinien pienmolekyylisen sitoutumiseen vaikuttavia tekijöitä ja lisää siten tietoa avidiinin rakenteen ja toiminnan välisestä suhteesta.

ABSTRACT

Avidin was first found in chicken, but is now known to be present in the avian egg-whites or egg-jellies. Avidin is thought to work as a defensive agent in the egg preventing the growth of unwanted organisms by creating a biotin drain. It binds its natural ligand, biotin, with very tight affinity. The determinants of this tight binding have been solved with the 3D structure of the avidin-biotin complex. It revealed that there are several hydrogen bonds between the avidin and biotin molecules and that the shape of the binding pocket is well optimized for biotin. Due to their tight affinity to biotin and good stability, avidin and its bacterial analogue streptavidin, have been used in a wide range of applications in biosciences. Collectively, applications based on avidin or streptavidin are referred to as strept(avidin)-biotin technology.

The aim of this study was to develop engineered forms of avidin and to characterise their properties. Avidin related proteins 1-7 (AVRs) belong to the avidin protein family. Their properties differ from each other and interestingly AVR4/5 was previously found to be more thermostable than avidin. To study the structural reasons behind the high thermostability of AVR4/5, we used genetic engineering to transfer parts of the sequence of AVR4/5 to avidin. As a result, we produced the chimeric avidin Ile117Tyr mutant (ChiAVD(I117Y)), the most thermostable avidin so far. Furthermore, we wanted to see if the high thermal stability of chimeric avidin correlated with stability in harsh organic solvents. Chimeric avidin was found equal or superior in stability when compared to avidin or streptavidin in most of the conditions studied. Therefore, chimeric avidin may prove useful in applications where high stability is required.

The loop region between the β -strands 3 and 4 (L3,4) of avidin is important for ligand binding. This region was removed by circular permutation and the properties of the resulting deletion mutant were then compared to those of wild type avidin. The deletion of L3,4 combined with the point mutation Asp118Met (cpAVD4 \rightarrow 3(N118M)) reduced biotin binding affinity and increased affinity to 4'-hydroxyazobenzene-2-carboxylic acid (HABA). Based on isothermal titration calorimetry (ITC) analysis the relative HABA/biotin binding affinity of

cpAVD4→3(N118M) increased over a billion fold (10^9) when compared to that of wt avidin.

The third part of the study aimed at characterising the structure and properties of a novel chicken biotin-binding protein A (BBP-A). The structure of the protein in complex with the ligand revealed that the ligand-binding site was occupied by d-biotin d-sulfoxide (BSO). ITC analysis confirmed that BBP-A binds both BTN and BSO with comparable affinities, a property unique among avidins. The structure of the protein resembles that of avidin, but its thermostability is lower.

In summary, this multidisciplinary study clarified factors affecting the ligand-binding properties of avidin and novel avidin-like proteins, and therefore, increased the knowledge of the structure-function relationship of avidins. Furthermore, this work may offer new tools for avidin-biotin technology.

1. Introduction

Proteins are an essential part of living organisms. They are made of amino acids attached to each other by peptide bonds (primary structure). The resulting biopolymer, a linear polypeptide chain of amino acid residues, is folded into a three-dimensional form (secondary structure). The final structure of the protein is achieved when secondary structures are organized into a three-dimensional form (tertiary structure). Further still, folded proteins can be arranged into multi-subunit complexes, referred to as quaternary structures. The function of a protein is defined by its three-dimensional structure and is based on the recognition of other molecules such as in enzyme-substrate, antibody-antigen or protein-ligand interactions. A reversible binding reaction is the sum of the formed and broken non-covalent interactions and the dissociation of water from the binding site. The binding force, or affinity, influences the function of the protein; if binding is too strong, the biological reaction might not proceed forward, and if too weak, specific binding might not occur. For example enzymes have to recognise their substrate efficiently and binding to the active site must be tight enough for an enzyme-substrate complex to form. After the substrate has been transformed to the product, it has to be released for the enzyme to perform the process again. Therefore, binding cannot be too tight either, or catalytic efficiency would be poor.

The chicken egg-white protein **avidin** together with the analogous protein **streptavidin** from the bacterium *Streptomyces avidinii* is known to non-covalently bind the water soluble vitamin **D-biotin**, with the strongest non-covalent protein-ligand interaction found in nature (Green 1975). The reason for this tight biotin binding is not known for sure, but is thought to prevent the growth of foreign bacteria. This remarkable affinity is one reason for using avidin-biotin binding in several life science applications and is at the origin of (strept)avidin-biotin technology (Wilchek, Bayer 1990, Wilchek, Bayer 1999). Other reasons that have expanded the (strept)avidin-biotin technology are the known structures of avidin (Livnah et al. 1993b) and streptavidin (Weber et al. 1989), new ligands (Livnah et

al. 1993a, Hofstetter et al. 2000, Skerra, Schmidt 2000, Korndorfer, Skerra 2002, Hirsch et al. 2002, Lamla, Erdmann 2004, Repo et al. 2006), modifiable structure (Nordlund et al. 2004, Hytönen et al. 2005b, Nordlund et al. 2005, Laitinen et al. 2006, Laitinen et al. 2007), and stable structure of the complex (Green 1963b, Green 1975).

In this study we modified the properties of avidin by circular permutation and rational mutagenesis yielding proteins with novel binding properties. We also characterized a new avidin like protein, the biotin-binding protein A (BBP-A). In addition, we were able to produce the most thermostable avidin made so far, named chimeric avidin (I117Y). The three-dimensional structures of these proteins were solved and ligand-binding analyses were done, to explain structure-function relationships. The structure of ChiAVD(I117Y) revealed that small differences in structure are responsible for the improved thermal and solvent stabilities when compared to those of avidin or AVR4/5. These novel avidin forms have revealed important factors affecting binding or other functional properties, like stability, and therefore make a valuable contribution to avidin-biotin technology.

2. Review of the literature

Understanding the thermodynamic principles of non-covalent interactions and examples of ligand binding are the main focuses of this thesis. In addition, I will review the physicochemical properties and some modified forms of the avidin protein, which was the principle subject of interest throughout the thesis study.

2.1 Covalent bonds and non-covalent interactions

Because atoms have different electronic structures they can become bonded together in different ways. In 1916 Walter Kössel noted that stable ions tend to form when atoms gain or lose electrons in sufficient numbers so that they attain the nearest noble-gas configuration (Fessenden, Fessenden & Logue 1998). Soon after Kössels work Gilbert N. Lewis published the concept of shared electron pairs, which after a couple of years was named the covalent bond by Irwin Langmuir (Langmuir 1919). Since then these findings have led to the subsequent theories of the ionic and covalent bond. To illustrate the concepts, an ionic bond is usually considered a bond where one atom donates its outermost electron or electrons to another atom or atoms, and the resulting positive cation and negative anion attract each other. Cations and anions can rotate relative to each other without affecting the bond, and therefore, ionic bonds are non-rigid and non-directional. In a covalent bond, instead of the transfer of electrons like in an ionic bond, electrons are shared between the two nuclei, and therefore, the bond is very directional. For example in the sigma bond (σ -bond) the two electrons of overlapping orbitals form the molecular orbital that is cylindrically symmetrical around an axis joining the two nuclei. Therefore, in a covalent bond, two atoms are always oriented in a certain way relative to each other. Covalent bonds can be classified as single, double and triple bonds, which all have different kinds of properties. They form the backbones of molecules but typically do not participate directly in the recognition of ligands or the interactions

between bio-molecules. Nor do covalent bonds alone define the shape of bio-molecules like proteins, because the shape is also determined by non-covalent bonds.

Non-covalent interactions differ from covalent and ionic bonds in that they do not share or transfer electrons in the bonding event. They can be seen as variations of different kinds of electromagnetic interactions. This means that non-covalent interactions have similar properties but their nomenclature is case-dependent. Non-covalent interactions include: hydrogen bonds, both attractive and repulsive van der Waals forces, entropic force, chemical polarity, mechanical bonds, halogen bonds, aurophilicity, intercalation and stacking ($\pi - \pi$ interactions).

2.2 Introduction to non-covalent interactions in protein-ligand binding

Proteins are organic compounds made of amino acid residues arranged in a linear chain and folded into a 3D structure. Their binding properties are the essence of many fields of biological sciences. A binding reaction is a two-step process; first, two components have to recognize each other before the actual binding can take place. Binding itself can be considered the sum of non-covalent interactions between molecules. The proper function of proteins is dependent on both selective recognition and sensible binding affinity. Therefore, many fields of research aim to characterize, modify, strengthen or destroy interactions between bio-molecules. For example, in functional genomics, which uses a vast set of data from genomic projects to describe protein functions and interactions, algorithms are used to predict potential protein-protein interactions to foresee previously unknown interactions between molecules (Kanaan et al. 2009). In medicine, the resolved structures of mutated proteins behind medical conditions are used to design new drugs or to improve existing ones. These modifications are usually made in hope of discovering new non-covalent interactions, which could improve the binding affinity of drug molecules. However, too many hydrogen bonds could also be unfavourable. The amounts of hydrogen bond donor and acceptor sites in a good drug molecule are generalised by the “Lipinski’s rule of Five” (Lipinski et al. 2001). It states that poor absorption or permeation is more likely if two of the following parameters are out of

range: there are more than five hydrogen-bond donors or 10 hydrogen-bond acceptor sites, or the complex has a molecular weight greater than 500 ($MW \geq 500$ g/mol) or Log P (CLogP, lipophilicity expressed as the ratio of octanol solubility to aqueous solubility) is over five ($ClogP \geq 5$) (Lipinski et al. 2001).

In almost every study on protein-ligand binding, it is of interest to understand the overall binding mechanism, but also to know, where ligands are bound in the protein. Every protein-ligand binding pair has a specific non-covalent interaction network. Factors thought to be involved in the network are first and foremost hydrogen bonds, van der Waals interactions and entropic/hydrophobic forces. These bonds have a typical interaction energy of 1~5 kcal/mol (4-21 kJ/mol) (Lodish et al. 2000), which is significantly weaker than that of a covalent bond, 35-230 kcal/mol (150-960 kJ/mol) (Fessenden, Fessenden & Logue 1998). However, non-covalent interactions act together to produce stable and specific associations between different parts of a large molecule or between different macromolecules. In some cases it is important that the bond is stable, whereas sometimes it might be necessary to form relatively weak interactions to enable biological function. These non-covalent interactions enable proteins to recognize and bind their ligands like enzymes do their substrates or antibodies their antigens.

2.3 Thermodynamics of ligand binding

The experimental observation that energy can be neither created nor destroyed is the basic postulate of thermodynamics. Usually this means the conversion of heat to mechanical, chemical or electrical energy or vice versa. Variables like temperature, pressure and volume are the basic elements of thermodynamics and changing one of these variables results in a change in the total energy (ΔU) of a system, which in thermodynamics is called the internal energy (U). The internal energy of a system is the total kinetic energy due to the motion of molecules and the total potential energy associated with the vibration and electric energy of atoms within molecules or crystals. The mathematical statement of the first law of thermodynamics is

$$\Delta U = q + w \tag{1}$$

where q is the energy transferred as heat to a system and w is the work done on the system. The SI unit of internal energy is the joule (J) and a change in the molar

internal energy is expressed in kilojoules per mole (kJ mol^{-1}) or occasionally in the old unit of measure - calories per mole (cal mol^{-1}) (Atkins 1998). Because calories are used in most of the original publications, they are also used in this thesis.

In chemical reactions, energy is needed to break chemical bonds in the starting substances, and energy is released when new bonds in the final product are formed. The sum of these tells us if the reaction consumes or releases energy. Like in chemical reactions, energy is needed in biological ligand-binding events. It is needed to break non-covalent bonds, to make conformational changes, to promote hydrophobic transfer of ligand from solution to binding site, in polyelectrolyte contributions for coupled ion release and to change the rotational/translational motion. Conversely, when non-covalent bonds are formed, energy is released. Together these events either release or absorb energy into or from the system. Binding reactions that release energy as heat to the solution are called **exothermic** reactions, and reactions that absorb heat energy from solution during the binding are called **endothermic** reactions.

2.3.1 Enthalpy, entropy - Gibbs energy of binding

Enthalpy (H) is a thermodynamic function of a system and defined as

$$H = U + pV \quad (2)$$

where U is the internal energy of the system, p is the pressure and V is the volume of the system. The change in enthalpy (ΔH) is independent of the path between the initial and final states and it is the sum of the heat change caused by the breaking and formation of bonds. Because the release of heat signifies a decrease in the enthalpy of a system, an exothermic process at constant pressure has a $\Delta H < 0$. Conversely, the absorption of heat results in an increase of enthalpy, meaning that an endothermic process at constant pressure has $\Delta H > 0$.

The enthalpy change can be measured calorimetrically by monitoring the temperature changes. In the calorimeter there is no net loss of heat from the calorimeter to the surroundings (adiabatic system). Thus the change in temperature (ΔT) in the calorimeter is proportional to the heat that the reaction releases or absorbs. Already in the 19th century chemists used calorimetry to measure chemical reactions (e.g. Pierre Eugène Berthelot, 1860), and since then new, more sensitive

technologies have made it possible to use calorimetry also for the meaningful study of biological molecules and interactions (Ladbury, Doyle 2004).

Entropy (S), a measure of the molecular disorder of a system, can be used to express the second law of thermodynamics: The entropy of an isolated system increases in the course of a spontaneous change.

$$\Delta S_{tot} > 0 \quad (3)$$

where S_{tot} is the total entropy of the system and its surroundings. A spontaneous change in an isolated system occurs with an increase in entropy. Accordingly, this means that when water molecules are usually displaced by the ligand from the binding site they are more randomly organized and entropy increases. In addition to enthalpy, entropy determines the free energy of the system. The thermodynamic definition of entropy is only valid for a system at equilibrium.

The Clausius inequality declares that for any change in a system:

$$dS \geq \frac{dq}{T} \quad (4)$$

where dq is the heat supplied to the system during the process and T is absolute temperature. When heat is transferred at constant pressure, and there is no other work than the expansion, $dq_p = dH$ and the Clausius equation can be rearranged to obtain:

$$TdS \geq dH \quad (5)$$

which can also be written:

$$dH - TdS \leq 0 \quad (6)$$

which can in turn be expressed more simply by a new thermodynamic quantitative, the famous Gibbs energy G :

$$G = H - TS \quad (7)$$

At constant temperature and pressure the change in Gibbs energy can be defined:

$$\Delta G = \Delta H - T\Delta S \quad (8)$$

An isobaric calorimeter measures q , and so the change in enthalpy (dH) is obtained. In other words $q = dH$ is the basis of a titration calorimetric experiment (Ladbury, Chowdhry 1998). Components of the Gibbs energy are well-known – the enthalpic term (ΔH) and the entropic term $T\Delta S$. The Gibbs energy (function), which is enthalpy minus the product of thermodynamic temperature and entropy, was formerly called free energy or free enthalpy. However, the International Union of

Pure and Applied Chemistry (IUPAC) recommends using the term Gibbs energy (McNaught, Wilkinson 1997). The enthalpic term consists of changes in bonding, van der Waals interactions, hydrogen bonding and charge interactions. The entropic term is respectively a sum of arrangement of the solvent together with rotational and transitional changes (Freire 2004). Figure 1 shows three examples of a binding reactions, where the thermodynamically important components differ significantly from each other.

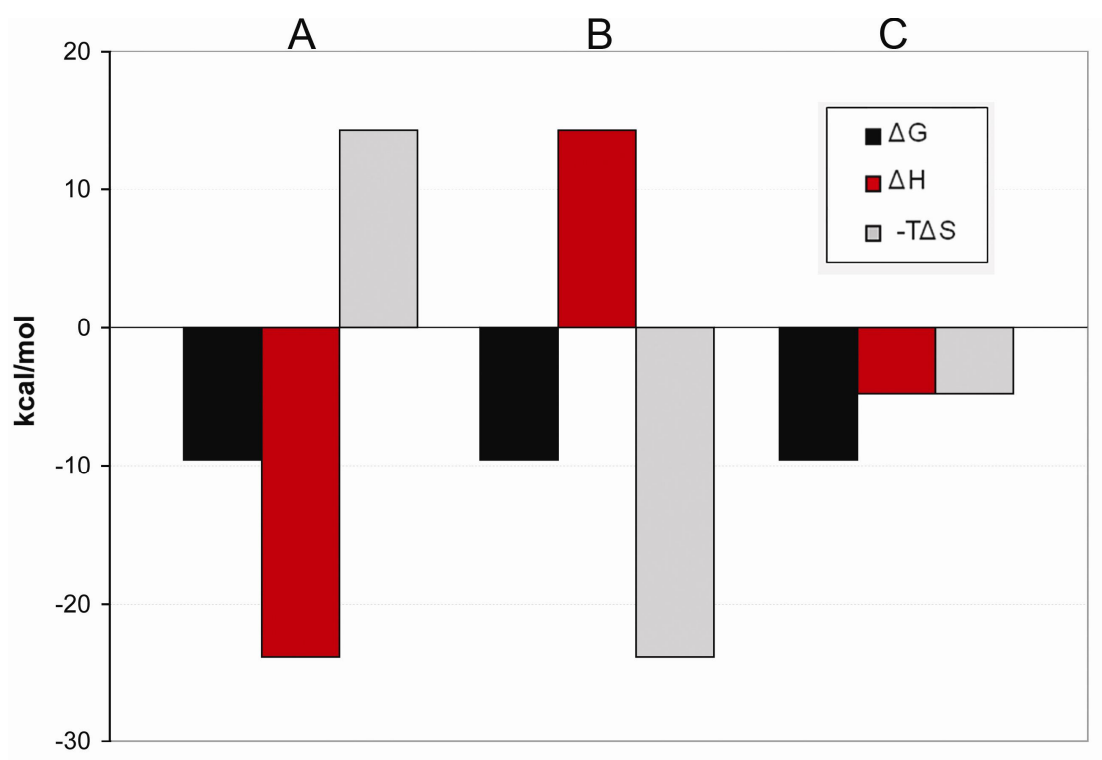


Figure 1. A schematic presentation to show the different imaginary thermodynamic profiles, which all have the same Gibbs binding energy ΔG : (A) example of an enthalpy-driven (ΔH) reaction, which is due to a high degree of hydrogen bonding in addition to conformational changes; (B) example of an entropy-driven (ΔS) reaction, which is dominated by the hydrophobic effect; (C) hydrogen bonding along with a hydrophobic contribution are favourable for a binding reaction.

2.4 Water, hydration and hydrophobic effect

Water participates in most chemical processes on earth. Among its unique properties is the flexibility to interact with a wide variety of molecules, determining, for instance, the three-dimensional structure, dynamics and function of biologically active macromolecules (Steinbach, Brooks 1993, Ooi 1994, Bryant 1996, Steinbach, Brooks 1996, Luecke et al. 1999, Ben-Amotz, Widom 2007). Nearly all protein-ligand interactions must compete with interactions with water. During complex formation water dissociates from the binding interface of both the protein and the ligand. At ligand binding sites, individual water molecules can stabilize ligand binding and have a considerable effect on binding affinity (Poornima, Dean 1995a, Poornima, Dean 1995b, Poornima, Dean 1995c, Bailly et al. 2003). Water is used as a classical example of hydrogen bonding and is also, as is apparent from its name, the main component of the hydrophobic effect. The properties of hydrogen-bonds and the hydrophobic effect are briefly discussed in the following sections.

2.4.1 Hydrogen bonding

Many unique properties of water arise from its ability to form hydrogen bonds (H-bond). In the H-bond, a hydrogen atom is covalently linked to an electronegative atom and therefore has a partial positive charge, and can be bonded to an electronegative acceptor as shown in figure 2. The H-bond is an interaction that has covalent, electrostatic, and van der Waals characteristics (Desiraju 2002).

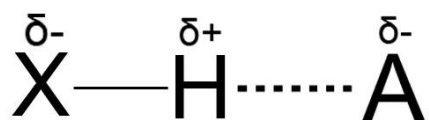


Figure 2. Hydrogen bond. A partially positive hydrogen atom of one molecule (X) is attracted to the unshared pair of electrons of the electronegative atom of another molecule (A).

The nature of X and A in figure 2 are important for the character of the H-bond. It has been proven that electronegative oxygen, nitrogen, fluorine or sulphur form H-bonds, but also weaker bonds like C—H...O and C—H... π are possible. The formation of the X—H...Y H-bond results in the weakening of the X—H bond. This weakening is accompanied by bond elongation and a concomitant decrease in the X—H stretch vibration frequency compared to the noninteracting atoms. A shift to lower frequencies, called a red shift, is the most easily detectable evidence of the formation of an H-bond. An improper, blue-shifting X—H...Y H-bond, is also known. It strengthens the X—H bond and causes a blue shift of the X—H stretch frequency. Overall, there are dozens of different types of X—H...A H-bonds that occur commonly in their condensed phases as well as in less common ones (Steiner 2002). All kinds of H-bonds probably participate in determining the shapes, properties, and functions of biomolecules (Derewenda, Lee & Derewenda 1995, Wahl, Sundaralingam 1997, Pierce, Sandretto & Bemis 2002, Jiang, Lai 2002). For example H-bonds are important in bio-molecular recognition and protein-ligand binding, but they also affect the physicochemical properties of a molecule, such as solubility, partitioning, distribution and permeability (Abraham et al. 2002).

2.4.1.1 *Furcation of hydrogen bonds*

Strong and weak H-bonds are ubiquitously present in protein-ligand interfaces (Pierce, Sandretto & Bemis 2002, Pierce et al. 2005, Panigrahi, Desiraju 2007). They are often furcated, which means that the donor can interact with several acceptors simultaneously or an acceptor can interact simultaneously with many donors as presented in figure 3. Donors and acceptors are furcated equally and the level of furcation can range from bifurcated to hexafurcated (Panigrahi, Desiraju 2007). Strong interactions like O—H...O tend to go towards non-furcated geometries, in contrast to weak interactions, which occur more often in the furcated form due to better interaction geometry and more efficient space filling (Panigrahi, Desiraju 2007).

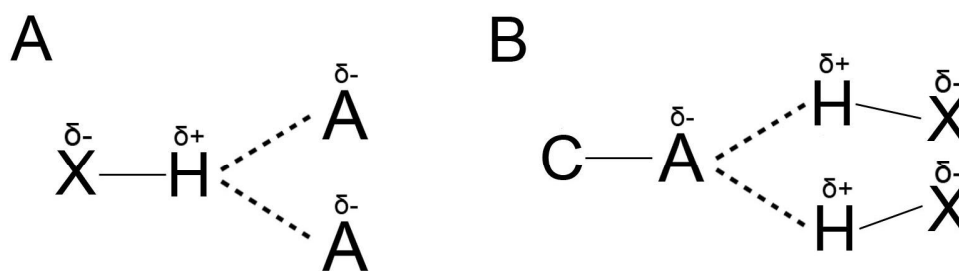


Figure 3. A schematic presentation of furcated hydrogen bonds. A; Bifurcated donor, B; Bifurcated acceptor.

2.4.2 Hydrophobic effect, desolvation of water molecules from the surface of the macromolecules

According to thermodynamics, matter seeks to be in a low-energy state, and bonding reduces chemical energy. Hydrophobes are electrically unpolarized molecules that are unable to form hydrogen bonds. Water repels them and they favour bonding with themselves. It is this effect that causes the hydrophobic “interaction”, which in itself is incorrectly named as the energetic force comes from the hydrophilic molecules. Therefore, the proper term instead of “hydrophobic interaction” would be hydrophobic effect.

Hydrophobic effect can be connected to various different terms. The dissolution of nonpolar solutes into an aqueous solution is referred to as hydrophobic hydration, the association of nonpolar substances in water is referred to as contact hydrophobic energy and associations, in which water intervenes, are referred to as solute separated hydrophobic forces (Franks, Eagland 1975). The Gibbs energy of hydration is defined as the transfer energy from the gas phase to the aqueous solution, which again is one form of the two contributions of the Gibbs energy of unfolding; the other one is the intramolecular interaction (Oobatake, Ooi 1993).

The stabilization and dynamics of macromolecules, protein-protein and protein-ligand binding are affected by the hydrophobic effect (Makhatadze, Privalov 1993, Privalov, Makhatadze 1993). The hydrophobic effect is also involved in self-assembly phenomena such as the formation of membranes and micelles (Tanford 1978, Cheng, Rossky 1998). Already sixty years ago Kauzmann demonstrated the

importance of the hydrophobic effect for protein folding (Kauzmann 1959) and later it was shown to be involved especially in early protein-folding events (Choe et al. 2000).

Nonpolar groups in proteins and ligands are covered by water molecules in aqueous solutions as schematically presented in figure 4. Water molecules in this shell are hydrogen bonded to each other and to the next shell. The movement of the water molecules decreases depending on how close they are to the macromolecule. If movement is restricted, also the freedom of water molecules is decreased and therefore the entropy of the system decreases. The entropy (ΔS) of the system increases if the water network is interrupted and water is removed from the shells of both protein and ligand as presented in figure 5. This also affects to the enthalpy ΔH of the system, because hydrogen bonds are broken.

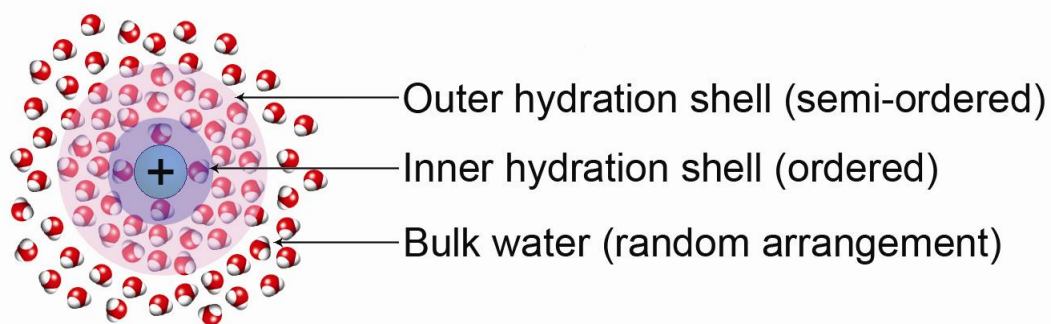


Figure 4. A schematic presentation of a hydration shell around a molecule. Water forms a hydration shell with three degrees of order: an inner shell with completely ordered water molecules, an outer hydration shell with semi-ordered water molecules and bulk water with only little order.

Due to the lower water solubility of small nonpolar molecules, nonpolar groups in proteins tend to associate with each other in an aqueous solution (Oobatake, Ooi 1993). This rearrangement of hydrophobic components, releases energy, because it frees water molecules from the contact site to form the common structure of water, i.e., the high-quality hydrogen bonding network between water molecules. Simultaneously, the entropy of the system increases, because the water has more freedom to rotate and move. This is described schematically in figure 5.

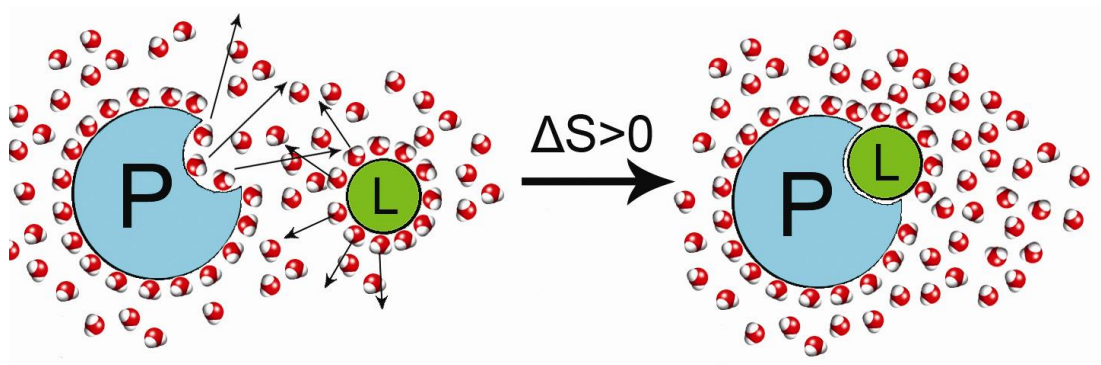


Figure 5. The release of ordered water molecules from a protein (P) binding site and from the inner hydration shell of a ligand (L). The dissociation of water molecules increases the entropy of the system.

Computational methods for the understanding of protein denaturation and hydration energy in an aqueous solution at the molecular level have improved (Mancera 2007, Rumfeldt et al. 2008, Sarupria et al. 2009). More sophisticated analyses have led to a better understanding of how solvents affect the Gibbs energy of denaturation or binding. For example, MD simulations are used for these purposes (Zhang et al. 1997). Protein hydration has been a central problem in structural biology for many years and despite increased knowledge, the influence of the solvent on ligand binding and protein folding is often not taken into account.

The hydration shell can be detected with cryogenic X-ray crystallography at the molecular and atomic level (Nakasako 2004) but the hydrodynamic radius of the protein-water complexes can also be measured by many scattering methods (Koch, Vachette & Svergun 2003) like dynamic light scattering (DLS) (Lisy et al. 2001). However, quantitative information about several aspects of protein hydration can only be achieved by magnetic relaxation dispersion (MRD). MRD can provide the number and residence times of long-lived water molecules and the mean rotational correlation time of water molecules at the protein surface (Halle 2004). Also different kinds of NMR methods based on incoherent magnetization transfer between water and protons in proteins by cross-relaxation and chemical exchange have been used to study protein hydration (Otting, Wuethrich 1989).

2.4.3 Bridging water in binding interfaces

Hydrophobic hydration is characterized by an increase in entropy. The release of water around non-polar groups into bulk solvent increases a system's entropy and is the reason for so-called entropy-driven binding events. Water can participate in the binding process also by forming a bridge between the macromolecule and the ligand. Therefore, fixed or transient water may have structural or functional roles and is not just an inert participant (Levitt, Park 1993, Wuthrich 1993, Billeter et al. 1996, Kosztin, Bishop & Schulten 1997, Nakasako 2004).

Water molecules have been found in the critical point of many protein systems (Smith et al. 1996, Hyre et al. 2002, Petrone, Garcia 2004, Hyre et al. 2006, Määttä et al. 2009). They can be classified as presented in figure 4 or in a more detailed way by four classes as suggested by Nakasako (Nakasako 2004): The “inside” class, the “contact” class and the “first- and second-layer” classes. In the two first classes, water molecules are either inside the solvent-accessible area (SAS) and occupy cavities in the protein molecules or they are located outside the SAS but still mediate intermolecular interactions between adjoining molecules. Water molecules outside the SAS (first-layer class) make hydrogen bonds and/or van der Waals contacts with the atoms of the protein surface, but molecules in the second-layer no longer directly interaction with the protein. Bound water molecules in the interface have two main tasks: (1) to fill empty spaces; or (2) to make bridging hydrogen bonds between the protein and the ligand. Such water molecules have been found for example in the class I MHC complex, the streptavidin-mutant – biotin complex (Hyre et al. 2006, Petrone, Garcia 2004) and recently also in the xenavidin-biotin complex as presented in figure 6 (Määttä et al. 2009).

2.5 Ligand binding kinetics

The starting point for most binding studies is the experimental determination of the **equilibrium association constant** (K_a) or the **dissociation constant** (K_d). These constants define the equilibrium between bound and free components in the reaction. They can be also called **affinity constants** or **equilibrium constants**. These constants can be determined accurately in a number of ways, usually by

measuring the concentrations of free and bound ligand as shown later. The relationship between the Gibbs energy and the equilibrium constant (described elsewhere (Atkins 1998)) can be determined by considering chemical potentials. As a result, the equilibrium constant can be related to the standard Gibbs energy change for the reaction at a given temperature and constant pressure by the equation (9):

$$\Delta_r G^\ominus = -RT \ln K_{eq} \quad (9)$$

where R is the universal gas constant, T is the absolute temperature in Kelvin, and K_{eq} is the equilibrium constant. The association constant and the dissociation constant, which is the reciprocal of the association constant, are specific types of equilibrium constants: The dissociation constant measures the propensity of a larger object to separate into two smaller objects (P,L) and conversely, the association or binding constant (K_a , K_b) measures the formation of a molecular complex (PL) as shown by equation (10) and in figure 7.

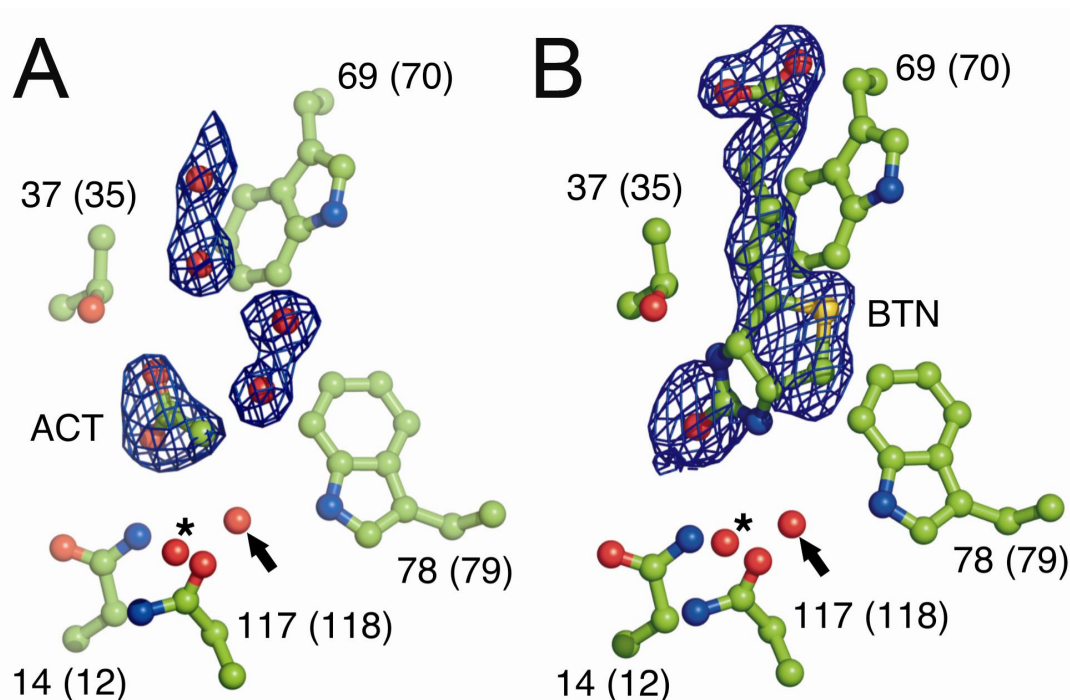


Figure 6. A simplified presentation of the ligand-binding site of xenavidin. A, unliganded xenavidin; B, xenavidin-BTN complex. Water molecules are shown as red spheres. An arrow pinpoints the water molecule in the binding pocket. Figure from Määttä et al. (2009).

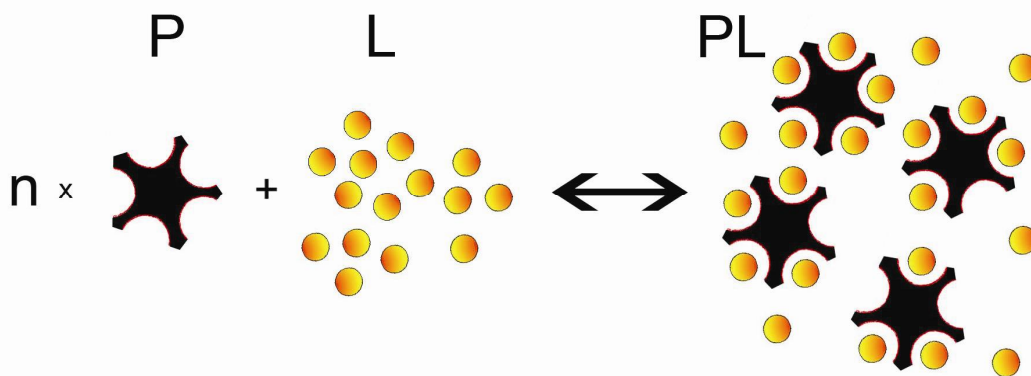


Figure 7. A schematic presentation of protein-ligand complex formation. The association and dissociation constants are derived from the concentrations of the protein and ligand, and the complex, respectively.



The association and dissociation constants that characterize a reversible equilibrium between a protein and a ligand are given by equation (11):

$$K_a = \frac{[PL]}{[P][L]}, K_d = \frac{1}{K_a} \quad (11)$$

Therefore, one way to determine K_a or K_d , is to measure the concentrations of free protein, free ligand, and protein-ligand complex in an equilibrated mixture. Another way to do this is to measure the change in Gibbs energy and calculate the equilibrium constant using equation (9).

2.5.1 Rate of reaction, kinetics

In a chemical reaction, the rate of the reaction refers to the change in the concentration of a reactant or a product in a given period of time. The rate can be determined by measuring the concentration of a certain component over time. The rate constant k whose unit depends on the particular way of expression, but is usually s^{-1} , includes all other factors (for example, temperature) that affect the reaction rate except concentration. The **reaction rate** (r) can be related to the concentration of the individual reactants by a rate equation in the form of:

$$r = k[A]^m[B]^n \quad (12)$$

where k is the rate constant, $[A]$ and $[B]$ are the concentrations of the reactants and

m and n are the orders of the reactions, respectively. The overall order of a reaction is the sum of the powers of the concentrations of the individual reactants.

In biological systems reagents often do not form an end-product. For example, protein (P) and ligand (L) do not form a fused product X . However they bind each other reversibly, forming a binary complex (PL). Typical examples of biologically important binding reactions are for instance the binding of agonist or antagonist to a receptor, protein-protein and protein-nucleic acid binding reactions, binding of substrate, activator or inhibitor to an enzyme, and binding of a metal ion or cofactor to a protein. The lifetime of the complex depends on the strength of non-covalent interactions, which eventually form an equilibrium between ligand-protein complexes and free ligands and proteins in the solution. This complex formation and dissociation was already described in formula (10). However, rate constants can be now included in this first order reaction formula:



where k_{ass} (equate with k_{on} and k_1) is the association rate constant and k_{diss} (equate with k_{off} and k_{-1}) is the dissociation rate constant. The equilibrium dissociation constant can therefore be described as the ratio of these constants:

$$K_d = \frac{k_{diss}}{k_{ass}} \quad (14)$$

This means that the association rate constant could also be determined by measuring the association rate over a range of different concentrations, because for reversible binding, it can be shown (Copeland 2000) that the value of k_{obs} is directly proportional to the concentration of ligand present as follows:

$$k_{obs} = k_{diss} + k_{ass} [L]_f \quad (15)$$

where $[L]_f$ is the concentration of free ligand. As illustrated in the following figure 8, the linear function can be fitted to the data. The values for k_{ass} and k_{diss} can then be estimated from the values of the slope and the y intercept, respectively. These estimations can be then be used to calculate the dissociation constant K_d .

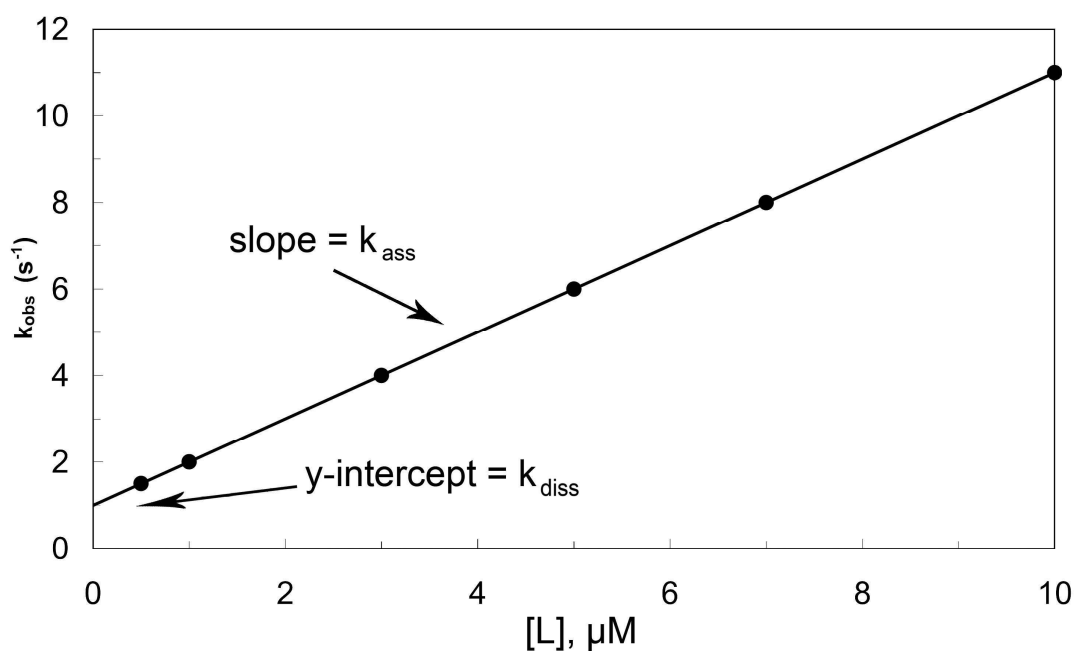


Figure 8. A schematic presentation of the observed association rate as the function of ligand concentration.

Dissociation, also called “off rate” can be measured as a function of time. Each ligand-protein complex dissociates stochastically at a given time, and the dissociation process is typically exponential decay. Exponential equations are used to model processes where the rate at which something happens is proportional to the amount of starting material left. For example, when ligands dissociate from protein, the number of molecules that dissociate in any short time interval is proportional to the number that were bound at the beginning of that interval. Equivalently, each individual molecule of ligand bound to a protein has a certain probability to dissociate from the protein in any small time interval. That probability does not increase when the ligand stays on the protein longer.

If we define Y to be the number of protein-ligand complexes still present at any given time X , the rate of change of Y is proportional to Y . This is expressed as a differential equation (16) and presented in figure 9:

$$\frac{\Delta Y}{\Delta X} = \frac{dY}{dX} = -k \times Y \quad (16)$$

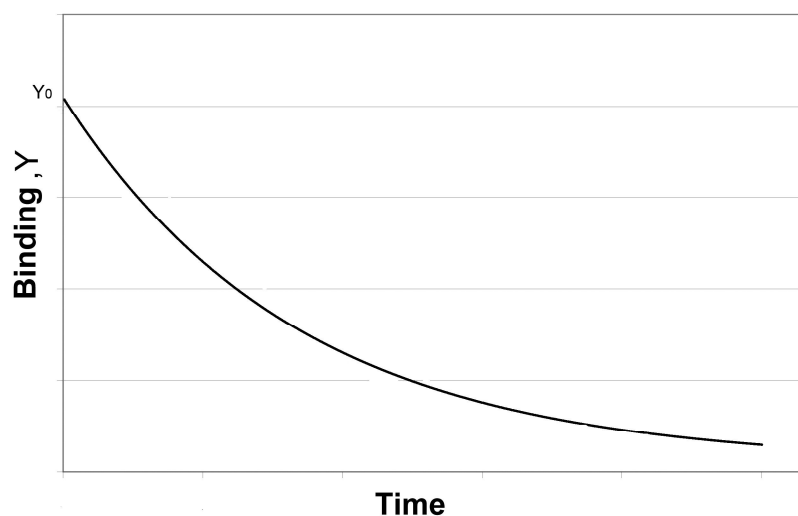


Figure 9. The exponential dissociation of specific protein-ligand complex.

Exponential function has a derivative that is proportional to Y and can be achieved by the integration of both sides of the equation. A new exponential equation (17) defines Y as a function of X , the rate constant k , and the value of Y at zero time, Y_0 .

$$Y = Y_0 \cdot e^{-k \cdot X} \quad (17)$$

Where the variable k is the dissociation constant (k_{off} or k_{-1}) expressed in units of inverse time. This term can be used to evaluate the binding affinity of proteins.

2.5.1.1 *Factors influencing the rate of the reaction*

The rate of the reaction is dependent on many things like concentration, temperature and solvent. The nature of the reactants also affects the reaction, for example gases tend to react much faster than solid materials. Concentration does not have an effect on reaction rate in a zero order reaction and only the concentration of one component has an influence on a one order reaction. The rate of the second-order reaction depends on the concentrations of one second-order reactant, or two first-order reactants as shown in following example:

$$r = k[A]^2 = k[A][B] = k[B]^2 \quad (18)$$

Usually, an increase in temperature is accompanied by an increase in the reaction rate. Temperature is a measure of the kinetic energy of a system, so a higher temperature implies a higher average kinetic energy of molecules and more collisions per unit time.

Catalyst or competitors (inhibitors) also have an effect on reaction rate. In biology enzymes are considered catalysts. They increase the reaction rate by lowering the activation energy.

2.5.1.2 *Methods to determine equilibrium constants*

Theories described in the previous chapters have been used in many different methods to determine interactions and affinities of protein-ligand or protein-protein pairs. The energies of these biological events are so small that building an instrument to measure them is challenging. However, the sensitivity and accuracy of such instruments have increased dramatically in the last few decades allowing the direct measurement of sub-millimolar to nanomolar binding constants (10^{-2} M to 10^{-9} M). Nowadays, the limit of binding constant measurements with isothermal titration calorimetry has reached the picomolar (10^{-12} M) level with the competitive method (Sigurskjold 2000). Increased sensitivity has also reduced the concentration of macromolecules needed for the assay, which was previously one of the obstacles of the measurements.

There are plenty of other methods for measuring binding. Stopped flow instruments allow the study of relatively fast binding reactions (Schechter 1970). The quartz microbalance measures the binding of molecules on the surface of a quartz crystal and this can be used to measure both association and dissociation events (Marx 2003, Pei et al. 2007, Pei et al. 2006). Ligands labelled with fluorescent groups or radio active molecules can be used for this purpose (Sharma, Schulman 1999). One powerful method for evaluating interactions is high-field solution nuclear magnetic resonance (NMR), which can also be used to estimate the structural, thermodynamic and kinetic aspects of a binding reaction (Skinner, Laurence 2008). Differential scanning calorimetry (DSC) has proven its applicability for poorly defined stoichiometries and low ligand solubilities and in instances where binding is more tight than normal ($10^9 \text{ M}^{-1} < K_b < 10^{20} \text{ M}^{-1}$)

(Brandts, Lin 1990, Plotnikov et al. 2002, D'Souza et al. 2009a, D'Souza et al. 2009b). However, two of the most popular techniques for measuring interactions between biomolecules are surface plasmon resonance (SPR) (Jonsson et al. 1991, Jonsson et al. 1993) and isothermal titration calorimetry (ITC) (Langerman, Biltonen 1979, Biltonen, Langerman 1979).

SPR-instruments use an optical method to measure the refractive index near a sensor surface, and in order to detect an interaction one molecule is immobilised onto the sensor's surface (figure 10A). When an analyte e.g. protein binds to the ligand on the surface, the accumulation of protein results in an increase in the refractive index also known as resonance units (RU). The glass surface of the instrument is coated with a thin film of a noble metal, which changes the normal reflection of this surface. Some of the light is "lost" into the metallic film and when the loss is greatest at a certain angle, the reflected light reaches its minimum. This angle is called the surface plasmon resonance angle. The loss of light is due to oscillation of mobile electrons, also known as plasma, at the surface of the metal film. These oscillating plasma waves are called surface plasmons and when they match with the wave vector of the light, the electrons resonate. The binding of molecules to the surface changes the refractive index and can be detected by a shift in the plasmon resonance angle. This change is measured in real time (figure 10B) and thus information about the association and dissociation kinetics can be obtained, which in turn can be used to obtain K_a and K_d using equation (11).

The calorimeter, on the other hand, measures the rate of heat flow, resulting from the heat effect induced by changing the composition of the sample by titrating near to a constant temperature. ITC measures the heat consumed or released in association to injection of a known amount of reaction component. The measurement allows the direct estimation of the enthalpy change (ΔH°) in the binding reaction. In addition, it may be possible to calculate the association constant (K_a), stoichiometry (n), Gibbs energy (ΔG°) and the entropy (ΔS°) of binding from a single experiment. To understand how these parameters are related to physical processes at the molecular level is difficult, but complex relationships can be argued for with structural data (Davies, Hubbard & Tame 1999).

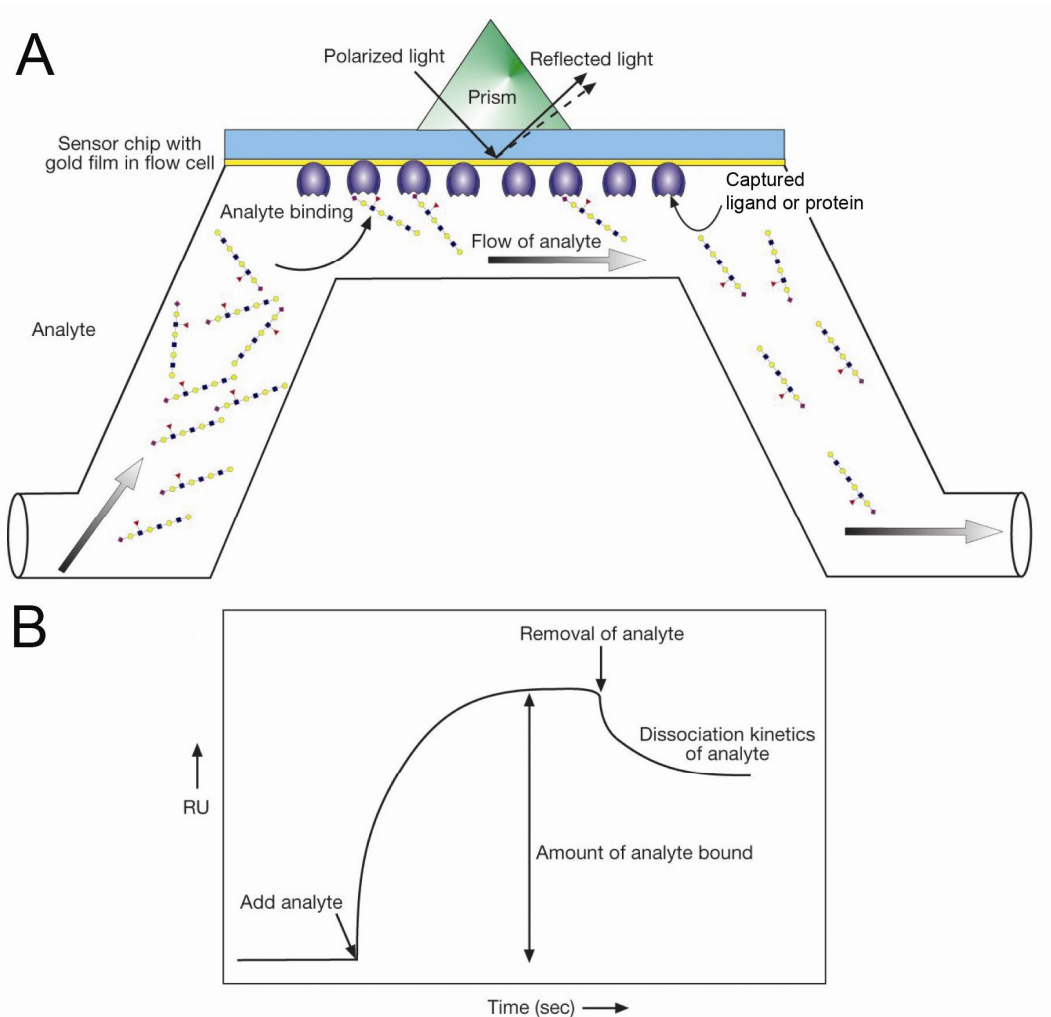


Figure 10. Surface plasmon resonance. A; In the SPR instrument, the binding of the analyte in the flow cell to the immobilized molecule causes a change in the reflected light that can be monitored. B; An example of a sensorgram showing the binding of the analyte to the ligand and the kinetics of binding and dissociation. RU indicates resonance units. The figure is adapted from the original (Varki et al. 2008).

The calorimeter measures the differential heat effects between a sample and a reference cell. The temperature difference between the two cells is constantly monitored and constant power is applied to the reference cell, which activates the feedback circuit to apply variable power to the sample cell in order to maintain a very small temperature difference between the cells. The feedback power is the baseline level in the absence of any reaction. When a reaction takes place in the sample cell, it either consumes energy (endothermic) or releases it (exothermic), and

has either an increasing or decreasing, respectively, effect on feedback power. The heat evolved, or absorbed by the reaction is then determined by integration of these deflections from the baseline, with respect to time (figure 11).

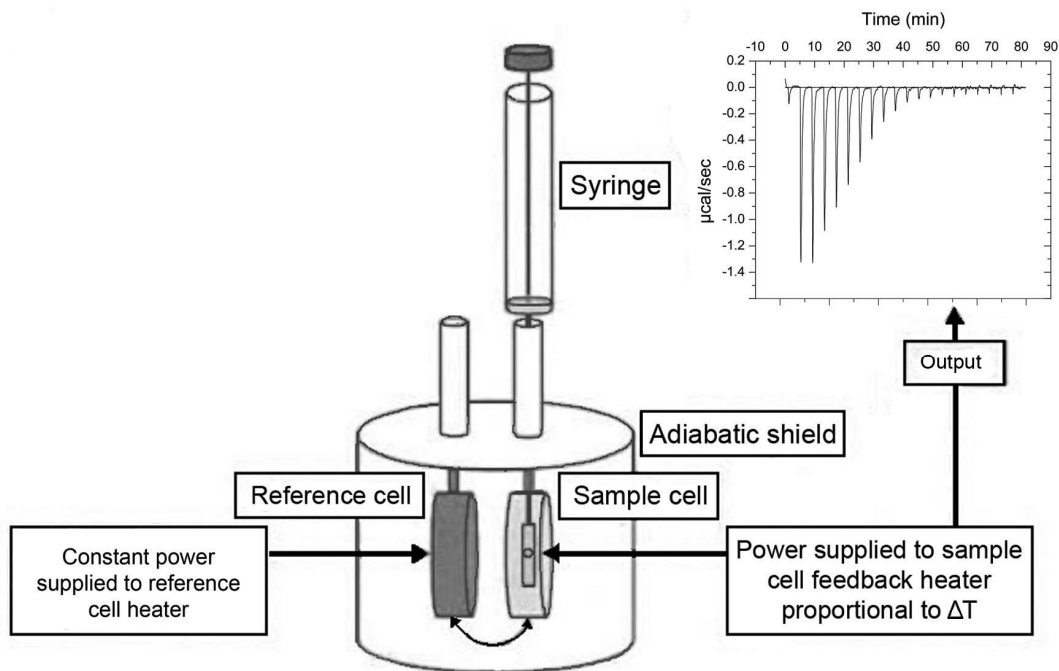


Figure 11. A schematic diagram of a typical cell arrangement in an ITC and the raw data output after the titration has been completed. The figure is adapted from the original (Ladbury, Doyle 2004).

2.6 Structural features in protein-ligand binding

In this section I will summarize the importance of larger protein structures on ligand-binding events, such as flexible loops and co-operative subunits. In many proteins, the binding of a small molecule at one site alters the structure of other regions of the protein. These changes can be relatively significant when two large domains on a single polypeptide move in relation to each other around a central flexible region. On the other hand, ligand binding may only cause small changes in loop regions in the active site. Often these changes enclose the ligand in a cage-like structure (Kempner 1993).

2.6.1 Mobile domains and flexible loop regions of proteins

The movement of massive protein domains is not usually directly required for ligand binding. However, large domains might make additional interactions, which can stabilize the structure. The two globular regions at the amino- and carboxy-termini of calmodulin can move and both can bind to myosin light chain kinase. Binding of ligand to one domain results in the movement of the other, this brings it near to ligand, and therefore, allows additional interactions that stabilize the structure (Ikura et al. 1992). In addition, domains in hexokinase (Bennett, Steitz 1978), alanine racemase (Galakatos, Walsh 1989), bacterial neutral proteases (Holland et al. 1992), phosphoglycerate kinase (Pickover et al. 1979) and cAMP-dependent protein kinase (Olah et al. 1993) have been found to move to form a flexible connecting region enclosing the substrate-binding region.

Movements more common to have an effect on ligand binding than those of massive domains, are subtle changes in loop regions. Although restricted, these conformational changes may have dramatic effects on biochemical activity. Usually a flexible loop closes the active site and protects the ligand-receptor complex or enzymatic reaction from surrounding water (Tanaka et al. 1992, Pompliano, Peyman & Knowles 1990). Another good example of loop movement influencing ligand binding is (strept)avidin-BTN complex formation. When biotin enters the binding pocket of avidin or streptavidin, the loop between $\beta 3$ and $\beta 4$ adopts a fixed conformation and seals the binding site (Livnah et al. 1993b, Weber et al. 1989). Deletion of this loop from wild-type streptavidin reduces the association constant approximately six orders of magnitude (Chu et al. 1998).

2.6.2 Co-operation between subunits

The binding of ligand is not always a matter of only one protein subunit or a single protein molecule. Many times neighbouring subunits participate in the binding of a ligand and may even strengthen the affinity. A classical example is haemoglobin, an iron-containing oxygen-transport metalloprotein in the red blood cells of vertebrates. Haemoglobin is a tetramer consisting of four globin subunits, which all have a haem group in their active centre. Haem is a prosthetic group, which has a ferrous iron atom in the centre of an organic heterocyclic ring named porphyrin.

Oxygen binds with high affinity to the iron complex of the haem group by the coordinate covalent bond (Fermi et al. 1984). This causes the iron atom to move back toward the centre of the plane of the porphyrin ring, which then causes the imidazole side chain of the histidine residue at the other side of the iron to move in same direction. Together these bend the shape of the tetramer so that the other monomers (in the tetramer) may bind oxygen more easily and with increased affinity (Perutz 1976).

Some members of the lipocalin family bind haem in a cooperative way (Montfort, Weichsel & Andersen 2000, Larsson, Allhorn & Kerström 2004, Allhorn, Klapyta & Akerström 2005). One such member α_1 -microglobulin, a yellow-brown lipocalin (Akerström et al. 2000), which binds haem at the entrance of the lipocalin cavity. The haem and other prosthetic groups are bonded to four amino acid residues from Cys34 (Escribano et al. 1991), Lys92, Lys118 and Lys130 (Berggard et al. 1999). These amino acid residues have been shown to act cooperatively (Allhorn, Klapyta & Akerström 2005).

The binding mechanism of avidin, which like lipocalins, is a member of the calycin superfamily, involves two so-called functional monomers in the binding of a ligand. The side chain of Trp110 participates in biotin binding from the neighbouring subunit and mutations to this important tryptophan reduce affinity by several orders of magnitude (Laitinen et al. 1999).

2.7 Avidin – a high affinity and highly stable ligand-binding protein

The basic, homotetrameric glycoprotein found in the chicken egg-white was named avidin due to its high biotin-binding capacity (Eakin, Snell & Williams 1941). Each avidin subunit consists of 128 amino acid residues and one sugar moiety attached to Asn17. The three-dimensional (3D) crystal structure of avidin revealed the presence of eight antiparallel successive β -strands, the structure of their interconnecting loop regions and the overall shape of a beta-barrel in the avidin monomer (Livnah et al. 1993b, Pugliese et al. 1993). The schematic presentation of the avidin tetramer in figure 12 illustrates the four identical subunits of the tetramer and one bound biotin molecule in each subunit.

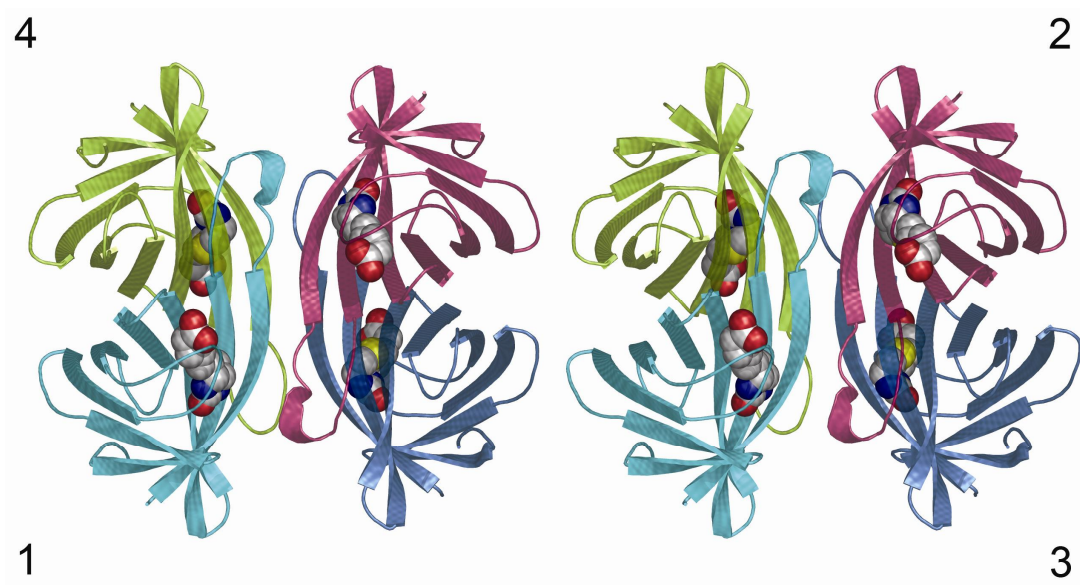


Figure 12. The stereo image of the avidin tetramer. A schematic presentation where four bound biotins are presented with CPK molecules in the binding pockets. Subunits of a different colour are numbered according to Livnah *et al.* (pdb: 2AVI, Livnah *et al.* 1993b)

2.7.1 Structure of the ligand-binding pocket of avidin – cooperative hydrogen bonds

The most remarkably feature of avidin is its extremely high affinity to D-biotin ($K_d \approx 6 \times 10^{-16}$ M, figure 13) (Green 1975, Green 1990).

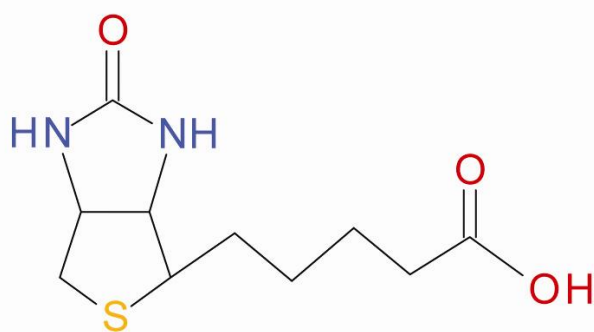


Figure 13. The structure of D-biotin. Synonyms for D-biotin include vitamin H, vitamin B7 and Hexahydro-2-oxo-1H-thieno[3,4-d]imidazole-4-pentanoic acid. MW 244.31 g/mol.

The high affinity is due to a network of eleven hydrogen bonds, hydrophobic interactions and near perfect structural complementarity between the ligand and the binding site. The binding of biotin to avidin is highly exothermic and enthalpy driven. The total entropy of the binding reaction is unfavourable or close to zero (Green 1966). It could, however, be thought to be favourable, because five structured water molecules that form a hydrogen-bonding network are removed from the binding pocket of an apoform of avidin by biotin binding (Wilchek, Bayer 1999), and the dehydration of biotin itself, also contributes to the binding entropy by increasing water movement. These positive contributions to binding are compensated by the negative effect rising from the increased conformational rigidity of biotin in the binding site, as well as the decreased movement of the three-dimensional structure of the avidin as presented in figure 14.

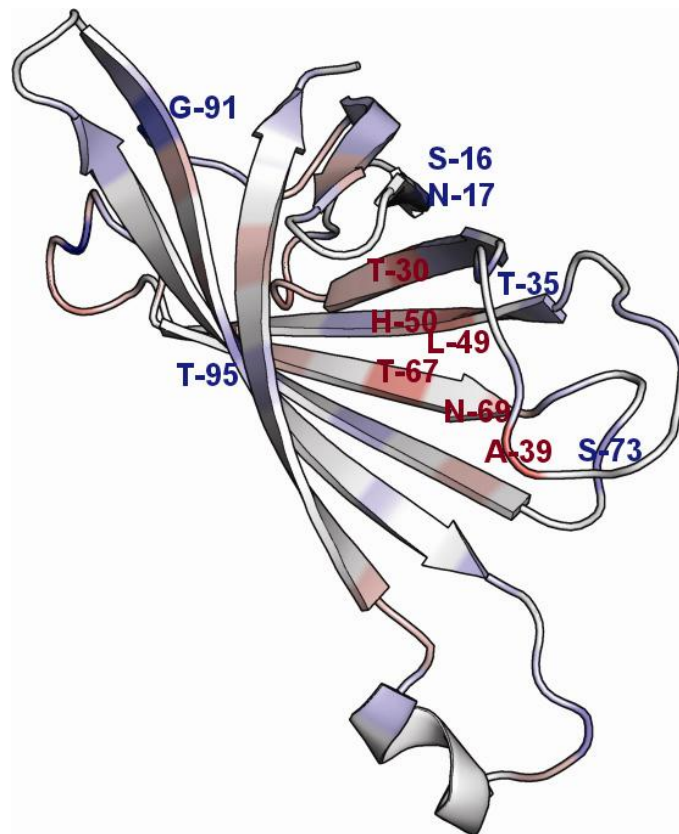


Figure 14. Average per-residue RMSD-F of side chains in wild type avidin with bound biotin in subunit D. Colors indicate where side chain mobilities differ between the protein-ligand complex and the apo-protein. Red indicates increased movement and blue indicates decreased movement. Figure adopted from (Kukkurainen 2009).

Eleven hydrogen bonds that bind biotin to avidin include both N—H...O and O—H...O hydrogen bonds. Streptavidin, which is an avidin like protein found in *Streptomyces avidinii*, has a structure homologous to avidin, but a 100-fold lower affinity to biotin. One explanation for this is the lower number of H-bonds in streptavidin compared to avidin as shown in figure 15.

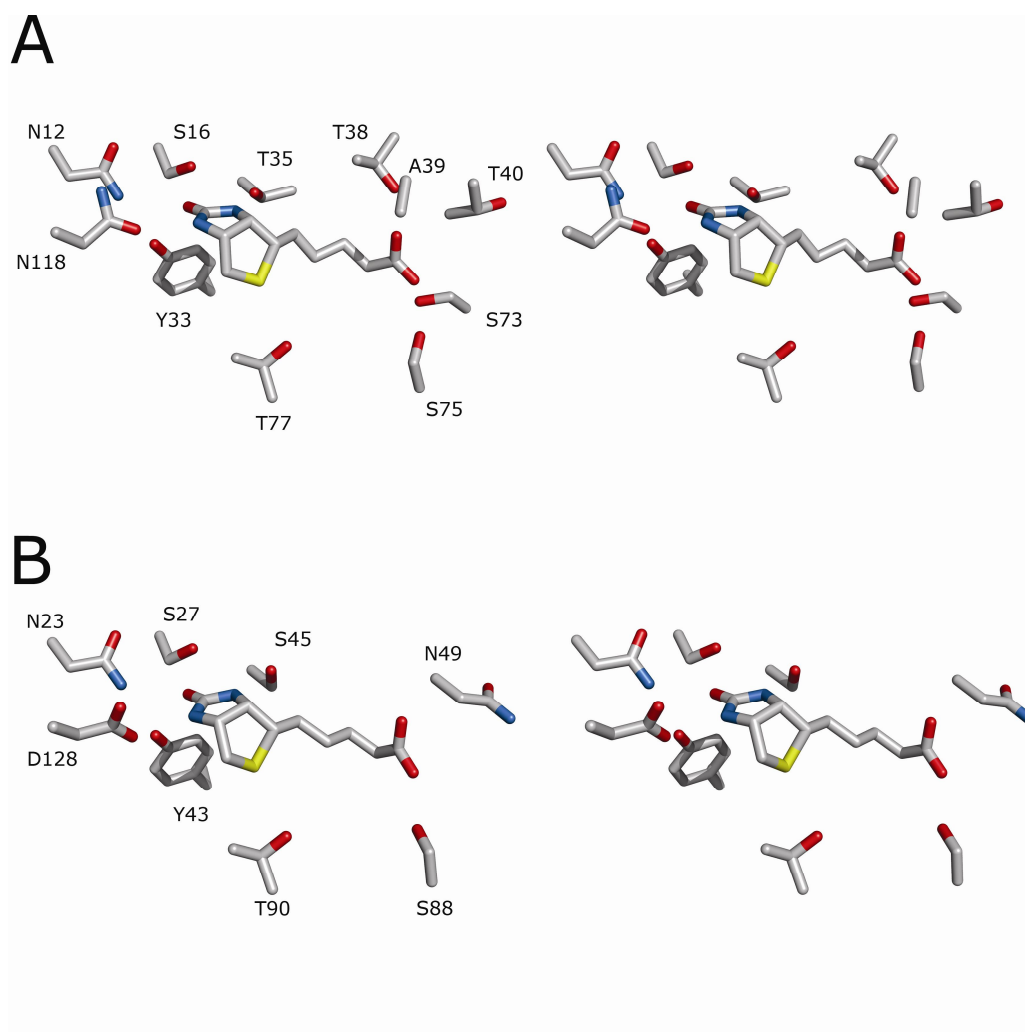


Figure 15. The stereo image of the amino acid residues, which form H-bonds with avidin (A, pdb: 2AVI) or streptavidin (B, pdb: 1MK5). Amino acid residues Thr35, Thr 38, Ala39 and Thr 40 of avidin are from L3,4-loop and forms H-bond with biotin. However, streptavidin has only two H-bonding amino acid residues from L3,4-loop, Ser45 and Asn49.

Cooperative hydrogen bonds are a network of hydrogen bonds, whose influence on stabilization is larger than the sum of individual hydrogen bond energies. A good example of cooperative hydrogen bonds is the ureido moiety of biotin in the bound

state in avidin and streptavidin, which forms five hydrogen bonds: three to carbonyl oxygen and one with each -NH group as shown in figure 15. For example, the charged Asp128 in streptavidin polarizes the urea of biotin and is the main reason for cooperativity. By removing the interaction by mutagenesis, cooperativity is diminished (DeChancie, Houk 2007). Cooperativity was studied with the streptavidin double mutant S45A/D128A by Hyre *et al.* 2006. The effect of removing both hydrogen bonds by mutagenesis evoked the loss of binding affinity by 11-fold more than predicted by a linear combination of the single-mutant energetic perturbations (Hyre *et al.* 2006).

Many hydrogen bonds between avidin and biotin are furcated. There is a trifurcated carbonyl oxygen donor in the ureido moiety of biotin and a bifurcated carbonyl oxygen donor and a trifurcated hydroxyl group acceptor in the carboxylic tail of biotin. If new properties and applications for avidin or streptavidin are required, one potential option is the mutagenesis of the amino acid residues involved in these furcated interactions. Alanine mutagenesis of Asn23Ala, Ser27Ala and Tyr43Ala in streptavidin, by Stayton and his co-workers revealed that destroying one member of this trifurcated hydrogen bond yields affinities 282, 114 and 67 times weaker, respectively, for biotin compared to wt streptavidin (Klumb, Chu & Stayton 1998). Similar results were obtained with avidin. Mutagenesis of Tyr33, which is equivalent to Tyr43 in streptavidin, did not have a significant effect on binding affinity, but it made it pH-dependent (Marttila *et al.* 2003). The site was used by Ting and co-workers to construct a triple mutant of a streptavidin (Asn23Ala, Ser27Asp, Ser45Ala) that binds biotin in the monovalent way (Howarth *et al.* 2006) and by Sano and Cantor to change ligand specificity to favour 2-iminobiotin (figure 17) over biotin with Asn23Ala and Ser27Asp mutations (Reznik *et al.* 1998).

2.7.1.1 *The effects of hydrated water on ligand binding in avidin*

Before a ligand binds to a protein, water molecules at the interacting regions of both molecules have to be removed. Desolvation of water molecules reduces the Gibbs energy of hydration. At the same time dehydrated water gains freedom to form the ordinary water structure that increases entropy.

The Gibbs energy of dehydration can be computationally estimated from the change of areas at the contact site (Ooi et al. 1987). Estimating the entropy change of the association is more difficult, but can be done simultaneously with determining enthalpy from the change in accessible surface areas (Ooi, Oobatake 1988). This method has not been used to estimate the interaction energy of avidin and biotin. However, the accessible surface area (ASA) of the apo- and holoforms could be easily calculated. The difference between the AVD-BTN complex and uncomplexed AVD is approximately 640 \AA^2 , and moreover, the solvent accessible area of biotin is 414 \AA^2 (calculated from the pdb:2AVI complex, Kukkurainen et al., unpublished). Therefore, the complex formation of AVD and BTN decreases the total solvent accessible area of the system by 1051 \AA^2 . If we consider that a water molecule has a radius of 1.4 \AA , and therefore covers over 6 \AA^2 of the surface of the proteins, in theory almost 170 water molecules could be released from both avidin and biotin to the surroundings when they form a complex. A tight interaction causes a decrease in molecular freedom. The positive effect of dehydration can not be seen in experimental measurements as a positive entropy value, because the movements of avidin and biotin are more restricted after complexation, which compensates for the entropy increase that arises from dehydration. For example, the loop L3,4 adopts a significantly more defined conformation after biotin binding (Livnah et al. 1993b).

2.7.2 Ligands of avidin

Besides its natural ligand biotin, avidin is also able to bind biotin analogs (Green 1975), as well as other small molecules like 4-hydroxyazobenzene-2-carboxylic acid (HABA) (Livnah et al. 1993a). HABA binds to the same binding pocket as biotin with significantly reduced affinity ($K_d = 6 \times 10^{-6} \text{ M}$) (Green 1965, Green 1970). This makes it an interesting molecule, because although biotin and HABA are quite similar in size, they have different structures as presented in figure 16. In addition it has been found that avidin does not bind only HABA, but also a wide variety of other azocompounds (Green 1963a, Green 1965, Repo et al. 2006).

HABA is used to measure the biotin-binding sites of avidin derivatives. When it binds to avidin it goes through a spectral shift from yellow to red and this shift can be reversed by D-biotin. Other azocompounds have slightly different kinds of

properties. Their binding affinities vary from weak (mM) to moderate (μM) and their spectroscopic properties differ as well.

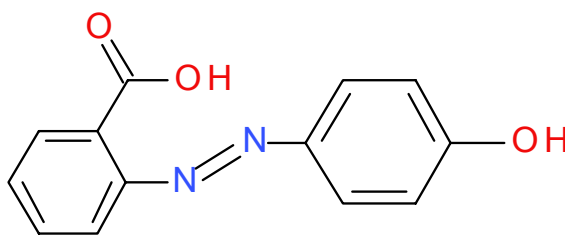


Figure 16. The molecular structure of 4-hydroxyazobenzene-2-carboxylic acid, HABA.
MW 242.23 g/mol

2.7.2.1 *Biotin and biotin analogues*

Biotin, also known as vitamin H (or vitamin B7), is an important component in several biochemical and cellular processes. This small water-soluble molecule is synthesised by most microbes, plants and fungi (Streit, Entcheva 2003) but for example humans need to obtain it from food.

The imidazolidone ring has been shown to be more important for binding to avidin than the thiophan ring (Dittmer, DU Vigneaud 1944) and that modifications to the carboxyl group have only a minor effect on binding to avidin (Wright et al. 1951). The affinities of biotin analogues like 2'-thiobiotin, 2'-iminobiotin and d-desthiobiotin (figure 17) were measured by spectrophotometric titrations and were all found to bind with tight affinity ($K_d < 10^{-8}$ M) (Green 1975).

Biotin analogues have an important role in (strept)avidin-technology. Of special importance are biotin analogues, which have a modified carboxyl group in their valeryl chain. This offers the possibility of attaching desired parts to biotin without affecting its affinity to avidin. An example of a more sophisticated application is the use of biotin analogues with various alkyl and poly(ethylene glycol) (PEG) chains coupled to DOTA. Such compounds have been used to monitor the survival of avidin-coated islets of Langerhans isolated from pancreas and used in transplantations i.e. to the liver (Blom, Langstrom & Velikyan 2009).

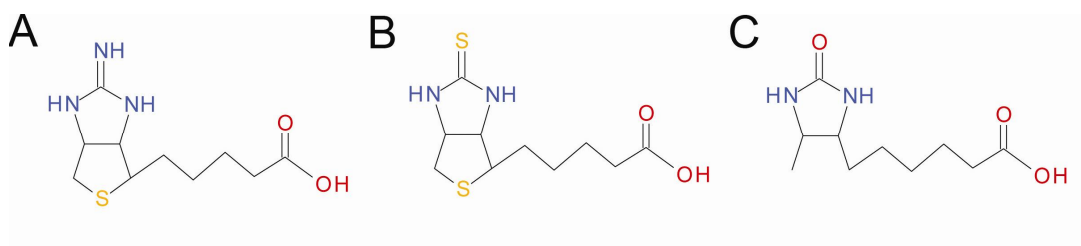


Figure 17. The molecular structures of A, 2'-iminobiotin; B, 2'-thiobiotin and C, d-desthiobiotin.

2.7.2.2 Peptide ligands of (strept)avidin

The size of the ligand-binding site in avidin and streptavidin also enables the entry and binding of small peptides. The binding affinities of the avidin binding peptides first discovered were significantly lower than that of biotin. Optimization has, however, increased the binding affinities of peptides and they have reached a nanomolar dissociation constant range (Voss, Skerra 1997). Peptides are therefore valuable for example in single-step purifications of fusion proteins (Cho et al. 2000, Keefe et al. 2001, Lamla, Erdmann 2004, Bessette, Rice & Daugherty 2004). Probably the most popular and commercially available product is Strept-tag (AWRHPQFGG) and its further improved version Strept-tagII (WSHPQFEK) (Schmidt, Skerra 1993, Schmidt et al. 1996). These peptides are used to extract recombinant proteins, for example. The sequence of the tag is inserted into the polypeptide chain of a recombinant protein. The tagged protein binds to the streptavidin column and can be easily eluted with biotin. The purification of recombinant proteins in this way has been described in several studies (Schmidt, Skerra 1993, Schmidt et al. 1996, Voss, Skerra 1997, Skerra, Schmidt 1999, Skerra, Schmidt 2000).

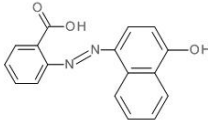
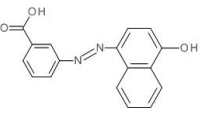
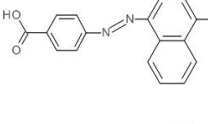
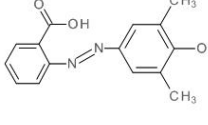
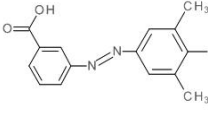
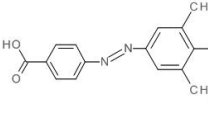
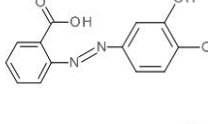
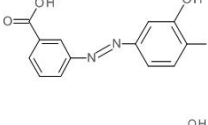
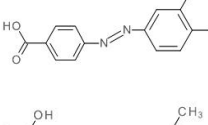
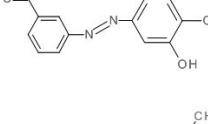
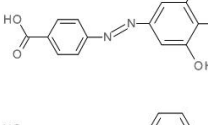
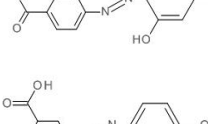
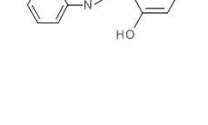
2.7.2.3 Azocompounds

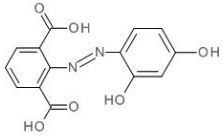
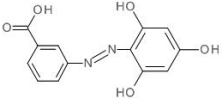
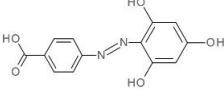
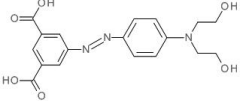
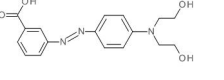
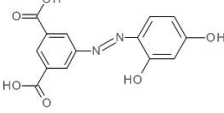
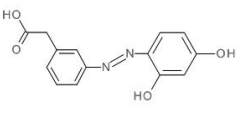
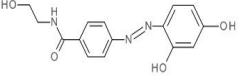
Avidin binds to a variety of different dyes and peptides, which have only little or no structural similarity to biotin (Tausig, Wolf 1964, Green 1975, Devlin, Panganiban & Devlin 1990). The sizes of these small molecules are similar to biotin and they appear to bind completely or partially to the same site, since biotin displaces them.

Avidin fluoresces when the weakly interacting HABA (4'-hydroxyazobenzene-2-carboxylic acid) is displaced by biotin. This phenomenon can be used to determine the amount and biotin-binding activity of avidin in solution with a spectroscopic method (Green 1965, Green 1970, Wilchek, Bayer 1990). HABA is also used in an assay where the amount of biotin linked to either proteins or nucleic acids is determined. The assay is based on fluorescence resonance energy transfer (FRET) and uses a complex of Alexa Fluor® 488 dye-labeled avidin with HABA as a quencher dye (Batchelor et al. 2007). The binding of HABA to avidin is an endothermic reaction and the moderate affinity and negative Gibbs energy of binding are due to the favourable entropy value of the binding. By modifying HABA, it has been possible to improve its affinity, and hereby the sensitivity of avidin-based applications (Repo et al. 2006). The binding of azocompounds to avidin was screened with ITC and the results are presented in Table 1.

Table 1. The structures of azocompounds and the determined dissociation constants for the azocompound-avidin complexes. The values are from ITC measurements carried out in a sodium phosphate buffer at 25 °C (Määttä *et al.* unpublished)

Ligand	Structural formula	K_d ($\times 10^{-6}$) M
4-Hydroxyazobenzene-2-carboxylic acid (HABA)		7.9
4-Hydroxyazobenzene-3-carboxylic acid		13.7
4-Hydroxyazobenzene-4-carboxylic acid		31.1
4-Hydroxy-3-nitroazobenzene-3-carboxylic acid		1.8
4-Hydroxy-3-nitroazobenzene-4-carboxylic acid		147.0

4-Hydroxynaphthaleneazobenzene-2-carboxylic acid		0.6
4-Hydroxynaphthaleneazobenzene-3-carboxylic acid		22.6
4-Hydroxynaphthaleneazobenzene-4-carboxylic acid		ND ^a
4-Hydroxy-3,5-dimethylazobenzene-2-carboxylic acid		4.3
4-Hydroxy-3,5-dimethylazobenzene-3-carboxylic acid		33.1
4-Hydroxy-3,5-dimethylazobenzene-4-carboxylic acid		22.8
3,4-Dihydroxyazobenzene-2-carboxylic acid		5.8
3,4-Dihydroxyazobenzene-3-carboxylic acid		12.3
3,4-Dihydroxyazobenzene-4-carboxylic acid		385.0
3,4-Dihydroxy-5-methylazobenzene-3-carboxylic acid		22.6
3,4-Dihydroxy-5-methylazobenzene-4-carboxylic acid		28.0
2,4-Dihydroxyazobenzene-4-carboxylic acid		9.9
2,4-Dihydroxyazobenzene-3-carboxylic acid		2.6

2,4-Dihydroxyazobenzene-2,6-dicarboxylic acid		10.1
2,4,6-trihydroxyazobenzene-3-carboxylic acid		15.2
2,4,6-trihydroxyazobenzene-4-carboxylic acid		18.6
4-(diethanolamine)-azobenzene-3,5-dicarboxylic acid		58.5
4-(diethanolamine)-azobenzene-3-carboxylic acid		158.0
2,4-Dihydroxyazobenzene-3,5-dicarboxylic acid		17.5
2,4-Dihydroxyazobenzene-4-ethanoic acid		120.0
2,4-Dihydroxyazobenzene-4-ethanolamide		ND ^a

^aNot determinable

2.7.3 Stability of the avidin protein

Avidin is a thermally stable protein. Its thermal denaturation melting point (T_m) measured with a scanning speed of ~ 1 °C/min by differential scanning calorimetry (DCS) was 85 °C or 118 °C with out and with biotin, respectively (Donovan, Ross 1973, Nordlund et al. 2003). In addition, avidin is stable at a wide pH range of 2-13 (Green 1975) and resists many detergents like 9 M urea (Fraenkel-Conrat, Snell & Ducay 1952, Ross, Carson & Fink 1986) or moderate concentrations of carbamimidoylazanium chloride, commonly known as guanidium chloride (GndCl, figure 18) (Green 1963b). Above 3.5 M GndCl the protein begins to unfold and lose its biotin binding activity, and at a concentration of 6 M GndCl the avidin tetramer dissociates (Green 1963b).

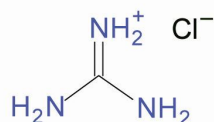


Figure 18. The molecular structure of carbamimidoylazanium chloride.

Ionic and hydrophobic interactions are usually thought to be major determinants of protein stability, but protein intersubunit interactions confer an increased resistance toward thermal stress (Marabotti et al. 2008). This is also true for avidin: several studies with one-to-three and one-to-four monomer-monomer interactions have shown reduced quaternary structure stability (Laitinen et al. 2001). The set of Met96, Val115 and Ile117 to alanine mutants on the one-to-three interface was thought to reduce hydrophobic forces. Asn54 and Asn69, for one, were mutated to alanines in order to eliminate the hydrogen-bonding potential from the one-to-four interface. Asn54Ala and Asn69Ala mutations destabilized forces in the interfaces in such a way that avidin became completely monomeric even at room temperature. However, the presence of biotin induced tetramerization (Laitinen et al. 2001). The combination of mutants Asn54Ala and Trp110Lys produced an avidin, which appeared to be monomeric in solution both in the absence and presence on biotin with a lower biotin binding ability ($K_d \approx 10^{-7}$) (Laitinen et al. 2003). Because Trp110 and Trp120 in adjacent monomers cause a significant increase in the binding affinity of avidin and streptavidin, respectively (Chilkoti, Tan & Stayton 1995, Laitinen et al. 1999), no high-affinity monomeric avidin or streptavidin has been developed. The closest attempt so far is a streptavidin where three out of four binding sites are blocked with mutations (Howarth et al. 2006). This protein has proved its potential in biochemical methods and it can be used to label specific cell surface proteins with biophysical probes on mammalian cells (Howarth, Ting 2008).

3. Aims of the study

Several years of biotin-binding protein research around the world has yielded hundreds of products for biotechnology and has led to the rise of the so-called avidin-biotin technology. The characterization of new avidin-like proteins has increased our basic knowledge of the avidin-biotin interaction. The current study was based on previous findings made in our group and was aimed at generating novel molecules to be added to the “tool box” of avidin-biotin technology.

The specific aims of the study were:

1. To produce chimeric forms of avidin and AVR4/5 to study the origins of the high stability of AVR4/5
2. To characterise the thermal stability and solvent stability properties of chimeric avidin ChiAVD(I117Y)
3. To determine the 3D structure of ChiAVD(I117Y)
4. To change the binding properties of avidin to prefer HABA over biotin by circular permutation and site-directed mutagenesis
5. Characterize the properties of a novel biotin-binding protein A

4. Materials and Methods

The materials and methods are described in detail in the original publications I-IV

4.1 Construction of expression vectors for avidins

The avidin cDNA core (Gope et al. 1987) or the fusion of the avidin cDNA core with the bacterial expression signal sequence (Hytönen et al. 2004) from *Bordetella avium* ompA protein (Gentry-Weeks et al. 1992) were used as primary templates for all avidin mutants (I, III). A cDNA clone of BBP-A [GenBank: BX930135] was obtained from the UK Chicken EST Consortium (ARK Genomics).

4.1.1 Production of the chimeric avidin mutant in insect cells

Chimeric avidin was originally produced using the Bac-To-Bac™ Baculovirus Expression System (Invitrogen) in insect cells (I) and later with an *E. coli* production system (IV). DNA encoding the chimeric avidin mutant I117Y (thereafter referred as ChiAVD(I117Y)) protein was produced by site-directed mutagenesis using the QuikChange (Stratagene, La Jolla, CA) method. The construct (ChiAVD) for QuickChange mutagenesis was made with three sequential PCR reactions in which the final product was obtained from partially overlapping DNA fragments by annealing them with the PCR reaction. The final product was cloned into the pFASTBAC1 vector that was used to produce recombinant baculoviruses as instructed by the manufacturer of the Bac-To-Bac™ system (Invitrogen).

4.1.2 *E. coli* expression vectors

In order to construct a DNA that encodes for circularly permuted cpAVD4→3, two PCR reactions were performed with two sets of oligonucleotides and the cDNA for wt avidin from the plasmid pGem+Av (Airenne et al. 1994) as template. The resulting DNA fragments were connected by overlapping primers: PCR products from the first and second PCR reactions where the overlapping DNA sequence encoded the polypeptide linker connecting the original termini. Mutagenesis of cpAVD4→3 was performed with the QuikChange mutagenesis kit (Stratagene, La Jolla, CA, USA). To create the expression construct for dcAVD2, DNA encoding cpAVD4→3 was digested with the restriction enzymes BamHI and HindIII and subcloned into a pGemTeasy vector that already contained the DNA sequence encoding cpAVD5→4 (Nordlund et al. 2004).

A cDNA clone for BBP-A [GenBank: BX930135] was obtained from the UK Chicken EST Consortium (ARKGenomics). In order to express BBP-A, dcAVD2, cpAVD4→3 and cpAVD4→3(N118M) proteins in *E. coli*, the signal peptide from *Bordetella avium* OmpA protein (Gentry-Weeks et al. 1992) was inserted by subcloning the DNA fragments into the pGemTeasy vector (Promega) containing the OmpA-signal peptide and flanking attL recombination cloning sequences (Hytönen et al. 2004) (II, III). For protein expression, the constructs were transferred to the pBVboostFG expression vector (Airenne et al. 2003, Laitinen et al. 2005) using LR recombination (Invitrogen). All DNA constructs were confirmed by DNA sequencing (I-III).

The chimeric avidin mutant I117Y was transformed from the insect cell production (I) system to the *E. coli* production system (IV). The OmpA secretion signal peptide was built into the ChiAVD(I117Y) by stepwise elongation of the sequence polymerase chain reaction (SES-PCR) in two separate PCR reactions using pFASTBAC1 (ChiAVD(I117Y)) as template. The PCR product was then subcloned into the pET101-D vector using directional topoisomerase (TOPO) cloning (Invitrogen) and confirmed by sequencing.

4.2 Protein expression and purification

The pFASTBAC1-cloned expression construct of ChiAVD(I117Y) was used to create recombinant baculovirus genomes by homologous recombination in *E. coli* DH10BAC cells. The recombinant viral genome was then used to transform Sf9 insect cells (derived from ovarian tissue of *Spodoptera frugiperda* pupa, Invitrogen). Recombinant baculoviruses were then collected from the growth medium and used to infect Sf9 cells in biotin-free medium. The infection was allowed to continue for 72 hours, after which cells were collected by centrifugation. Soluble protein was isolated from the supernatant after suspension into lysis buffer and sonication, in a single step by affinity chromatography using 2-iminobiotin-Sepharose 4 Fast flow (Affiland) columns as described previously (I) (Airenne et al. 1997). ChiAVD(I117Y) (IV) as well as BBP-A, dcAVD2 and cpAVD4→3 (II-IV) and their mutants were later produced in *E. coli* BL21-AI cells (Invitrogen).

Because of the toxicity of avidin proteins (Beaumont 1833, Steinitz 1898), a tight induction control is needed for expression in *E. coli*. This was achieved with the arabinose-inducible T7 polymerase gene promoter, whose expression was further repressed by glucose. The chemically competent BL21-AI cells (Invitrogen) were transformed using the manufacturers protocol and allowed to grow in LB plates supplemented with 0.1% glucose, and the appropriate antibiotics. Tetracycline and gentamycin (II-III) were used to maintain pBVboostFG whereas tetracycline and ampicillin (IV) were used to maintain pET101-D. After an incubation overnight at 37 °C, colonies were transformed to 10 ml of LB medium containing the appropriate antibiotics and 0.1% glucose and then incubated again overnight at 28 °C in an orbital shaker. The bacterial suspension was diluted to the final growth volume containing only ampicillin or gentamycin and 0.1% glucose and allowed to grow until it reached an optical density of 0.1 at A600. Cultures were then induced with 0.2% L-(+)-arabinose (w/v) and 0.1 mM IPTG. After overnight incubation of at 28 °C, cells were collected by centrifugation. The cell pellet was then suspended into hypertonic SET buffer (0.5 M sucrose, 1 mM EDTA and 200 mM Tris, pH 7.4) supplemented with lysozyme. Sonication of the cell suspension at +4 °C was followed by dilution to HilloI buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Triton-X100, 2 mM EDTA) and centrifugation. Soluble protein was recovered from the supernatant by affinity chromatography using either 2-iminobiotin- or biotin-

Sepharose 4 Fast flow (Affiland) columns as has been previously described (Airenne et al. 1997). The protein concentration was determined using a UV/Vis spectrophotometer by measuring the absorbance at 280 nm.

4.2.1 Fermentation

ChiAVD(I117Y) was produced in pilot scale in a bioreactor to carry out protein consuming assays like for example affinity measurements with isothermal titration calorimetry. The fermentation was started in the same way as the Erlenmayer flask productions by transferring colonies from agar plates to the LB medium followed by overnight incubation at 26-28 °C. One ml of this culture was used to start a new, approximately 300 ml culture, in a medium more suitable for fermentation (50 mM potassium phosphate, 30 mM sodium phosphate, 2 mM MgSO₄, 30 mM glucose, yeast extract, trace metals and glycerol). After overnight incubation at 26-28 °C in an orbital shaker with a stirring speed of 200 rpm this liquid culture was used for inoculation.

The fermentation was carried out in a Labfors 3 (Infors, Bottmingen, Switzerland) bioreactor, which has a working volume of 5 l. The fermentation medium was the same as used for inoculation, but did not contain tetracycline. The fermentation chamber was sterilized together with the fermentation medium, and the cell culture, antibiotics, glucose, trace metals, magnesium sulfoxide, feed solution and induction substrates were added using aseptic methods. The oxygen concentration during fermentation was controlled with the stirring speed and air flow control. The fermentation process started with a stirring speed of 500 rounds per minute (rpm). It had a feed-back loop to control the stirring speed between 150 and 900 rpm to keep the oxygen concentration (pO₂) at 20%. At the beginning the optical density (OD) was approximately 0.3. When the OD reached 0.8 – 1.2, feed solution was pumped into the fermentation solution with a flow speed of 5%. Protein production was induced at OD 1.0 – 1.5 by adding 0.25 mM IPTG and 0.2% L-arabinose (w/v). The temperature was kept at 28 °C and the fermentation was stopped after 24 h by collecting the cells by centrifugation. The cells were stored at -20 °C until protein isolation was performed as described in the previous section.

4.3 Characterization of ligand binding

4.3.1 Isothermal titration calorimetry

Isothermal titration calorimetry ITC has been used in our laboratory since 2005 to measure binding parameters for avidin and avidin like proteins (Helppolainen et al. 2007, Helppolainen et al. 2008). The thermodynamic parameters for the binding of avidin to biotin as well as to a number of other small ligands have been characterized (II, III, IV, Table 1).

Proteins were dialyzed against a sodium phosphate (II, III, IV) or potassium acetate buffer (III), both containing 100 mM sodium chloride. After careful weighting, small ligands, were diluted with the last dialysate to assure identical buffer conditions of protein and ligand. The energetic parameters of ligand binding were determined with a VP-ITC Isothermal Titration Calorimeter (Microcal). Titrations were typically performed at 25 °C or at various temperatures to determine the heat capacity (C_pL) of ligand binding (IV, Fig. 5).

Both protein and ligand solution were degassed before injections to avoid air bubbles in the samples. Absorbed or released heat in the ligand-binding reaction was measured to analyze binding parameters, i.e. the binding association constant K_b , change in enthalpy ΔH , stoichiometry n and change in entropy ΔS . Protein and ligand concentrations were carefully adjusted to achieve a proper saturation curve. This was achieved by calculating the unitless c value:

$$c = n \times [M_l] \times K_b \quad (19)$$

where, n is stoichiometry, $[M_l]$ is molar concentration and K_b the binding affinity constant of the macromolecule. The concentration of the macromolecule was adjusted so that the c value was between 5 and 50 using an approximation of the affinity constant. If a measurement was not satisfying, it was repeated with different concentrations. The ligand concentration was approximately ten times higher than that of the macromolecule.

The data were analyzed with the Origin 7.0 software using the “One Set of Sites” method provided by the manufacturer. The titration curve (heat change $\mu\text{cal}/\text{injection}$) resulting from the injections was analyzed by fitting the data into a nonlinear least square curve. If estimation of the binding constant of the ligand to

the protein was impossible due to tight binding over the limit of the instrument ($K_b > 10^{12}$, without using competitive method $K_b > 10^9$), only the enthalpy and the stoichiometry of the binding were determined. Exceeding of the instrument's limit can be seen from a high error (>15%) in the binding constant value, from the shape of the fitting curve and from a high Chi-square value.

4.3.2 Fluorescence emission spectroscopy assay

Fluorescence spectroscopy was used to measure the dissociation rate of fluorescent biotin from the proteins (I-III). The analysis was performed by measuring the reversal of quenching of the biotin coupled fluorescent probe ArcDia™ BF560 (ArcDia Ltd., Turku, Finland) after addition of a 100-fold molar excess of free biotin. Analyses were made in a 50 mM sodium phosphate (pH 7.0) buffer containing 650 mM NaCl. Due to the slow dissociation of some proteins at 25 °C, measurements were also done at 50 °C to increase the speed of complex dissociation. Measurements were done as previously described using a PerkinElmer LS55 luminometer (Hytönen et al. 2004). The fluorescence data were interpreted using the “single-phase dissociation model”, as described elsewhere (Hytönen et al. 2004) and the dissociation rate constants (k_{diss}) were determined using the equation:

$$-k_{diss}t = \ln \frac{B}{B_0} \quad (20)$$

where B_0 is the maximal binding (100%) measured (the difference between the fluorescence of the free dye and that of the protein–dye complex) and B is the amount of complex measured as a function of time. The release of fluorescent biotin was followed for 1 h to ensure total equilibrium of measurements.

4.3.3 Radiobiotin dissociation assay

The dissociation of radioactive biotin (d -[8,9- 3 H]biotin, Amersham) from proteins (I-III) was determined at different temperatures as described by Klumb and co-workers (Klumb, Chu & Stayton 1998). Briefly, the displacement of radioactive biotin by a 1000-fold excess of d -biotin as a function of time was followed and the dissociation rate constant was determined from the fraction of biotin bound at each

time-point using the global fit method (I) according to Hyre et al. (2000) or from the slope of the line fit to the plot of $\ln(\text{fraction biotin bind})$ versus time according to the integrated expression for a first-order reaction rate (II, III) as presented by Klumb, Chu & Stayton (1998):

$$\ln\left[\frac{(x_t - x)}{(x_t - x_0)}\right] = \ln(\text{fraction bound}) = -k_{diss}t \quad (21)$$

4.3.4 Surface plasmon resonance (SPR)

The ligand-binding characteristics of avidins were analyzed with a surface plasmon resonance optical biosensor (I). The IAsys instrument was used to analyze association and dissociation from 2-iminobiotin, which was covalently linked to the sensor chip as previously described (Laitinen et al. 2002).

4.4 Thermostability and solvent stability

4.4.1 Differential scanning calorimetry

The thermostabilities of avidin proteins (I, II) were determined by comparing the transition midpoints (T_m) of heat denaturation using a Calorimetry Sciences Corporation Nano II (Lindon, UT, USA) differential scanning calorimeter (DSC) in collaboration with the group lead by Professor Slotte (Åbo Akademi University, Turku, Finland) as described in detail earlier (Gonzalez, Argarana & Fidelio 1999, Nordlund et al. 2003). Proteins were analyzed both in the absence and presence of D-biotin.

4.4.2 SDS-PAGE based transition temperature analysis

The stability of proteins (II, III) was also studied with a SDS-PAGE -based thermostability assay (Bayer, Ehrlich-Rogozinski & Wilchek 1996). Chemically acetylated proteins in the presence or absence of ligands were subsequently subjected to thermal treatment at many different temperatures for 20 min in the

presence of SDS and 2-mercaptoethanol. After treatment the oligomeric state of the protein was analysed by SDS-PAGE and Coomassie blue staining.

4.4.3 Microplate assay

A microplate assay was used to determine at which temperature proteins lose their ligand-binding activity. Proteins were held at 99.9 °C for 32 min and the remaining activity was probed by measuring the ability of the proteins to bind biotinylated alkaline phosphatase (BTN-AP) (I). Heat-treated protein samples were coated onto the wells of a microplate and the amount of bound BTN-AP that was determined by following the colorimetric reaction on the microplate with a spectrophotometer (Nordlund et al. 2003).

4.4.4 Solvent stability assay with resin linked proteins and ³H-labelled biotin

The solvent stability of proteins was analyzed by incubating proteins, which were covalently attached to the sepharose 4FF resin, in different organic solvents (IV). The stability of proteins was determined by measuring their ability to bind ³H-labelled biotin after solvent treatment (IV, Fig. 4A) or in the presence of solvent (figure 21). The amount of bound tritium-labelled biotin could be determined by measuring the amount of tritium-labelled biotin remaining in the supernatant after separation from the functional resin-coupled protein by filtration. The solvent used for control samples was water or formic acid. Water was assumed to have no effect on protein and conversely formic acid was assumed to destroy all activity of the proteins analyzed.

4.5 Structural analysis by X-ray crystallography

4.5.1 Circular permuted avidin Loop3-4 and Biotin Binding Protein A (II,III)

The crystal structures of cpAVD4→3 (II) and BBP-A (III) were solved by X-ray crystallographic analysis in collaboration with the group led by Professor Mark Johnson (Åbo Academy University, Turku, Finland). The optimal conditions for crystallization were screened by the sitting drop or vapour diffusion method either with a HamiltonSTAR robot in the Institute of Biotechnology at the University of Helsinki or with the ClassicsTM screen (Nextal Biotechnology, Qiagen, Valencia, CA, USA). After optimization, suitable crystals were harvested and diffraction data sets were collected at the MAX-lab beam line I711 using a MarCCD detector (Lund, Sweden). A more detailed experimental description can be found in the original publications (II, III).

4.5.2 Chimeric avidin (I117Y) (IV)

The X-ray crystallographic analysis was also used to solve the crystal structure of ChiAVD(I117Y) in collaboration with Professor Oded Livnah (The Hebrew University of Jerusalem, Jerusalem, Israel). Initial crystals of the ChiAVD were obtained by the hanging-drop vapour diffusion method. Later crystallization conditions were screened using the Douglas Instruments ORYX6 automatic protein crystallization system with the Hampton Research screening kits for better quality crystals. Crystals were obtained at 293 K using first the microbatch and then the hanging drop method. Diffraction data sets were collected at the ESRF beam line ID14-4 using an ADSC Q315R CCD detector. The methods and procedures are described in detail in the original publication (IV).

4.6 Mass spectrometry

Mass spectrometric analysis (III, IV) was performed in collaboration with Dr. Janne Jänis (University of Joensuu, Joensuu, Finland). The experiment was done on a

Fourier transform (FT) ion cyclotron resonance (ICR) mass spectrometer (Bruker Daltonics, Billerica, MA, USA) that had a mass-selective quadrupole interface and electro spray ionization (ESI) for ion source. Data was processed using the Bruker XMASS 7.0.8 software.

4.7 Fast Protein Liquid Chromatography Gel Filtration

The size and quality of the purified proteins (I, II, III) were analyzed using gel filtration chromatography. Proteins were analyzed using a Superose 6 10/300GL (Tricorn) (III) or SuperDex 200 10/300GL (Tricorn) (II) column in an ÄKTÄ™ purifier 10 FPLC instrument (Amersham Biosciences) as previously described (Nordlund et al. 2003). Proteins were diluted with 50 mM sodium phosphate buffer containing 650 mM NaCl (pH 7.0). Biotin-complexed proteins were prepared by incubating the sample for 10 min in the presence of an excess of biotin prior to analysis.

4.8 Computer programs

The molecular representations were created with the PyMOL molecular Graphics System (DeLano 2008) (II,III) or MIDAS(Ferrin et al. 1988) and edited with the Corel Draw 11 (II) and X3 (III) or Photoshop CS3 program suite (III, IV). The multiple sequence alignment of BBP-A and avidin was made using the program Malign implemented in BODIL (Lehtonen et al. 2004). The solvent-accessible surface areas were calculated with the program Areaimol (Lee, Richards 1971) of the CCP4i suite (II,III) (Collaborative Computational Project, Number 4 1994, Potterton et al. 2003). The figures in “Review of the literature” section were created with the PyMOL molecular Graphics System (DeLano 2008) and/or prepared with Photoshop CS3. All the programmes used in the analyses are mentioned in the original communications (I-IV)

5. Review of the Results

5.1 Chimeric avidin mutants (I, IV)

The avidin gene family in chicken consists of avidin and up to seven avidin related proteins (AVRs) (Ahlroth et al. 2000). AVRs have been produced as recombinant proteins and their functional and structural properties have been carefully characterized (Laitinen et al. 2002). AVR proteins have many interesting features when compared to avidin. One of them, the identical product of the genes *AVR4* and *AVR5*, and therefore, named AVR4/5, was found to be the most stable biotin-binding protein encoded by a natural gene with a T_m of 106.4 °C (Hytönen et al. 2004). The interest to explore the structural origins of the high stability of AVR4 was the starting point for studies leading to the hyperthermostable protein ChiAVD(I117Y) (I).

5.1.1 Expression vectors, fermentation and purification of recombinant proteins

A 21 amino acid segment from AVR4 was transferred to avidin and the resulting chimeric protein was named ChiAVD. In another approach, effects of the mutation I117Y were studied, because this residue is located on the structurally important interface between subunits and has been shown to contribute to the stability of avidin by previous studies (Nordlund et al. 2003, Laitinen et al. 2001). These two modifications were combined in ChiAVD(I117Y) as presented in figure 19.

The ChiAVD(I117Y) was first produced using a baculovirus expression vector in insect cells (I) and later in *E. coli* (IV) using the signal peptide from the ompA protein as previously described (Hytönen et al. 2004). Avidin proteins were purified effectively by 2-iminobiotin-agarose affinity chromatography. In order to get large amounts of protein for X-ray crystallography (IV) and for more detailed binding and stability assays, ChiAVD(I117Y) was also produced with pilot scale fermentation

(IV). A typical yield was approximately 50 mg per one litre of fermentation of *E. coli* bacteria. Optimization the fermentation protocol by applying dissolved oxygen controlled feed rise the protein yield above 200 mg/l of fermentation medium. The produced protein was confirmed to by mass spectrometry (figure 20). The measured most abundant isotopic mass (14380.31 Da) was very close to the calculated value (14380.51). The result confirms that the protein is as expected and that the sulphur bridge between cysteine 4 and 81 is formed.

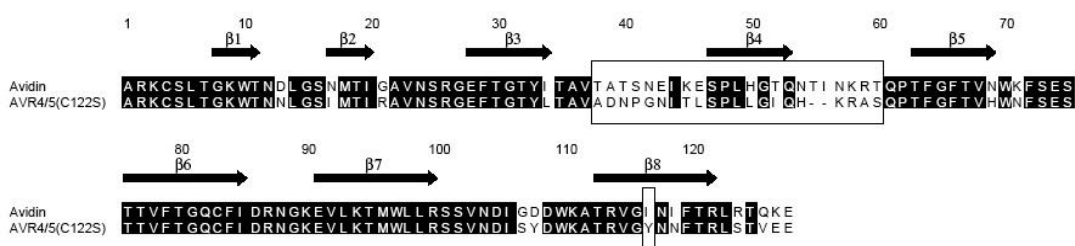


Figure 19. An alignment of chicken avidin and AVR4. Boxed areas indicate the parts of AVR4 which were transferred to avidin, to form ChiAVD(I117Y).

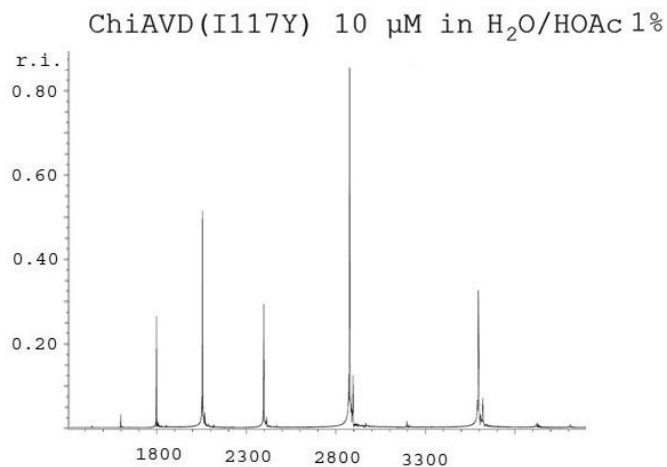


Figure 20. The mass spectrometry result of ChiAVD(I117Y). The theoretical most abundant isotopic mass for ChiAVD(I117Y) is 14380.50 Da. The measured most abundant isotopic mass for ChiAVD(I117Y) was 14380.31 Da.

5.1.2 Ligand-binding analyses

The chimeric avidin mutant I117Y showed biotin binding properties similar to those of avidin or AVR4. The binding properties of the proteins were analyzed by SPR (I, Table 3), as well as radiobiotin (I, Fig. 4) and fluorescent biotin dissociation assays

(I, Table 4). The SPR analysis showed a slightly decreased affinity to the 2-aminobiotin surface for ChiAVD(I117Y) when compared to avidin, $(6.7 \pm 1.2) \times 10^{-8}$ M and $(2.1 \pm 0.6) \times 10^{-8}$ M. ChiAVD and ChiAVD(I117Y) resembled AVR4 in their behaviour in the radioactive biotin dissociation assay, but they had significantly higher dissociation rate constants than wt avidin. The fluorescent biotin dissociation assay showed that the I117Y mutation seemed to tighten the binding of the biotin conjugate to avidin. At 25 °C ChiAVD(I117Y) ($1.94 \times 10^{-5} \text{ s}^{-1}$) had a dissociation rate similar to wt avidin ($2.04 \times 10^{-5} \text{ s}^{-1}$) and ChiAVD(I117Y) and wt avidin both had a lower dissociation rate than AVR4 ($2.77 \times 10^{-5} \text{ s}^{-1}$) (I, Table 4). The enthalpy and stoichiometry of the biotin-binding reaction for ChiAVD(I117Y) determined with ITC (IV, Fig. 5), were 25.7 kcal/mol and 0.93, respectively. The determination of biotin-binding affinity by ITC proved to be impossible because of the limitations of the method ($K_d < 10^{-9}$ M). The binding properties of the different avidins are shown in Table 2.

Table 2. Binding properties of avidins

ITC						
Biotin 25 °C	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)	n	K_d (M)	$L\Delta C_p$ (cal/mol °C)
AVD	-25.7	ND	ND	1.0	$<10^{-9}$	-461
ChiAVD(I117Y)	-26.9	ND	ND	1.1	$<10^{-9}$	-270
HABA 25 °C						
AVD	1.8	9.1	-7.3	0.7	4.8×10^{-6}	ND
ChiAVD(I117Y)	0.8	6.7	-5.8	0.9	5.2×10^{-5}	ND
AVR4	1.7	8.8	-7.1	1.0	5.9×10^{-6}	ND
SPR						
	K_d (M)	k_{ass} ($M^{-1}s^{-1}$)	k_{diss} ($10^{-5} s^{-1}$) 25 °C	Release 1 h (%) 25 °C	k_{diss} ($10^{-4} s^{-1}$) 50 °C	Release 1 h (%) 50 °C
AVD	$(2.1 \pm 0.6) \times 10^{-8}$	$(2.6 \pm 0.2) \times 10^4$	2.04	10.7	2.74	71.5
AVD(I117Y)	$(6.3 \pm 1.2) \times 10^{-8}$	$(2.3 \pm 0.3) \times 10^4$	0.81	3.0	1.61	56.5
ChiAVD	$(1.1 \pm 0.4) \times 10^{-7}$	$(1.7 \pm 0.1) \times 10^4$	1.76	13.1	8.00	93.9
ChiAVD(I117Y)	$(6.7 \pm 1.2) \times 10^{-8}$	$(1.6 \pm 0.2) \times 10^4$	1.94	9.5	7.33	94.8
AVR4	$(1.4 \pm 0.4) \times 10^{-7}$	$(9.3 \pm 0.4) \times 10^3$	2.77	12.5	7.88	93.7

ND = not determined

5.1.3 Thermal and solvent stability of avidins

The thermal stability of proteins was studied by DSC analysis and also with the microplate assay (Nordlund et al. 2003). In the DSC experiments the holoforms of the proteins had higher T_m values than the apoforms (I, Table 2). ChiAVD(I117Y) is the most heat-stable avidin analyzed so far ($T_m = 111.1$ °C). ChiAVD(I117Y) bound to biotin had a T_m value higher than the upper limit of the instrument, i.e. >130 °C.

The microplate assay revealed that ChiAVD(I117Y) retains 98% of its activity after treatment at 99.9 °C for 32 min, whereas the other proteins analyzed lost a significant portion of their activity during a similar treatment (I, Table 2). An SDS-PAGE based analysis of the transition temperature (T_r) was in line with previous findings. ChiAVD(I117Y) had an approximately 20 °C higher T_r in both its apo- and holoforms than avidin (IV, Fig. 3B). The stability of the avidins is presented in Table 3.

Table 3. Structural properties of avidins

Protein	SDS-PAGE		DSC ^a			Microplate assay	
	T_r (-biotin) (°C)	T_r (+biotin) (°C)	T_m (-biotin) (°C)	T_m (+biotin) (°C)	ΔT_m^b (°C)	ΔH^c (kJ/mol)	Activity ^d (%)
AVD	65	85	83.5 ± 0.1	117.0 ± 0.7	33.5	329 ± 5	4
AVD(I117Y)	ND	ND	97.5 ± 0.4	123.7 ± 0.1	26.2	536 ± 6	47
ChiAVD	ND	ND	96.5 ± 0.2	124.4 ± 0.2	27.9	527 ± 10	47
ChiAVD(I117Y)	95	110	111.1 ± 0.2	~130 ^e	~19	659 ± 38	98
AVR4	ND	ND	106.4 ± 0.8	125.4 ± 0.8	19.5	575 ± 33	72
SA	65	90	75.5	112.2	36.7	629	ND

^aThe results for AVD and AVR4/5(C122S) are from (Laitinen et al. 2002) and the results for SA are from (Gonzalez et al. 1997).

^b ΔT_m is the change in T_m upon addition of a three-fold molar excess of biotin.

^c ΔH value obtained from a sample without biotin. The value could not be determined accurately from samples saturated with biotin because the T_m values were too close to the temperature limit of the DSC instrument.

^dBiotin-binding activity after a 32 min treatment at 99.9 °C

^eCould not be determined accurately due to upper temperature limit of the DSC instrument.

ND = not determined

5.1.3.1 Solvent stability assay with resin linked proteins and tritium labelled biotin

High thermostability of ChiAVD(I117Y) gave us the idea to investigate its stability in different kinds of organic solvents. If the protein could stay active in the presence of organic solvents, it could possibly be used in applications which require the capacity to resist exotic chemical environments or long storage.

ChiAVD(I117Y) showed higher stability in most of the organic solvents when compared to avidin or streptavidin (IV, Fig. 4). ChiAVD(I117Y) was able to bind radiobiotin despite of treatments with methanol, acetonitrile or DMSO at an elevated temperature. For example, streptavidin lost 52% of its activity after treatment with 100% methanol for 4 h at 50 °C whereas ChiAVD(I117C) lost only 16% of its activity in such a treatment. Biotin increases the stability of ChiAVD(I117Y) as expected as shown in figure 21. Based on the earlier assays it was also expected that ChiAVD(I117Y) would not lose any of its binding capacity in water (figure 21).

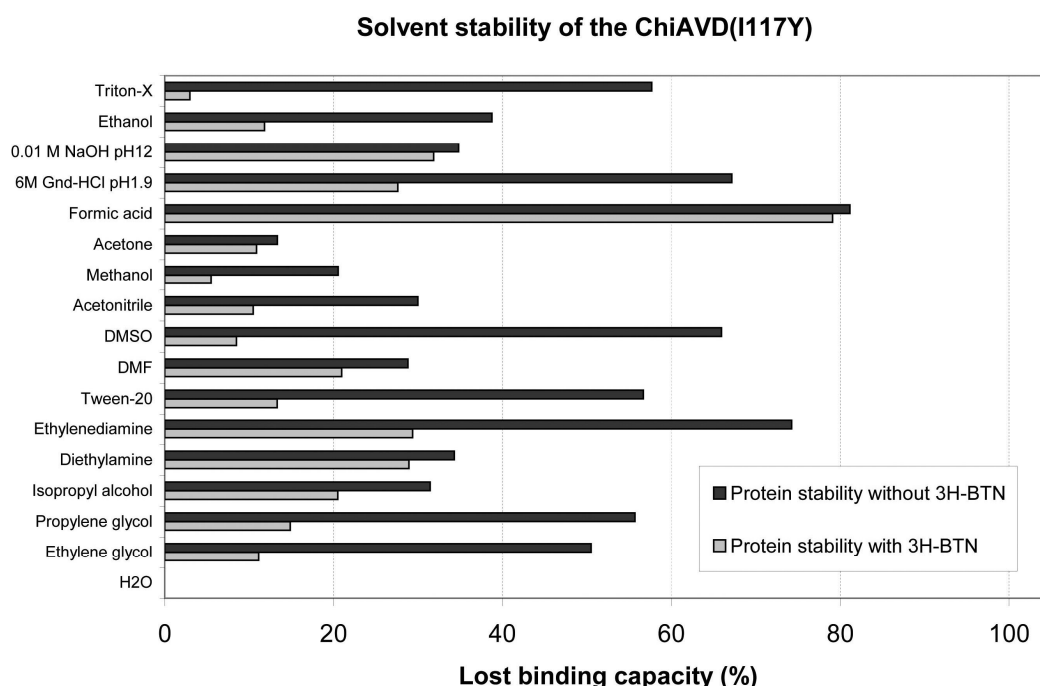


Figure 21. The loss of the binding capacity of ChiAVD(I117Y) after four hours treatment in organic solvents and at an elevated temperature (50 °C) with or without ³H-BTN (Määttä *et al.* unpublished result).

5.1.4 Structural analysis by X-ray crystallography

The three dimensional crystal structure of ChiAVD(I117Y) (resolution: 2.7 Å, figure 22) was solved by X-ray crystallography. The structure of the chimeric avidin mutant was, as expected, an eight-stranded up-and-down β -barrel, the other parts of which were similar to avidin, but the loop between β strands three and four (L3,4 loop) was shorter, like in AVR4. This shorter loop has a modest effect on biotin binding. In streptavidin, the corresponding loop has an evident influence, and it is the main reason for weaker biotin binding (Green 1990). The point mutation (I117Y) in the subunit interface of avidin was thought to lead to the formation of an π - π stack between two neighbouring monomers like in the case of AVR4 (pdb: 2FHL). This is indeed what occurs in the area between intermonomeric interfaces of ChiAVD(I117Y) (IV, Fig. 2 and figure 22). Overall, the structure of ChiAVD(I117Y) closely resembles that of AVR4/5, and the 3D structure revealed no simple structural explanations for the increased stability.

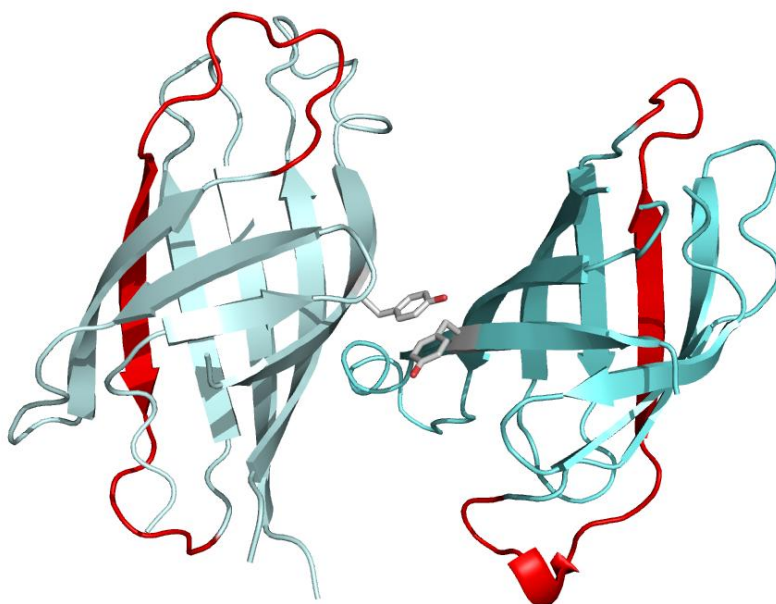


Figure 22. A Schematic presentation of two subunits of ChiAVD(I117Y). The 3D structure determined by X-ray crystallography was used to generate the figure (pdb: 3MMO). The red part represents the region taken from AVR4 (described in figure 19). Tyr117 residues are presented as sticks to illustrate the stabilizing π - π interaction, which is formed between the interfaces of the two subunits.

5.2 Structure, stability and ligand binding mode of BBP-A (II)

Avidin is not the only biotin-binding protein found in chicken egg. It has been reported that the egg contains biotin-binding proteins I and II (BBP-I, BBP-II) in its yolk (White et al. 1976, Meslar, Camper & White 1978, White, Whitehead 1987, Bush, McGahan & White 1988, Subramanian, Adiga 1995). Although it is not known for sure whether the cDNAs encoding the proteins *BBP-A* and *BBP-B* (Niskanen et al. 2005) are same as the BBP I and II proteins characterized earlier, the chicken genome project made it possible to produce them as recombinant proteins and characterize them more carefully. According to the studies, BBP-A can be classified as a new member of the AVD family.

BBP-A was expressed both in *E. coli* (bBBP-A) and in insect cells (iBBP-A) and isolated using 2-iminobiotin affinity chromatography. According to gel filtration chromatography (II, Table 1), BBP-A is tetramer, and this was also supported by the 3D crystal structures. BBP-A has one N-glycosylation site (II, Fig. 2), which is glycosylated in the iBBP-A recombinant protein, and the size of the carbohydrate is similar to that of wt AVD.

The tetrameric BBP-A was less stable ($T_r = 70$ °C in the presence of biotin) in an SDS-PAGE-based analysis than wtAVD ($T_r = 90$ °C in the presence of biotin) (II, Table 1). The reason for this may be the 10% larger solvent accessible surface area, which was calculated from the crystal structures of BBP-A (pdb: 2C1Q, 2.10 Å; 2C1S, 1.75Å) and AVD-BTN complex (pdb: 1AVD; 2AVI). The crystal structure also revealed that BBP-A has a more positively charged surface than avidin (II, Fig. 5c-f) and a few differences in ligand binding mode. The crystallographic analysis was done by adding diluted biotin to the crystallization solution before crystallization. To our surprise, after determining the structure, we detected a D-biotin D-sulfoxide (BSO) molecule in the crystal, which is presented in figure 23. The binding mode of BSO was almost identical to that of BTN. BSO has not been found in any other protein structures presently in the Protein Data Bank (PDB) (Berman et al. 2002).

The affinity of BBP-A towards BTN is quite similar to that of streptavidin, and therefore, about 50 times lower than that of egg-white avidin. Interestingly, according to the determined dissociation rate constants, BBP-A binds BSO even

more tightly than BTN, whereas AVD has a dissociation rate constant 20 times greater when bound to BTN compared to BSO.

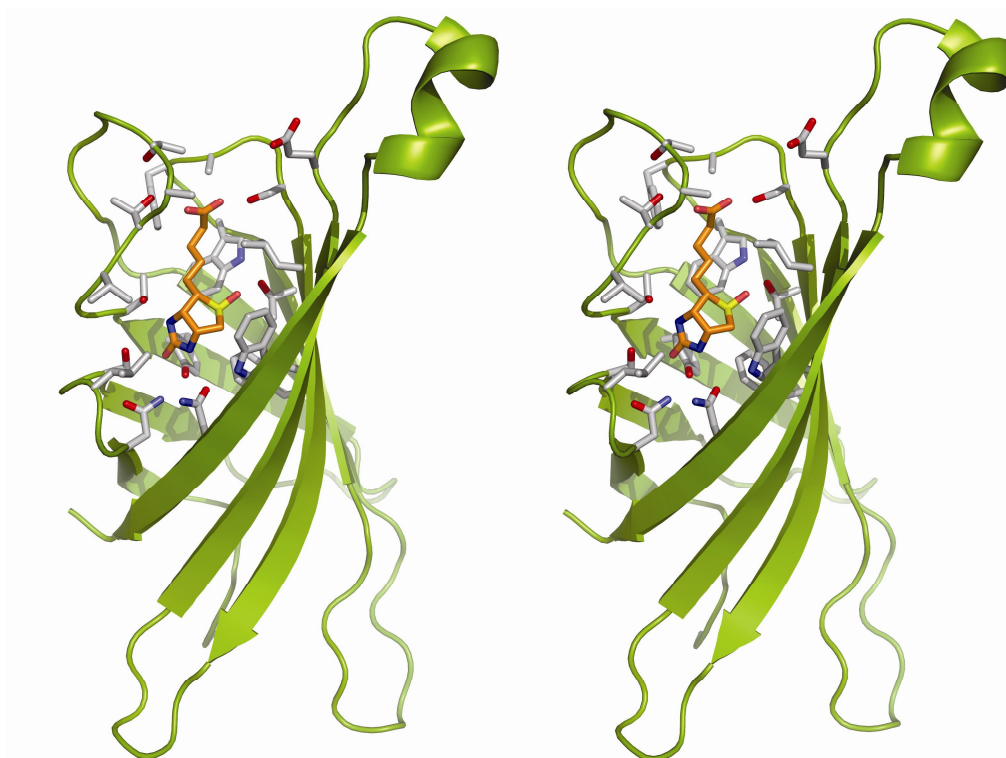


Figure 23. The X-ray structure of the chicken BBP-A binding pocket and D-biotin D-sulfoxide. Amino acid residues within 4 Å of BSO in the BBP-A – BSO structure are presented in gray.

5.3 Rational modification of avidin by circular permutation and mutagenesis (III)

The influence of the loops to the ligand affinity of avidin and streptavidin has been carefully studied (Pazy et al. 2002). These flexible loops near the active sites can be viewed as gates that are open to allow ligand association, and close to prevent solvent access and to form bonds, thereby minimizing dissociation. In particular, the shorter L3,4-loop and the longer L5,6-loop of streptavidin together with the absence of a few other hydrogen bonds as compared to avidin, have been shown to be the main reasons for weaker BTN-binding affinity. The same structural differences are also thought to have an effect on the HABA-binding affinity of AVD and SA. Avidin binds HABA more tightly ($K_d \approx 6 \times 10^{-6}$ M) than SA ($K_d \approx 1 \times 10^{-4}$ M)

(Green 1990). Interestingly, the L3,4-loop has an open conformation when avidin is complexed with HABA. Therefore, the L3,4-loop does not seem to contribute significantly to the binding of HABA. In theory, the removal of the loop could even increase affinity to HABA and significantly decrease affinity to BTN. In fact, the decrease in the biotin-binding affinity of streptavidin by L3,4-loop removal has been shown in an earlier study (Chu et al. 1998). This is also supported by previous findings, where Chu and co-workers found AVD to be more susceptible to proteolysis by proteinase K in the presence of HABA, suggesting enhanced loop mobility in avidin due to HABA binding (Ellison et al. 1995).

In order to test this hypothesis, we designed and produced a circularly permuted form of avidin, where the loop between β -strands 3 and 4 is deleted (cpAVD4 \rightarrow 3). The protein was successfully expressed in *E. coli* BL21-AI cells and characterized by spectrophotometric methods. Interestingly, the ITC experiments together with ligand-dissociation analyses using fluorescent biotin and radiobiotin assays showed that the biotin-binding affinity of mutant was significantly reduced ($K_d = 1.4 \pm 0.3 \times 10^{-8}$ M) while its HABA-binding affinity was increased by 1.5-fold ($K_d = 5.4 \pm 0.7 \times 10^{-6}$ M) (III, Table 2; Fig. 2: Table 3).

The structure of the HABA-avidin complex provides a good opportunity for rationally modifying avidin. To increase the HABA-binding affinity of avidin, Asn118, which does not make a favourable interaction with HABA in wt AVD (Livnah et al. 1993a), was mutated to Met (III, Fig. 6). Another aim was to reduce BTN's affinity to avidin. Because Asn118 makes an important hydrogen bond with BTN (pdb: 2AVI, figure 15A), it was a rational target for mutagenesis. Mutagenesis of a corresponding site in streptavidin also led to a decrease in biotin affinity (Qureshi et al. 2001), supporting our strategy. Our experiments proved that the extremely strong femtomolar biotin-dissociation constant of wt AVD was reduced to a nanomolar value ($K_d = 4.2 \times 10^{-9}$ M) and at the same time the HABA affinity was slightly increased ($K_d = 5.2 \times 10^{-6}$ M).

We assumed that by combining circular permutation and point mutagenesis, the resulting cpAVD4 \rightarrow 3(N118M) construct would be able to bind HABA tighter than BTN. This was not the case, as cpAVD4 \rightarrow 3(N118M) binds BTN approximately four times tighter than HABA. However, cpAVD4 \rightarrow 3(N118M) has the highest affinity towards HABA measured so far ($K_d = 1.0 \times 10^{-6}$ M) and the relative biotin-binding affinity was reduced over one billion-fold to $K_d = 2.4 \times 10^{-7}$ M.

The dual chain concept of avidin made it possible to modify the binding sites of avidin independently (Nordlund et al. 2004, Hytönen et al. 2005b). cpAVD4→3(N118M) was therefore attached to cpAVD5→4, which is another circularly permuted form of avidin (Nordlund et al. 2004). The result, dual-chain avidin 2 (dcAVD2), has two different biotin-binding sites: cpAVD5→4 that has wt AVD-like binding properties, and cpAVD4→3(N118M) that has only moderate BTN-binding affinity and increased HABA-affinity. With further mutations, the specific binding of HABA binding in each site may be increased. Preliminary results from a Thr77 random mutagenesis experiment indicate that the HABA affinity of the Thr77Asp avidin mutant may be even tighter than the HABA affinity of the Asn118Met avidin mutant (Määttä *et al.* unpublished).

6. Discussion

6.1.1 Water molecules in binding events

Osmotic stress has been used to measure the forces of interactions between large molecular surfaces, and the measurements have shown that the energy involved in dehydration varies from 1.5 to 15 kcal/mol per 100 Å of surface area (Rand, Parsegian 1989). The same method has been used to probe the energetics of the hydration of single macromolecules. The binding of oxygen to haemoglobin has been shown to release 60 water molecules (Colombo, Rau & Parsegian 1992). How many water molecules are released when biotin binds to avidin? Based on calculations of molecular surface areas, we estimated, that the binding of biotin to avidin could release 111 water molecules from avidin ($\Delta A = 637 \text{ \AA}^2$) and 72 water molecules from biotin ($\Delta A = 414 \text{ \AA}^2$), assuming a 5.7 \AA^2 area for a single water molecule (Nicholls 2000). This calculation appears to be in line with that of Bennett and Steitz (Bennett, Steitz 1980), who calculated from their hexokinase A structure that on closing there is approximately a 163 \AA^2 decrease in area in the cleft, which is involved in the removal of approximately 26 water molecules.

So far the discussion has focused on water molecules on the surface of the avidin-biotin complex, but water molecules in vicinal layers may also influence entropy. The number of these water molecules is difficult to count, and it is even more difficult to estimate how much impact they have on entropy. For hexokinase, it has been calculated that the number of water molecules in the second and third layers could be equivalent to those on the protein surface (Rand, Parsegian 1989). This would mean that another hundred water molecules, which are organized to some extent, could be released when biotin binds to avidin. Water molecules in the outer hydration shell have a hydrogen bond network more or less similar to bulk water, but it has been shown that hydrophilic surfaces significantly perturb water well beyond the first layer (Rand, Parsegian 1989). Therefore, the increase in entropy due to the dissociation of water molecules from vicinal layers is less than

when water molecules are removed from the inner hydration shell. However, the simultaneous removal of a large numbers of weak bonds may have a significant influence on the energetics of the binding event.

How big is the Gibbs energy contribution from the hydrophobic transfer process? This can be estimated based on both heat capacity changes and changes in solvent-accessible surface area (Spolar, Ha & Record 1989, Ha, Spolar & Record 1989, Livingstone, Spolar & Record 1991, Sharp et al. 1991a, Sharp et al. 1991b, Spolar, Livingstone & Record 1992, Spolar, Record 1994, Janin 1997). Heat capacity change can be used to calculate the Gibbs energy change of hydration (ΔG_{hyd}) from the following equation (Ha, Spolar & Record 1989):

$$\Delta G_{\text{hyd}} = 80 (\pm 10) \times \Delta C_p L \quad (22)$$

And because ITC measurements state that $\Delta C_p L$ for avidin is $-0.46 \text{ kcal mol}^{-1} \text{K}^{-1}$ the corresponding Gibbs energy for hydration is $\Delta G_{\text{hyd}} = -(36.8 \pm 4.6) \text{ kcal/mol}$. This value is large, larger by far than the Gibbs energy change observed in the binding process, and therefore it must be counterbalanced by obligatory conformational changes together with the entropic cost of complex formation.

It is known that the strept(avidin) binding pocket contains water molecules in the absence of a ligand (Weber et al. 1992, Rosano, Arosio & Bolognesi 1999). The organization of amino acid residues in the ligand-binding pockets are thought to be the reason for the tighter binding of biotin to avidin versus streptavidin, because the placement of Asp13 reduces the large desolvation penalty resulting from a first-shell charged residue, and provides good hydrogen-bond cooperativity via juxtaposition with Asn118 in avidin (DeChancie, Houk 2007).

Overall, water molecules are a very important factor in ligand binding, both in the binding pocket and outside of the protein. DeChancie and Houk even state that “The evolutionary driving force for avidin to bind biotin with greater affinity is not achieved by strengthening interactions with the ligand, but by weakening binding interactions with water”, and therefore, their contribution to the binding should be taken into account more carefully (DeChancie, Houk 2007).

Altogether, because the ligand-binding interaction is a very complex event, and the sum of many things like non-covalent interactions and the dehydration of water, it is important to solve the complete thermodynamic profile (ΔG , ΔH , ΔS , $\Delta C_p L$) of the event. This helps in understanding the binding as a process better, and in obtaining knowledge on the driving force of the binding. The information gained

will become valuable when designing new mutants or new ligands with desired properties.

6.2 Development of a stable chimeric avidin (I117Y)

(Strept)avidin appears to be a suitable target for protein engineering. Structural modifications (Nordlund et al. 2004, Hytönen et al. 2005b, Nordlund et al. 2005, Chu et al. 1998), point mutations (Laitinen et al. 2006, Laitinen et al. 2007, Chilkoti, Tan & Stayton 1995, Aslan et al. 2005, Avrantinis et al. 2002, Chilkoti et al. 1995, Le Trong et al. 2003, Pazy et al. 2003, Pazy et al. 2001, Dixon, Kollman 1999, Freitag et al. 1999) and the attachment of fusion partners (Hu et al. 2009, Lesch et al. 2009) have been successfully applied to (strept)avidin without destroying the functional barrel-structure, and therefore, the primary function: its high affinity to biotin.

Avidin-biotin technology is popular in numerous bioanalytical applications (Wilchek, Bayer 1988, Wilchek, Bayer & Livnah 2006) and more await discovery. Although avidin and streptavidin are reasonably thermostable proteins $T_m = 85.5$ °C (Nordlund et al. 2003) and $T_m = 75.5$ °C (Gonzalez et al. 1997), respectively, it could be useful to have an avidin protein, which is stable in conditions where wt avidin would not be sufficiently durable. The discovery of genes encoding avidin related proteins (Ahlroth et al. 2000, Keinänen, Laukkanen & Kulomaa 1988, Keinänen et al. 1994) and their characterization (Laitinen et al. 2002) as recombinant proteins, revealed many novel properties compared to wt avidin. Besides a variety of biotin-binding affinities, ranging from the avidin-like tight affinity of AVR4 (Hytönen et al. 2004) to the moderate affinity of AVR2 (Hytönen et al. 2005a), they also exhibited varying degrees of stability. It was found that compared to wt avidin, AVR4 is a very thermostable protein ($T_m = 106.4$ °C). However, the reason for the increased thermostability was unclear, and a role for Ile117Tyr mutation in the one-to-three interface and the shorter L4,5-loop were given as potential explanations (Hytönen et al. 2004).

Could the properties of different proteins be transferred inside the gene family? We thought that if structurally important parts of AVR4/5 could be transferred to avidin, we could also change its stability, and thereby learn about the origins of the

high stability of AVR4/5. As the result, we made the ChiAVD, Avd(I117Y) and ChiAVD(I117Y) constructs. To make ChiAVD, β -strand 4 and parts of its surrounding loops from AVR4/5 were transformed to wt avidin. Avd(I117Y) in turn has an AVR4/5-inspired point mutation in the one-to-three subunit interface (I, Fig. 1). Both these changes were found to be beneficial for the stability of avidin. Therefore, a protein version containing both modifications was made and named ChiAVD(I117Y). It proved to be the most thermostable avidin protein seen so far ($T_m = 111.1$ °C) (I, Table 2). Despite its altered structure, ChiAVD(I117Y) carries wt-AVD like binding properties.

This result gave us hope that ChiAVD(I117Y) could stay active in high temperatures longer than (strept)avidin. This property could be of advantage in, for example, PCR-based applications involving biotinylated molecules and thermostable biotin-binding proteins. ChiAVD(I117Y) seemed to maintain its binding properties indefinitely through a typical PCR temperature cycling whereas avidin and streptavidin lost their binding properties relatively early (IV, Fig. 3C).

Experiments *in vitro* are often carried out in physiological water surroundings, e.g. in an isotonic sodium phosphate buffer. However, many industrially important enzymes catalyze processes in non-aqueous solvents (Klibanov, Koskinen 1996, Serdakowski, Dordick 2008). The function of enzymes (activity, specificity and stability) in organic solvents is strongly depended on the nature of the solvent (Zaks, Klibanov 1988). This drove us to study the properties of ChiAVD(I117Y) in some of the most commonly used solvents. We found that ChiAVD(I117Y) remained functional in environments and circumstances where avidin and streptavidin lost their binding ability to some extent (IV, Fig. 3 and 4).

6.3 Structure-function relationships of avidin like proteins

The crystal structures of proteins provide good possibilities for understanding structure-function relationships and for further designing and engineering proteins. The structures of BBP-A (II), cpAVD4 \rightarrow 3(N118M) (III), ChiAVD(I117Y) (IV) together with other crystal structures of biotin-binding proteins: streptavidin (SA), tamavidins, avidin (AVD), avidin related protein 2 (AVR2), avidin related protein 4

and 5 (AVR4/5) (Weber et al. 1989, Livnah et al. 1993b, Hytönen et al. 2004, Hytönen et al. 2005a, Helppolainen et al. 2008, Meir et al. 2009, Määttä et al. 2009, Takakura et al. 2009), have offered a reliable general view of the biotin-binding site and taught us how to modify these proteins to meet the needs of biotechnology.

Although BBP-A is biochemically and structurally clearly distinguishable from chicken avidin and the AVRs, its biotin-binding mode is highly similar to that observed for the known structures. The rational mutagenesis of avidins benefits from the solved structure of BBP-A. Sequence comparison could be used as a guide to mutate amino acids in other avidins and thereby change their properties. For example pI could be changed without unwanted changes to secondary or tertiary structures (Marttila et al. 1998). The potential of such an approach has also been shown with the putative avidin-like protein from the sea urchin – fibropellin – and streptavidin (Pazy et al. 2001). Switching tryptophan 110 in avidin or 120 in streptavidin to lysine, according to the fibropellin sequence, reduced the affinities of both avidin and streptavidin towards biotin and led to the dissociation of their tetrameric structures into stable dimers (Laitinen et al. 1999). Later it was found that in avidin the Trp110Lys mutation combined with the Asn54Ala mutation weakened the bonding of interfaces between subunits and resulted in the dominance of the monomeric form of the protein (Laitinen et al. 2001).

Rational mutagenesis has uncovered the nature of the biotin-binding process of streptavidin and avidin. Dozens of streptavidin and avidin mutants have been made as reviewed by our group (Laitinen et al. 2006). Together with the known crystal structures and molecular modelling, it is now possible to design new avidins with tailored binding properties. One target of the present study was to develop an avidin with increased HABA-affinity. To achieve this, we used crystallographic structures to design an optimized binding pocket. The known crystal structure (Livnah et al. 1993a) and the existing potential of HABA in labelling applications (Hofstetter et al. 2000) encouraged us to try whether the affinity of HABA to avidin could be increased over BTN. As the result, cpAVD4→3(N118M), which has a loop missing between β strands 3 and 4 and one point mutation resulting in a more optimal HABA interaction, binds HABA tighter than any known avidin protein. Preliminary results from random mutagenesis experiments have also shown that further mutagenesis of threonine 77 to aspartate could have a positive effect on HABA binding (Määttä *et al.* unpublished results). We have also tried to find better HABA

binders using a phage display method of random mutagenesis in particular sites or even in entire loop regions (Paldanius et al., unpublished). Our final aim is to develop, using either rational or random mutagenesis, an avidin that binds HABA over BTN. This “proof of principle” could be then expanded to develop avidins, with specificity to all kinds of molecules. Moreover, specific binding properties could then be combined into dual chain or single chain avidin scaffolds (Nordlund et al. 2004, Hytönen et al. 2005, Nordlund et al. 2005b) to produce multifunctional avidins. A previous study (Hytönen et al. 2005) already showed that the dual affinity dual chain avidin, which has two tight and two reversible biotin-binding sites, could be used for the affinity purification of biotinylated enzymes. Other new possibilities in avidin-biotin technology are novel bioseparation tools (Bayer, Wilchek 1990), carrier proteins (Holmberg et al. 2005), or nanoscale adapters as demonstrated by Zhen et al. (2004).

Circular permutation has been used to modify avidin and streptavidin (Chu et al. 1998, Nordlund et al. 2004, Hytönen et al. 2005b, Nordlund et al. 2005). For example in avidin, circular permutation has been used to cleave the loops between β -strands L3,4; L4,5; and L5,6. This method allows the design of versatile novel avidins. In the future, circular permutation will be used to study how other loop regions affect the binding properties of avidin and whether they could be made to bind different kinds of small molecules.

The known structures of avidin and AVR4 did give us hints on what the ChiAVD(I117Y) might look like. The X-ray structure of ChiAVD(I117Y) showed that beta strand 4 and the surrounding loops were more compactly folded in ChiAVD(I117Y) than in wt avidin. A more compact structure means that the loops of the protein are less exposed to the surroundings, which is one possible reason for the increased thermostability (Thompson, Eisenberg 1999). Other reasons, which commonly appear to have effect on the thermostability of proteins, are interactions in the interface area and between aromatic pairs (Serrano, Bycroft & Fersht 1991, Clackson, Wells 1995, Hu et al. 2000, DeLano 2002). Indeed, the successful mutation of I117Y on the subunit interface had a positive outcome on the thermostability of the chimeric avidin protein, and from the X-ray structure it could be seen that two tyrosines 117s from neighbouring subunits form a stabilizing $\pi - \pi$ interaction as presented in figure 22. In spite of this, the reason for the higher

thermostability of ChiAVD(I117Y) as compared to AVR4 remains somewhat unclear.

In the future we would like to evaluate the potential of ChiAVD(I117Y) in life science applications. The successful pilot-scale fermentation proved that ChiAVD(I117Y) is also suitable for industry-scale production and ongoing studies of ChiAVD(I117Y) imply that it could be a industrially valuable new tool in avidin-biotin technology. By now we have already produced several modified forms of ChiAVD(I117Y), for example one which enables us to link it to surfaces via the cysteine in its C-terminus (Määttä et al. unpublished results) and another which has a cysteine in the bottom of the binding pocket in order to form a covalent bond with the ligand (Leppiniemi et al. unpublished results). Moreover, a cysteine bridged form, ChiAVD(I117C), has also been produced (Hytönen and Nordlund, unpublished). In the future phage display will probably offer new ligand-binding motifs (Paldanius et al. unpublished results) that could be combined with the strong structure of ChiAVD(I117Y) and applied to diagnostic methods. Another series of ongoing experiments are related to BBP-A, which has been used as source material for the development of DNA libraries of recombined biotin-binding proteins (Niederhauser et al. unpublished).

7. Summary and conclusion

This thesis describes the structure function relationships of two novel avidins and one novel avidin-like protein. First we designed chimeric avidin, which is a combination of the avidin related protein 4 (AVR4) and wild type avidin. The part taken from AVR4 covers β -strand 4 and parts of its surrounding loops and one mutated amino acid residue, isoleucine 117. The X-ray structure of chimeric avidin resembles that of wt avidin and AVR4 – all three are homotetrameric proteins, which have two cysteine residues in each subunit that form a disulphide bridge inside the subunits. The ligand binding properties of chimeric avidin also resemble those of the parent proteins. However, it came as a surprise to us that chimeric avidin is clearly more thermalstable than either parent. We found that the thermal stability correlates with stability against many organic solvents, and therefore, chimeric avidin is the most durable avidin protein developed so far. Thus, due to its high biotin affinity and good stability in harsh conditions we are now studying whether chimeric avidin could be used in novel applications.

The second avidin protein was circularly permuted avidin, cpAVD4 \rightarrow 3. Its original terminuses were connected with a small polypeptide linker and new terminuses were placed after and before the loop region between β -strands 3 and 4. Because cpAVD4 \rightarrow 3 lacks the L3,4-loop, its biotin binding affinity was reduced significantly and HABA binding affinity became better when compared to wt avidin. Through the additional point mutation Asn118Met we were successfully able to further reduce biotin binding affinity with one magnitude and increase HABA binding affinity five-fold. The X-ray structure and binding properties of cpAVD4 \rightarrow 3(N118M) provided us with useful information about the structure-function relationship of the L3,4 loop. Furthermore, we combined the cpAVD4 \rightarrow 3 and the wt avidin-like binding subunit cpAVD5 \rightarrow 4 in a dual-chain avidin called dcAVD2. Only one dual affinity avidin has been produced before, and therefore, dcAVD2 is a valuable addition to the dual chain avidin protein family, which has potential use in self-assembly nanotechnology applications.

Structures and binding properties of new avidin-like proteins have expanded our knowledge of the significance of individual amino acid residues for the biotin-binding affinity. The novel biotin-binding protein A (BBP-A) is an interesting addition to the growing family of avidin-like proteins. X-ray structures with D-biotin and D-biotin D-sulfoxide (BSO) revealed that BBP-A shares a structural fold with the avidin protein family. BBP-A binds free biotin with high affinity, but interestingly it binds D-biotin D-sulfoxide with even higher affinity. The high affinity towards BSO may have use in biotechnological applications.

In conclusion, this work presents the characterization of three novel proteins: ChiAVD(I117Y), cpAVD4→3(N118M) and BBP-A. They all are new members of the avidin protein family, and provide novel information on the expression, binding properties, biochemical properties and structure-function relationships of avidins.

Acknowledgements

This work began at the University of Jyväskylä, at the Department of Biological and Environmental Science, Division of Molecular Biology, but most of the work was carried out at the Institute of Medical Technology, University of Tampere, Division of Molecular Biotechnology, during the years 2005-2010. The work has been supported by the National Graduate School in Informational and Structural Biology (ISB) and The Academy of Finland.

First, I wish to express my sincere gratitude to my supervisor Professor Markku Kulomaa. I want to thank Markku for offering me the opportunity to come and work in his group in Tampere and above all, for his kind way of supervising me.

I owe a great expression of gratitude to my supervisor Dr. Vesa Hytönen, who pushed me forward in this study. I also want to thank him for his insatiable desire to discuss scientific and related issues with me. I can't forget my late supervisor Docent Henri Nordlund who taught me to be a more independent scientist, and how he planted in me the faith in the practical significance of our research. Henkka, rest in peace.

I would like to give special thank to my nearest and dearest colleague Tiina Paldanius. We have walk along the same path for so long, and we have sheared so many good laughs. It has been a pleasure to work with you; you have saved my day more than once. I also want to thank Tiia Koho for enjoyable conversations and good times during these years; you have become a friend to me. I'd like to thank Satu Helppolainen for collaboration and shearing interest toward calorimetry. I owe a dept of gratitude to Ulla Kiiskinen for her endless will to help. I thank Jarkko Valjakka for all the scientific conversations and opinions. I would then like to thank all our group members. You all have made my work here more pleasant: Olli Laitinen, Soili Hiltunen, Jenni Leppiniemi, Sampo Kukkurainen, Barbara Niederhauser, Suvi Varjonen, Selina Mäkinen, Toni Grönroos, Tuomas Mäntylä, and Joanna Zmurko. I am also grateful for all my "slaves", your efforts in my research have been priceless: Katariina Tolvanen, Ville Niskanen, Katja

Veneskoski, Deirdre O' Malley, Katja Häkkinen, Aoife Kearns, Joonas Siivonen, Mikko Soikkeli, Ilmari Tamminen and especially Jukka Lehtonen who has delighted me also in our free-time..

I thank my thesis committee, Professor Kari Rissanen and Docent Juhani Huuskonen. My interest towards organic chemistry is greatly to your credit.

I want to thank all my other collaborators and co-authors for their valuable contributions: Thomas Nyholm, Yael Eisenberg-Domovich, David Hyre, Jarno Hörhä, tuomas Kulomaa, Eevaleena Porkka, Patric S Stayton, Olli Laitinen, Kaisa Helttunen, Katrin Halling, Mark Johnson, Tiina Salminen, Janne Jänis, Pirjo Vainiotalo and Ruhama Hayouka. Also I'd like to specially thank co-authors Dr. Tomi Airene and Dr. Oded Livnah for their collaboration in the X-ray crystallography.

I'd like to thank my reviewers Professors Reijo Lahti and Antti Poso for improving this thesis with their valuable comments,

I haven't forgotten my background and colleagues at the University of Jyväskylä. Huge thank to Einari Niskanen and Johanna Laakkonen, for our unforgettable trips and times together. I would also like to thank Teemu Ihalainen for the exhilarating time when we were just young scientist.

The balance in my life has come from my friends. Thank you all for cheering me up during these years: Lauri, Annukka, Marjaana, Joni, Mari, Jarno, Emma, Maisa and Joni. Thanks to Esa for early morning discussions about life and soccer, it has been pleasure. I express my gratitude to my parents, your care has carried me this far. Kiitos kaikesta antamastanne tuesta! I would like to thank my sister and nephews, your band have brought pleasure to my life countless times. I would also want to thank my parents-in-law for hospitality and enjoyable times.

Finally, my deepest gratitude goes to my wife Anna-Leena. Your love, care and friendship mean everything to me. You make me whole. OX

References

- Abraham, M.H., Ibrahim, A., Zissimos, A.M., Zhao, Y.H., Comer, J. & Reynolds, D.P. 2002, "Application of hydrogen bonding calculations in property based drug design", *Drug discovery today*, vol. 7, no. 20, pp. 1056-1063.
- Ahlroth, M.K., Kola, E.H., Ewald, D., Masabanda, J., Sazanov, A., Fries, R. & Kulomaa, M.S. 2000, "Characterization and chromosomal localization of the chicken avidin gene family", *Anim Genet*, vol. 31, no. 6, pp. 367-75.
- Airenne, K.J., Oker-Blom, C., Marjomäki, 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. 9326, pp. 100-108.
- Airenne, K.J., Peltomaa, E., Hytönen, V.P., Laitinen, O.H. & Ylä-Herttuala, S. 2003, "Improved generation of recombinant baculovirus genomes in *Escherichia coli*", *Nucleic Acids Res*, vol. 31, no. 17, pp. e101.
- Airenne, K.J., Sarkkinen, P., Punnonen, E.L. & Kulomaa, M.S. 1994, "Production of recombinant avidin in *Escherichia coli*", *Gene*, vol. 144, no. 1, pp. 75-80.
- Akerström, B., Logdberg, L., Berggard, T., Osmark, P. & Lindqvist, A. 2000, "alpha(1)-Microglobulin: a yellow-brown lipocalin", *Biochimica et biophysica acta*, vol. 1482, no. 1-2, pp. 172-184.
- Allhorn, M., Klapysa, A. & Akerström, B. 2005, "Redox properties of the lipocalin alpha1-microglobulin: reduction of cytochrome c, hemoglobin, and free iron", *Free radical biology & medicine*, vol. 38, no. 5, pp. 557-567.
- 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.
- Atkins, P.W. 1998, *Physical Chemistry*, Sixth edition edn, Oxford University Press, New York.
- Avrantinis, S.K., Stafford, R.L., Tian, X. & Weiss, G.A. 2002, "Dissecting the streptavidin-biotin interaction by phage-displayed shotgun scanning", *Chembiochem*, vol. 3, no. 12, pp. 1229-134.
- Bailly, C., Chessari, G., Carrasco, C., Joubert, A., Mann, J., Wilson, W.D. & Neidle, S. 2003, "Sequence-specific minor groove binding by bis-benzimidazoles: water molecules in ligand recognition", *Nucleic acids research*, vol. 31, no. 5, pp. 1514-1524.
- Batchelor, R.H., Sarkez, A., Cox, W.G. & Johnson, I. 2007, "Fluorometric assay for quantitation of biotin covalently attached to proteins and nucleic acids", *BioTechniques*, vol. 43, no. 4, pp. 503-507.
- Bayer, E.A., 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*, vol. 17370, pp. 1319-1324.
- Bayer, E.A. & Wilchek, M. 1990, "Application of avidin-biotin technology to affinity-based separations", *J Chromatogr*, vol. 51076, pp. 3-11.
- Beaumont, W. 1833, *Experiments and observations on the gastric juice, and the physiology of digestion*. 1 st edn, F. P. Allen, Plattsburgh.
- Ben-Amotz, D. & Widom, B. 2007, "Nonideal gas solvation thermodynamics", *The Journal of chemical physics*, vol. 126, no. 10, pp. 104502.

- Bennett, W.S., Jr & Steitz, T.A. 1980, "Structure of a complex between yeast hexokinase A and glucose. II. Detailed comparisons of conformation and active site configuration with the native hexokinase B monomer and dimer", *Journal of Molecular Biology*, vol. 140, no. 2, pp. 211-230.
- Bennett, W.S., Jr & Steitz, T.A. 1978, "Glucose-induced conformational change in yeast hexokinase", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 75, no. 10, pp. 4848-4852.
- Berggard, T., Cohen, A., Persson, P., Lindqvist, A., Cedervall, T., Silow, M., Thogersen, I.B., Jonsson, J.A., Enghild, J.J. & Akerström, B. 1999, "Alpha1-microglobulin chromophores are located to three lysine residues semiburied in the lipocalin pocket and associated with a novel lipophilic compound", *Protein science : a publication of the Protein Society*, vol. 8, no. 12, pp. 2611-2620.
- 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 Crystallogr D Biol Crystallogr*, vol. 58, no. Pt 6 No 122032454, pp. 899-907.
- Bessette, P.H., Rice, J.J. & Daugherty, P.S. 2004, "Rapid isolation of high affinity protein binding peptides using bacterial display", *Protein Eng Des Sel*, vol. 17, no. 10, pp. 731-739.
- Billeter, M., Guntert, P., Luginbuhl, P. & Wuthrich, K. 1996, "Hydration and DNA recognition by homeodomains", *Cell*, vol. 85, no. 7, pp. 1057-1065.
- Biltonen, R.L. & Langerman, N. 1979, "Microcalorimetry for biological chemistry: experimental design, data analysis, and interpretation", *Methods in enzymology*, vol. 61, pp. 287-318.
- Blom, E., Langstrom, B. & Velikyan, I. 2009, "68Ga-labeling of biotin analogues and their characterization", *Bioconjugate chemistry*, vol. 20, no. 6, pp. 1146-1151.
- Brandts, J.F. & Lin, L.N. 1990, "Study of strong to ultratight protein interactions using differential scanning calorimetry", *Biochemistry*, vol. 29, no. 29, pp. 6927-6940.
- Bryant, R.G. 1996, "The dynamics of water-protein interactions", *Annual Review of Biophysics and Biomolecular Structure*, vol. 25, pp. 29-53.
- 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.
- Cheng, Y.K. & Rosicky, P.J. 1998, "Surface topography dependence of biomolecular hydrophobic hydration", *Nature*, vol. 392, no. 6677, pp. 696-699.
- Chilkoti, A., Schwartz, B.L., Smith, R.D., Long, C.J. & Stayton, P.S. 1995, "Engineered chimeric streptavidin tetramers as novel tools for bioseparations and drug delivery", *Bio/Technology*, vol. 13, no. 12, pp. 1198-1204.
- 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", *Proc Natl Acad Sci U S A*, vol. 92, no. 5, pp. 1754-178.
- Cho, G., Keefe, A.D., Liu, R., Wilson, D.S. & Szostak, J.W. 2000, "Constructing high complexity synthetic libraries of long ORFs using in vitro selection", *J Mol Biol*, vol. 297, no. 2, pp. 309-319.
- Choe, S.E., Li, L., Matsudaira, P.T., Wagner, G. & Shakhnovich, E.I. 2000, "Differential stabilization of two hydrophobic cores in the transition state of the villin 14T folding reaction", *Journal of Molecular Biology*, vol. 304, no. 1, pp. 99-115.
- Chu, V., Freitag, S., Le Trong, I., Stenkamp, R.E. & Stayton, P.S. 1998, "Thermodynamic and structural consequences of flexible loop deletion by circular permutation in the streptavidin-biotin system", *Protein Sci*, vol. 7, no. 4, pp. 848-859.
- Clackson, T. & Wells, J.A. 1995, "A hot spot of binding energy in a hormone-receptor interface", *Science*, vol. 267, no. 5196, pp. 383-386.

- Collaborative Computational Project, Number 4 1994, "The CCP4 suite: programs for protein crystallography", *Acta crystallographica. Section D, Biological crystallography*, vol. 50, no. Pt 5, pp. 760-763.
- Colombo, M.F., Rau, D.C. & Parsegian, V.A. 1992, "Protein solvation in allosteric regulation: a water effect on hemoglobin", *Science (New York, N.Y.)*, vol. 256, no. 5057, pp. 655-659.
- Copeland, R.A. 2000, *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, 2 nd edn, A John Wiley & Sons, INC., New York, NY.
- Davies, T.G., Hubbard, R.E. & Tame, J.R. 1999, "Relating structure to thermodynamics: the crystal structures and binding affinity of eight OppA-peptide complexes", *Protein science : a publication of the Protein Society*, vol. 8, no. 7, pp. 1432-1444.
- DeChancie, J. & Houk, K.N. 2007, "The origins of femtomolar protein-ligand binding: hydrogen-bond cooperativity and desolvation energetics in the biotin-(strept)avidin binding site", *Journal of the American Chemical Society*, vol. 129, no. 17, pp. 5419-5429.
- DeLano, W.L. 2008, *The PyMOL Molecular Graphics System.*, Delani Scientific LLC, Palo Alto, CA, USA.
- DeLano, W.L. 2002, "Unraveling hot spots in binding interfaces: progress and challenges", *Curr Opin Struct Biol*, vol. 12, no. 1, pp. 14-20.
- Derewenda, Z.S., Lee, L. & Derewenda, U. 1995, "The occurrence of C-H...O hydrogen bonds in proteins", *Journal of Molecular Biology*, vol. 252, no. 2, pp. 248-262.
- Desiraju, G.R. 2002, "Hydrogen bridges in crystal engineering: interactions without borders", *Accounts of Chemical Research*, vol. 35, no. 7, pp. 565-573.
- Devlin, J.J., Panganiban, L.C. & Devlin, P.E. 1990, "Random peptide libraries: a source of specific protein binding molecules", *Science (New York, N.Y.)*, vol. 249, no. 4967, pp. 404-406.
- Dittmer, K. & DU Vigneaud, V. 1944, "Antibiotins", *Science (New York, N.Y.)*, vol. 100, no. 2589, pp. 129-131.
- Dixon, R.W. & Kollman, P. 1999, "The free energies for mutating S27 and W79 to alanine in streptavidin and its biotin complex: the relative size of polar and nonpolar free energies on biotin binding", *Proteins*, vol. 36, no. 4, pp. 471-473.
- Donovan, J.W. & Ross, K.D. 1973, "Increase in the stability of avidin produced by binding of biotin. A differential scanning calorimetric study of denaturation by heat", *Biochemistry*, vol. 12, no. 3, pp. 512-517.
- D'Souza, C., Kanyalkar, M., Joshi, M., Coutinho, E. & Srivastava, S. 2009a, "Probing molecular level interaction of oseltamivir with H5N1-NA and model membranes by molecular docking, multinuclear NMR and DSC methods", *Biochimica et biophysica acta*, vol. 1788, no. 2, pp. 484-494.
- D'Souza, C., Kanyalkar, M., Joshi, M., Coutinho, E. & Srivastava, S. 2009b, "Search for novel neuraminidase inhibitors: Design, synthesis and interaction of oseltamivir derivatives with model membrane using docking, NMR and DSC methods", *Biochimica et biophysica acta*, vol. 1788, no. 9, pp. 1740-1751.
- Eakin, R.E., Snell, E.E. & Williams, R.J. 1941, "The concentration and assay of avidin, the injury-producing protein in raw egg-white", *J.Biol.Chem.*, vol. 140, pp. 535-543.
- Ellison, D., Hinton, J., Hubbard, S.J. & Beynon, R.J. 1995, "Limited proteolysis of native proteins: the interaction between avidin and proteinase K", *Protein Sci*, vol. 4, no. 7, pp. 1337-1345.
- Escribano, J., Grubb, A., Calero, M. & Mendez, E. 1991, "The protein HC chromophore is linked to the cysteine residue at position 34 of the polypeptide chain by a reduction-resistant bond and causes the charge heterogeneity of protein HC", *The Journal of biological chemistry*, vol. 266, no. 24, pp. 15758-15763.
- Fermi, G., Perutz, M.F., Shaanan, B. & Fourme, R. 1984, "The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution", *Journal of Molecular Biology*, vol. 175, no. 2, pp. 159-174.

- Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. 1988, "The MIDAS Display System", *J.Mol.Graphics*, vol. 6, no. 1, pp. 36-37.
- Fessenden, R.J., Fessenden, J.S. & Logue, M.W. 1998, *Organic chemistry*, 6th ed. edn, Brooks/Cole, Pacific Grove (CA).
- Fraenkel-Conrat, H., Snell, N.S. & Ducay, E.D. 1952, "Avidin. II. Composition and mode of action of avidin A", *Archives of Biochemistry and Biophysics*, vol. 39, no. 1, pp. 97-107.
- Franks, F. & Eagland, D. 1975, "The role of solvent interactions in protein conformation", *CRC critical reviews in biochemistry*, vol. 3, no. 2, pp. 165-219.
- Freire, E. 2004, "Isothermal titration calorimetry: controlling binding forces in lead optimization", *Drug Discovery Today: Technologies*, vol. 1, no. 3, pp. 295-299.
- Freitag, S., Le Trong, I., Klumb, L.A., Chu, V., Chilkoti, A., Stayton, P.S. & Stenkamp, R.E. 1999, "X-ray crystallographic studies of streptavidin mutants binding to biotin", *Biomol Eng*, vol. 16, no. 1-4, pp. 252453, pp. 13-9.
- Galakatos, N.G. & Walsh, C.T. 1989, "Mutations at the interdomain hinge region of the DadB alanine racemase: effects of length and conformational constraint of the linker sequence on catalytic efficiency", *Biochemistry*, vol. 28, no. 20, pp. 8167-8174.
- 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*", *J Bacteriol*, vol. 174, no. 23, pp. 7729-7742.
- Gonzalez, M., Argarana, C.E. & Fidelio, G.D. 1999, "Extremely high thermal stability of streptavidin and avidin upon biotin binding", *Biomol Eng*, vol. 16, no. 1-4, pp. 67-72.
- Gonzalez, M., Bagatolli, L.A., Echabe, I., Arrondo, J.L., Argarana, C.E., Cantor, C.R. & Fidelio, G.D. 1997, "Interaction of biotin with streptavidin. Thermostability and conformational changes upon binding", *The Journal of biological chemistry*, vol. 272, no. 17, pp. 11288-11294.
- Gope, M.L., Keinanen, R.A., Kristo, P.A., Conneely, O.M., Beattie, W.G., Zarucki-Schulz, T., O'Malley, B.W. & Kulomaa, M.S. 1987, "Molecular cloning of the chicken avidin cDNA", *Nucleic Acids Res*, vol. 15, no. 8, pp. 3595-3606.
- Green, N.M. 1990, "Avidin and streptavidin", *Methods Enzymol*, vol. 184, pp. 51-67.
- Green, N.M. 1975, "Avidin", *Adv Prot Chem*, vol. 29, pp. 85-133.
- Green, N.M. 1970, "Spectrophotometric determination of avidin and streptavidin", *Methods Enzymol.*, vol. 18, pp. 418-424.
- Green, N.M. 1966, "Thermodynamics of the binding of biotin and some analogues by avidin", *The Biochemical journal*, vol. 101, no. 3, pp. 774-780.
- Green, N.M. 1965, "A Spectrophotometric Assay for Avidin and Biotin Based on Binding of Dyes by Avidin", *Biochem J*, vol. 94, pp. 23C-24C.
- Green, N.M. 1963a, "Avidin. 3. the Nature of the Biotin-Binding Site", *The Biochemical journal*, vol. 89, pp. 599-609.
- Green, N.M. 1963b, "Avidin. 4. Stability at Extremes of Ph and Dissociation into Sub-Units by Guanidine Hydrochloride", *The Biochemical journal*, vol. 89, pp. 609-620.
- Ha, J.H., Spolar, R.S. & Record, M.T., Jr 1989, "Role of the hydrophobic effect in stability of site-specific protein-DNA complexes", *Journal of Molecular Biology*, vol. 209, no. 4, pp. 801-816.
- Halle, B. 2004, "Protein hydration dynamics in solution: a critical survey", *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 359, no. 1448, pp. 1207-23; discussion 1223-4, 1323-8.
- 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., Määttä, J.A., Halling, K.K., Slotte, J.P., Huhtala, T., Liimatainen, T., Ylä-Herttuala, S., Airene, K.J., Narvanen, A., Jänis, 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.
- Hirsch, J.D., Eslamizar, L., Filanoski, B.J., Malekzadeh, N., Haugland, R.P. & Beechem, J.M. 2002, "Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation", *Anal Biochem*, vol. 308, no. 2, pp. 343-357.
- Hofstetter, H., Morpurgo, M., Hofstetter, O., Bayer, E.A. & Wilchek, M. 2000, "A labeling, detection, and purification system based on 4-hydroxyazobenzene-2-carboxylic acid: an extension of the avidin-biotin system", *Anal Biochem*, vol. 284, no. 2, pp. 354-366.
- Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K.M., Stark, W., Jansonius, J.N., McKay, D.B. & Matthews, B.W. 1992, "Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis", *Biochemistry*, vol. 31, no. 46, pp. 11310-11316.
- Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J. & Uhlen, M. 2005, "The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures", *Electrophoresis*, vol. 26, no. 3, pp. 501-510.
- 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.
- Howarth, M. & Ting, A.Y. 2008, "Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin", *Nature protocols*, vol. 3, no. 3, pp. 534-545.
- Hu, Z., Ma, B., Wolfson, H. & Nussinov, R. 2000, "Conservation of polar residues as hot spots at protein interfaces", *Proteins*, vol. 39, no. 4, pp. 331-342.
- Hu, Z., Tan, W., Zhang, L., Liang, Z., Xu, C., Su, H., Lu, J. & Gao, J. 2009, "A Novel Immunotherapy for Superficial Bladder Cancer by Intravesical Immobilization of GM-CSF", *Journal of Cellular and Molecular Medicine*.
- Hyre, D.E., Amon, L.M., Penzotti, J.E., Le Trong, I., Stenkamp, R.E., Lybrand, T.P. & Stayton, P.S. 2002, "Early mechanistic events in biotin dissociation from streptavidin", *Nat Struct Biol*, vol. 9, no. 822140693, pp. 582-55.
- 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 Sci*, vol. 9, no. 520306970, pp. 878-85.
- 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*, vol. 15, no. 3, pp. 459-467.
- Hytönen, V.P., Laitinen, O.H., Airene, T.T., Kidron, H., Meltola, N.J., Porkka, E., Hörhä, J., Paldanius, T., Määttä, J.A., Nordlund, H.R., Johnson, M.S., Salminen, T.A., Airene, K.J., Ylä-Herttuala, S. & Kulomaa, M.S. 2004, "Efficient production of active chicken avidin using a bacterial signal peptide in *Escherichia coli*", *Biochem J*, 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. & Airene, T.T. 2005a, "Avidin related protein 2 shows unique structural and functional features among the avidin protein family", *BMC biotechnology*, vol. 5, pp. 28.
- 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., Pentikainen, 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", *J Biol Chem*, vol. 279, no. 10, pp. 9337-9343.
- Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. & Bax, A. 1992, "Solution structure of a calmodulin-target peptide complex by multidimensional NMR", *Science*, vol. 256, no. 5057, pp. 632-638.

- Janin, J. 1997, "Angstroms and calories", *Structure*, vol. 5, no. 4, pp. 473-479.
- Jiang, L. & Lai, L. 2002, "CH...O hydrogen bonds at protein-protein interfaces", *The Journal of biological chemistry*, vol. 277, no. 40, pp. 37732-37740.
- Jonsson, U., Fagerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Lofas, S., Persson, B., Roos, H. & Ronnberg, I. 1991, "Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology", *BioTechniques*, vol. 11, no. 5, pp. 620-627.
- Jonsson, U., Fagerstam, L., Lofas, S., Stenberg, E., Karlsson, R., Frostell, A., Markey, F. & Schindler, F. 1993, "Introducing a biosensor based technology for real-time biospecific interaction analysis", *Annales de Biologie Clinique*, vol. 51, no. 1, pp. 19-26.
- Kanaan, S.P., Huang, C., Wuchty, S., Chen, D.Z. & Izaguirre, J.A. 2009, "Inferring protein-protein interactions from multiple protein domain combinations", *Methods in molecular biology*, vol. 541, pp. 43-59.
- Kauzmann, W. 1959, "Some factors in the interpretation of protein denaturation", *Advances in Protein Chemistry*, vol. 14, pp. 1-63.
- Keefe, A.D., Wilson, D.S., Seelig, B. & Szostak, J.W. 2001, "One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-Tag", *Protein Expr Purif*, vol. 23, no. 3, pp. 440-446.
- Keinänen, R.A., Laukkanen, M.L. & Kulomaa, M.S. 1988, "Molecular cloning of three structurally related genes for chicken avidin", *J Steroid Biochem*, vol. 30, no. 1-688260103, pp. 17-21.
- Keinänen, 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", *Eur J Biochem*, vol. 220, no. 294170814, pp. 615-621.
- Kempner, E.S. 1993, "Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity", *FEBS letters*, vol. 326, no. 1-3, pp. 4-10.
- Klibanov, A.M. & Koskinen, A.M.P. 1996, *Enzymatic reactions in organic media*, Blackie Academic & Professional, London.
- 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.
- Koch, M.H., Vachette, P. & Svergun, D.I. 2003, "Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution", *Quarterly reviews of biophysics*, vol. 36, no. 2, pp. 147-227.
- Korndorfer, I.P. & Skerra, A. 2002, "Improved affinity of engineered streptavidin for the Strep-tag II peptide is due to a fixed open conformation of the lid-like loop at the binding site", *Protein Sci*, vol. 11, no. 4, pp. 883-893.
- Kosztin, D., Bishop, T.C. & Schulten, K. 1997, "Binding of the estrogen receptor to DNA. The role of waters", *Biophysical journal*, vol. 73, no. 2, pp. 557-570.
- Kukkurainen, S. 2009, *A molecular dynamics study of ligand bindings in mutant avidins*.
- Ladbury, J.E. & Chowdhry, B.Z. (eds) 1998, *Biocalorimetry: Applications of calorimetry in the biological science*, Wiley, Chichester, England.
- Ladbury, J.E. & Doyle, M.L. (eds) 2004, *Biocalorimetry 2: Applications of calorimetry in the biological sciences*, Wiley, Chichester, England.
- Laitinen, O.H., Airene, K.J., Hytönen, V.P., Peltomaa, E., Mähönen, A.J., Wirth, T., Lind, M.M., Mäkelä, K.A., Toivanen, P.I., Schenkwein, D., Heikura, T., Nordlund, H.R., Kulomaa, M.S. & Ylä-Herttuala, S. 2005, "A multipurpose vector system for the screening of libraries in bacteria, insect and mammalian cells and expression in vivo", *Nucleic acids research*, vol. 33, no. 4, pp. e42 1-10.
- Laitinen, O.H., Airene, 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 Lett*, vol. 461, no. 1-2, pp. 52-58.
- Laitinen, O.H., Hytönen, V.P., Ahlroth, M.K., Pentikäinen, O.T., Gallagher, C., Nordlund, H.R., Ovod, V., Marttila, A.T., Porkka, E., Heino, S., Johnson, M.S., Airene, K.J. &

- Kulomaa, M.S. 2002, "Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin", *Biochem J*, vol. 363, no. Pt 3, pp. 609-617.
- 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., Marttila, A.T., Airene, K.J., Kulik, T., Livnah, O., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 2001, "Biotin induces tetramerization of a recombinant monomeric avidin. A model for protein-protein interactions", *J Biol Chem*, vol. 276, no. 11, pp. 8219-8224.
- 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., Airene, K.J., Livnah, O., Bayer, E.A., Wilchek, M. & Kulomaa, M.S. 2003, "Rational design of an active avidin monomer", *J Biol Chem*, vol. 278, no. 622450663, pp. 4010-4014.
- Lamla, T. & Erdmann, V.A. 2004, "The Nano-tag, a streptavidin-binding peptide for the purification and detection of recombinant proteins", *Protein Expr Purif*, vol. 33, no. 1, pp. 39-47.
- Langerman, N. & Biltonen, R.L. 1979, "Microcalorimeters for biological chemistry: applications, instrumentation and experimental design", *Methods in enzymology*, vol. 61, pp. 261-286.
- Langmuir, I. 1919, "The Structure of Atoms and the Octet Theory of Valence", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 5, no. 7, pp. 252-259.
- Larsson, J., Allhorn, M. & Kerström, B. 2004, "The lipocalin alpha(1)-microglobulin binds heme in different species", *Archives of Biochemistry and Biophysics*, vol. 432, no. 2, pp. 196-204.
- Le Trong, I., Freitag, S., Klumb, L.A., Chu, V., Stayton, P.S. & Stenkamp, R.E. 2003, "Structural studies of hydrogen bonds in the high-affinity streptavidin-biotin complex: mutations of amino acids interacting with the ureido oxygen of biotin", *Acta Crystallogr D Biol Crystallogr*, vol. 59, no. Pt 9, pp. 1567-1573.
- Lee, B. & Richards, F.M. 1971, "The interpretation of protein structures: estimation of static accessibility", *Journal of Molecular Biology*, vol. 55, no. 3, pp. 379-400.
- Lehtonen, J.V., Still, D.J., Rantanen, V.V., Ekholm, J., Bjorklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikainen, O., Nyronen, T., Salminen, T., Gyllenberg, M. & Johnson, M.S. 2004, "BODIL: a molecular modeling environment for structure-function analysis and drug design", *Journal of computer-aided molecular design*, vol. 18, no. 6, pp. 401-419.
- Lesch, H.P., Pikkarainen, J.T., Kaikkonen, M.U., Taavitsainen, M., Samaranayake, H., Lehtolainen-Dalkilic, P., Vuorio, T., Määttä, A.M., Wirth, T., Airene, K.J. & Ylä-Herttuala, S. 2009, "Avidin fusion protein-expressing lentiviral vector for targeted drug delivery", *Human Gene Therapy*, vol. 20, no. 8, pp. 871-882.
- Levitt, M. & Park, B.H. 1993, "Water: now you see it, now you don't", *Structure (London, England : 1993)*, vol. 1, no. 4, pp. 223-226.
- Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. 2001, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings", *Advanced Drug Delivery Reviews*, vol. 46, no. 1-3, pp. 3-26.
- Lisy, V., Brutovsky, B., Zatonovsky, A.V. & Zvelindovsky, A.V. 2001, "On the theory of light and neutron scattering from droplet microemulsions", *Journal of Molecular Liquids*, vol. 93, no. 1-3, pp. 113-118.
- Livingstone, J.R., Spolar, R.S. & Record, M.T., Jr 1991, "Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area", *Biochemistry*, vol. 30, no. 17, pp. 4237-4244.

- Livnah, O., Bayer, A., Wilchek, M. & Sussman, J.L. 1993a, "The structure of the complex between avidin and the dye, 2-(4'-hydroxyazobenzene) benzoic acid (HABA)", *FEBS*, vol. 328165, pp. 165-168.
- Livnah, O., Bayer, E.A., Wilchek, M. & Sussman, J.L. 1993b, "Three-dimensional structures of avidin and the avidin-biotin complex", *Proc Natl Acad Sci U S A*, vol. 90297, pp. 5076-5080.
- Lodish, H., Berk, A., Zipursky, L.S., Matsudaira, P., Baltimore, D. & Darnell, J. 2000, *Molecular Cell Biology*, 4th edn, W. H. Freeman and Company, New York, England.
- Luecke, H., Schobert, B., Richter, H.T., Cartailler, J.P. & Lanyi, J.K. 1999, "Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution", *Science*, vol. 286, no. 5438, pp. 255-261.
- Määttä, J.A., Helppolainen, S.H., Hytönen, V.P., Johnson, M.S., Kulomaa, M.S., Airene, T.T. & Nordlund, H.R. 2009, "Structural and functional characteristics of xenavidin, the first frog avidin from *Xenopus tropicalis*", *BMC structural biology*, vol. 9, pp. 63.
- Makhatadze, G.I. & Privalov, P.L. 1993, "Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration", *Journal of Molecular Biology*, vol. 232, no. 2, pp. 639-659.
- Mancera, R.L. 2007, "Molecular modeling of hydration in drug design", *Current opinion in drug discovery & development*, vol. 10, no. 3, pp. 275-280.
- Marabotti, A., Scire, A., Staiano, M., Crescenzo, R., Aurilia, V., Tanfani, F. & D'Auria, S. 2008, "Wild-type and mutant bovine odorant-binding proteins to probe the role of the quaternary structure organization in the protein thermal stability", *Journal of proteome research*, vol. 7, no. 12, pp. 5221-5229.
- Marttila, A., Airene, K., Laitinen, O., Kulik, T., Bayer, E., Wilchek, M. & Kulomaa, M. 1998, "Engineering of chicken avidin: a progressive series of reduced charge mutants", *FEBS Lett*, vol. 441, pp. 313-317.
- Marttila, A.T., Hytönen, 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", *Biochem J*, vol. 369, no. Pt 2, pp. 249-254.
- Marx, K.A. 2003, "Quartz Crystal Microbalance: A Useful Tool for Studying Thin Polymer Films and Complex Biomolecular Systems at the Solution-Surface Interface", *Biomacromolecules*, vol. 4, no. 5, pp. 1099-1120.
- McNaught, A.D. & Wilkinson, A. 1997, *Compendium of Chemical Terminology, The Gold Book*, Second Edition edn, Blackwell Science, Oxford.
- Meir, A., Helppolainen, S.H., Podoly, E., Nordlund, H.R., Hytönen, V.P., Määttä, 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.
- Montfort, W.R., Weichsel, A. & Andersen, J.F. 2000, "Nitrophorins and related antihemostatic lipocalins from *Rhodnius prolixus* and other blood-sucking arthropods", *Biochimica et biophysica acta*, vol. 1482, no. 1-2, pp. 110-118.
- Nakasako, M. 2004, "Water-protein interactions from high-resolution protein crystallography", *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 359, no. 1448, pp. 1191-204; discussion 1204-6.
- Nicholls, P. 2000, "Introduction: the biology of the water molecule", *Cellular and molecular life sciences : CMLS*, vol. 57, no. 7, pp. 987-992.
- 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. 1, pp. 41.
- Nordlund, H.R., Hytönen, V.P., Hörhå, J., Määttä, J.A., White, D.J., Halling, K., Porkka, E.J., Slotte, J.P., Laitinen, O.H. & Kulomaa, M.S. 2005, "Tetravalent single-chain

- avidin: from subunits to protein domains via circularly permuted avidins", *The Biochemical journal; The Biochemical journal*, vol. 392, no. Pt 3, pp. 485-491.
- 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 Lett*, vol. 555, no. 3, pp. 449-454.
- Nordlund, H.R., Laitinen, O.H., Hytönen, V.P., Uotila, S.T., Porkka, E. & Kulomaa, M.S. 2004, "Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins", *J Biol Chem*, vol. 279, no. 35, pp. 36715-36719.
- Nordlund, H.R., Laitinen, O.H., Uotila, S.T., Nyholm, T., Hytönen, V.P., Slotte, J.P. & Kulomaa, M.S. 2003, "Enhancing the thermal stability of avidin. Introduction of disulfide bridges between subunit interfaces", *J Biol Chem*, vol. 278, no. 422421427, pp. 2479-283.
- Olah, G.A., Mitchell, R.D., Sosnick, T.R., Walsh, D.A. & Trewella, J. 1993, "Solution structure of the cAMP-dependent protein kinase catalytic subunit and its contraction upon binding the protein kinase inhibitor peptide", *Biochemistry*, vol. 32, no. 14, pp. 3649-3657.
- Oobatake, M. & Ooi, T. 1993, "Hydration and heat stability effects on protein unfolding", *Progress in biophysics and molecular biology*, vol. 59, no. 3, pp. 237-284.
- Ooi, T. 1994, "Thermodynamics of protein folding: effects of hydration and electrostatic interactions", *Advances in Biophysics*, vol. 30, pp. 105-154.
- Ooi, T. & Oobatake, M. 1988, "Intermolecular interactions between protein and other molecules including hydration effects", *Journal of Biochemistry*, vol. 104, no. 3, pp. 440-444.
- Ooi, T., Oobatake, M., Nemethy, G. & Scheraga, H.A. 1987, "Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 84, no. 10, pp. 3086-3090.
- Otting, G. & Wuethrich, K. 1989, "Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules", *Journal of the American Chemical Society*, vol. 111, no. 5, pp. 1871-1875.
- Panigrahi, S.K. & Desiraju, G.R. 2007, "Strong and weak hydrogen bonds in the protein-ligand interface", *Proteins*, vol. 67, no. 1, pp. 128-141.
- Pazy, Y., Kulik, T., Bayer, E.A., Wilchek, M. & Livnah, O. 2002, "Ligand exchange between proteins. Exchange of biotin and biotin derivatives between avidin and streptavidin", *J Biol Chem*, vol. 277, no. 34, pp. 30892-30900.
- Pazy, Y., Laitinen, O.H., Ravoy, B., Kulomaa, M.S., Wilchek, M., Bayer, E.A. & Livnah, O. 2001, "Crystallization and preliminary X-ray analysis of W120K mutant of streptavidin", *Acta Crystallogr D Biol Crystallogr*, vol. 57, no. Pt 12, pp. 1885-1886.
- 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", *J Biol Chem*, vol. 278, no. 9, pp. 7131-7134.
- Pei, Y., Yu, H., Pei, Z., Theurer, M., Ammer, C., Andre, S., Gabius, H., Yan, M. & Ramstrom, O. 2007, "Photoderivatized Polymer Thin Films at Quartz Crystal Microbalance Surfaces: Sensors for Carbohydrate-Protein Interactions", *Analytical Chemistry*, vol. 79, no. 18, pp. 6897-6902.
- Pei, Z., Larsson, R., Aastrup, T., Anderson, H., Lehn, J. & Ramström, O. 2006, "Quartz crystal microbalance bioaffinity sensor for rapid identification of glycosyldisulfide lectin inhibitors from a dynamic combinatorial library", *Biosensors and Bioelectronics*, vol. 22, no. 1, pp. 42-48.
- Perutz, M.F. 1976, "Structure and mechanism of haemoglobin", *British medical bulletin*, vol. 32, no. 3, pp. 195-208.
- Petrone, P.M. & Garcia, A.E. 2004, "MHC-peptide binding is assisted by bound water molecules", *Journal of Molecular Biology*, vol. 338, no. 2, pp. 419-435.

- Pickover, C.A., McKay, D.B., Engelman, D.M. & Steitz, T.A. 1979, "Substrate binding closes the cleft between the domains of yeast phosphoglycerate kinase", *The Journal of biological chemistry*, vol. 254, no. 22, pp. 11323-11329.
- Pierce, A.C., Sandretto, K.L. & Bemis, G.W. 2002, "Kinase inhibitors and the case for CH...O hydrogen bonds in protein-ligand binding", *Proteins*, vol. 49, no. 4, pp. 567-576.
- Pierce, A.C., ter Haar, E., Binch, H.M., Kay, D.P., Patel, S.R. & Li, P. 2005, "CH...O and CH...N hydrogen bonds in ligand design: a novel quinazolin-4-ylthiazol-2-ylamine protein kinase inhibitor", *Journal of medicinal chemistry*, vol. 48, no. 4, pp. 1278-1281.
- Plotnikov, V., Rochalski, A., Brandts, M., Brandts, J.F., Williston, S., Frasca, V. & Lin, L.N. 2002, "An autosampling differential scanning calorimeter instrument for studying molecular interactions", *Assay and drug development technologies*, vol. 1, no. 1 Pt 1, pp. 83-90.
- Pompliano, D.L., Peyman, A. & Knowles, J.R. 1990, "Stabilization of a reaction intermediate as a catalytic device: definition of the functional role of the flexible loop in triosephosphate isomerase", *Biochemistry*, vol. 29, no. 13, pp. 3186-3194.
- Poornima, C.S. & Dean, P.M. 1995a, "Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions", *Journal of computer-aided molecular design*, vol. 9, no. 6, pp. 500-512.
- Poornima, C.S. & Dean, P.M. 1995b, "Hydration in drug design. 2. Influence of local site surface shape on water binding", *Journal of computer-aided molecular design*, vol. 9, no. 6, pp. 513-520.
- Poornima, C.S. & Dean, P.M. 1995c, "Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins", *Journal of computer-aided molecular design*, vol. 9, no. 6, pp. 521-531.
- 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. Pt 7, pp. 1131-1137.
- Privalov, P.L. & Makhatadze, G.I. 1993, "Contribution of hydration to protein folding thermodynamics. II. The entropy and Gibbs energy of hydration", *Journal of Molecular Biology*, vol. 232, no. 2, pp. 660-679.
- 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 Å resolution", *Journal of Molecular Biology*, vol. 231, no. 3, pp. 698-710.
- 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", *J Biol Chem*, vol. 276, no. 2, pp. 1029-1039.
- Rand, R.P. & Parsegian, V.A. 1989, "Hydration forces between phospholipid bilayers.", *Biochim. Biophys. Acta*, vol. 988, pp. 351-351-376.
- Repo, S., Paldanius, T.A., Hytönen, V.P., Nyholm, T.K., Halling, K.K., Huuskonen, J., Pentikäinen, O.T., Rissanen, K., Slotte, J.P., Airene, T.T., Salminen, T.A., Kulomaa, M.S. & Johnson, M.S. 2006, "Binding properties of HABA-type azo derivatives to avidin and avidin-related protein 4", *Chemistry & biology*, vol. 13, no. 10, pp. 1029-1039.
- Reznik, G.O., Vajda, S., Sano, T. & Cantor, C.R. 1998, "A streptavidin mutant with altered ligand-binding specificity", *Proc Natl Acad Sci U S A*, vol. 95, no. 23, pp. 13525-1330.
- Rosano, C., Arosio, P. & Bolognesi, M. 1999, "The X-ray three-dimensional structure of avidin", *Biomol Eng*, vol. 16, no. 1-4, pp. 5-12.
- Ross, S.E., Carson, S.D. & Fink, L.M. 1986, "Effects of detergents on avidin-biotin interaction", *BioTechniques*, vol. 4, no. 3, pp. 350-354.
- Rumfeldt, J.A., Galvagnion, C., Vassall, K.A. & Meiering, E.M. 2008, "Conformational stability and folding mechanisms of dimeric proteins", *Progress in biophysics and molecular biology*, vol. 98, no. 1, pp. 61-84.
- Sarupria, S., Ghosh, T., Garcia, A.E. & Garde, S. 2009, "Studying pressure denaturation of a protein by molecular dynamics simulations", *Proteins*, vol. 78, no. 7, pp. 1641-1651.

- Schechter, A.N. 1970, "Measurement of fast biochemical reactions", *Science (New York, N.Y.)*, vol. 170, no. 955, pp. 273-280.
- Schmidt, T.G., Koepke, J., Frank, R. & Skerra, A. 1996, "Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin", *J Mol Biol*, vol. 255, no. 5, pp. 753-766.
- Schmidt, T.G. & Skerra, A. 1993, "The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment", *Protein Eng*, vol. 6, no. 1, pp. 109-122.
- Serdakowski, A.L. & Dordick, J.S. 2008, "Enzyme activation for organic solvents made easy", *Trends in biotechnology*, vol. 26, no. 1, pp. 48-54.
- Serrano, L., Bycroft, M. & Fersht, A.R. 1991, "Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles", *J Mol Biol*, vol. 218, no. 2, pp. 465-475.
- Sharma, A. & Schulman, S.G. 1999, *Introduction to Fluorescence Spectroscopy*, 1 st edn, Wiley-Interscience, New York.
- Sharp, K.A., Nicholls, A., Fine, R.F. & Honig, B. 1991a, "Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects", *Science (New York, N.Y.)*, vol. 252, no. 5002, pp. 106-109.
- Sharp, K.A., Nicholls, A., Friedman, R. & Honig, B. 1991b, "Extracting hydrophobic free energies from experimental data: relationship to protein folding and theoretical models", *Biochemistry*, vol. 30, no. 40, pp. 9686-9697.
- Sigurskjold, B.W. 2000, "Exact analysis of competition ligand binding by displacement isothermal titration calorimetry", *Analytical Biochemistry*, vol. 277, no. 2, pp. 260-266.
- Skerra, A. & Schmidt, T.G. 2000, "Use of the Strep-Tag and streptavidin for detection and purification of recombinant proteins", *Methods Enzymol*, vol. 326, pp. 271-304.
- Skerra, A. & Schmidt, T.G. 1999, "Applications of a peptide ligand for streptavidin: the Strep-tag", *Biomol Eng*, vol. 16, no. 1-4, pp. 79-86.
- Skinner, A.L. & Laurence, J.S. 2008, "High-field solution NMR spectroscopy as a tool for assessing protein interactions with small molecule ligands", *Journal of pharmaceutical sciences*, vol. 97, no. 11, pp. 4670-4695.
- Smith, K.J., Reid, S.W., Harlos, K., McMichael, A.J., Stuart, D.I., Bell, J.I. & Jones, E.Y. 1996, "Bound water structure and polymorphic amino acids act together to allow the binding of different peptides to MHC class I HLA-B53", *Immunity*, vol. 4, no. 3, pp. 215-228.
- Spolar, R.S., Ha, J.H. & Record, M.T., Jr 1989, "Hydrophobic effect in protein folding and other noncovalent processes involving proteins", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 21, pp. 8382-8385.
- Spolar, R.S., Livingstone, J.R. & Record, M.T., Jr 1992, "Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water", *Biochemistry*, vol. 31, no. 16, pp. 3947-3955.
- Spolar, R.S. & Record, M.T., Jr 1994, "Coupling of local folding to site-specific binding of proteins to DNA", *Science (New York, N.Y.)*, vol. 263, no. 5148, pp. 777-784.
- Steinbach, P.J. & Brooks, B.R. 1996, "Hydrated myoglobin's anharmonic fluctuations are not primarily due to dihedral transitions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 1, pp. 55-59.
- Steinbach, P.J. & Brooks, B.R. 1993, "Protein hydration elucidated by molecular dynamics simulation", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 19, pp. 9135-9139.
- Steiner, T. 2002, "The hydrogen bond in the solid state", *Angewandte Chemie (International ed. in English)*, vol. 41, no. 1, pp. 49-76.
- Steinitz, F. 1898, "Ueber das Verhalten phosphorhaltiger Eiweisskörper im Stoffwechsel", *Pflügers Archiv European Journal of Physiology*, vol. 72, pp. 75-104.

- Streit, W.R. & Entcheva, P. 2003, "Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production", *Applied Microbiology and Biotechnology*, vol. 61, no. 1, pp. 21-31.
- 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. Pt 2, pp. 573-577.
- 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.
- Tanaka, T., Kato, H., Nishioka, T. & Oda, J. 1992, "Mutational and proteolytic studies on a flexible loop in glutathione synthetase from *Escherichia coli* B: the loop and arginine 233 are critical for the catalytic reaction", *Biochemistry*, vol. 31, no. 8, pp. 2259-2265.
- Tanford, C. 1978, "The hydrophobic effect and the organization of living matter", *Science*, vol. 200, no. 4345, pp. 1012-1018.
- Tausig, F. & Wolf, F.J. 1964, "Streptavidin--a substance with avidin-like properties produced by microorganisms", *Biochemical and biophysical research communications*, vol. 14, pp. 205-209.
- Thompson, M.J. & Eisenberg, D. 1999, "Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability", *J Mol Biol*, vol. 290, no. 2, pp. 595-604.
- Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W. & Etzler, M.E. (eds) 2008, *Essentials of Glycobiology*, 2nd edn, Cold Spring Harbor Laboratory Press, New York.
- Voss, S. & Skerra, A. 1997, "Mutagenesis of a flexible loop in streptavidin leads to higher affinity for the Strep-tag II peptide and improved performance in recombinant protein purification", *Protein Eng*, vol. 10, no. 8, pp. 975-982.
- Wahl, M.C. & Sundaralingam, M. 1997, "C-H...O hydrogen bonding in biology", *Trends in biochemical sciences*, vol. 22, no. 3, pp. 97-102.
- Weber, P.C., Ohlendorf, D.H., Wendoloski, J.J. & Salemme, F.R. 1989, "Structural origins of high-affinity biotin binding to streptavidin", *Science*, vol. 24322, pp. 85-88.
- Weber, P.C., Wendoloski, J.J., Pantoliano, M.W. & Salemme, F.R. 1992, "Crystallographic and Thermodynamic Comparison of Natural and Synthetic Ligands Bound to Streptavidin", *J.Am.Chem.Soc.*, vol. 114, pp. 3197-3200.
- 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.
- Wilchek, M. & Bayer, E. 1990, "Introduction to avidin-biotin technology", *Methods in enzymology*, vol. 18429, pp. 5-13.
- Wilchek, M. & Bayer, E.A. 1999, "Foreword and introduction to the book (strept)avidin-biotin system", *Biomol Eng*, vol. 16, no. 1-4, pp. 1-4.
- Wilchek, M. & Bayer, E.A. 1988, "The avidin-biotin complex in bioanalytical applications", *Anal Biochem*, vol. 171, no. 188307879, pp. 1-32.
- Wilchek, M., Bayer, E.A. & Livnah, O. 2006, "Essentials of biorecognition: the (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction", *Immunology letters*, vol. 103, no. 1, pp. 27-32.
- Wright, L.D., Cresson, E.L., Skeggs, H.R., Peck, R.L., Wolf, D.E., Wood, T.R., Valiant, J. & Folkers, K. 1951, "The elucidation of biocytin", *Science*, vol. 114, no. 2972, pp. 635-636.
- Wuthrich, K. 1993, "Hydration of biological macromolecules in solution: surface structure and molecular recognition", *Cold Spring Harbor symposia on quantitative biology*, vol. 58, pp. 149-157.

- Zaks, A. & Klivanov, A.M. 1988, "Enzymatic catalysis in nonaqueous solvents", *The Journal of biological chemistry*, vol. 263, no. 7, pp. 3194-3201.
- Zhang, C., Vasmatzis, G., Cornette, J.L. & DeLisi, C. 1997, "Determination of atomic desolvation energies from the structures of crystallized proteins", *Journal of Molecular Biology*, vol. 267, no. 3, pp. 707-726.
- Zhen, G., Egli, V., Voros, J., Zammaretti, P., Textor, M., Glockshuber, R. & Kuennemann, E. 2004, "Immobilization of the enzyme beta-lactamase on biotin-derivatized poly(L-lysine)-g-poly(ethylene glycol)-coated sensor chips: a study on oriented attachment and surface activity by enzyme kinetics and in situ optical sensing", *Langmuir: the ACS journal of surfaces and colloids*, vol. 20, no. 24, pp. 10464-10473.

Design and Construction of Highly Stable, Protease-resistant Chimeric Avidins*

Received for publication, December 17, 2004, and in revised form, January 10, 2005
Published, JBC Papers in Press, January 13, 2005, DOI 10.1074/jbc.M414196200

Vesa P. Hytönen‡, Juha A. E. Määttä‡§, Thomas K. M. Nyholm¶, Oded Livnah||, Yael Eisenberg-Domovich||, David Hyre**, Henri R. Nordlund‡§, Jarno Hörhå‡, Einari A. Niskanen‡, Tiina Paldanius‡§, Tuomas Kulomaa‡, Eevaleena J. Porkka‡, Patrick S. Stayton**, Olli H. Laitinen‡‡, and Markku S. Kulomaa‡§§

From the ‡Department of Biological and Environmental Science, P.O. Box 35 (YAB) FIN-40014, University of Jyväskylä, Finland, ¶Department of Biochemistry and Pharmacy, Åbo Akademi University, P.O. Box 66, FIN-20521 Turku, Finland, ||Department of Biological Chemistry, The Institute of Life Sciences, The Wolfson Centre for Applied Structural Biology, The Hebrew University of Jerusalem, Givat Ram, Jerusalem 91904, Israel, **Department of Bioengineering, University of Washington, Seattle, Washington 98195, and ‡‡A. I. Virtanen Institute, Department of Molecular Medicine, University of Kuopio, P.O. Box 1627, FIN-70211 Kuopio, Finland

The chicken avidin gene family consists of avidin and seven separate avidin-related genes (AVRs) 1–7. Avidin protein is a widely used biochemical tool, whereas the other family members have only recently been produced as recombinant proteins and characterized. In our previous study, AVR4 was found to be the most stable biotin binding protein thus far characterized ($T_m = 106.4^\circ\text{C}$). In this study, we studied further the biotin-binding properties of AVR4. A decrease in the energy barrier between the biotin-bound and unbound state of AVR4 was observed when compared with that of avidin. The high resolution structure of AVR4 facilitated comparison of the structural details of avidin and AVR4. In the present study, we used the information obtained from these comparative studies to transfer the stability and functional properties of AVR4 to avidin. A chimeric avidin protein, ChiAVD, containing a 21-amino acid segment of AVR4 was found to be significantly more stable ($T_m = 96.5^\circ\text{C}$) than native avidin ($T_m = 83.5^\circ\text{C}$), and its biotin-binding properties resembled those of AVR4. Optimization of a crucial subunit interface of avidin by an AVR4-inspired point mutation, I117Y, significantly increased the thermostability of the avidin mutant ($T_m = 97.5^\circ\text{C}$) without compromising its high biotin-binding properties. By combining these two modifications, a hyperthermostable ChiAVD(I117Y) was constructed ($T_m = 111.1^\circ\text{C}$). This study provides an example of rational protein engineering in which another member of the protein family has been utilized as a source in the optimization of selected properties.

Avidin is a homotetrameric glycoprotein isolated from chicken egg white. Each eight-stranded β -barrel subunit of avidin consists of 128 amino acids and has one ligand-binding site. Avidin, like its bacterial analogue streptavidin from *Streptomyces avidinii*, is able to form a tight and specific complex

* This work was supported by the ISB (National Graduate School in Informational and Structural Biology) and by a grant from the Academy of Finland. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Institute of Medical Technology, University of Tampere, FIN-33014 Tampere, Finland.

§§ To whom correspondence should be addressed. Tel.: 358-3-2158499; Fax: 358-3-2157710; E-mail: markku.kulomaa@uta.fi.

with a water-soluble vitamin, *d*-biotin ($K_d \approx 10^{-15}$ M) (1, 2). This special property of avidin together with its tetrameric nature and high stability have made it one of the most widely exploited protein tools in the life sciences across a range of biochemical, pharmaceutical, and biophysical applications (3, 4).

The avidin gene family consists of avidin and seven avidin-related genes (AVRs) (5). Although avidin protein is expressed in various tissues (6), the other members of the gene family so far have not been found in the form of proteins in the chicken. However, mRNA molecules of some AVRs are detected because of inflammatory reaction (7). To study their functional and structural properties, AVR proteins were recently produced by a baculovirus insect cell expression system (8). Avidin and AVR4¹ are ~80% identical in amino acid sequence, and almost all of the residues involved in biotin binding in avidin are conserved in AVR4. It was found that recombinant AVR4 bound biotin almost as tightly ($K_d \approx 3.6 \times 10^{-14}$ M) as avidin ($K_d \approx 1.1 \times 10^{-16}$ M). However, it was shown to be significantly more thermostable ($T_m = 106.4^\circ\text{C}$) than avidin ($T_m = 83.5^\circ\text{C}$) (9).

It has been proposed that protein oligomerization in nature serves to obtain more stable structures (10, 11). In addition, stable proteins tend to have only a few intrinsic water clefts in their structures (12, 13). Moreover, the role of ionic bonds in establishing the high thermal stability of proteins has been studied by Szilágyi and Závodszy (14), who performed a statistical analysis of high quality protein structures obtained from mesophilic and thermophilic organisms. They observed a correlation between the number of ion pairs and growth temperature of the organism and hypothesized that ion pairs have structural importance especially at high temperatures. The ionic bonds found in thermostable proteins have successfully been transferred to their analogues from mesophilic organisms to stabilize them (15). The importance of aromatic pairs in thermostable proteins has also been noticed (16), and these pairs have successfully been transferred between proteins to improve the thermal stability (17).

In this study, we compared the structures of avidin and AVR4. Our aim was to determine the importance of the differences in the primary and three-dimensional structures² of

¹ The abbreviations used are: AVR4, Avidin-related protein encoded by genes AVR4/5 bearing mutation C122S; DSC, differential scanning calorimeter.

² Eisenberg-Domovich, Y., Hytönen, V. P., Wilchek, M., Bayer, E. A., Kulomaa, M. S., and Livnah, O. (2005) *Acta Crystallog. Sect. D. Biol. Crystallog.*, in press.

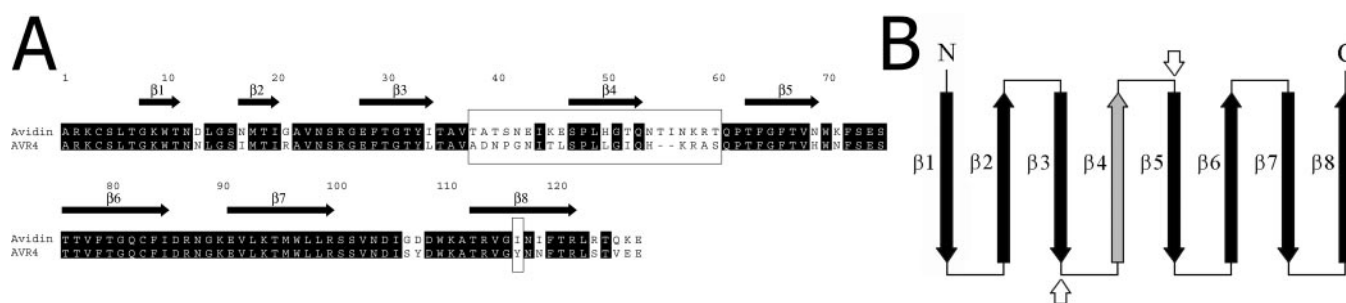


FIG. 1. *A*, Needleman-Wunsch alignment of chicken avidin and AVR4. The identity of the proteins is 77.8%. The area covering β -strand 4 and parts of its surrounding loops are boxed as is the mutated amino acid residue Ile¹¹⁷. The alignment was made using MALIGN (38). *B*, topology diagram of the chimeric protein. The segments taken from avidin are shown in black, and the part inserted from AVR4 is shown in gray. The boundaries between these segments are highlighted by white arrows.

these proteins for their biotin-binding and stability properties (9). On the basis of the results, we first engineered a chimeric avidin, in which β -strand 4 and its adjacent loops were replaced by the corresponding region from AVR4. In another mutant, a point mutation (I117Y) was introduced into avidin from the AVR4 sequence. Furthermore, we combined this mutation to the chimeric avidin also. The analysis of these new forms of avidin yields novel knowledge about the avidin protein family and in general provides an insight into how to improve the stability of proteins by combining characteristics from their known analogues.

EXPERIMENTAL PROCEDURES

Mutagenesis of Avidin and the Purification of Expressed Recombinant Proteins—AVR4/5(C122S) was previously found to have properties similar with the wild type protein but lacking the tendency to form higher order oligomers (9). In this article, term AVR4 denotes the C122S mutant of identical proteins coded by the genes *AVR4* and *AVR5*.

Site-directed mutagenesis of avidin encoding cDNA was performed by using the QuikChange (Stratagene, La Jolla, CA) method of mutagenesis. DNA encoding chimeric ChiAVD protein was produced by three sequential PCR reactions in which the final product was obtained from partially overlapping megaprimer (18) products. The final product was then cloned to the pFASTBAC1 vector. Recombinant baculoviruses coding for ChiAVD forms, AVR4 and AVD(I117Y), were generated as instructed by the manufacturer of the Bac-To-Bac™ system (Invitrogen). Proteins were produced in baculovirus-infected Sf9 insect cells in biotin-free medium as reported earlier (8). AVR4 was also produced in *Escherichia coli* as previously reported (named AVR4-6) (19), exploiting the bacterial signal peptide from *Bordetella avium* OmpA protein. The proteins were then purified by affinity chromatography using 2-iminobiotin-agarose, as described previously (20). The protein forms are summarized in Table I.

Proteinase K Assay—The proteolytic resistance of the proteins was studied using proteinase K as described previously (8). A protein sample (4 μ g) was incubated in the presence of proteinase K (1:25, w/w) at 37 °C for a predetermined time period, denatured by boiling in sample buffer (SDS, 2-mercaptoethanol), and subjected to SDS-PAGE followed by Coomassie Blue staining.

Gel Filtration Analysis—The oligomeric state of the proteins was assayed with fast protein liquid chromatography gel filtration as described previously (9). Sodium carbonate buffer (50 mM, pH 11) with 150 mM NaCl was used as the liquid phase. Protein samples of 5–10 μ g were used in the analysis.

Optical Biosensor Studies—The biotin-binding characteristics of the different avidins were studied by a surface plasmon resonance optical biosensor (IASys). The binding affinities to a fully 2-iminobiotin-covered surface in 50 mM borate buffer (pH 9.5, 1 M NaCl) were measured as reported previously (8).

Radiobiotin Dissociation Assay—The dissociation rate of [³H]biotin from avidin, AVR4, and the avidin mutants was determined as described in Ref. 21 at various temperatures. The activation thermodynamic parameters for AVR4 and avidin were determined by analysis of the dependence of the dissociation rate upon temperature using the global fit of all data as described in Ref. 22.

Fluorescent Biotin Dissociation Assay—The binding of labeled biotin to avidins was analyzed by a method based on the quenching of a biotin-coupled fluorescent probe ArcDia™ BF560 (ArcDia Ltd., Turku,

Finland) caused by binding to avidin as described previously (19). The measurements were performed using a PerkinElmer Life Sciences LS55 luminometer with thermostated cuvette (25 or 50 °C). The signal measured for 3600 s (25 °C) or 2400 s (50 °C) after the addition of a 100-fold amount of unlabeled biotin (5 μ M) was used to determine the dissociation rate constant.

Differential Scanning Calorimetry—The transition midpoint of the heat denaturation (T_m) of the avidin proteins was studied using a Calorimetry Sciences Corporation Nano II DSC as in previous reports (23, 24). Proteins (~0.5 mg/ml) were analyzed both in the absence and presence of biotin (3:1 molar ratio, biotin:avidin monomer).

Microplate Assay—The inactivation of the proteins during heat treatment was analyzed by using a microplate assay (24). The proteins were heated to 99.9 °C in 50 mM phosphate buffer containing 100 mM NaCl (pH 7.0) for 32 min. The remaining activity was probed by measuring the ability of the proteins to bind biotinylated alkaline phosphatase by coating the microplate wells with samples of the heated proteins.

RESULTS

Sequence Analysis—The sequence identity between avidin and AVR4 is 77.8%. 15 of 28 differences between these two proteins are found on the relatively short amino acid segment between the end of β 3 and the beginning of β 5 (23 residues in avidin and 21 residues in AVR4, Fig. 1). All residues showing contact with biotin (25) are conserved, excluding the Thr³⁸-Ala³⁹-Thr⁴⁰-sequence located in the L3,4 loop (connecting β 3 and β 4 strands), which is replaced by Ala³⁸-Asp³⁹-Asn⁴⁰ in AVR4. Subunit interface residues (41 residues) (9) are also well conserved with the only amino acid differences being T38A, A39D, H50L, T52I, N54H, and I117Y (numbering according to avidin sequence).

Expression and Purification of Recombinant Proteins—To study the significance of the differing segment between β 3 and β 5, a ChiAVD was constructed in which this segment was transferred from AVR4 to avidin. Furthermore, we mutated isoleucine 117 in avidin to tyrosine according to AVR4 (Fig. 1). The mutagenesis work was first planned using information obtained from the primary sequence analysis, known three-dimensional structure of avidin, and homology modeling. During the work, the determination of the structure of AVR4² confirmed the selected rationale. The mutated avidins were produced using a baculovirus expression system (8). Non-glycosylated AVR4-b was produced using the *E. coli* expression system (19)³ to study the influence of the carbohydrate chains on the properties of the protein (Table I). Avidin proteins were purified effectively by 2-iminobiotin-agarose affinity chromatography. Isolated proteins showed high purity in SDS-PAGE analysis (Fig. 2). The glycosylation patterns of the purified proteins differed because AVR4 has three potential glycosyla-

³ Laitinen, O. H., Airenne, K. J., Hytönen, V. P., Peltomaa, E., Mähönen, A., Wirth, T., Lind, M. M., Mäkelä, K. A., Toivanen, P. I., Schenkwein, D., Heikara, T., Nordlund, H. R., Kulomaa, M. S., and Ylä-Herttuala, S. (2005) *Nucleic Acids Res.*, in press.

TABLE I
Description of proteins analyzed in this study

Protein	Modifications	Source
AVD		Chicken ^a
AVD(I117Y)	Mutation I117Y in 1–3 subunit interface.	BEVS ^b
ChiAVD	Residues 38–58 moved from AVR4 to avidin.	BEVS ^b
ChiAVD(I117Y)	Residues 38–58 moved from AVR4 to avidin and mutation I117Y in 1–3 subunit interface.	BEVS ^b
AVR4	Cysteine residue forming intermonomeric disulphide bridges in AVR4 mutated to serine.	BEVS ^b
AVR4-b	As above but non-glycosylated.	<i>E. coli</i> ^c

^a Chicken avidin obtained from S. A. Belovo (Bastogne, Belgium).

^b Recombinant protein produced by baculovirus expression vector system in insect cells.

^c Recombinant protein produced by bacterial expression system. This form contains three additional residues (Gln-Thr-Val) in the N terminus from the bacterial signal peptide (19).

tion sites, whereas avidin has only one (1, 8). One of these sites (Asn⁴³) of AVR4 was transferred to the chimeric protein, resulting in more extensive glycosylation of the chimera when compared with that of native avidin. Bacterially produced AVR4-b was non-glycosylated as expected (Fig. 2).

Fast Protein Liquid Chromatography Gel Filtration—All the engineered proteins showed tetrameric appearance when subjected to gel filtration analysis. Bacterially produced AVR4 showed a slightly lower apparent molecular weight as compared with the other proteins, as expected, because of the lack of the carbohydrate moiety (Table II).

Proteinase K Assay—Avidin and avidin mutant AVD(I117Y) were found to be 50% digested after treatment for 16 h with proteinase K (Fig. 3). When these proteins were saturated with biotin before treatment, no cleavage was observed. AVR4, ChiAVD, ChiAVD(I117Y), and AVR4-b, however, displayed total resistance to the proteolytic activity of proteinase K even without biotin. This indicates that the conformation of L3,4 of the AVR4 apoprotein² protects it from digestion because there are putative cleavage sites for proteinase K close to the known cleavage sites in avidin (26) in the primary sequence (data not shown). Furthermore, glycosylation in residue Asn⁴³ (9) did not seem to play a role in the protease resistance.

Differential Scanning Calorimetry—In the DSC experiments the ChiAVD protein and AVD(I117Y) showed significantly better thermal stability than avidin both as apoforms and holoforms (Table II). When these modifications were combined, the resultant ChiAVD(I117Y) was found to be even more stable than AVR4. In all cases holoforms were clearly more stable than apoforms. We observed a significant increase in the unfolding enthalpy with an increasing melting temperature. AVR4-b showed similar behavior as compared with the protein produced in insect cells in the DSC analysis, but the determined ΔH of the unfolding was significantly lower, possibly reflecting the role of glycosylation moieties in the unfolding process (37).

Microplate Assay—The inactivation of the proteins during heat treatment was microplate-assayed as in our previous study (24). The remaining activity after treatment for 32 min is shown in Table II. These results are in line with DSC analyses showing that ChiAVD(I117Y) is the most thermally stable of the proteins characterized.

Biotin-binding Analyses—The binding kinetics of the fluorescent biotin conjugate BF560 to different avidins were compared by measuring the dissociation rate constants at 25 °C and 50 °C (Table IV). Avidin showed a lower dissociation rate when compared with AVR4. ChiAVD showed characteristics similar to AVR4 in this assay, *i.e.* faster dissociation at 50 °C. Interestingly, mutation Ile117Tyr seemed to tighten the binding of the biotin conjugate to avidin.

Ligand-binding analyses done with an IAsys optical biosensor showed a slightly decreased affinity to the 2-*iminobiotin* surface in the case of ChiAVD as compared with avidin (Table III). However, the affinity was nonetheless high resembling the

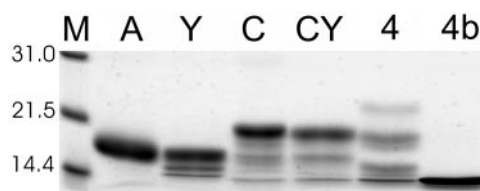


FIG. 2. **SDS-PAGE analysis of the purified proteins.** A, avidin; Y, AVD(I117Y); C, ChiAVD; CY, ChiAVD(I117Y); 4, AVR4; 4b, AVR4-b produced in bacteria. The marker proteins (M) with molecular masses of 14.4, 21.5, and 31 kDa are shown in figure. The multiple bands are caused by N-linked glycosylation occurring in insect cells (9).

values found for AVR4 (9). We also observed a slight decrease in the association rate constants of both ChiAVD forms as compared with avidin and AVD(I117Y). The surface used in this assay was fully 2-*iminobiotin*ylated and therefore did not exclude multiple binding events.

AVR4 showed long lived biotin binding in the radiobiotin dissociation assay (Fig. 4). However, the measured dissociation rate constants were significantly higher than those of avidin. ChiAVD and ChiAVD(I117Y) resembled AVR4 in this assay. Interestingly, bacterial AVR4-b showed a slightly slower dissociation rate in this assay than AVR4 produced in insect cells. Avidin mutant AVD(I117Y) showed very slow, avidin-like biotin dissociation. The activation thermodynamic parameters were calculated from the data as described elsewhere (22) and combined with the thermodynamic parameters obtained for AVR4 and avidin in a previous study (9). The values obtained are shown in Fig. 5. The analysis indicates the increase in both in the transition state and bound state free energies of AVR4 relative to those of avidin.

Three-dimensional Structure Analysis—The information obtained from the avidin (25) and AVR4² x-ray structures suggests that the substitution of $\beta 4$ and its adjacent L3,4 and L4,5 loops from AVR4 to avidin cause no crucial change in the overall shape of the resulting protein. The different conformation of the L3,4 loop of AVR4 (Fig. 6) when compared with that of avidin should be analogously reflected in the properties of ChiAVD, namely in a lower number of hydrogen bonds to the bound ligand (avidin, 11 hydrogen bonds; AVR4, 8 hydrogen bonds).² Furthermore, the interchanged sequence includes the two-residue deletion in loop L4,5, which was observed not to change the structural properties of the surrounding region of the loop in AVR4.²

DISCUSSION

In the present study we sought to clarify the features that make AVR4 a so much more thermostable protein than chicken avidin (9) and to transfer this higher stability to avidin. Another objective of the study was to explore and compare the biotin-binding properties of avidin, AVR4, and the chimeric proteins produced in this study. To accomplish the first objective, molecular modeling and the solved three-dimensional structure of avidin were utilized, and the results used to trans-

TABLE II
Structural properties of avidins

Fast protein liquid chromatography gel filtration elution times and calculated molecular weights of the proteins. Heat-induced unfolding of proteins determined by DSC (average \pm S.D).

Protein	Gel filtration		DSC ^a				Microplate assay
	Elution time <i>min</i>	Molecular mass <i>kDa</i>	$T_m(-\text{biotin})$ $^{\circ}\text{C}$	$T_m(+\text{biotin})$ $^{\circ}\text{C}$	ΔT_m^b $^{\circ}\text{C}$	ΔH^c <i>kJ/mol</i>	Activity ^d $\%$
AVD	29.3	53.1	83.5 \pm 0.1	117.0 \pm 0.7	33.5	329 \pm 5	4
AVD(I117Y)	29.6	49.5	97.5 \pm 0.4	123.7 \pm 0.1	26.2	536 \pm 6	47
ChiAVD	29.0	56.9	96.5 \pm 0.2	124.4 \pm 0.2	27.9	527 \pm 10	47
ChiAVD(I117Y)	29.0	56.9	111.1 \pm 0.2	\sim 130 ^e	\sim 19	659 \pm 38	98
AVR4	29.5	50.6	106.4 \pm 0.8	125.4 \pm 0.8	19.5	575 \pm 33	72
AVR4-b	30.3	42.1	109.9	127.1	17.2	460	56

^a The results of AVD and AVR4 are from Ref. 9.

^b ΔT_m is the change in T_m upon the addition of a 3-fold molar excess of biotin.

^c ΔH value obtained from sample without biotin. The value could not be determined accurately from samples saturated with biotin because the T_m values were too close to the temperature limit of the DSC instrument.

^d Biotin-binding activity after 32 min of treatment at 99.9 $^{\circ}\text{C}$.

^e Could not be determined accurately due to the upper limit of temperature of the DSC instrument.

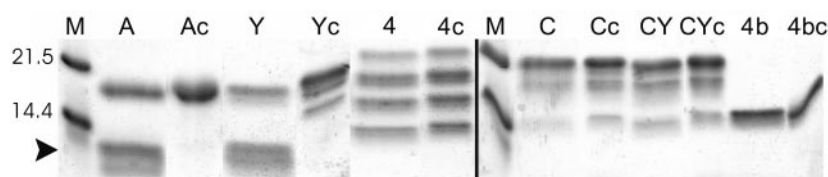


FIG. 3. SDS-PAGE analysis of protein samples incubated with proteinase K. The arrow shows the location of digested protein forms. The samples marked with “c” are control samples without protease treatment. A, avidin + proteinase K; Y, AVD(I117Y); 4, AVR4; C, ChiAVD; CY, ChiAVD(I117Y); 4b, AVR4-b. The marker proteins (M) with molecular weights of 14.4 and 21.5 are shown in figure.

TABLE III
2-Iminobiotin-binding properties of proteins
determined by IAsys optical biosensor

Iminobiotin-binding properties of the proteins analysed by an IAsys optical biosensor at 20 $^{\circ}\text{C}$. Affinities to 2-iminobiotin surface are determined from the equilibrium response data.

	K_d^a	k_{ass}
	<i>M</i>	$\text{M}^{-1} \text{s}^{-1}$
AVD	$(2.1 \pm 0.6) \times 10^{-8}$	$(2.6 \pm 0.2) \times 10^4$
AVD(I117Y)	$(6.3 \pm 1.2) \times 10^{-8}$	$(2.3 \pm 0.3) \times 10^4$
ChiAVD	$(1.1 \pm 0.4) \times 10^{-7}$	$(1.7 \pm 0.1) \times 10^4$
ChiAVD(I117Y)	$(6.7 \pm 1.2) \times 10^{-8}$	$(1.6 \pm 0.2) \times 10^4$
AVR4	$(1.4 \pm 0.4) \times 10^{-7}$	$(9.3 \pm 0.4) \times 10^3$

^a Apparent dissociation constant calculated from equilibrium values.

for the stabilizing elements from AVR4 to avidin. On the basis of the sequence comparison made between avidin and AVR4, the highly variable segment between the two proteins is located between L3,4 and L4,5 (Fig. 1). Thus, our initial approach was to substitute this segment in avidin for the corresponding sequence in AVR4. Another substitution (I117Y) was made on the basis of the modeling (9) This mutation rationale was later confirmed when the three-dimensional x-ray structure of AVR4 came available during the present study,² which was thought to play an important role in the intersubunit interactions and thus contribute to the thermal stability of AVR4 (Table I).

As expected, a significantly more stable protein was obtained when a segment 21 amino acids long (residues 38–58) taken from AVR4 was transferred to avidin, replacing a 23-residue segment. The shortening of the L4,5 loop by two residues in ChiAVD avidin may provide a partial explanation for the higher stability. By comparing genomes of mesophiles, thermophiles, and extremophiles, Thompson and Eisenberg (27) found shorter exposed loops in temperature-resistant proteins when compared with those of their mesophilic analogues. The L3,4 loop, however has the same length in avidin and AVR4, yet its amino acid composition is very different (Fig. 1). The three-dimensional structure of L3,4 clearly shows that it has a dif-

ferent conformation in AVR4 than in avidin (Fig. 6).² We assume that ChiAVD and AVR4 share a similar L3,4 loop conformation. The presence of the Pro⁴¹-Gly⁴² stretch and the salt bridge between Asp³⁹ and Arg¹¹² induces stability in the L3,4 loop in both the apo and biotin complex forms.²

The importance of certain stability “hot spot” residues in protein interfaces has been reported (28–30). Aromatic pairs are known to form stabilizing pairs in protein structures, which have also been studied experimentally (17). In the present study, we were able to optimize the subunit interface of avidin by replacing the Ile¹¹⁷ residue with tyrosine according to the AVR4 sequence. The previous modeling analysis (9) suggested that tyrosine in this location likely contributes to stability of the AVR4 tetramer as compared with that of avidin. The three-dimensional structure of AVR4 indicates the presence of π - π stacking between two tyrosine residues from neighboring monomers,² and experimental data support the improved stability caused by interactions at this site (AVD(I117Y), $T_m = 97.5$ $^{\circ}\text{C}$; avidin, $T_m = 83.5$ $^{\circ}\text{C}$). Kannan and Vishveshwara (16) compared the aromatic clusters in proteins from thermophilic and mesophilic organisms. They found that residues comprising aromatic clusters in proteins from thermophiles are preferably replaced by Leu or Ile in proteins from mesophilic organisms.

Regarding stability, the most interesting result was that combination of the chimera approach and the I117Y point mutation produced a protein that was even more thermostable than AVR4; the sum is greater than the parts. This indicates that we have successfully recognized and transferred the structural factors that account for the difference in stability between avidin and AVR4. It has been proposed earlier that recombination inside the avidin gene family is a frequent event (31). Our results indicate that such recombination might produce functional chimeric proteins inside the gene family because building blocks moved from AVR4 seem to be able to function as part of the avidin structure without negative implications.

We found that AVR4 binds biotin almost as tightly as avidin.

FIG. 4. Radiobiotin dissociation rate constants measured for the proteins as a function of temperature. The models for dissociation rate constants obtained by global fit analysis are shown by lines and the determined individual dissociation rates by symbols.

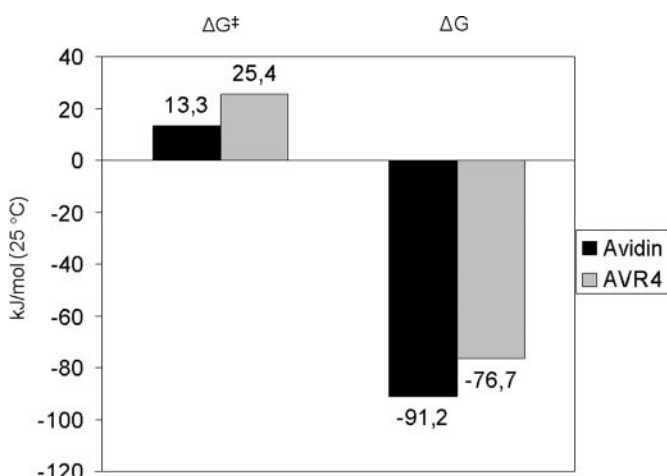
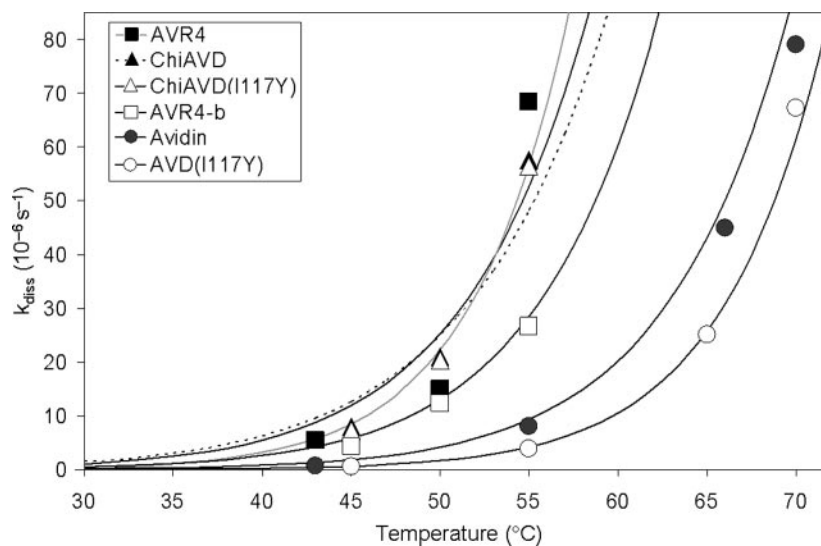


FIG. 5. Reaction thermodynamic coordinates for biotin binding to avidin and AVR4. The activation energy of dissociation obtained from the Eyring equation fitted to the dissociation data were used to calculate the transition state energy (ΔG^\ddagger) for the reaction, using the total binding energy values (ΔG) previously determined for binding (9).

The analysis of the [^3H]biotin dissociation data measured at different temperatures revealed that the energy barrier between unbound and bound states in AVR4 is somewhat smaller than in avidin (Fig. 5). The higher free energy of the transition state might also explain the slower association rate to the 2-iminobiotin surface of AVR4 as compared with avidin (Fig. 5, Table III). Because the free energy of the binding is lower in the case of AVR4 (9), the biotin dissociation barrier is still lower for AVR4 despite the higher transition state free energy. A potential explanation for the differences in biotin-binding characteristics between AVR4 and avidin lies in the differences in the L3,4 loop. This region has been found to be an important factor in biotin binding to streptavidin (32). Both the ChiAVD forms showed biotin-binding properties similar to those of AVR4 when measured by various methods, therefore supporting this hypothesis. We found that glycosylation may play a minor role in biotin binding because AVR4-b produced in bacteria showed slightly slower dissociation rates in both the radiobiotin and fluorescent biotin dissociation analyses (Fig. 4, Table IV).

The dissociation rates observed in the fluorescent biotin assay were significantly faster than those obtained from the radiobiotin analysis, but the proteins nevertheless showed similar trend in relative characteristics (Table IV). Proteins bear-

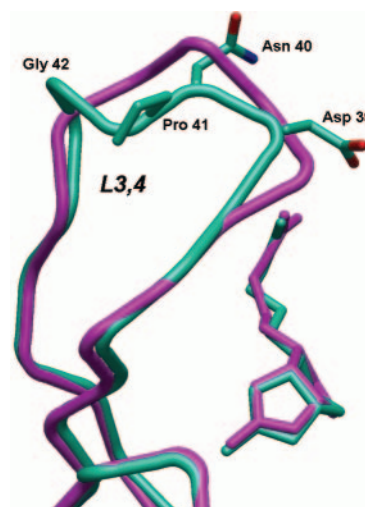


FIG. 6. Tube representation of the superposition of the avidin (magenta) and AVR4 (cyan) in the L3,4 loop region with biotin molecules in the binding site for reference. The side chains of residues 39–42 are shown for AVR4 and indicate the kink in the loop induced by the Pro-Gly tandem. Although the L3,4 loops in both proteins are of the same size, they adopt an entirely different conformation. The closed conformation of L3,4 in the AVR4 is apparent in both the apo and biotin complex forms. The figure was constructed using MIDAS (39).

TABLE IV
Biotin conjugate dissociation kinetics

Dissociation rate constants measured for biotin-BF560 conjugate by a spectrofluorometer at 25 °C and at 50 °C. The total releases of the probe after measurement for 1 h are also shown.

	k_{diss} (25 °C)	Release 1 h (25 °C)	k_{diss} (50 °C)	Release 1 h (50 °C)
	10^{-5} s^{-1}	%	10^{-4} s^{-1}	%
AVD	2.04	10.7	2.74	71.5
AVD(I117Y)	0.81	3.0	1.61	56.5
ChiAVD	1.76	13.1	8.00	93.9
ChiAVD(I117Y)	1.94	9.5	7.33	94.8
AVR4	2.77	12.5	7.88	93.7
AVR4-b	2.31	10.9	6.20	88.3

ing $\beta 4$ and its adjacent loops from AVR4 showed a higher dissociation rate constants when compared with avidin. The difference in radiobiotin analysis was ~ 6 -fold at 50 °C (Fig. 4), and the fluorescent BF560-biotin showed ~ 3 -fold higher dissociation rate constant at 50 °C (Table IV). Therefore, the differ-

ences in the structure of the L3,4 between AVR4 and avidin could explain this minor change in the binding properties of conjugated biotin relative to free biotin (Fig. 6).² Notably, the stabilizing mutation I117Y decreased the dissociation rate of the fluorescent biotin conjugate from avidin. This may indicate that the biotin dissociation event is at least partially determined by the tetrameric integrity of avidin. Although the binding of biotin to avidin is non-cooperative (for a review, see Ref. 1), the neighboring subunit is important for the structure (25) and function (33, 34) of the binding site. Therefore, the stabilization of the tetramer can enhance the binding of biotin conjugate in certain conditions. We have previously noticed the importance of movement of water associated with biotin dissociation from streptavidin (35). Because the residue 117 lies in this region in avidin, the change in the local environment caused by the mutation I117Y may also reflect to the molecular events linked with biotin movement.

Proteinase K cleaves avidin in only one region in the loop between β -strands 3 and 4 (26). It was also found that biotin efficiently inhibits the cleavage. On the other hand, AVR4 and streptavidin are resistant to cleavage by proteinase K even without bound biotin (9, 34). Because proteinase K cleaves a variety of sequences, the explanation for the resistance to cleavage might lie in the conformation of the L3,4 loop of AVR4. The proteinase K resistance of ChiAVD supports these results (Fig. 3). The different loop structure can be seen in the structure of AVR4 (Fig. 6).² Proline in this loop seems to cause bending in the middle of the loop. Accordingly, the corresponding loop in streptavidin is three residues shorter (36) and may not, therefore, be accessible to the protease. We did not observe any difference between the glycosylated and non-glycosylated form of AVR4 in this analysis; hence, the sugar moiety in this loop in AVR4 cannot explain the resistance to the protease.

In conclusion, we were able to define two structural elements that make AVR4 a notably more thermostable protein than avidin. First, AVR4 has a better set of interface contacts around position 115 (tyrosine-tyrosine) than avidin has in the equivalent position (117, isoleucine-isoleucine). Second, the region between β_3 and β_5 seems, at least partially, to explain the better stability properties of AVR4 when compared with those of avidin. This segment might also cause the slight differences observed in biotin-binding properties between avidin and AVR4. The chimeras thus obtained will find a role in applications utilizing extreme conditions. For example, these proteins can resist high temperatures used in PCR-based applications utilizing biotinylated molecules and a biotin-binding protein, and they can be used with crude cell extracts and other solutions containing proteolytic enzymes.

Acknowledgments—We thank Irene Helkala, Eila Korhonen, Pirjo Käpylä, Sonja Koistinen, and Philip Duffy for first class technical assistance. We thank Peter Slotte for access to the calorimeter laboratory facilities. We acknowledge Michael Freeman for proofreading the manuscript.

REFERENCES

- Green, N. M. (1975) *Adv. Protein Chem.* **29**, 85–133
- Green, N. M. (1990) *Methods Enzymol.* **184**, 51–67
- Wilchek, M., and Bayer, E. A. (1999) *Biomol. Eng.* **16**, 1–4
- Wilchek, M., and Bayer, E. (1990) *Methods Enzymol.* **184**, 5–13
- Ahloth, M. K., Kola, E. H., Ewald, D., Masabanda, J., Sazanov, A., Fries, R., and Kulomaa, M. S. (2000) *Anim. Genet.* **31**, 367–375
- Tuohimaa, P., Joensuu, T., Isola, J., Keinänen, R., Kunnas, T., Niemela, A., Pekki, A., Wallen, M., Ylikomi, T., and Kulomaa, M. (1989) *Int. J. Dev. Biol.* **33**, 125–134
- Kunnas, T. A., Wallen, M. J., and Kulomaa, M. S. (1993) *Biochim. Biophys. Acta* **1216**, 441–445
- Laitinen, O. H., Hytönen, V. P., Ahlroth, M. K., Pentikäinen, O. T., Gallagher, C., Nordlund, H. R., Ovod, V., Marttila, A. T., Porkka, E., Heino, S., Johnson, M. S., Airene, K. J., and Kulomaa, M. S. (2002) *Biochem. J.* **363**, 609–617
- 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., and Kulomaa, M. S. (2004) *J. Biol. Chem.* **279**, 9337–9343
- Backmann, J., Schafer, G., Wyns, L., and Bonisch, H. (1998) *J. Mol. Biol.* **284**, 817–832
- Villerey, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N., and Van Beeumen, J. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 2801–2806
- Knapp, S., de Vos, W. M., Rice, D., and Ladenstein, R. (1997) *J. Mol. Biol.* **267**, 916–932
- Britton, K. L., Yip, K. S., Sedelnikova, S. E., Stillman, T. J., Adams, M. W., Ma, K., Maeder, D. L., Robb, F. T., Tolliday, N., Vetricani, C., Rice, D. W., and Baker, P. J. (1999) *J. Mol. Biol.* **293**, 1121–1132
- Szilágyi, A., and Závodszy, P. (2000) *Structure Fold Des.* **8**, 493–504
- Vetricani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., Robb, F. T. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12300–12305
- Kannan, N., and Vishveshwara, S. (2000) *Protein Eng.* **13**, 753–761
- Serrano, L., Bycroft, M., and Fersht, A. R. (1991) *J. Mol. Biol.* **218**, 465–475
- Sarkar, G., and Sommer, S. S. (1990) *BioTechniques* **8**, 404–407
- Hytönen, V. P., Laitinen, O. H., Airene, K. J., Kidron, H., Meltola, N. J., Porkka, E., Hörhå, J., Paldanius, T., Määttä, J. A., Nordlund, H. R., Johnson, M. S., Salminen, T. A., Airene, K. J., Ylä-Herttua, S., and Kulomaa, M. S. (2004) *Biochem. J.* **384**, 385–390
- Laitinen, O. H., Marttila, A. T., Airene, K. J., Kulik, T., Livnah, O., Bayer, E. A., Wilchek, M., and Kulomaa, M. S. (2001) *J. Biol. Chem.* **276**, 8219–8224
- Klumb, L. A., Chu, V., and Stayton, P. S. (1998) *Biochemistry* **37**, 7657–7663
- Hyre, D. E., Le Trong, I., Freitag, S., Stenkamp, R. E., and Stayton, P. S. (2000) *Protein Sci.* **9**, 878–885
- Gonzalez, M., Argarana, C. E., and Fidelio, G. D. (1999) *Biomol. Eng.* **16**, 67–72
- Nordlund, H. R., Laitinen, O. H., Uotila, S. T., Nyholm, T., Hytönen, V. P., Slotte, J. P., and Kulomaa, M. S. (2003) *J. Biol. Chem.* **278**, 2479–2483
- Livnah, O., Bayer, E. A., Wilchek, M., and Sussman, J. L. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5076–5080
- Ellison, D., Hinton, J., Hubbard, S. J., and Beynon, R. J. (1995) *Protein Sci.* **4**, 1337–1345
- Thompson, M. J., and Eisenberg, D. (1999) *J. Mol. Biol.* **290**, 595–604
- Clackson, T., and Wells, J. A. (1995) *Science* **267**, 383–386
- DeLano, W. L. (2002) *Curr. Opin. Struct. Biol.* **12**, 14–20
- Hu, Z., Ma, B., Wolfson, H., and Nussinov, R. (2000) *Proteins* **39**, 331–342
- Ahloth, M. K., Ahlroth, P., and Kulomaa, M. S. (2001) *Biochem. Biophys. Res. Commun.* **288**, 400–406
- Chu, V., Freitag, S., Le Trong, I., Stenkamp, R. E., and Stayton, P. S. (1998) *Protein Sci.* **7**, 848–859
- Chilkoti, A., Tan, P. H., and Stayton, P. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1754–1758
- Laitinen, O. H., Airene, K. J., Marttila, A. T., Kulik, T., Porkka, E., Bayer, E. A., Wilchek, M., and Kulomaa, M. S. (1999) *FEBS Lett.* **461**, 52–58
- Hyre, D. E., Amon, L. M., Penzotti, J. E., Le Trong, I., Stenkamp, R. E., Lybrand, T. P., and Stayton, P. S. (2002) *Nat. Struct. Biol.* **9**, 582–585
- Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J., and Salemme, F. R. (1989) *Science* **243**, 85–88
- Wang, C., Eufemi, M., Turano, C., and Giartosio, A. (1996) *Biochemistry* **35**, 7299–7307
- Johnson, M. S., Overington, J. P., and Blundell, T. L. (1993) *J. Mol. Biol.* **231**, 735–752
- Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics* **6**, 36–37

Research article

Open Access

Structure and characterization of a novel chicken biotin-binding protein A (BBP-A)

Vesa P Hytönen^{1,6}, Juha AE Määttä^{†1,3}, Einari A Niskanen^{†1}, Juhani Huuskonen², Kaisa J Helttunen^{1,2}, Katrin K Halling⁴, Henri R Nordlund^{1,3}, Kari Rissanen², Mark S Johnson⁴, Tiina A Salminen⁴, Markku S Kulomaa^{1,3}, Olli H Laitinen⁵ and Tomi T Airene^{*4}

Address: ¹NanoScience Center, Department of Biological and Environmental Science, P.O. Box 35 (YAB), FI-40014 University of Jyväskylä, Finland, ²NanoScience Center, Department of Chemistry, P.O. Box 35, FI-40014 University of Jyväskylä, Finland, ³Institute of Medical Technology, FI-33014 University of Tampere, Finland, ⁴Department of Biochemistry and Pharmacy, Åbo Akademi University, Tykistökatu 6 A, FI-20520, Turku, Finland, ⁵A. I. Virtanen Institute, Department of Molecular Medicine, University of Kuopio, P.O. Box 1627, FI-70211 Kuopio, Finland and ⁶Department of Materials, ETH Zürich, CH-8093 Zürich, Switzerland

Email: Vesa P Hytönen - veshyto@mat.ethz.ch; Juha AE Määttä - juha.maatta@uta.fi; Einari A Niskanen - einnisk@cc.jyu.fi; Juhani Huuskonen - juhani.huuskonen@jyu.fi; Kaisa J Helttunen - kjhelttu@cc.jyu.fi; Katrin K Halling - khalling@abo.fi; Henri R Nordlund - henri.nordlund@uta.fi; Kari Rissanen - kari.rissanen@jyu.fi; Mark S Johnson - johnson4@abo.fi; Tiina A Salminen - tiina.salminen@abo.fi; Markku S Kulomaa - markku.kulomaa@uta.fi; Olli H Laitinen - olli.laitinen@uku.fi; Tomi T Airene* - tomi.airene@abo.fi

* Corresponding author †Equal contributors

Published: 7 March 2007

Received: 8 February 2007

BMC Structural Biology 2007, 7:8 doi:10.1186/1472-6807-7-8

Accepted: 7 March 2007

This article is available from: <http://www.biomedcentral.com/1472-6807/7/8>

© 2007 Hytönen et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The chicken genome contains a *BBP-A* gene showing similar characteristics to avidin family genes. In a previous study we reported that the *BBP-A* gene may encode a biotin-binding protein due to the high sequence similarity with chicken avidin, especially at regions encoding residues known to be located at the ligand-binding site of avidin.

Results: Here, we expand the repertoire of known macromolecular biotin binders by reporting a novel biotin-binding protein A (BBP-A) from chicken. The BBP-A recombinant protein was expressed using two different expression systems and purified with affinity chromatography, biochemically characterized and two X-ray structures were solved – in complex with D-biotin (BTN) and in complex with D-biotin D-sulfoxide (BSO). The BBP-A protein binds free biotin with high, "streptavidin-like" affinity ($K_d \sim 10^{-13}$ M), which is about 50 times lower than that of chicken avidin. Surprisingly, the affinity of BBP-A for BSO is even higher than the affinity for BTN. Furthermore, the solved structures of the BBP-A – BTN and BBP-A – BSO complexes, which share the fold with the members of the avidin and lipocalin protein families, are extremely similar to each other.

Conclusion: BBP-A is an avidin-like protein having a β -barrel fold and high affinity towards BTN. However, BBP-A differs from the other known members of the avidin protein family in thermal stability and immunological properties. BBP-A also has a unique ligand-binding property, the ability to bind BTN and BSO at comparable affinities. BBP-A may have use as a novel material in, e.g. modern bio(nano)technological applications.

Background

Several biotin-binding proteins have been characterized from egg-laying vertebrates. The best known of these proteins is chicken avidin (AVD) [1-4]. This tetrameric ~60 kDa egg-white protein, together with its bacterial homologs streptavidin and bradavidin [5-7], has a fascinating feature, the ability to bind a small water soluble vitamin, D-biotin (BTN), with tremendous affinity ($K_d \approx 10^{-15}$ M for AVD) [5]. The formation of the extraordinary strong protein-ligand complex is a result of "perfect" structural complementarity between the (strept)avidin ligand-binding site and BTN [8,9], and optimized "packing" of the overall tertiary and quaternary structure, too [5,10]. Although BTN is thought to be the natural ligand of (strept)avidin, these proteins are known to bind naturally-occurring as well as synthetic BTN analogous and their derivatives, which include the biotechnologically valuable 2-iminobiotin, 4'-hydroxyazobenzene-2-carboxylic acid (HABA) and desthiobiotin (for a review, see [1,5]). However, (strept)avidin has clearly weaker affinity for ligands other than BTN. The biological role of AVD is still partially unclear, but it has been postulated to function as an antimicrobial defence protein in chicken by ensuring that no free biotin is present in egg white; a vitamin required for the growth of bacteria [1]. Nevertheless, it is the numerous bio(nano)technological applications, where AVD's unique biotin-binding property has been utilized, which have made AVD one of the most well-known proteins.

In addition to AVD of chicken egg-white, other proteins capable of binding BTN tightly, biotin-binding protein I (BBP-I) [11-13] and II (BBP-II) [14,15] of chicken egg-yolk, have been reported. Nothing is known about the affinity of BBP-I/II for ligands other than BTN. Although immunologically similar and having comparable N-terminal sequences [15,16], BBP-I and BBP-II differ in their thermal stability, BBP-I being more stable than BBP-II [14]. The biological functions of the two BBP forms are only partly resolved, but they are known to be synthesized in the liver and believed to transport BTN via plasma to the egg-yolk [14,16,17]. More specifically, it is believed that BBP-I has a general role as a transport protein in hen plasma, whereas BBP-II may be needed for the efficient deposition of BTN into the yolk of maturing oocytes [14].

There is a substantial evidence that the quaternary structure of BBP-I/II is tetrameric, but it is not clear whether the tetramers are formed as a result of assembling four monomers or if the tetramers result from limited proteolysis of a single polypeptide containing all the binding sites [13,16]. Evidence for the former hypothesis became available when the putative genes and cDNAs encoding the BBP-I/II proteins were reported in 2005 by Niskanen and

co-workers [18]: the exon-intron structures of the BBP genes [18], named as *BBP-A* and *BBP-B*, mimic those of AVD and avidin-related genes (AVRs) [19-21] and they encode proteins with a single ligand binding site per polypeptide chain. It is worth mentioning that it is not yet completely clear whether the *BBP-A* and *BBP-B* genes really encode for either of the earlier characterized BBP-I and BBP-II proteins; *BBP-A* in particular seems to encode a novel protein, BBP-A, not characterized before [18].

In the present study, we show that the chicken *BBP-A* gene encodes a functional homotetrameric protein. We report two different X-ray structures of BBP-A, one in complex with BTN and an other one in complex with D-biotin D-sulfoxide (BSO), at 2.1 Å and 1.75 Å resolution, respectively. To our knowledge, no other structures of any proteins in complex with BSO have been reported before. Using several biochemical methods, we show that BBP-A binds BSO even tighter than BTN. A comparative study of the ligand-binding and physicochemical properties of BBP-A and AVD are presented together with data from site-directed mutagenesis.

Results

Biochemical characterization of BBP-A

BBP-A, expressed both in *E. coli* (bBBP-A) and in insect cells (iBBP-A), was isolated and purified using 2-iminobiotin affinity chromatography. A typical yield was 5 mg of pure protein per one litre of culture medium for both expression systems. The mutations A74S and T118F, which were made in order to study the differences in the molecular origin of the biotin-binding affinity and the thermal stability of BBP-A as compared to chicken AVD, had no significant effect on the protein yields. The isolated proteins were shown to be over 95% pure using SDS-PAGE analysis (data not shown). A molecular weight of 13999.0 Da was determined for the expressed bBBP-A using mass spectrometry, and it matched the molecular weight calculated from the expression construct.

The oligomeric state of bBBP-A, bBBP-A(A74S), bBBP-A(T118F), iBBP-A, bAVD and wtAVD in solution was analysed using gel filtration chromatography. All of the proteins eluted as single peaks in this analysis, demonstrating the homogeneity of the samples. The apparent molecular weights of iBBP-A (60.8 kDa) and wtAVD (62.5 kDa), based on gel filtration analysis, corresponded well with the theoretical molecular weights for their respective tetramers. As expected, the proteins produced in bacterial cells showed apparent molecular weights lower than the corresponding proteins expressed in insect cells, at least partially due to the absence of glycosylation (bBBP-A, 42.7 kDa and bAVD, 55.3 kDa). The mutations A74S and T118F had no effect on the appearance of BBP-A in the gel filtration analysis. Nor did the addition of BTN to the

samples significantly alter the apparent molecular masses of the analysed proteins. This, together with the results showing that BTN significantly stabilizes tetramers of bBBP-A (see below; Table 1 and Figure 1), supports the idea that the quaternary structure of bBBP-A is tetrameric, an idea that is also supported by the crystal structures of bBBP-A – BSO and bBBP-A – BTN (see below).

The amino acid sequence of BBP-A contains one putative N-glycosylation site (Figure 2), which corresponds to an N-glycosylation site known to be present and glycosylated in AVD [1]. To see whether iBBP-A was also glycosylated, the protein was treated with Endoglycosidase H and analysed on SDS-PAGE (Figure 3). After the deglycosylation treatment, the iBBP-A protein had a molecular weight comparable to that of bBBP-A, a molecular weight clearly lower than the molecular weight of the untreated iBBP-A sample. This result, together with the control data (wtAVD), indicates that the N-glycosylation site of BBP-A was indeed glycosylated in insect cells and the size of the attached carbohydrate is comparable to that present on wtAVD.

The immunological cross-reactivity of BBP-A and wtAVD was studied with two polyclonal antibodies made against chicken AVD (Figure 4). Neither of the antibodies did cross-react with BBP-A in a Western blotting analysis. The reactivity of both antibodies was significantly weaker with iBBP-A (10 µg) than in the case of a 100-fold dilution of wtAVD (0.1 µg). In dot-blot analysis, only one antibody, the TdaVIII antibody, was used to test whether it detects iBBP-A and bBBP-A. In agreement with the immunoblot analysis, no or only negligible cross-reactivity was observed.

Stability of the tetrameric assembly of BBP-A

The thermal stability of the tetrameric form of BBP-A was analysed in the presence of 2-mercaptoethanol using an SDS-PAGE-based method described in [22]. Both bBBP-A and iBBP-A appeared mainly in monomeric forms on SDS-PAGE gels already at room temperature in the absence of BTN, whereas addition of BTN stabilised the tetrameric form, which was stable until the temperature was raised to 70 °C (Table 1 and Figure 1). In the presence of BTN, glycosylated iBBP-A had slightly better thermal stability ($T_r = 75$ °C) compared to bBBP-A ($T_r = 70$ °C). Both bBBP-A and iBBP-A were clearly more thermally labile than wtAVD. The mutations A74S and T118F had no significant effect on the stability of the tetrameric forms of bBBP-A (Table 1, Figure 1).

Differential scanning calorimetry (DSC) was not applicable to the bBBP-A forms in the absence of ligands, since a clear, single denaturation temperature could not be measured for these proteins. For bBBP-A, two peaks at around 51 °C and 68 °C were obtained. In the presence of BTN, however, the value of T_m was around 100 °C for all bBBP-A forms. These values are significantly lower in comparison to those measured for wtAVD (Table 1). Furthermore, the thermal stability of bBBP-A and AVD was analysed in the presence of BSO and D-biotin sulfone (Table 2). Only negligible differences were, however, found between the effects of these two ligands and BTN on the thermal stabilities of bBBP-A and AVD.

The overall X-ray structures of BBP-A

The crystal structures of bBBP-A in complex with BTN and BSO were solved at 2.1 Å and 1.75 Å resolution, respectively. The statistics for the structure determinations are summarized in Table 3. The overall BBP-A – BTN and

Table 1: Biochemical properties of different BBP-A forms.

	Gel filtration				SDS-PAGE-based thermostability assay		Dissociation of fluorescent BTN (25°C)		DSC	
	Elution time (min)		Molecular mass (kDa) ^a		T_r (°C)		k_{diss} ($10^{-5} s^{-1}$)	Release 1 h (%)	T_m (°C)	
	-BTN	+BTN	-BTN	+BTN	-BTN	+BTN			-BTN	+BTN
bBBP-A	29.9	30.1	42.7	40.9	N. D.	70	54.8	86	N. A.	103.4
bBBP-A (A74S)	30.0	30.1	41.9	40.8	N. D.	75	47.1	83	N. A.	103.2
bBBP-A (T118F)	30.1	29.9	41.3	42.7	N. D.	70	69.1	94	N. A.	102.3
iBBP-A	28.4	28.4	60.8	59.7	N. D.	75	20.4	53	N. M.	N. M.
AVD	28.2	28.2	62.5	63.6	60 ^b	90 ^b	2.0	10	83.5 ^c	117.0
bAVD	28.8	28.8	55.3	55.1	60 ^b	90 ^b	1.1	5	N. M.	N. M.

N. D. Protein appeared mainly in monomeric form already at room temperature; N. A. Not applicable; N. M. Not measured.

^aMolecular mass calculated from the measured elution times of thyroglobulin (670 kDa), gamma-globulin (158 kDa), BSA (67 kDa), ovalbumin (44 kDa) and myoglobin (17 kDa).

^bResults from ref. [30].

^cResults from ref. [33].

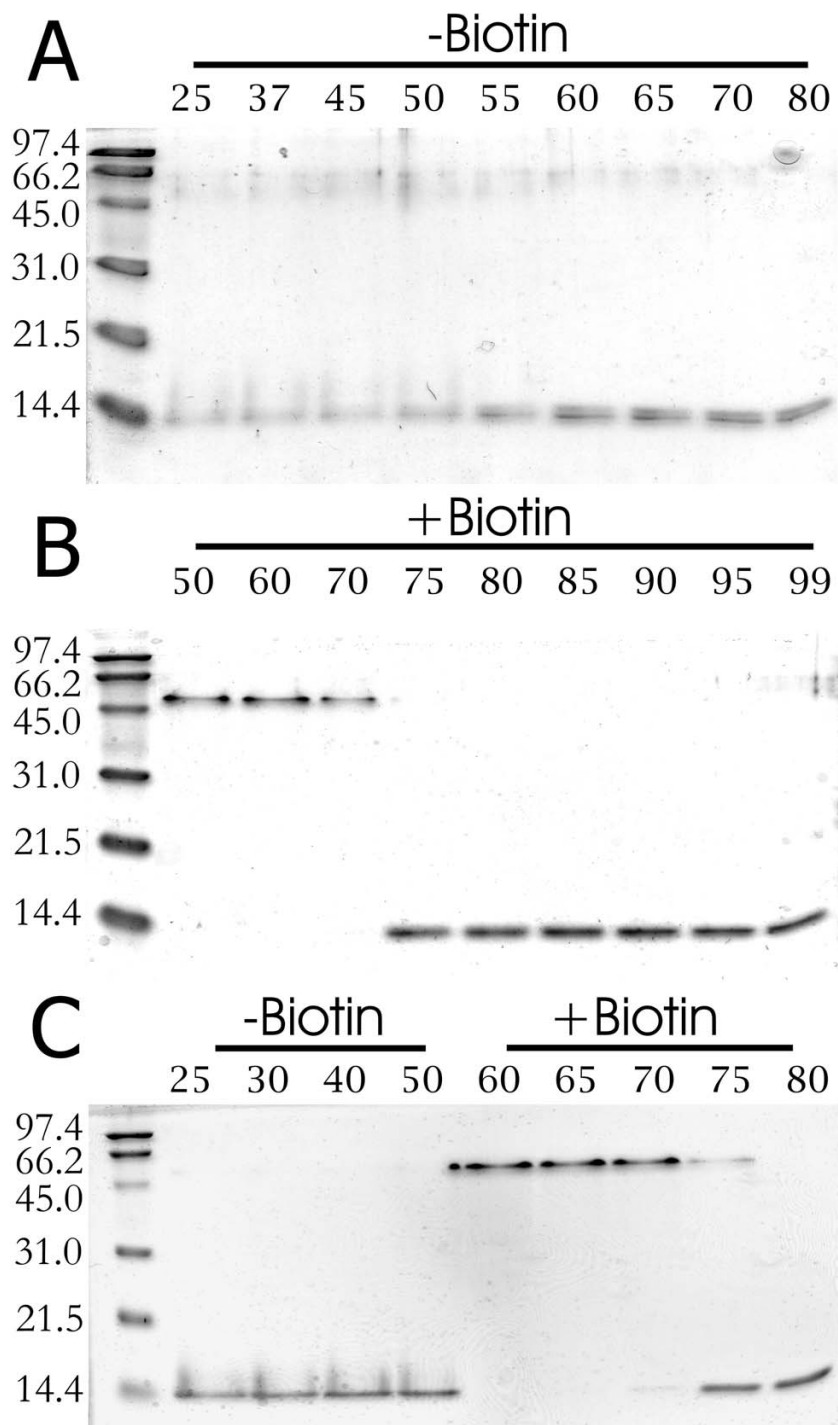
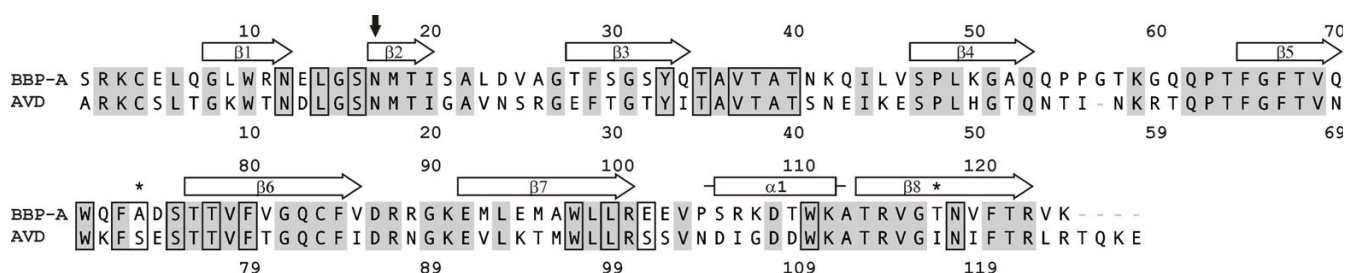
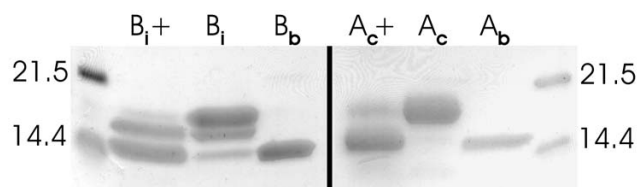


Figure 1
SDS-PAGE based analysis of the thermal stability of the tetramers of BBBP-A. Samples saturated with BTN (+Biotin) prior to analysis as well as samples without added BTN (-Biotin) are shown. (a) bBBP-A, (b) bBBP-A in the presence of BTN and (c) bBBP-A(A74S) in the absence and presence of BTN. After acetylation *in vitro*, the samples were subjected to thermal treatment in SDS-PAGE sample buffer containing 2-mercaptoethanol and SDS. The different temperatures used in the analysis are indicated in the upper part of the figure (°C). The molecular weights of the standard proteins (Bio-Rad) are shown on the left.

**Figure 2**

Sequence alignment of BBP-A and AVD. Identical residues are shown in gray shading. The signal peptides are omitted from the alignment. Secondary structure elements are indicated according to the BBP-A – BSO structure (chain A). The amino acids within 4 Å of BSO in the BBP-A – BSO structure (and the equivalent residues in AVD) are boxed. The residues mutated in this study are indicated by an asterisk. The glycosylation site of BBP-A is indicated by an arrow.

BBP-A – BSO structures are practically identical to each other; the C α atoms of these structures superimpose with an *rmsd* of < 0.2 Å. The BBP-A structures consist of four identical subunits, each subunit adopting a β -barrel fold of eight anti-parallel β -strands with an α -helix positioned between the β 7- and β 8-strands (Figure 5a,b). The BTN/BSO-binding site is located at the open end of each β -barrel. The overall fold of BBP-A is similar to that seen for the known structures of the AVD family complexes, i.e. streptavidin-BTN [8], AVD-BTN [9,23], AVR2-BTN [24] and AVR4-BTN [25] complexes. The solvent accessible surface area is, however, over 10% larger in the BBP-A structures (> 21000 Å²) compared to the AVD-BTN structures (< 19000 Å² [PDB: 1AVD and 2AVI]). The electrostatic surface properties of BBP-A and AVD differ, too, BBP-A having a seemingly more positively charged surface than AVD (Figure 5c–f).

**Figure 3**

Deglycosylation analysis of BBP-A. The BBP-A produced in insect cells was treated with Endo H_f glycosidase (B_i+) and analysed with SDS-PAGE. Untreated control sample (B_i) and BBP-A produced in bacteria (B_b) were also analysed. For comparison, the same analysis was performed for AVD isolated from chicken (A_c+). Chicken AVD control sample (A_c) as well as AVD produced in bacteria (A_b) [30] are also shown. The molecular weight markers are shown on both sides of the gel (14.4 and 21.5 kDa).

Mode of BTN and BSO binding

The biotin-binding site of the BBP-A – BTN structure mimics the biotin-binding site of AVD [9,23]. The amino acid residues within 4 Å of BTN in the complex with BBP-A (Asn12, Leu14, Ser16, Tyr33, Thr35, Val37, Thr38, Ala39, Thr40, Trp71, Phe73, Ala74, Ser76, Thr78, Phe80, Trp98, Leu100, Glu102, Asn119 and Trp111) are not only highly conserved but also have conformations similar to those seen in AVD (Figure 6a). Actually, only one polar (Glu102) and one hydrophobic residue (Ala74) within 4 Å of BTN are not conserved between the BBP-A – BTN and AVD-BTN complex [PDB: 2AVI] structures. The side-chain oxygen of Glu102 of BBP-A is hydrogen bonded to one of the valeric oxygen atoms of BTN, whereas in AVD the equivalent residue, Ser101, can not form a hydrogen bond to the valeric oxygen because the distance between the two atoms is too long (5.8 Å). The methyl carbon of Ala74 of BBP-A, on the other hand, is positioned only 3.7 Å away from the valeric O10B atom of BTN (numbering according to [26]), but is not likely to be hydrogen bonded with BTN unlike the side-chain oxygen atom of the corresponding Ser73 in AVD, which is hydrogen bonded to BTN.

In addition to the BBP-A – BTN complex structure, we solved the X-ray structure of BBP-A – BSO, too. To our knowledge, BSO is not found in any other protein structure presently within the Protein Data Bank (PDB) [27]. Electron density for the BSO ligand was clearly seen in the difference map of the BBP-A – BSO structure (Figure 7). The binding mode of BSO to BBP-A is surprisingly similar to that of BTN (Figure 6b) despite the additional oxygen atom that is covalently linked to the sulphur atom of the bicyclic ring system of BSO but is missing from BTN. The distance between the sulfur atom of BTN and the OG1 atom of Thr78 (3.4 Å), for example, is very close to the BSO – Thr78 distance (3.6 Å). In the known AVD-BTN

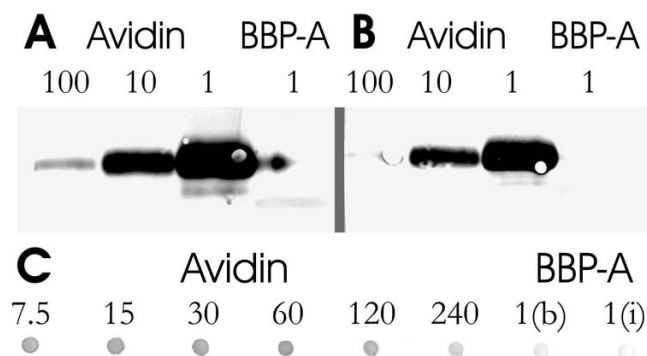


Figure 4
Immunological cross-reactivity of chicken BBP-A and AVD. (a) Western blot analysis of 1/100, 1/10 and 1/1 diluted samples of chicken AVD (10 µg) and undiluted bBBP-A protein (10 µg). The polyclonal antibody TdaVII against AVD [51] was used as the primary antibody. (b) The same analysis as in A using another polyclonal antibody (University of Oulu, Finland). (c) Dot-blot analysis of BBP-A (10 µg) produced in bacteria 'b' and in insect cells 'i'. The polyclonal antibody TdaVII against AVD was used as the primary antibody. Serial dilutions of AVD (10 µg) were used as controls (the dilution factors are indicated in the figure).

complex structures [PDB: [2AVI](#) and [1AVD](#)], the corresponding distances are 3.2 Å and 3.8 Å. The only noticeable difference around the ligands in the two BBP-A structures was found at Glu102, which in the BBP-A – BTN structure is hydrogen bonded to the valeric O10A atom of BTN but in the BBP-A – BSO structure was observed either to form a salt bridge with Arg115 (chains B and D) or to hydrogen bond to a structural water molecule (chains A and C). The difference in the conformation of Glu102 is not likely to result from the varying crystal contacts of the two BBP-A structures that crystallized as different space groups (data not shown). Moreover, the overall conformations of the BTN and BSO ligands bound to BBP-A were very similar, too.

Table 2: Dissociation of BTN and its oxidized forms from BBP-A and AVD. Fluorescence spectroscopy and radiobiotin dissociation analysis data at 40°C are shown. Binding enthalpies were measured by ITC at 25°C. bBBP-A and commercial chicken AVD was used in the analyses.

Ligand	$k_{diss} \times 10^{-4} s^{-1}$	Enthalpy (kcal/mol) (ITC)		T_m (°C) (DSC)	
		BBP-A	AVD	BBP-A	AVD
BSO	1.3 ^a	-29.3 ± 0.1	-25.9 ± 0.1	116.9 ± 0.2	100.4 ± 0.2
D-biotin sulfone	3.3 ^a	-25.1 ± 0.1	-21.4 ± 0.1	117.0 ± 0.0	101.4 ± 0.2
BTN	5.4 ^b (4.4 ^c)	-22.6 ± 0.1	-20.5 ± 0.1	117.0 ± 0.7	103.4 ± 0.1

^aFrom fluorescence spectroscopic analysis, excess of BTN was used as a competitive ligand.

^bFrom fluorescence spectroscopic analysis, excess of BSO was used as a competitive ligand.

^cValue obtained from [³H]-biotin dissociation analysis, excess of BTN was used as a competitive ligand.

Subunit interfaces of BBP-A

Comparing BBP-A with AVD, we saw differences at subunit interfaces 1–3 and 1–4, whereas the amino acid residues found at the 1–2 interface were conserved and had similar conformations (numbering of subunit interfaces according to [9]).

Of the three residues (Val116, Gly117 and Thr118) within 5 Å from each other at the 1–3 subunit interface of the BBP-A – BSO structure (Figure 8a), valine and glycine are conserved between BBP-A and AVD, while Thr118 is replaced by Ile117 in the AVD structure. In BBP-A there is no interaction equivalent to the hydrophobic or van der Waals interactions introduced by the two Met96 residues of AVD at the 1–3 subunit interface: in BBP-A, Met96 is replaced by Ala97 and the distance between Ala97 in the two subunits is about 11 Å. Generally, the 1–3 subunit interface seems to be more loosely packed in BBP-A compared to AVD, AVR4 [25] and even AVR2 [24] (data not shown).

The total contact surface area of the 1–4 interface of BBP-A – BSO is 2428 Å², whereas the corresponding contact surface area in AVD is 2319 Å². Twenty-six of the 49 amino acid residues found at the 1–4 interface within a contact radius of 4 Å are different between the BBP-A – BSO and AVD-BTN [PDB: [2AVI](#)] structures. Hydrophobic residues located at the core of the 1–4 subunit interfaces are well conserved in BBP-A and AVD, whereas several differences among the polar residues were observed, especially at the interacting loop regions (Figure 8b,c). The hydrogen bond found in AVD between the OE2 atom of Glu28 of the β3 sheet and the NE2 nitrogen of His50 of the β4 sheet, for example, is not seen in BBP-A where Glu28 and His50 are respectively replaced by Thr28 and Lys50. The hydrogen bond formed between the side-chain atoms of Gln61 (L4,5 loop) and Glu103 (L7,8 loop) in BBP-A is, in turn, missing from AVD, where the residues are replaced by Thr60 and Ser102. Yet another example of the differences at the 1–4 subunit region between BBP-A and AVD was seen for the only α-helix and the β7-strand

Table 3: Data collection and structure determination statistics for BBP-A

Data collection ^a	BBP-A – BTN	BBP-A – BSO
Wavelength (Å)	1.063	1.063
Beamline	I711 (MAX-lab)	I711 (MAX-lab)
Detector	MarCCD 165	MarCCD 165
Resolution (Å)	25-2.1 (2.2-2.1)	25-1.75 (1.85-1.75)
Unique observations	10571 (1351)	26447 (3993)
I/sigma	14.06 (4.31)	12.78 (3.24)
R _{factor} (%) ^b	10.5 (48.6)	7.7 (48.2)
Completeness	99.9 (100)	99.9 (100)
Redundancy	8.6 (8.7)	4.8 (4.8)
Refinement		
Space group	I4 ₁ 22	P2 ₁ 2 ₁ 2
Unit cell:		
a, b, c (Å)	61.7, 61.7, 179.7	79.7, 56.0, 57.1
α, β, γ (°)	90, 90, 90	90, 90, 90
Monomers (asymmetric unit)	1	2
Resolution (Å)	25-2.1	25-1.75
R _{work} (%)	19.5	19.3
R _{free} (%)	22.0	23.3
Protein atoms	969	1965
Heterogen atoms	22	34
Solvent atoms	61	212
R.m.s.d:		
Bond lengths (Å)	0.014	0.014
Bond angles (°)	1.6	1.6
Ramachandran plot:		
Residues in most favored regions	93.3	93.3
Residues in additional allowed regions	6.7	6.7
Residues in generously allowed regions	0	0
Residues in disallowed regions	0	0

^aThe numbers in parenthesis refer to the highest resolution bin.

^bObserved R-factor from XDS [53].

of these structures: in BBP-A the side-chain oxygen atom of Glu95 can form hydrogen bonds with the NE and NH2 atoms of the side chain of Arg107, whereas the atoms of the respective residues, Lys94 and Ile106, of AVD may introduce only van der Waals contacts. These and other differences (Figure 8b,c) reflect the varying architecture of the 1–4 interfaces in BBP-A and AVD.

Ligand-binding analyses

The dissociation rate of [³H]biotin from BBP-A was measured using a competitive dissociation assay (Figure 9; Table 2). The analysis showed that bBBP-A has a dissociation rate about 100-fold higher than that of AVD [28], indicating that the affinity for BTN is lower in bBBP-A compared to AVD, but similar to that of streptavidin [29]. The dissociation rate of the fluorescent BTN conjugate ArcDia™ BF560-biotin was, in turn, about 30-fold higher for bBBP-A compared to wtAVD (Table 1). Interestingly, iBBP-A had an almost three-fold slower dissociation rate compared to bBBP-A. WtAVD showed a higher dissociation rate compared to the bAVD form produced in *E. coli*,

which is in line with previously reported observations [30]. The A74S and T118F mutations of BBP-A did not significantly affect the biotin-binding properties of BBP-A (Table 1, Figure 9).

The biotin-binding characteristics of bBBP-A and wtAVD studied by ITC analysis were similar for both proteins. Because the binding of BTN was tight ($K_d < 10^{-9}$ M) in both cases, the binding constants could not be determined directly from the ITC data (Figure 10). The binding enthalpy could, however, be accurately determined, and the obtained values are listed in Table 2. In addition to BTN, the binding of BSO and D-biotin sulfone to wtAVD and bBBP-A were analysed using ITC (Table 2). The oxidised BTN forms exhibited a slightly higher binding enthalpy compared to BTN for both proteins studied. Overall, ligand binding to AVD released more heat, but a rather similar variation in the binding enthalpies of the studied ligands was observed in BBP-A and AVD. Therefore, the ITC analysis of the enthalpy of binding does not suggest different ligand preferences for BBP-A and AVD.

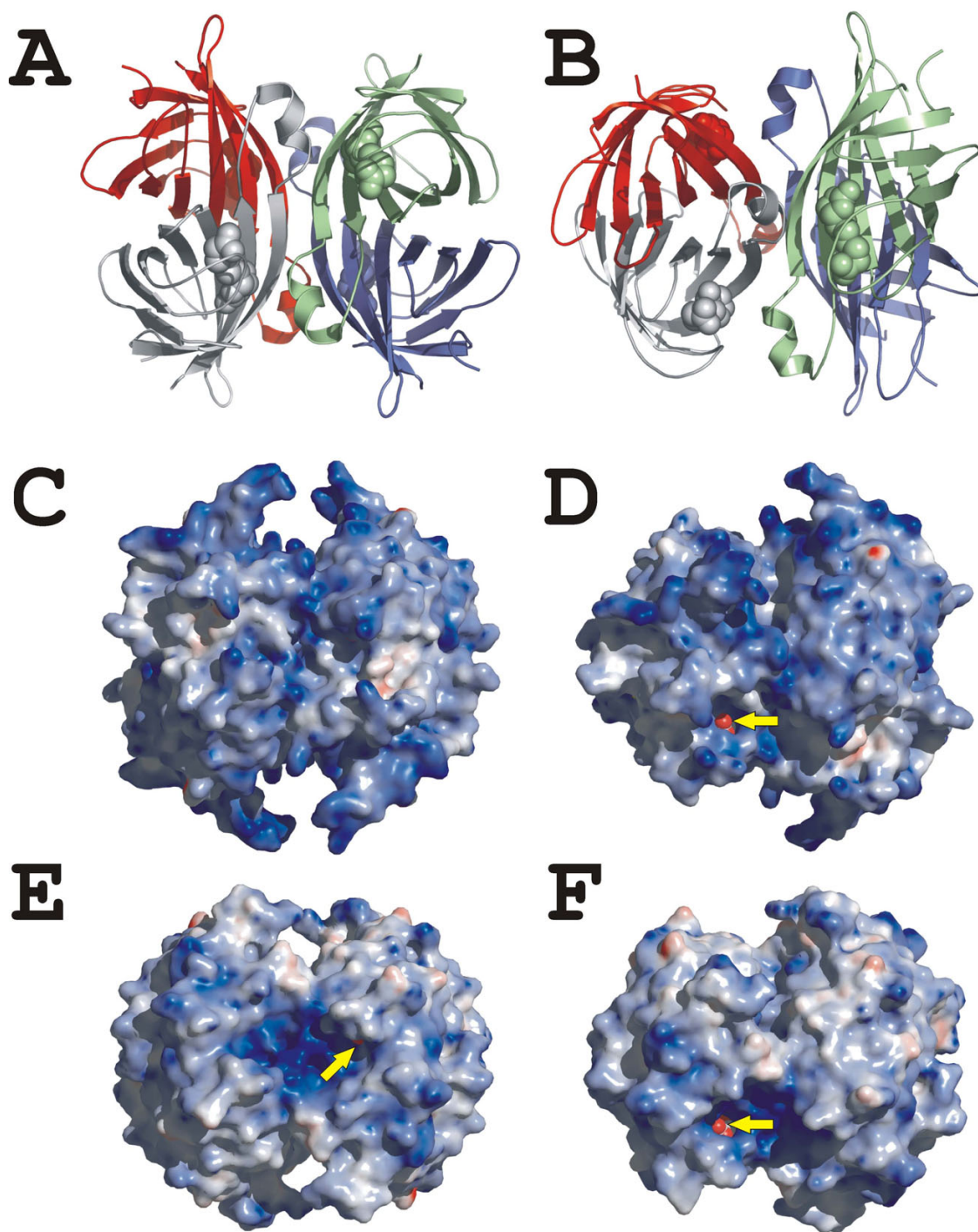


Figure 5

The overall X-ray structure of chicken BBP-A. (a, b) A tetrameric ribbon model of the BBP-A – BSO structure. (c, d) Electrostatic potentials mapped onto the molecular surface of the BBP-A – BSO structure. (e, f) Electrostatic potentials mapped onto the molecular surface of the AVD [PDB: [2AVI](#)] [23] structure after superimposition on the BBP-A structure. The views shown in (a), (c) and (e) are rotated 45° around the X-axis in (b), (d) and (f). The BSO and BTN ligands of the BBP-A – BSO and AVD structures, respectively, are shown as spheres. The yellow arrows pinpoint visible parts of the valeric acid tails of BSO and BTN seen in (d), (e) and (f).

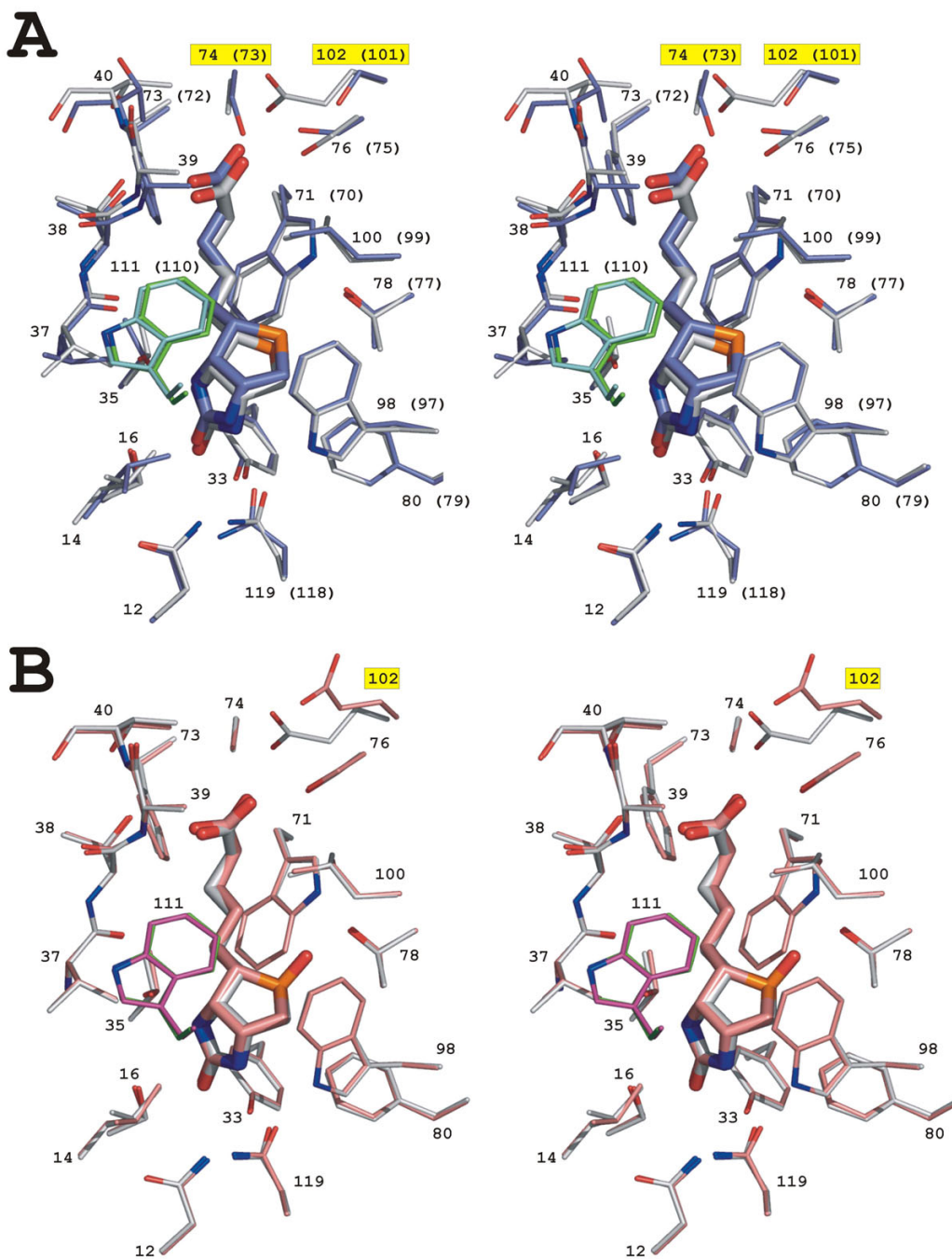


Figure 6
The binding mode of BTN and BSO. (a) BTN (thick sticks) and the amino acids within 4 Å of BTN in the BBP-A – BTN and AVD [PDB: 2AVI] structures are shown. The carbon atoms of residues from subunit I of the BBP-A and AVD structures are coloured light grey and blue, respectively. The carbon atoms of Trp111 (Trp110) from subunit 2 of BBP-A (AVD) are coloured green (cyan). The amino acids of BBP-A (AVD) are numbered. The labels of the non-conserved residues are indicated with a yellow background. (b) BTN and BSO (thick sticks), and the amino acids within 4 Å of the ligands in the BBP-A – BTN and BBP-A – BSO structures are shown. The carbon atoms of residues from subunits I of the BBP-A – BTN and BBP-A – BSO structures are coloured light grey and light red, respectively. The carbon atoms of Trp111 from subunit 2 of the BBP-A – BTN and BBP-A – BSO structure are coloured green and dark red, respectively. The amino acids are numbered. Glu102, whose side chain has a different conformation in the BBP-A – BTN and BBP-A – BSO structures, is labelled with a yellow background.

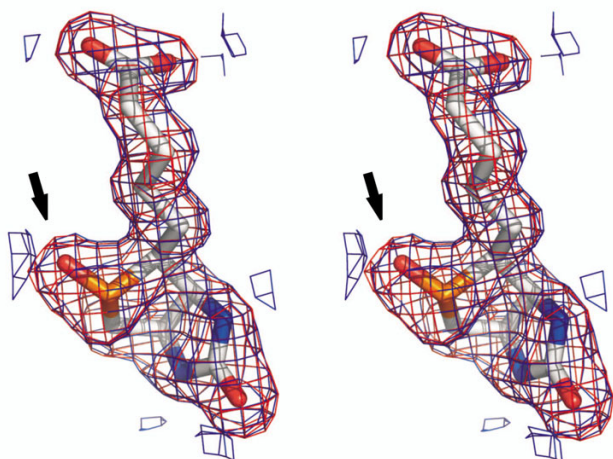


Figure 7
Stereo image of electron density around the BSO ligand and of the BBP-A – BSO structure. The additional oxygen atom on the BSO ligand (in comparison to BTN) is indicated by an arrow. The 2Fo-Fc (blue) and Fo – Fc (red) maps calculated without the ligand but drawn around BSO of the final model are shown at the level of 1.0 and 3.0 sigma, respectively.

The ligand-induced changes to the emission spectra of bBBP-A and wtAVD were studied, too (Figure 11). All biotin forms, BTN, BSO and D-biotin sulfone, enhanced the emission intensity of BBP-A. In contrast, only D-biotin sulfone increased the emission intensity of AVD, whereas BSO and especially BTN decreased the emission intensity of AVD. Compared to the ligand-free proteins, for each of the ligand-protein combinations the maximum in the emission spectra moved towards the shorter blue wavelengths. Altogether, these results indicate differences in the ligand-protein interactions between BBP-A and AVD, particularly reflecting variation in the local environment of tryptophan residues [31].

EST database search

Chicken EST databases were searched using NCBI blastn [32] and the BBP-A and AVD cDNAs as query sequences. The search resulted in 9 significant (E-value < 1×10^{-100} and score > 400) hits in the case of BBP-A, whereas the cDNA of AVD yielded 94 hits. Further analysis of the BBP-A hits showed that they correspond to mRNAs isolated from kidney and the adrenal glands (3 hits), the trunks of chicken stage 36 embryos (4 hits) and multiple-tissue preparations (2 hits). The AVD hits, in turn, corresponded to mRNAs isolated from lymphoid tissues (1), intestine (1), ovary (2), the chondrocytes of cartilage (6), multiple-tissue preparations (4), splenic T cells (2), pituitary gland, hypothalamus and pineal gland (1), and PBL macrophages (77).

Discussion

Here, we report the biochemical and structural characterization of BBP-A, produced efficiently both in *E. coli* and *Spodoptera frugiperda* cells. This study demonstrates that BBP-A can be classified as a new member of the AVD family: BBP-A binds BTN with high affinity like all the other known members of the family [4,24,33] and has the β -barrel fold characteristic of AVD and of the more heterogeneous calycin superfamily of proteins [34]. The binding mode of BTN to BBP-A (Figure 6) is also highly similar to that observed in the known structures of AVD [9,23], AVR2 [24] and AVR4 [25]. The BBP-A protein is, however, biochemically and structurally clearly distinguishable from chicken AVD and the AVRs [1,4,18,24,25]. There are also differences in the immunological properties of BBP-A and AVD, since immunoblot and dot-blot analyses with polyclonal AVD antibodies did not show cross-reactivity between BBP-A and AVD (Figure 4).

Does BBP-A represent the earlier reported BBPs, BBP-I or BBP-II? Based on the biochemical characterization of BBP-I [11,35-37], and BBP-II [14,15], this seems not to be the case. The [^3H]-biotin dissociation rate constant of BBP-I reported in 1978 by Meslar and co-workers [11] is 2–4 times higher than that determined for bBBP-A and even 5–12 times higher in comparison to the glycosylated iBBP-A, which has a dissociation rate of [^3H]biotin (Figure 9) in the same range as reported for streptavidin [29]. The pI (4.6) of the BBP-I protein differs from that calculated for BBP-A (pI = 9.75), too. Moreover, the reported N-terminal sequences of BBP-I and BBP-II do not match the sequence of BBP-A [18]. In conclusion, the BBP-A protein reported here seems to be a novel protein not described in any previous publications.

In order to study the molecular details of ligand recognition by BBP-A and compare the ligand-binding properties of BBP-A with the previously determined structures of AVD [9], streptavidin [8], AVR4 [25] and AVR2 [24], we crystallized bBBP-A in complex with BTN. Surprisingly, in one of the two BBP-A crystals that were analyzed, BSO was found bound to BBP-A even though only BTN was added in the co-crystallisation experiments. To our knowledge, this complex is unique, as no other protein structure has been reported in complex with BSO. What can be the source of the bound BSO? Based on the crystallization conditions (see Methods), it is not easy to say why one structure bound BTN and the other BSO. It is known that some BTN preparations may carry minor amounts of BSO [38], but based on mass spectrometry analysis (data not shown) no detectable amounts of BSO were found in the diluted BTN solution that was used for crystallization. One possibility is that the source of the BSO ligand seen in the BBP-A – BSO structure is BTN that had undergone oxidation during crystallization. Another possibility is

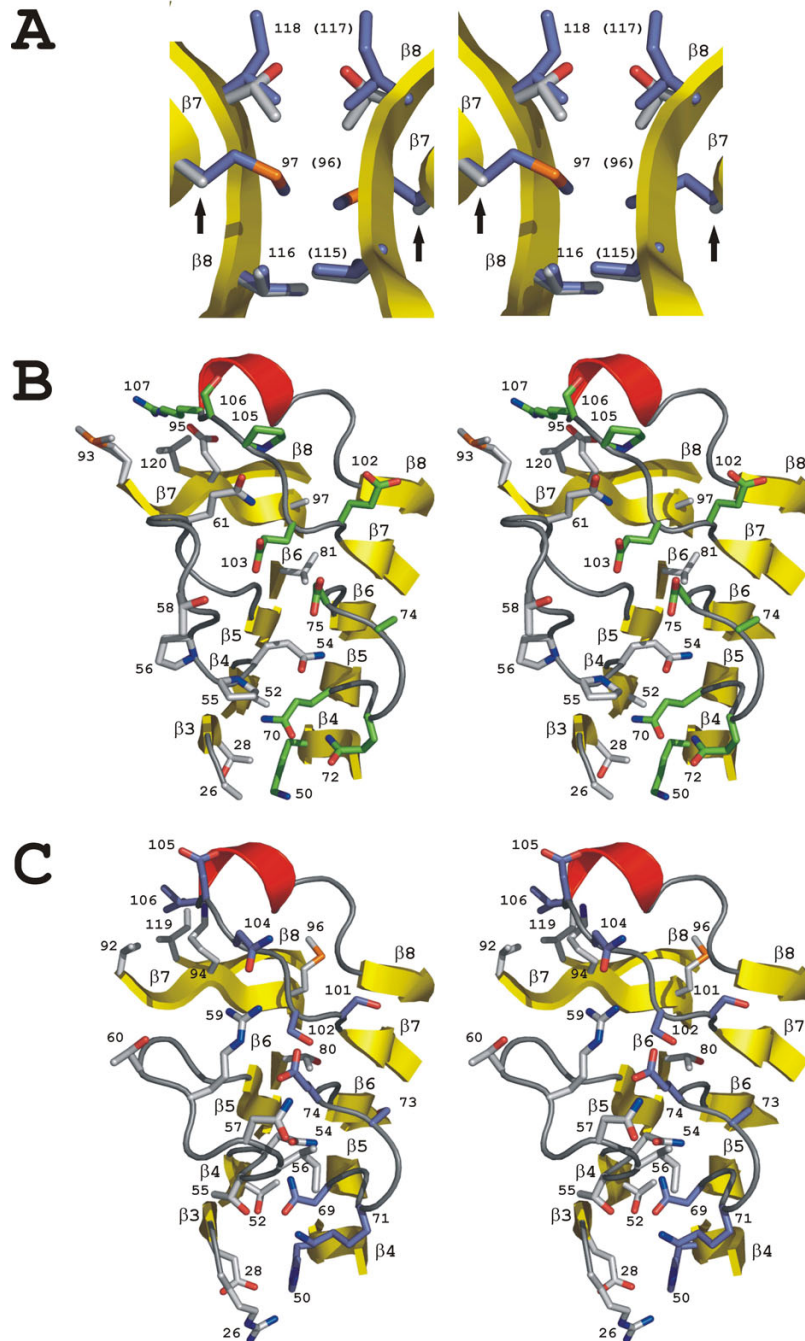


Figure 8

Stereo images of the I–3 and I–4 subunit interfaces of BBP-A and AVD. (a) The I–3 subunit interface of the AVD [PDB: [1VYQ](#)] [68] structure superimposed with the corresponding interface of the BBP-A – BSO structure. The carbon atoms of amino acids from BBP-A and AVD are coloured light grey and blue, respectively. Amino acids and β -sheets are numbered. Two occurrences of Ala97 in the BBP-A structure are indicated by arrows. (b, c) A simplified view of the I–4 interface of the BBP-A – BSO (b) and AVD [PDB: [2AVI](#)] [9] (c) structure after superimposition. The atoms of the non-conserved side chains of amino acids within 4 Å of subunit I of either the BBP-A or AVD structure are shown. For clarity, Met96 and Thr110 of BBP-A as well as the equivalent residues of AVD (Thr95 and Asp109) are not shown, and only one half of the symmetrical I–4 interface is shown. The carbon atoms of residues from subunit I and 4 are respectively coloured light grey and green in BBP-A, and grey and blue in AVD. Amino acids and β -sheets are numbered.

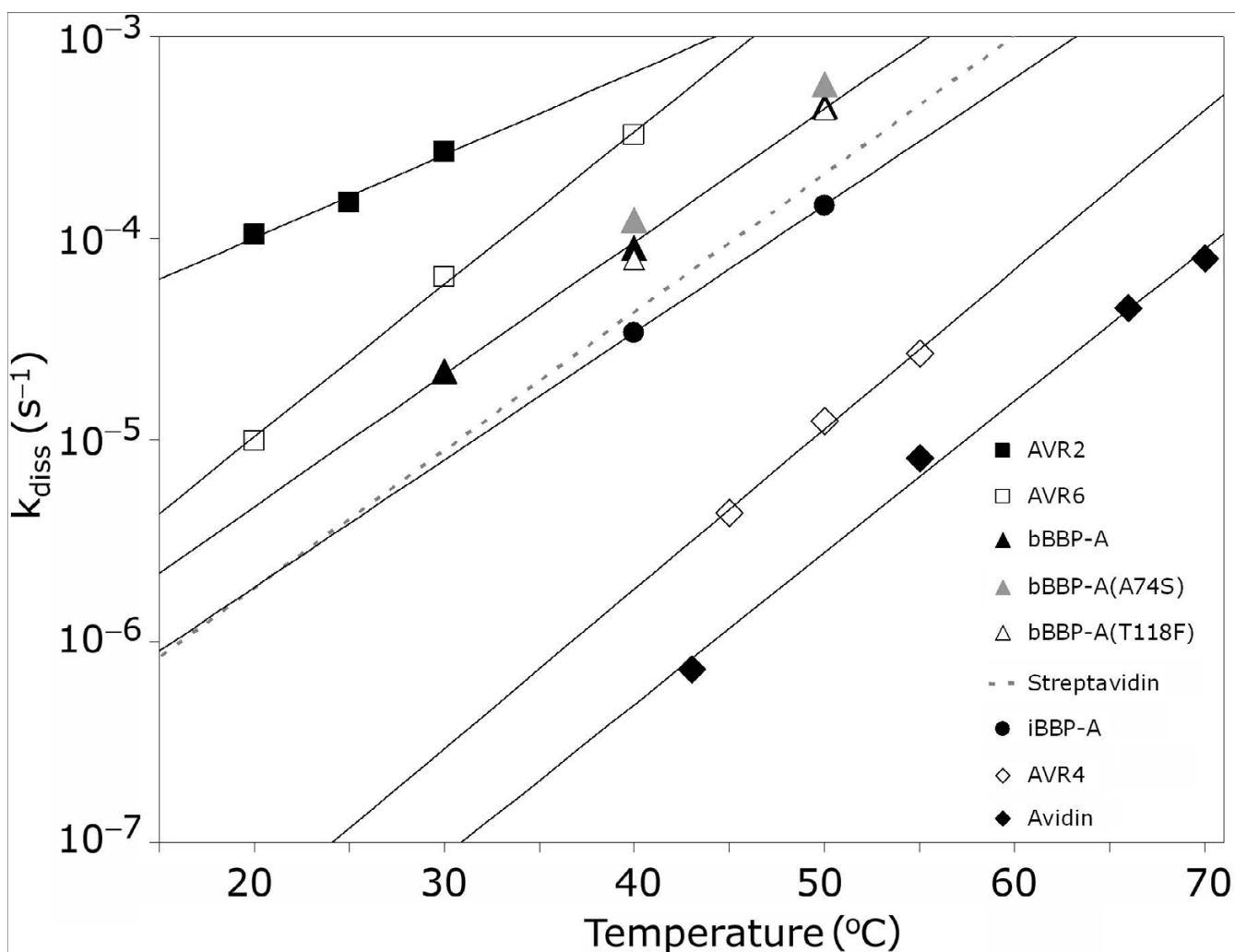


Figure 9

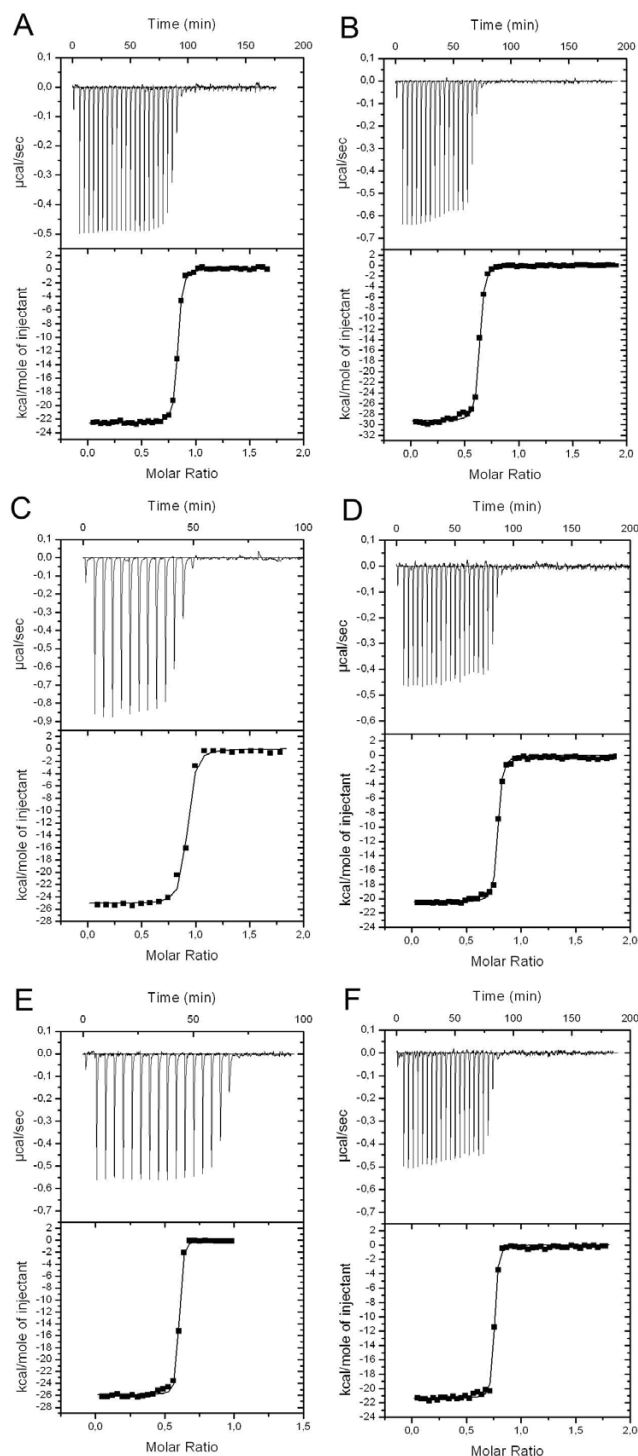
Biotin dissociation analysis. The dissociation of [^3H]biotin from various proteins was determined over time. The individual dissociation rates of [^3H]biotin measured for bBBP-A (black triangles), bBBP-A(A74S) (grey triangles), bBBP-A(T118F) (white triangles) and iBBP-A (black circles) are shown. Solid lines represent linear fit to data sets. Data measured previously for AVR2 (black rectangles), AVR6 (white rectangles) and AVR4 (white tilted rectangles), all produced in bacteria, are also shown [24, 28]. Data measured for chicken AVD (tilted black rectangles) are presented, too [42]. AVR6 and AVR4 carried the mutations C58S and C122S, respectively, which prevented the formation of intersubunit disulphide bridges [33]. The dissociation rates for streptavidin are shown as a dotted line according to [29].

that BBP-A had itself converted BTN to BSO by some yet unknown catalytic mechanism, but experimental proof for this hypothesis is lacking and the final explanation for the presence of BSO in one of the BBP-A structures remains to be studied.

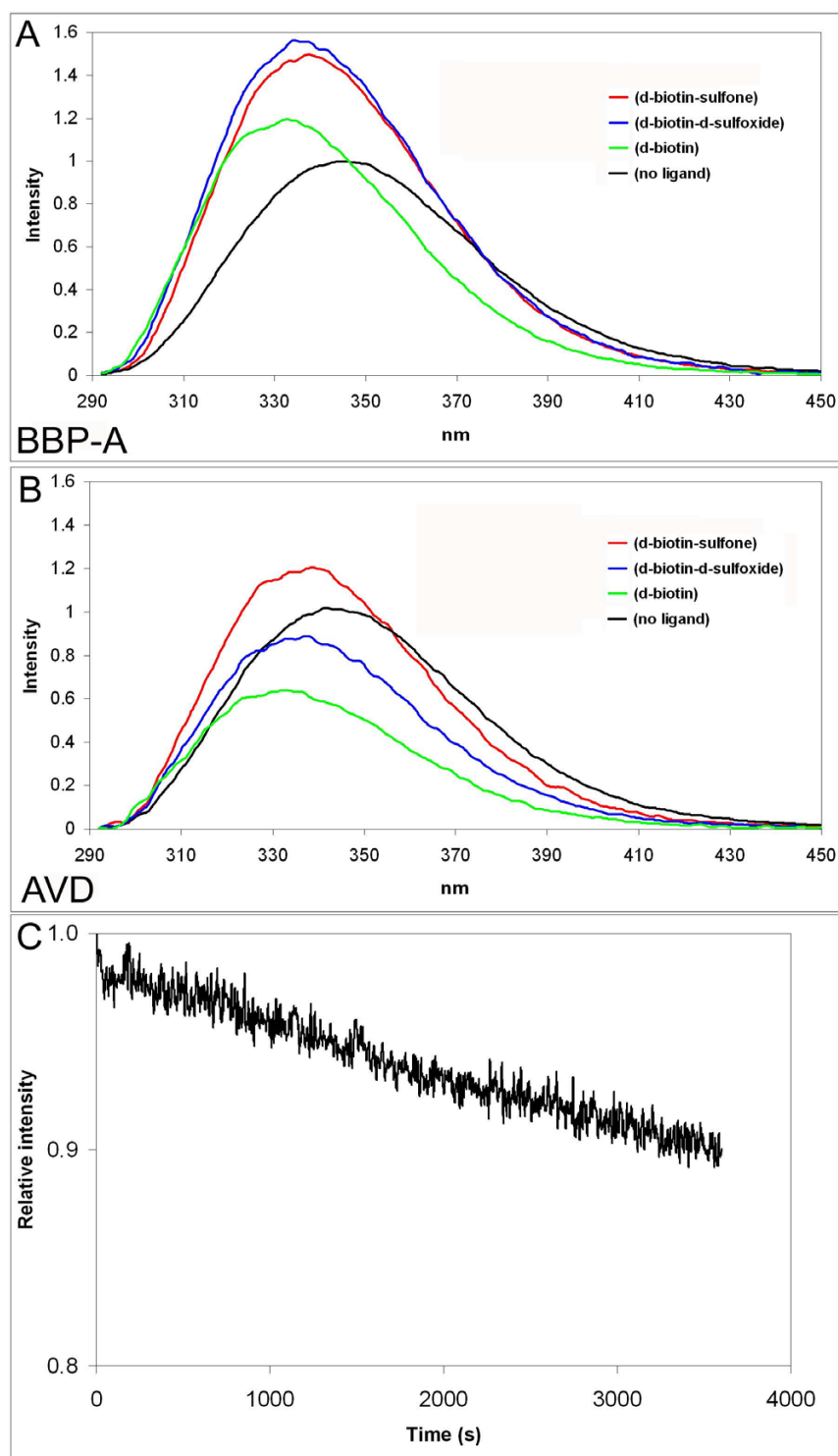
The BBP-A – BSO structure inspired us to investigate whether BBP-A recognizes BSO with altered affinity in comparison to BTN. We analysed the binding of the fully oxidised sulfone form of BTN, too. To our surprise, the dissociation rate of BSO (fluorometric analysis at 40°C,

Table 2) from BBP-A was similar or even slower than the dissociation rate of BTN (fluorometric and radiobiotin analysis at 40°C), whereas in the case of AVD the dissociation rate of BSO was more than 20 times higher in comparison to BTN (for AVD, the dissociation rate of BSO was determined using fluorometric analysis at 50°C (data not shown), whereas the dissociation of BTN was determined using radiobiotin assay at 50°C (Figure 9)).

This is the first reported case, to our knowledge, where the product of a naturally occurring gene encoding a BTN

**Figure 10**

Isothermal titration calorimetry analysis of bBBP-A and AVD. Binding thermograms measured for (a) AVD-BTN, (b) AVD-BSO, (c) AVD – D-biotin sulfone, (d) bBBP-A – BTN, (e) bBBP-A – BSO and (f) bBBP-A – D-biotin sulfone complexes are shown at the upper part of the graphs. Non-linear least square curves (at the lower part of each graph) were fitted to enthalpies integrated from the individual titrations. For each titration, 0.5 mM ligand solution was used. The reaction volume was 10 μ l except for the first titration, which was done in 2 μ l.

**Figure 11**

Fluorescence spectroscopy analysis of BBP-A and AVD. (a) Emission spectra of BBP-A in the absence of ligand (black) and with BTN (green), BSO (blue) or D-biotin sulfone (red) are shown. (b) Emission spectra of AVD. (c) Dissociation of D-biotin sulfone from BBP-A displaced with excess of BTN at 40°C. The measured emission intensity is plotted over time and corrected using the emission decay data from control measurements, which were made for protein saturated with BTN. Excitation at 280 nm was used.

binder shows equally high affinity to a ligand other than BTN. Based on DSC analysis the studied ligands did not differ significantly in their effect on the thermal stability of BBP-A and AVD (Table 2). Nor did the measured binding enthalpies vary significantly over ligand sets between AVD and BBP-A (Table 2). Overall, these ligand binding analyses suggest that all the studied ligands are efficiently associated (high negative binding enthalpy) with BBP-A and AVD. However, the determined dissociation rate of BTN from AVD was significantly lower than the dissociation rates of BSO and D-biotin sulfone, whereas the observed dissociation rates of the same ligands from BBP-A were rather similar to each other, the slowest observed for BSO. As a conclusion, it is the different dissociation rather than association rates that determines the ligand-binding preferences of BBP-A and AVD; out of the three ligands studied, AVD prefers BTN, whereas BBP-A seems to prefer BSO. Moreover, it is well known that the extremely slow dissociation of BTN from AVD is the main determinant of the high affinity binding [1].

One possible explanation for the differences in the dissociation rates of the studied ligands is the weaker structural stability of BBP-A as compared to AVD (Table 1). It has been experimentally shown, that binding of BTN to streptavidin lowers the rate of H/D exchange in large parts of the structure [10] at least partially due to the positive structural cooperativity in the binding process (thoroughly explained in [39]), *i.e.* improved packing and compactness of not only the ligand-binding site but of entire subunits, too. The weaker stability, or lower level of "compactness", of the entire BBP-A tetramer can therefore explain the differences in the observed ligand-binding properties of AVD and BBP-A. The BBP-A barrel fold may also be considered more flexible or dynamic than the AVD barrel, which could lead to a higher rate at which BTN is dislodged from the binding pocket, possibly by a mechanism similar to that presented by Hyre *et al.* for streptavidin [40].

What are the structural determinants that specify the recognition of BSO and why does the affinity of BSO meet that of BTN for BBP-A? The BBP-A – BSO structure is strikingly similar to the BBP-A – BTN structure – even the additional D-sulfoxide moiety of the ligand in the BBP-A – BSO structure does not seem to alter the binding pocket of BBP-A as compared to the BBP-A – BTN structure (Figure 6). One clear difference can, however, be detected: the hydrogen bond present in the BBP-A – BTN complex structure between Glu102 and BTN is missing in the BBP-A – BSO structure due to the different conformation of Glu102 in the two structures. Moreover, in the BBP-A – BSO structure, the carboxylate side-chain of Glu102 forms a salt-bridge with the guanidinium group of the Arg115 side-chain within the same subunit (in chains B and D) or

is hydrogen bonded to a structural water molecule (in chains A and C), interactions that are not seen in the BBP-A – BTN complex. Arg115 is one of the conserved amino acids found at the 1–2 subunit interfaces of all known structures of members of the AVD family, but its conformation varies within the family (data not shown). This suggests that, in addition to the residues at the ligand-binding pocket, the type and conformation of residues at the subunit interfaces may regulate ligand binding, *e.g.* by allowing or disabling alternative hydrogen bonding networks.

Conclusion

Our present study reveals that chicken has an avidin-like protein, BBP-A in its biotin-binding repertoire. Even though similar to AVD according to its three-dimensional structure and biotin-binding properties, *i.e.* the high affinity for BTN, BBP-A has several unique features – the thermal stability and immunological cross-reactivity of BBP-A is very different as compared to AVD. Furthermore, BBP-A is the first example of a protein from the AVD family that naturally has equally high affinity for BTN and for a non-BTN ligand. The biological function of BBP-A is not known, but since it binds both BSO and BTN with high affinity, it may have a function in the storage or delivery of BTN to an embryo as has been suggested in earlier studies for BBP-I and BBP-II [14,17]. In addition to providing clues of biological function and information about biotin-binding determinants in proteins, the high-resolution structure and biochemical analysis of BBP-A makes this novel protein attractive material for protein engineering and bio(nano)technological applications. For example, BBP-A could be covalently combined with AVD or AVR polypeptide using a circular permutation strategy [41,42] or used as a scaffold in the development of artificial enzymes [43].

Methods

Preparation of expression constructs

A cDNA clone of BBP-A [GenBank: [BX930135](#)] was obtained from the UK Chicken EST Consortium (ARK-Genomics). For insect cell production the cDNA, which included the sequence encoding for a putative signal peptide [18], was PCR amplified and cloned into a pFASTBAC1 vector (Invitrogen). The following primers were used: 5'BBP-A+signal (5'-AAAAGATCTATGGAG-CACCTCCGCTG) and 3'BBP-A (5'-ATTTAAGCTTACTT-GACACG GGTG), in which the restriction enzyme cleavage sites for BglII and HindIII, respectively, are shown in italics. In order to express BBP-A in *E. coli*, the original signal sequence of BBP-A was replaced with the OmpA signal peptide from *Bordetella avium* [44] by cloning into the pGemTeasy vector (Promega) containing the OmpA-signal and flanking attL recombination cloning sequences [30]. The construct was then transferred to the

pBVboostFG expression vector [45,46] using LR recombination (Invitrogen). 5'BBP-A-core (5'-AAAGGTACCAG-GAAGTGCAGC; KpnI) and 3'BBP-A were used as the primers for PCR. The nucleotide sequences of the final expression vectors were confirmed by DNA sequencing.

The mutagenesis of BBP-A cloned in the pGemTeasy plasmid was performed using the QuikChange mutagenesis method (Stratagene, La Jolla, CA, USA). The BBP-A mutants were transferred to the pBVboostFG expression vector essentially as described previously [46].

Protein expression and purification

BBP-A and its mutated forms were produced in *E. coli* BL-21(AI) cells (Invitrogen) as previously described [30]. The mutation A74S was created in order to study the significance of this residue for BTN binding (in AVD the side chain of the equivalent Ser73 forms a hydrogen bond with BTN). T118F, in turn, was created in order to increase the subunit contact area and consequently the stability of BBP-A; in AVD, I117Y was found to increase the stability of the protein [28]. iBBP-A, employing its natural signal peptide, was expressed in the eukaryotic host, *Spodoptera frugiperda*, using the Bac-to-Bac baculovirus expression system (Invitrogen). The proteins were isolated using 2-aminobiotin affinity chromatography (Affiland S. A., Liege, Belgium) as described elsewhere [30,47]. AVD isolated from chicken egg-white (wtAVD; Belovo S. A., Bastogne, Belgium) and AVD produced in *E. coli* (bAVD) [30] were used as control proteins throughout this study.

Chemical synthesis of BSO and D-biotin sulfone

The syntheses of BSO and D-biotin sulfone were performed using a procedure described by Melville [48]. The melting points of the BTN derivatives were 201–202 °C (lit. 200–203 °C [48]) for BSO and 275–277 °C (lit. 274–275 °C [49]) for D-biotin sulfone. From ESI-MS analysis (see below) the following values were obtained: BSO, calculated molecular weight ($C_{10}H_{16}N_2O_4S_1$) = 260.31, [M-H]⁻ m/z = 259.0753 and measured m/z = 259.0388; D-biotin sulfone, calculated molecular weight ($C_{10}H_{16}N_2O_5S_1$) = 276.31, [M-H]⁻ m/z = 275.0702 and measured m/z = 275.0184.

Stability analysis

The stability of BBP-A was analysed by SDS-PAGE as described previously [22]. Prior to analysis, the protein sample was acetylated *in vitro* and subsequently subjected to thermal treatment for 20 min in the presence of SDS and 2-mercaptoethanol. The oligomeric state of the treated protein was assessed by SDS-PAGE using Bio-Safe Coomassie (Bio-Rad) staining.

Differential scanning calorimetry

Differential scanning analysis was performed in 50 mM NaPO₄ buffer (pH 7.0) containing 100 mM NaCl as previously described [33]. The concentration of the analyzed proteins was between 0.2 and 0.6 mg/ml. In samples containing a ligand, the molar ligand concentration was three times as high as the protein subunit concentration. The samples were scanned from 25 to 130 °C at a rate of 0.92 °C/min.

Gel filtration chromatography

Gel filtration analysis was performed with a ÄKTA™ purifier HPLC instrument (Amersham Biosciences) equipped with Superdex 200 10/300 GL column (Tricorn) as previously described [50]. A buffer containing 50 mM NaPO₄ and 650 mM NaCl (pH 7.0) was used as the liquid phase. Biotin-complexed proteins were prepared by incubating the sample in the presence of 0.22 mM BTN 15 min prior to analysis.

Mass spectrometry

The mass spectrometric studies were performed with a Micromass LCT ESI-TOF instrument equipped with a Z geometry electrospray ion source TOF detector. The analysis of BBP-A was performed as previously described [30]. Before analysis, the sample was dialysed against distilled water and lyophilised. MS analysis of the BTN used in the crystallization experiments with BBP-A was performed to determine whether BSO was present in the ligand solution. The amount of BSO in the sample was below the detection limit of the method (less than 1%). The detection limit was determined by making mixtures of 10 µg/ml (41 µM) of BTN with 1 (3.8 µM), 0.5 (1.9 µM) or 0.1 µg/ml (0.38 µM) of BSO.

Ligand-binding analyses

The dissociation rate of [³H]biotin (Amersham) was measured at various temperatures with a competition assay as described previously [29]. Measurements were carried out in 50 mM NaPO₄ buffer containing 100 mM NaCl.

The dissociation rate of the fluorescent biotin conjugate ArcDia™ BF560-biotin (ArcDia Ltd., Turku, Finland) was measured as reported previously [30]. The measurements were made at 25 °C in 50 mM NaPO₄ buffer (pH 7.0) containing 650 mM NaCl using a PerkinElmer LS55 luminescence meter.

The energetics of biotin binding was determined with a VP-ITC (MicroCal™) Isothermal Titration Calorimeter. Measurements were performed at 25 °C in degassed 50 mM NaPO₄ buffer (pH 7.0) containing 100 mM NaCl. In order to calculate the enthalpies of binding (ΔH) for BBP-A and AVD, the cell used for measurements was filled with

30 μM protein solution. BTN, BSO or D-biotin sulfone (0.5 mM) was then added to the measurement cell using 15 equal volume injections (10 μl) and at 240 second intervals. The data were analysed with Origin 7.0 software using the "One Set of Sites" method. The titration curve (heat change $\mu\text{cal}/\text{injection}$) resulting from the 15 injections was analysed by fitting the data to a nonlinear least square curve. We only determined the enthalpy of binding from these experiments, since the estimation of binding constants directly from the data was impossible due to the tight binding.

In order to study the intrinsic fluorescence of the proteins, the emission spectra of bBBP-A and wtAVD in 50 mM NaPO_4 buffer (pH 7.0) containing 650 mM NaCl were measured using a PerkinElmer LS55 spectrofluorometer and excitation at 280 nm (slit 2.5 nm). During the analysis, protein solutions (100 nM) were continuously mixed using an integrated magnetic stirrer and maintained at 25°C using a circulating water bath. The emission spectra were also measured after addition of 200 nM ligand (BTN or its oxidised forms) to the protein solutions.

Fluorescence spectroscopy was also used to measure the rate of protein-ligand dissociation. These experiments were made at 40°C in order to measure the dissociation events within an experimentally applicable timescale. Firstly, the emission intensity of bBBP-A or wtAVD (50 nM) was measured at 350 nm (slit 10 nm) in the presence of BTN, BSO or D-biotin sulfone (100 nM). Secondly, in order to detect and quantify the dissociation events, a 1000-fold molar excess (100 μM) of BTN (BSO in case of determination of BTN dissociation) was added to the samples and the measurements were recorded for 3600 seconds. The measured spectral properties of each protein-ligand complex (*i.e.* the emission intensity of the protein-ligand complex) were used to create a single-phase dissociation model, which was fitted to the data. For example, in the case of the dissociation of BSO from BBP-A and binding of BTN to BBP-A, the dissociation process was observed as a decrease in the emission intensity (Figure 11c). The decrease in the fluorescence signal obtained in the control measurement performed in the presence of a 1000 molar excess of BTN was corrected during data analysis.

Deglycosylation analysis

The BBP-A protein produced in insect cells was treated with Endoglycosidase H (New England Biolabs) to see if it was glycosylated. Wild-type AVD was used as a control protein. Prior to treatment, the analysed proteins were denatured by boiling them in the presence of 2-mercaptoethanol and SDS. Deglycosylation was performed overnight at 37°C. The samples were then boiled and

subjected to SDS-PAGE analysis followed by staining with Bio-Safe Coomassie (Bio-Rad).

Immunological analysis

The cross-reactivity of polyclonal AVD antibodies with BBP-A was studied using immunoblot and dot-blot analyses. The polyclonal rabbit antibodies TdaVIII [51] and an AVD antibody (University of Oulu, Finland) were used (dilution 1:5000) in Western blotting to analyse a 10 μg sample of BBP-A produced in *E. coli*. Three different amounts, 0.1, 1 and 10 μg of wtAVD, were used as controls in this experiment.

Dot-blot analysis was performed using only the TdaVIII antibody and 10 μg of BBP-A produced either in *E. coli* or in insect cells. AVD from chicken was used as a control (0.04–1.33 μg).

Crystallization and diffraction data collection

Random and sparse matrix screens [52] prepared with the HamiltonSTAR robot in the Institute of Biotechnology at the University of Helsinki were initially used to search for suitable conditions for crystallization of BBP-A. Sitting drops of equal volumes (100 nl) of sample and well solution were automatically prepared by the Cartesian MicroSys robot on 96-well Greiner 3-SQ plates at 20°C. For optimization, the drop size was increased to 2 μl and crystallization was performed on conventional 24-well crystallization plates (Nextal/Hampton Research) using the vapour diffusion method and either sitting or hanging drops. In order to prepare BBP-A – BTN complexes, BTN (Sigma) diluted in buffer containing 5 mM Tris (pH 8.8) and 8 mM CHES (pH 9.5) was added to the protein samples in an approximate 1:10 molar ratio before crystallization. Two crystals were used to collect diffraction data and were obtained from conditions where 1 μl of protein solution (~ 0.4 mg/ml) containing 50 mM sodium acetate (pH 4.0) and 100 mM sodium chloride, and 1 μl of well solution containing either 2 M ammonium sulphate and 5% isopropanol (v/v) (BBP-A – BTN crystal) or 0.2 M sodium acetate, 0.1 M Tris (pH 8.6) and 30% (v/v) PEG 4000 (BBP-A – BSO crystal) were used. The diffraction data were collected at the MAX-lab beam line I711 (Lund, Sweden) at 100 K using a MarCCD detector. The BBP-A – BTN and BBP-A – BSO crystals were cryoprotected by adding 0.8 μl of 100% glycerol and 1 μl of 4 M sodium formate, respectively, to the crystallization drops just prior to flash-freezing in a 100 K liquid nitrogen stream (Oxford Cryosystem). Diffraction data were processed with programs of the XDS program package [53]. The data collection statistics are summarized in Table 3.

Structure determination

The X-ray structures of BBP-A – BTN and BBP-A – BSO were solved using the molecular replacement program

Amore [54] from the CCP4i suite [55,56]. A monomer of AVR2 [PDB: [1WBI](#)] [24] was used as a trial model in Amore to solve the BBP-A – BTN structure. The BBP-A – BSO structure, in turn, was solved using the BBP-A – BTN structure as a search model in Amore. The best solutions from molecular replacement were selected as input for automatic model building with ARP/wARP [57]. The models were refined with Refmac5 [58], and modified and rebuilt with O [59]. Solvent atoms were added to the model with an automatic procedure in ARP/wARP [60] and other non-protein atoms were built manually in O. The coordinate file of the BSO ligand was obtained from the Cambridge Structural Database (CSD version 5.26) and molecular topologies of BSO were created with PRODRG [61]. The BBP-A structures were analyzed with the programs PROCHECK [62] and WHATIF [63]. The structure determination statistics are summarized in Table 3. The coordinates and structure factors of the BBP-A – BTN and BBP-A – BSO structures have been deposited in the Protein Data Bank with entry codes 2C1Q and 2C1S, respectively.

Miscellaneous methods

Figures 2, 3, 4, 5 were created with the PyMOL Molecular Graphics System [64] and edited with the Corel Draw11 program suite. The multiple sequence alignment shown in Figure 2 was created using the program Malign implemented in BODIL [65]. Electrostatic potentials were calculated using the ABPS [66] plugin of PyMOL. The programs Contact and Areaimol [67] of the CCP4i suite were used to calculate the solvent accessible surface areas and to identify residues at the subunit interfaces, respectively.

Expressed sequence tag (EST) databases at NCBI were searched (blastn [32]) using the cDNA of BBP-A [GenBank: [BX930135](#)] and AVD [GenBank: [X05343](#)] as query sequences.

Abbreviations

AVD, avidin; bAVD, AVD produced in bacteria; BBP, biotin-binding protein; bBBP-A, BBP-A expressed in bacteria; iBBP-A, BBP-A expressed in insect cells; BTN, D-biotin; BSO, D-biotin D-sulfoxide; DSC, differential scanning calorimetry; ESI-TOF, electrospray ionization time-of-flight; HABA, 4'-hydroxyazobenzene-2-carboxylic acid; ITC, isothermal titration calorimetry; wt, wild-type

Authors' contributions

The biochemical studies of BBP-A and AVD were mainly carried out in the group of MSK at the Univ. of Jyväskylä and Tampere, whereas the structural analysis was mostly done in the groups of TAS and MSJ in the Structural Bioinformatics Laboratory, Åbo Akademi University. VPH carried out partially the protein expression and analysis experiments, JAEM and KKH carried out the calorimetric

analyses, EAN was involved in the ligand-binding analyses and protein expression, JH synthesized BSO and D-biotin sulfone and carried out the mass spectrometric analyses, and KJH expressed and purified BBP-A and participated to the protein analyses. VPH, HRN, KR, OHL and MSK designed and supervised the experimental work related to biochemical analyses. TTA designed and carried out the 3D-structure analyses of BBP-A. VPH, MSJ, MSK, OHL and TTA were mainly responsible for writing and editing the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Irene Helkala, Eila Korhonen and Mirja Lahtiperä for excellent technical assistance. We thank Professor Meir Wilchek for helpful discussions. This work was supported by grants from the Academy of Finland, the Emil Aaltonen Foundation, the Sigrid Jusélius Foundation, and the Foundation of Åbo Akademi (Center of Excellence in Cell Stress). It was also supported by ARK Therapeutics Group Plc, Kuopio, Finland, and by the National Graduate School in Informational and Structural Biology (ISB), Turku, Finland. We would like to thank the staff at the beam line I711 for excellent support. We acknowledge the support by the European Community – Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme (through the Integrated Infrastructure Initiative "Integrating Activity on Synchrotron and Free Electron Laser Science").

References

- Green NM: **Avidin**. *Adv Prot Chem* 1975, **29**:85-133.
- Botte V, Granata G: **Induction of avidin synthesis by RNA obtained from lizard oviducts**. *J Endocrinol* 1977, **73**:535-536.
- Korpela JK, Kulomaa MS, Elo HA, Tuohimaa PJ: **Biotin-binding proteins in eggs of oviparous vertebrates**. *Experientia* 1981, **37**:1065-1066.
- Laitinen OH, Hytönen VP, Ahlroth MK, Pentikäinen OT, Gallagher C, Nordlund HR, Ovod V, Marttila AT, Porkka E, Heino S, Johnson MS, Airene KJ, Kulomaa MS: **Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin**. *Biochem J* 2002, **363**:609-617.
- Green NM: **Avidin and streptavidin**. *Method Enzymol* 1990, **184**:51-67.
- Bayer EA, Kulik T, Adar R, Wilchek M: **Close similarity among streptavidin-like, biotin-binding proteins from Streptomyces**. *Biochim Biophys Acta* 1995, **1263**:60-66.
- Nordlund HR, Hytönen VP, Laitinen OH, Kulomaa MS: **Novel avidin-like protein from a root nodule symbiotic bacterium, Bradyrhizobium japonicum**. *J Biol Chem* 2005, **280**:13250-13255.
- Weber PC, Ohlendorf DH, Wendoloski JJ, Salemme FR: **Structural origins of high-affinity biotin binding to streptavidin**. *Science* 1989, **243**:85-88.
- Livnah O, Bayer EA, Wilchek M, Sussman JL: **Three-dimensional structures of avidin and the avidin-biotin complex**. *Proc Natl Acad Sci USA* 1993, **90**:5076-5080.
- Williams DH, Stephens E, Zhou M: **Ligand binding energy and catalytic efficiency from improved packing within receptors and enzymes**. *J Mol Biol* 2003, **329**:389-399.
- Meslar HW, Camper SA, White HB III: **Biotin-binding protein from egg yolk. A protein distinct from egg white avidin**. *J Biol Chem* 1978, **253**:6979-6982.
- Murthy CV, Adiga PR: **Purification of biotin-binding protein from chicken egg yolk and comparison with avidin**. *Biochim Biophys Acta* 1984, **786**:222-230.
- Subramanian N, Adiga PR: **Simultaneous purification of biotin-binding proteins-I and -II from chicken egg yolk and their characterization**. *Biochem J* 1995, **308**:573-577.
- White HB III, Whitehead CC: **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* 1987, **241**:677-684.
15. Bush L, McGahan TJ, White HB III: **Purification and characterization of biotin-binding protein II from chicken oocytes.** *Biochem J* 1988, **256**:797-805.
 16. Bush L, White HB III: **Conversion of domains into subunits in the processing of egg yolk biotin-binding protein I.** *J Biol Chem* 1989, **264**(10):5741-5745.
 17. White HB III: **Biotin-binding proteins and biotin transport to oocytes.** *Ann N Y Acad Sci* 1985, **447**:202-211.
 18. Niskanen EA, Hytönen VP, Grapputo A, Nordlund HR, Kulomaa MS, Laitinen OH: **Chicken genome analysis reveals novel genes encoding biotin-binding proteins related to avidin family.** *BMC Genomics* 2005, **6**:41.
 19. Keinänen RA, Laukkanen ML, Kulomaa MS: **Molecular cloning of three structurally related genes for chicken avidin.** *J Steroid Biochem* 1988, **30**:17-21.
 20. Keinänen RA, Wallén MJ, Kristo PA, Laukkanen MO, Toimela TA, Helenius MA, Kulomaa MS: **Molecular cloning and nucleotide sequence of chicken avidin-related genes I-5.** *Eur J Biochem* 1994, **220**:615-621.
 21. Ahlroth MK, Kola EH, Ewald D, Masabanda J, Sazanov A, Fries R, Kulomaa MS: **Characterization and chromosomal localization of the chicken avidin gene family.** *Anim Genet* 2000, **31**:367-375.
 22. Bayer EA, Ehrlich-Rogozinski S, Wilchek M: **Sodium dodecyl sulfate-polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin.** *Electrophoresis* 1996, **17**:1319-1324.
 23. Pugliese L, Coda A, Malcovati M, Bolognesi M: **Three-dimensional structure of the tetragonal crystal form of egg-white avidin in its functional complex with biotin at 2.7 Å resolution.** *J Mol Biol* 1993, **231**:698-710.
 24. Hytönen VP, Määttä JA, Kidron H, Halling KK, Hörhå J, Kulomaa T, Nyholm TK, Johnson MS, Salminen TA, Kulomaa MS, Airene TT: **Avidin related protein 2 shows unique structural and functional features among the avidin protein family.** *BMC Biotechnol* 2005, **5**:28.
 25. Eisenberg-Domovich Y, Hytönen VP, Wilchek M, Bayer EA, Kulomaa MS, Livnah O: **High-resolution crystal structure of an avidin-related protein: insight into high-affinity biotin binding and protein stability.** *Acta Crystallogr D* 2005, **61**:528-538.
 26. DeTitta GT, Edmonds JW, Stallings W, Donohue J: **Molecular structure of biotin. Results of two independent crystal structure investigations.** *J Am Chem Soc* 1976, **98**:1920-1926.
 27. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucleic Acids Res* 2000, **28**:235-242.
 28. Hytönen VP, Määttä JA, Nyholm TK, Livnah O, Eisenberg-Domovich Y, Hyre D, Nordlund HR, Hörhå J, Niskanen EA, Paldanius T, Kulomaa T, Porkka EJ, Stayton PS, Laitinen OH, Kulomaa MS: **Design and construction of highly stable, protease-resistant chimeric avidins.** *J Biol Chem* 2005, **280**:10228-10233.
 29. Klumb LA, Chu V, Stayton PS: **Energetic roles of hydrogen bonds at the ureido oxygen binding pocket in the streptavidin-biotin complex.** *Biochemistry* 1998, **37**:7657-7663.
 30. Hytönen VP, Laitinen OH, Airene TT, Kidron H, Meltola NJ, Porkka E, Hörhå J, Paldanius T, Määttä JA, Nordlund HR, Johnson MS, Salminen TA, Airene KJ, Ylä-Herttua S, Kulomaa MS: **Efficient production of active chicken avidin using a bacterial signal peptide in Escherichia coli.** *Biochem J* 2004, **384**:385-390.
 31. Edwards K, Chan RY, Sawyer WH: **Interactions between fatty acids and lipoprotein lipase: specific binding and complex formation.** *Biochemistry* 1994, **33**:13304-13311.
 32. **NCBI blastn** [<http://www.ncbi.nlm.nih.gov/blast/>]
 33. Hytönen VP, Nyholm TK, Penttinen OT, Vaarno J, Porkka EJ, Nordlund HR, Johnson MS, Slotte JP, Laitinen OH, Kulomaa MS: **Chicken avidin-related protein 4/5 shows superior thermal stability when compared with avidin while retaining high affinity to biotin.** *J Biol Chem* 2004, **279**:9337-9343.
 34. Flower DR: **The lipocalin protein family: structure and function.** *Biochem J* 1996, **318**:1-14.
 35. White HB III, Dennison BA, Della Fera MA, Whitney CJ, McGuire JC, Meslar HW, Sammelwitz PH: **Biotin-binding protein from chicken egg yolk. Assay and relationship to egg-white avidin.** *Biochem J* 1976, **157**:395-400.
 36. Mandella RD, Meslar HW, White HB III: **Relationship between biotin-binding proteins from chicken plasma and egg yolk.** *Biochem J* 1978, **175**:629-633.
 37. Seshagiri PB, Adiga PR: **Identification and molecular characterization of a biotin-binding protein distinct from avidin of chicken egg white and comparison with yolk biotin-binding protein.** *Biochim Biophys Acta* 1987, **926**:321-330.
 38. Melville DB, Genghof DS, Lee JM: **Biological properties of biotin d- and l-sulfoxides.** *J Biol Chem* 1954, **208**:503-512.
 39. Williams DH, Stephens E, O'Brien DP, Zhou M: **Understanding noncovalent Interactions: Ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes.** *Angew Chem Int Ed* 2004, **43**:6596-6616.
 40. Hyre DE, Amon LM, Penzotti JE, Le Trong I, Stenkamp RE, Lybrand TP, Stayton PS: **Early mechanistic events in biotin dissociation from streptavidin.** *Nat Struct Biol* 2002, **9**:582-585.
 41. Nordlund HR, Laitinen OH, Hytönen VP, Uotila ST, Porkka E, Kulomaa MS: **Construction of a dual chain pseudotetrameric chicken avidin by combining two circularly permuted avidins.** *J Biol Chem* 2004, **279**:36715-36719.
 42. Hytönen VP, Nordlund HR, Hörhå J, Nyholm TK, Hyre DE, Kulomaa T, Porkka EJ, Marttila AT, Stayton PS, Laitinen OH, Kulomaa MS: **Dual-affinity avidin molecules.** *Proteins* 2005, **61**:597-607.
 43. Thomas CM, Ward TR: **Artificial metalloenzymes: proteins as hosts for enantioselective catalysis.** *Chem Soc Rev* 2005, **34**:337-346.
 44. Gentry-Weeks CR, Hultsch AL, Kelly SM, Keith JM, Curtiss R III: **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* 1992, **174**:7729-7742.
 45. Airene KJ, Peltomaa E, Hytönen VP, Laitinen OH, Ylä-Herttua S: **Improved generation of recombinant baculovirus genomes in Escherichia coli.** *Nucleic Acids Res* 2003, **31**:e101.
 46. Laitinen OH, Airene KJ, Hytönen VP, Peltomaa E, Mähönen AJ, Wirth T, Lind MM, Mäkelä KA, Toivanen PI, Schenkwein D, Heikura T, Nordlund HR, Kulomaa MS, Ylä-Herttua S: **A multipurpose vector system for the screening of libraries in bacteria, insect and mammalian cells and expression in vivo.** *Nucleic Acids Res* 2005, **33**:e42.
 47. Airene KJ, Oker-Blom C, Marjomäki VS, Bayer EA, Wilchek M, Kulomaa MS: **Production of biologically active recombinant avidin in baculovirus-infected insect cells.** *Prot Exp Pur* 1997, **9**:100-108.
 48. Melville DB: **Biotin sulfoxide.** *J Biol Chem* 1954, **208**:495-501.
 49. Hofmann K, Melville DB, du Vigneaud V: **Characterization of the functional groups of biotin.** *J Biol Chem* 1941, **141**:207-214.
 50. Nordlund HR, Hytönen VP, Laitinen OH, Uotila ST, Niskanen EA, Savolainen J, Porkka E, Kulomaa MS: **Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding.** *FEBS Lett* 2003, **555**:449-454.
 51. Kulomaa MS, Elo HA, Tuohimaa PJ: **A radioimmunoassay for chicken avidin. Comparison with a [¹⁴C]biotin-binding method.** *Biochem J* 1978, **175**:685-690.
 52. Jancarik J, Scott WG, Milligan DL, Koshland DE Jr., Kim SH: **Crystallization and preliminary X-ray diffraction study of the ligand-binding domain of the bacterial chemotaxis-mediating aspartate receptor of Salmonella typhimurium.** *J Mol Biol* 1991, **221**:31-34.
 53. Kabsch W: **Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants.** *J Appl Crystallogr* 1993, **26**:795-800.
 54. Navaza J: **Amore - an automated package for molecular replacement.** *Acta Crystallogr A* 1994, **50**:157-163.
 55. Collaborative computational project, number 4 : **The CCP4 suite: programs for protein crystallography.** *Acta Crystallogr D* 1994, **50**:760-763.
 56. Potterton E, Briggs P, Turkenburg M, Dodson E: **A graphical user interface to the CCP4 program suite.** *Acta Crystallogr D* 2003, **59**:1131-1137.
 57. Perrakis A, Morris R, Lamzin VS: **Automated protein model building combined with iterative structure refinement.** *Nat Struct Biol* 1999, **6**:458-463.

58. Murshudov GN, Vagin AA, Dodson EJ: **Refinement of macromolecular structures by the maximum-likelihood method.** *Acta Crystallogr D* 1997, **53**:240-255.
59. Jones TA, Zou JY, Cowan SW, Kjeldgaard: **Improved methods for building protein models in electron density maps and the location of errors in these models.** *Acta Crystallogr A* 1991, **47**:110-119.
60. Lamzin VS, Wilson KS: **Automated refinement of protein models.** *Acta Crystallogr D* 1993, **49**:129-147.
61. Schuttelkopf AW, van Aalten DM: **PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.** *Acta Crystallogr D* 2004, **60**:1355-1363.
62. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **Procheck - a program to check the stereochemical quality of protein structures.** *J Appl Crystallogr* 1993, **26**:283-291.
63. Vriend G: **WHAT IF: a molecular modeling and drug design program.** *J Mol Graph* 1990, **8**:52-56.
64. DeLano WL: **The PyMOL Molecular Graphics System.** [<http://pymol.sourceforge.net/>].
65. Lehtonen JV, Still DJ, Rantanen VV, Ekholm J, Björklund D, Iftikhar Z, Huhtala M, Repo S, Jussila A, Jaakkola J, Pentikäinen OT, Nyrönen T, Salminen TA, Gyllenberg M, Johnson MS: **BODIL: a molecular modeling environment for structure-function analysis and drug design.** *J Comput Aided Mol Des* 2004, **18**(6):401-419.
66. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA: **Electrostatics of nanosystems: application to microtubules and the ribosome.** *Proc Natl Acad Sci USA* 2001, **98**:10037-10041.
67. Lee B, Richards FM: **The interpretation of protein structures: estimation of static accessibility.** *J Mol Biol* 1971, **55**:379-400.
68. Repo S, Paldanius TA, Hytönen VP, Nyholm TK, Halling KK, Huuskonen J, Pentikäinen OT, Rissanen K, Slotte JP, Airene TT, Salminen TA, Kulomaa MS, Johnson MS: **Binding properties of HABA-type azo derivatives to avidin and avidin-related protein 4.** *Chem Biol* 2006, **13**:1029-1039.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



DOI: 10.1002/cbic.200((will be filled in by the editorial staff))

Dual chain avidin 2

RATIONAL MODIFICATION OF LIGAND-BINDING PREFERENCE OF AVIDIN BY CIRCULAR PERMUTATION AND MUTAGENESIS

Juha A. E. MÄÄTTÄ^[a], Tomi T. AIRENNE^[b], Henri R. NORDLUND^[a], Janne JÄNIS^[c], Tiina A. PALDANIUS^[a], Pirjo VAINIOTALO^[c], Mark S. JOHNSON^[b], Markku S. KULOMAA^[a] and Vesa P. HYTÖNEN^{*[a]}

Chicken avidin is a key component used in a wide variety of biotechnological applications. Here we present a circularly permuted avidin (cpAvd4→3) which lacks the loop between beta strands 3 and 4. Importantly, the deletion of the loop has a positive effect on the binding of 4'-hydroxyazobenzene-2-carboxylic acid (HABA) to avidin. In order to increase the HABA-affinity of cpAvd4→3 even further, we mutated asparagine 118 on the bottom of the ligand-binding pocket to methionine, which simultaneously caused a significant drop in biotin-binding affinity. The X-ray structure of cpAvd4→3(N118M) allows an

understanding of the effect of mutation to biotin-binding, whereas isothermal titration calorimetry revealed that the relative binding affinity of biotin and HABA had changed by over one billion-fold between wild-type avidin and cpAvd4→3(N118M). In order to demonstrate the versatility of the cpAvd4→3 construct, we have shown that it is possible to link cpAvd4→3 and cpAvd5→4 to form the dual chain avidin called dcAvd2. These novel avidins may serve as a basis for the further development of self-organizing nanoscale avidin building blocks.

Introduction

Avidin is a tetrameric glycoprotein which has four identical subunits, each composed of eight antiparallel β -strands.^[1] Avidin and its prokaryotic analogues streptavidin,^[2] bradavidin^[3] and rhizavidin^[4] bind D-biotin with extremely high affinity ($K_d \approx 10^{-15}$ M for avidin). This strong interaction is widely exploited in so-called avidin-biotin technology.^[5, 6]

Avidin and streptavidin, (strept)avidin, have previously been genetically engineered in various ways, recently reviewed by Kulomaa and associates.^[7, 8] Recently, a circular permutation strategy was used to join avidin subunits within a single polypeptide chain, forming both dual chain avidin (dcAvd)^[9] and single chain avidin (scAvd),^[10] which respectively display two and four ligand-binding sites, each of which can be independently modified to bind different ligands with varying affinities. A dual-affinity avidin has already demonstrated the feasibility of this concept.^[11] Streptavidin has also been subjected to circular permutation^[12] and subunit fusion.^[13]

When D-biotin is bound to streptavidin, only 18 Å² of the ligand surface is accessible to a solvent. In a circularly permuted

form of streptavidin where the loop connecting β -strands 3 and 4 was deleted the solvent-accessible surface increased to 57 Å², and the carboxylate group of D-biotin became more exposed.^[12] In avidin, proteinase K cleaves the analogous loop, and protease K-treated avidin has been reported to display enhanced affinity towards the azo-compound HABA (4'-hydroxyazobenzene-2-carboxylic acid).^[14] Streptavidin has a lower affinity towards HABA ($K_d = 1.4 \times 10^{-4}$ M)^[15] than avidin ($K_d = 6.0 \times 10^{-6}$ M),^[16, 17] which difference can at least partially be explained by the fact that the L3,4 loop, known to be important for ligand binding, is shorter in streptavidin than in avidin.^[18, 19] Accordingly, it has been observed that AVR4, an avidin-related protein,^[20, 21] also exhibits lower affinity for HABA in comparison to avidin, most probably because AVR4 has a kinked and therefore shorter L3,4 loop.^[22]

In the present study, we constructed a novel circularly-permuted form of avidin, cpAvd4→3, in order to alter the ligand-

- [a] MSc J. A. E. Määttä, Dr H. Nordlund,* MSc T. Paldanius, Prof. Dr. M. S. Kulomaa, Dr. V. P. Hytönen
Institute of Medical Technology
University of Tampere
FI-33014 University of Tampere (Finland)
Fax: (+358)3-3551-7710
E-mail: vesa.hytonen@uta.fi
- [b] Dr. T. Airene, Prof. Dr. M.S. Johnson
Department of Biochemistry and Pharmacy
Åbo Akademi University
FI-20520 Turku (Finland)
- [c] Dr. J. Jänis, Prof. Dr. P. Vainiotalo
Department of Chemistry
University of Joensuu
FI-80101 Joensuu (Finland)
- [*] Present address: NEXT Biomed Technologies NBT Oy
FI-00790 Helsinki (Finland)

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author. ((Please delete if not appropriate))

binding preferences of avidin and to increase the potential of avidin to bind ligands larger than biotin. A point mutation N118M (numbering according to wt avidin), was introduced into cpAvd4→3 in order to increase affinity for HABA and to reduce that for biotin.

Further, cpAvd4→3 was fused to the C-terminus of cpAvd5→4, another circularly permuted form of avidin,^[9] creating a novel fusion of avidin subunits, designated dual chain avidin 2 (dcAvd2), possessing advanced structural and functional characteristics.

Avidin and streptavidin are examples of widely used protein bio-tools. They are utilized in technologies ranging from the biosciences and medicine to material science.^[6, 23, 24, 25] The extreme affinity for the small biotin ligand and the high stability of these proteins has encouraged scientists to develop them further to meet the growing needs of bio- and nanotechnology.

Results and Discussion

Novel circularly permuted avidins

CpAvd4→3 and dcAvd2 were successfully purified by 2-aminobiotin and cpAvd4→3(N118M) by biotin affinity chromatography. The final products were highly pure and had correct size according to SDS-PAGE analysis (Figure 1).

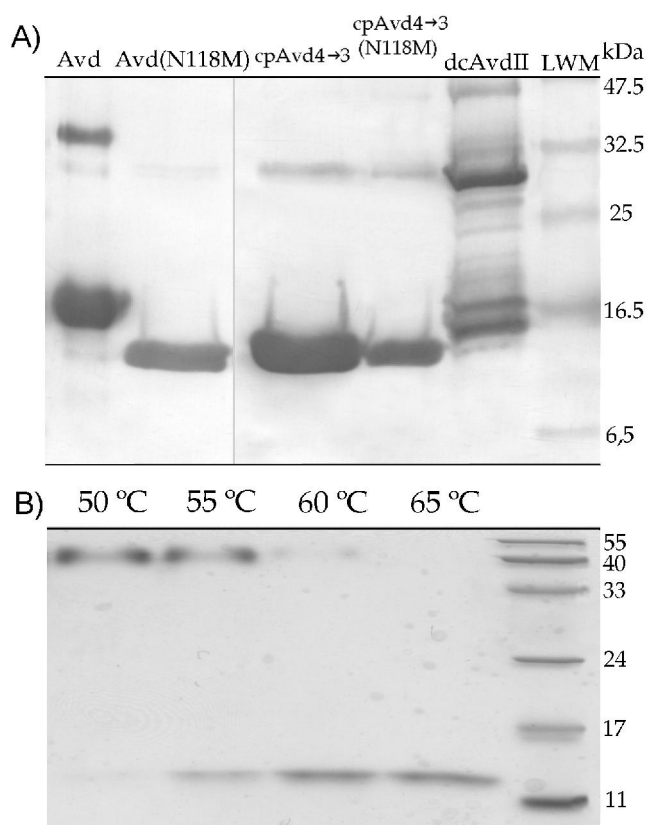


Figure 1. Western blot and SDS-PAGE-based thermal stability assay. A, Western blot analysis of the purified proteins, avidin, avidin(N118M) mutant, cpAvd4→3, cpAvd4→3(N118M), dcAvd2 and low molecular weight marker. B, thermal stability assay of cpAvd4→3 protein at temperatures of 50, 55, 60, 65 °C and low molecular weight marker. The loss of quaternary structure of cpAvd4→3 was observed at 55 °C.

Table 1. Oligomeric size of the proteins.^[a]

Protein	Elution time ^[b]		MW (kDa)	
	-BTN	+BTN	-BTN	+BTN
Avidin	28.23	27.97	52.9	56.1
cpAvd4→3	29.03	28.78	44.3	46.9
cpAvd4→3(N118M)	28.69	28.70	47.8	47.7
dcAvd2	27.91	27.95	56.8	56.3

[a] The apparent molecular weights (MW) obtained by gel filtration analysis are indicated. [b] Elution time in minutes.

However, in the case of dcAvd2, a significant amount of an ~20 kDa form was observed in addition to the protein of expected size (~30 kDa). The smaller polypeptide chain detected in the dcAvd2 samples was likely due to a proteolytic cleavage. A typical yield after purification was ~5 mg of protein per litre of bacterial culture. Gel filtration chromatography showed the homogeneity of the purified proteins, all being detected as single peaks in the analysis (not shown).

The calculated molecular masses of the protein forms (cpAvd4→3, cpAvd4→3(N118M) and dcAvd2) produced in *Escherichia coli* (*E. coli*) were 47-56 kDa, which are lower than the theoretical mass values of the corresponding tetramers (57-58 kDa) (Table 1). However, this is a typical behaviour of avidin, as observed in our previous studies.^[11]

Mass spectrometry

ESI FT-ICR mass spectrometry was used to confirm the amino acid sequence and proper folding of the proteins (mass spectra data not presented). For some unknown reason dcAvd2 was the only protein which could not be identified by this method. The most abundant isotopic masses were determined to be 14685.52 ± 0.07 Da for Avid(N118M), 14285.24 ± 0.01 Da for cpAvd4→3 and 14302.18 ± 0.03 Da for cpAvd4→3(N118M). The determined mass of Avid(N118M) was as expected (*calc.* 14685.50 Da), but the masses of both cpAvd4→3 and cpAvd4→3(N118M) were about 14 Da higher as compared to their expected, calculated values of 14271.18 and 14288.17 Da, respectively. This result would suggest protein methylation (*theor.* + 14.016 Da).

When the disulfide bridges were reduced, the most abundant isotopic masses of the proteins increased by two Daltons as expected, indicating proper folding. This also further supports the conception that the mass shift of ~14 Da could be a result of methylation. Post-translational protein methylation is a common amino acid modification, known to occur at the side chains of many amino acids.^[26] We have further identified and localized post-translational methylation in cpAvd4→3(N118M) using on-line digestion combined with ESI FT-ICR tandem mass spectrometry (Jänis J. *et al.*, unpublished results). A single methylation site in one of the serine residues in cpAvd4→3(N118M) was unambiguously observed, revealing a previously unknown protein modification.

Deletion of critical residues in the L3,4 loop by circular permutation reduces biotin affinity

The biotin-binding properties of the engineered avidins were studied by measuring the dissociation of fluorescent biotin conjugate BF-560-biotin. At 25 °C, cpAvd4→3 showed an eight-fold higher dissociation rate compared to wt avidin. In comparison, dcAvd2, having dissimilar subunit pairs, showed a three-fold increased dissociation rate compared to avidin. Half of the binding sites were assumed to behave as in wt avidin (in dcAvd2 the circularly permuted subunit cpAvd5→4 should according to our knowledge have wt-like biotin-binding properties),^[9] thus a four-fold increased dissociation rate was attributed to the cpAvd4→3 subunit of dcAvd2.

At 50 °C, cpAvd4→3 showed the highest dissociation rate for biotin, evincing about 20-fold weaker binding compared to avidin. DcAvd2 had an about five-fold faster dissociation rate compared to avidin. When the data were analyzed as described above, subtracting the contribution due to the wt-like subunit, the dissociation rate of biotin from the cpAvd4→3 subunit of dcAvd2 was found to be about ten-fold faster as compared to avidin (Table 2).

The dissociation rate of α -[8,9-³H]biotin from the protein samples was measured at different temperatures (Figure 2). Avidin had the slowest dissociation rate of α -[8,9-³H]biotin over the entire temperature range analyzed, and as assumed, dcAvd2 had a slower dissociation rate for biotin than cpAvd4→3.

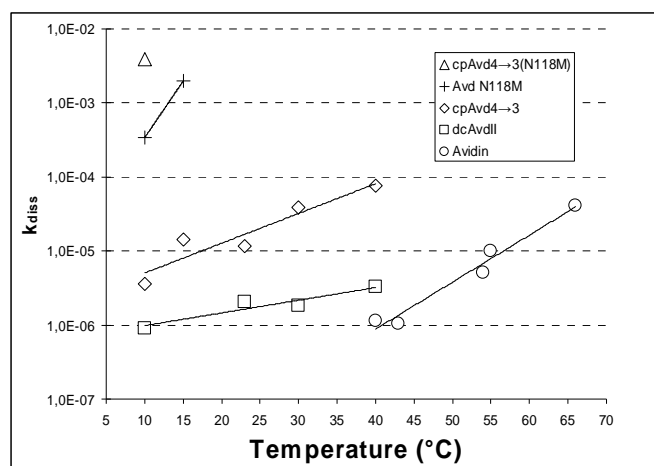


Figure 2. Radiobiotin assay. Dissociation rates determined using α -[8,9-³H]biotin at various temperatures are shown.

The affinity of avidin for biotin was too tight to be measured by isothermal titration calorimetry (ITC), but the binding constant (K_b) and dissociation constant ($K_d = 1/K_b$) could be solved for the modified avidins (Figure 3). If the wild type binding site (cpAvd5→4) of dcAvd2 is assumed to bind biotin with femtomolar (10^{-15} M; wt-like) affinity, the biotin affinity of the other circularly permuted binding site (cpAvd4→3) can be directly determined from the data. The cpAvd4→3 subunit in fusion with cpAvd5→4

Table 2. Dissociation analysis of fluorescent biotin.

Protein	k_{diss} ($\times 10^{-5}$) $s^{-1}M^{-1}$ at 25 °C	Release (%)	k_{diss} (10^{-5}) $s^{-1}M^{-1}$ at 50 °C	Release (%)
Avidin	2.3 ^[a]	14.1 ^[a]	27.4 ^[b]	71.5
Avidin(N118M)	>300 ^[e]			
cpAvd4→3	18.7	46.8	449.9 ^[c]	
cpAvd4→3(N118M)	>300 ^[e]			
dcAvd2	5.3/7.8 ^[d]	17.7	135.3 ^[c] /200 ^[d]	71.0
dcAvd	1.5	6.0	41	77.9

[a] Ref.^[27] [b] Ref.^[28] [c] Determined during the first 500 seconds of the measurement. [d] Half of the binding sites were assumed to behave like wt avidin. The fitting was performed only for the first 500 seconds. [e] The dissociation rate was already too fast to be measured accurately at 10 °C.

of dcAvd2 seemed to have a slightly higher affinity for biotin as compared to the tetrameric cpAvd4→3. Moreover, the N118M point mutation reduces the biotin-binding affinity of the cpAvd4→3 protein about five-fold (Table 3).

Novel circularly permuted avidin has altered thermodynamics and increased affinity for HABA

Our ITC experiment revealed that avidin binds HABA with moderate affinity ($K_d = 7.9 \times 10^{-6}$ M), which is in good agreement with the previously published result of $K_d = 6 \times 10^{-6}$ M (Green 1975), whereas its bacterial analogue streptavidin has a weaker HABA-binding affinity ($K_d = 1.4 \times 10^{-4}$ M).^[15] The different HABA-binding affinities of avidin and streptavidin are thought to result from the varying architecture of the L3,4 loop, i.e., its conformation and primary structure, as also noted in the case of AVR4.^[22] Our working hypothesis here was that by deleting the L3,4 loop from wt avidin using a circular permutation strategy, the HABA-binding affinity of avidin could be increased. This was tested by measuring the HABA-binding affinities of cpAvd4→3 and wt avidin using ITC. It emerged that the cpAvd4→3 construct had indeed stronger affinity to HABA ($K_d = 5.4 \times 10^{-6}$ M) as compared to wt avidin ($K_d = 7.9 \times 10^{-6}$ M).

Table 3. Thermodynamic analysis of ligand binding using ITC.

Protein	ΔH kcal/mol	$-T\Delta S$ kcal/mol	ΔG kcal/mol	K_d M
BTN				
Avidin	-23.0±0.2	4.4 ^[b]	-18.6 ^[b]	1.3×10^{-15} ^[a]
Avidin(N118M)	-16.4±0.1	5.0	-11.4	$(4.2 \pm 0.2) \times 10^{-9}$
cpAvd4→3	-17.6±0.1	7.2	-10.4	$(1.4 \pm 0.3) \times 10^{-8}$
cpAvd4→3(N118M)	-8.3±0.1	-0.7	-9.0	$(2.4 \pm 0.4) \times 10^{-7}$
HABA				
Avidin	1.9±0.03	-8.8	-7.0	$(7.9 \pm 0.1) \times 10^{-6}$
Avidin(N118M)	-0.8±0.02	-6.4	-7.2	$(5.2 \pm 0.5) \times 10^{-6}$
cpAvd4→3	-3.5±0.06	-3.6	-7.1	$(5.4 \pm 0.7) \times 10^{-6}$
cpAvd4→3(N118M)	-5.1±0.05	-3.0	-8.1	$(1.0 \pm 0.2) \times 10^{-6}$

[a] Ref.^[29] [b] Calculated from the K_d value.

The reaction between avidin and biotin is strongly exothermic ($\Delta H = -23.0 \pm 0.2$ kcal/mol). In contrast, that between HABA and avidin is endothermic ($\Delta H = 1.9 \pm 0.03$ kcal/mol) and thus the moderate binding affinity is entropy-driven. Our ITC analysis showed the interaction of HABA with Avd(N118M), cpAvd4→3 or cpAvd4→3(N118M) to be exothermic. The most markedly exothermic reaction with HABA was measured for cpAvd4→3(N118M), which has about six-fold more favourable enthalpy compared to Avd(N118M) and one and a half-fold more favourable compared to cpAvd4→3 (Table 3). The calculated entropy term was found to be beneficial for binding in every HABA-binding reaction analyzed here.

A problem arises in calorimetric measurements in that the wt biotin-binding site of dcAvd2, i.e., the cpAvd5→4 subunit, binds HABA in an endothermic manner, whereas the cpAvd4→3 subunit binds HABA in an exothermic manner. Since it is possible to see only the sum of the energies, the net sum of the observed heat of the binding reaction for dcAvd2 was almost zero (Figure 4) and therefore the affinity of the two different kinds of binding sites of dcAvd2 to HABA could not be determined.

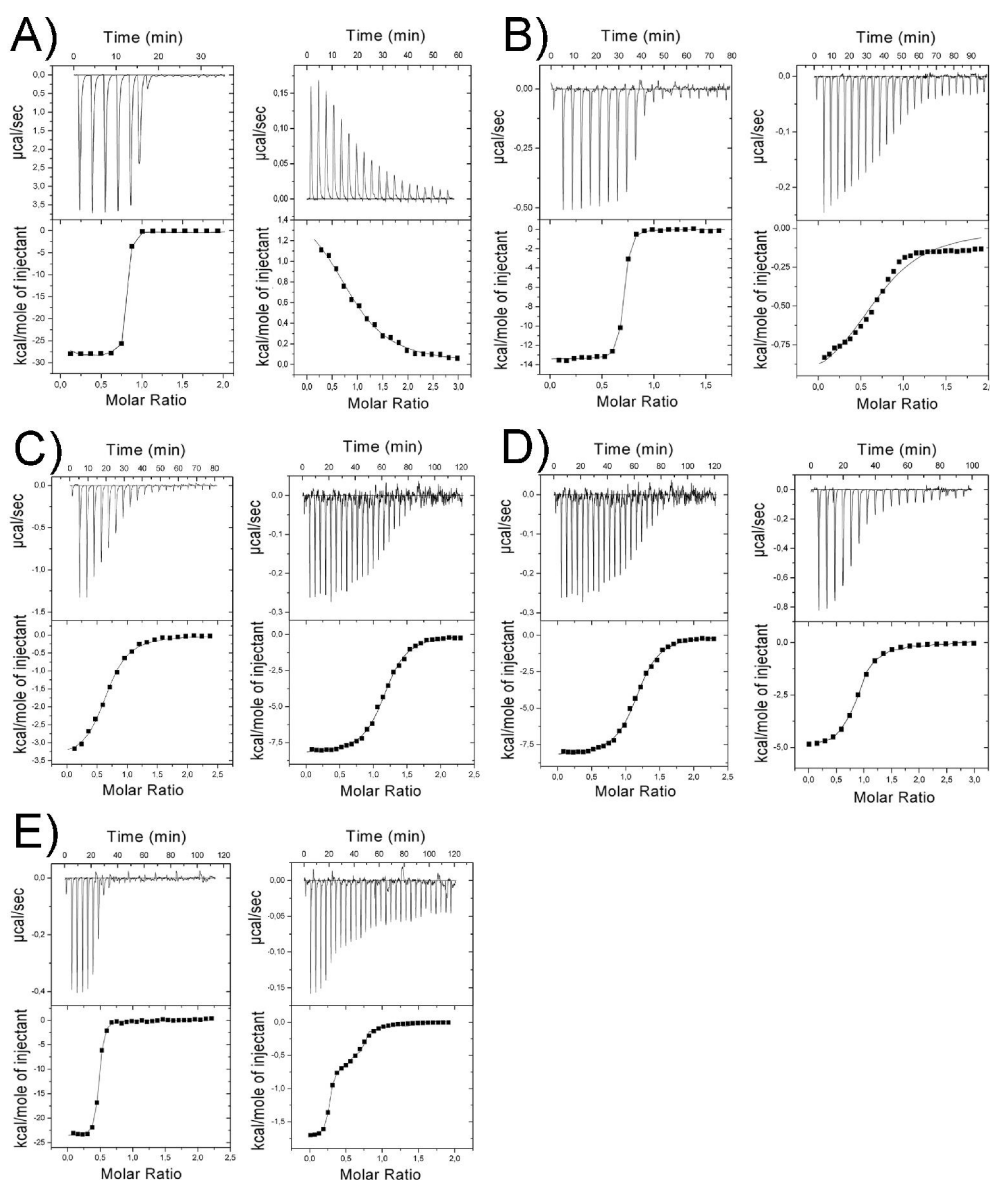
Mutation N118M increases HABA-binding and reduces the biotin-binding affinity of avidin

The mutation N118M was designed based on the known X-ray structures of the avidin-biotin (PDB entry 2AVI)^[1] and avidin-HABA complexes.^[19] In the structure of the avidin-biotin complex, the side chain of Asn118 is hydrogen bonded to biotin (Figure 7(a)), whereas in the avidin-HABA structure there is no significant interaction between the side chain of Asn118 of the avidin and HABA (Figure 7(b)). Therefore, a mutation at Asn118 which eliminates the hydrogen bond should lead to a decrease in the biotin-binding affinity of avidin. Methionine was chosen to replace asparagine, because its long and flexible side chain was assumed to constitute steric hindrance for biotin. Also preliminary experiments between different kinds of mutations made by Paldanius *et al.* (manuscript) revealed that methionine caused largest changes in both biotin and HABA binding. Indeed, using ITC, a significant decrease in the biotin-binding affinity of avidin was observed for an Avd(N118M) mutant. As a result of the N118M mutation, the extremely strong femtomolar biotin-dissociation constant of wt avidin^[29] was reduced to a nanomolar value ($K_d = 4.17 \times 10^{-9}$ M). Furthermore, the enthalpy of the Avd(N118M)-biotin binding ($\Delta H = -16.4 \pm 0.1$ kcal/mol) was only two-thirds that of the Avd-biotin binding ($\Delta H = -23.0 \pm 0.2$ kcal/mol) (Table 3). The point mutation N118M seemed to have

a clear effect on the biotin dissociation rate: Avd(N118M) has an ~2500-times higher α -[8,9-³H]biotin dissociation rate than wt avidin. The N118M mutation was also introduced into the cpAvd4→3 construct (cpAvd4→3(N118M)). It had, in turn, a very fast dissociation rate ($k_{\text{diss}} \approx 4 \times 10^{-3}$ s⁻¹) even at a low temperature of 10 °C. The corresponding dissociation rate of avidin would be about 10⁻⁸ s⁻¹ if extrapolated from the measured data (see Figure 2). At higher temperatures, the dissociation rate of cpAvd4→3(N118M) for biotin was too fast to be determined accurately (Figure 2).

In contrast to biotin, HABA has been observed to bind to Avd(N118M) slightly more strongly than to wt avidin: analysis using fluorescence spectrometry showed an apparent $K_d = 6.3 \times 10^{-5}$ M for the avidin-HABA complex and a $K_d = 0.6 \times 10^{-5}$ M for Avd(N118M)-HABA (Paldanius *et al.* manuscript). Our ITC measurements were in agreement with these results – the HABA-binding affinity was almost the same for both Avd(N118M) ($K_d = 5.2 \times 10^{-6}$ M) and cpAvd4→3 ($K_d = 5.4 \times 10^{-6}$ M), i.e., both have higher affinity for HABA than wt avidin ($K_d = 7.9 \times 10^{-6}$ M; Table 3).

Figure 3. ITC diagrams. A, avidin; B, Avd(N118M); C, cpAvd4→3; D, cpAvd4→3(N118M); E, dcAvd2. Biotin titrations are shown on the left and HABA titrations on the right panel. All measurements were made at 25 °C.



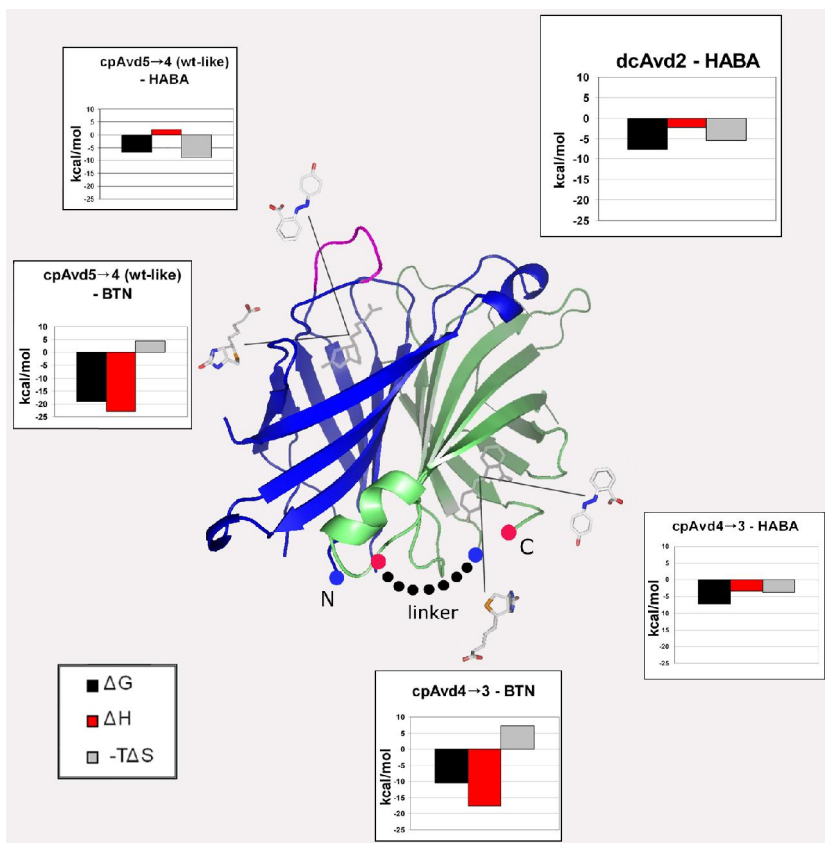


Figure 4. The ligand binding sites of dcAvd2. A modelled structure of dcAvd2 with two different kinds of binding sites is presented. The cpAvd5→4 (blue) of dcAvd2 has a wt-like binding site. It binds biotin with tight affinity ($K_d = 10^{-15}$ M) and HABA with moderate affinity ($K_d = 10^{-6}$ M). The cpAvd4→3 subunit (green) of dcAvd2 lacks the L3,4 loop (pink) and binds both HABA and biotin with moderate affinity. The new N- and C-termini of cpAvd4→3 are marked with purple dots. Diagrams represent thermodynamic values obtained from ITC measurements – Gibbs free energy (ΔG , black bar), enthalpy (ΔH , red bar) and the entropy term in measurement temperature ($-T\Delta S$, grey bar). DcAvd2 – HABA diagram shows combined energies of both cpAvd5→4 and cpAvd4→3 HABA binding events. The model was made by deleting by hand the L3,4 loop from two of subunits from the X-ray structure of avidin.

Similarly to the result of mutation in the case of wt avidin, the N118M mutation in cpAvd4→3 exerted the following effects: 1) it reduced the affinity to biotin about 16-fold and 2) it increased the affinity to HABA about five-fold compared to cpAvd4→3. Thus, of all the avidin forms examined to date, the cpAvd4→3(N118M) construct has the highest affinity for HABA ($K_d = 1.0 \times 10^{-6}$ M; Table 3).

Overall X-ray structure of cpAvd4→3(N118M)

The crystal structure of cpAvd4→3(N118M) in complex with D-biotin was determined at 1.9 Å resolution. The structure determination statistics are summarized in Table 4. Judging from the 3D structure, the circular permutation strategy had not affected the folding of cpAvd4→3(N118M). The overall structure of cpAvd4→3(N118M) was tetrameric and highly similar to that of chicken avidin^[1] and bacterial streptavidin.^[18, 30] However, in each of the four β -barrel monomers of cpAvd4→3(N118M), the biotin-binding pocket was less deeply buried and much more exposed to solvent in comparison to wt avidin.^[1, 31] This is due to the lack of residues 39-44 of the L3,4 loop (numbering according to wt avidin^[1]) in the cpAvd4→3 construct and the artificial locations of the N- and C-termini (Figure 5). In fact, the biotin-binding pocket of cpAvd4→3(N118M) was even more open than the biotin-binding site of streptavidin.^[30] The average solvent-accessible surface area of the bound biotin molecules in the avidin-biotin (PDB entry 2AVI), streptavidin-biotin (PDB entry 1MK5) and

cpAvd4→3(N118M)-biotin complexes were calculated to be 7, 18 and 22 Å², respectively.

The C α atoms of the four subunits of cpAvd4→3(N118M) were superimposed (subunits B, C and D on A) with an *rmsd* (root mean squared deviation) of 0.3 Å. Only slight variations were detected in the fine architecture of the different chains, mainly within the loop regions, and probably caused by crystal packing effects. Out of the 130 residues comprising cpAvd4→3(N118M), residues 1-3 (chains A, C, D), 45-47 (A), 81 (A, C) and 82-95 (A-D) could not be included in the final X-ray model due to the lack of interpretable electron density around these residues (high thermal motion). Residues 82-95 correspond to a linker region, which was introduced in order to connect the N- and C-terminus of wt avidin in forming the cpAvd4→3(N118M) construct. In addition, the subunit interfaces seen in cpAvd4→3(N118M) were highly similar to those present in wt avidin.

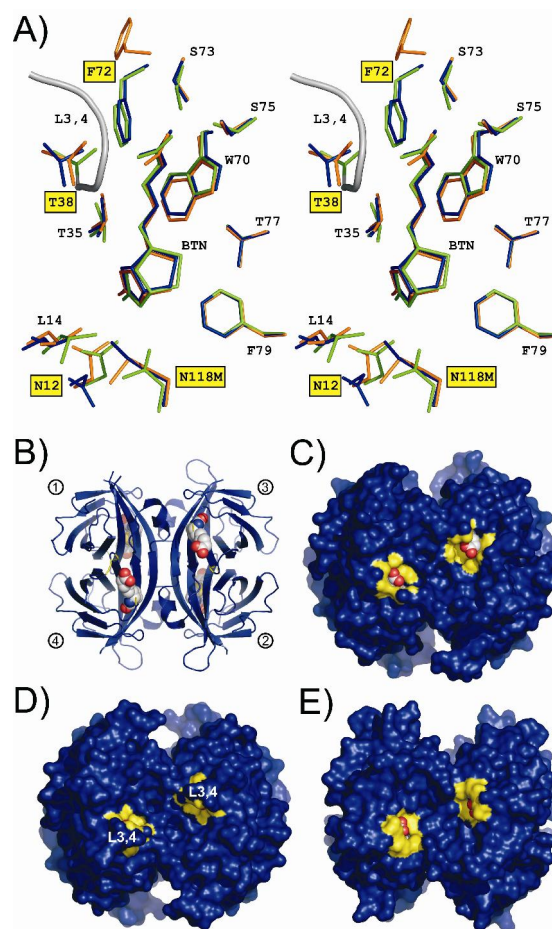


Figure 5. A stereoview of the biotin-binding site of cpAvd4→3(N118M) and X-ray structure of cpAvd4→3(N118M). Selected residues at the biotin-binding site of cpAvd4→3(N118M) (chain C, orange; chain D, blue) and avidin-biotin complex structure (PDB code 2AVI; chain A; green) are shown as stick models. Only side chain atoms are shown. The amino acids are numbered according to Ref 1. Bound biotin (BTN) ligands are shown as sticks. The labels for residues

N12 and F72, which have clearly different conformations in chains C and D of cpAvd4→3(N118M), as well as the N118M mutation and the C-terminal residue (T38) of cpAvd4→3 are boxed and have yellow background. A part of the L3,4-loop of avidin extending from T38 and not present in cpAvd4→3(N118M) is drawn as a grey ribbon. B, structure of cpAvd4→3(N118M) showing the secondary structure elements (ribbon model) and connecting loops. The subunits are numbered according to ref. 1. C, molecular surface of cpAvd4→3(N118M); D, avidin-biotin complex (PDB code 2AVI); and E, streptavidin-biotin complex (PDB code 1MK5). The surfaces are colored blue except for areas corresponding to atoms 6 Å from the C10 atom (naming according to ref. 32) of biotin. These atoms are colored yellow to pinpoint the entrances to the biotin-binding sites. The bound biotin molecules are shown as spheres (oxygen atoms, red; carbon atoms, white; nitrogen atoms, blue; sulfur atoms, yellow).

Table 4. Data collection and structure determination statistics for cpAvd4→3(N118M).

Data collection ^[a]	
Wavelength (Å)	0.900
Beamline	ID29 (ESRF)
Detector	ADSC
Resolution (Å)	25 – 1.9 (2.0 – 1.9)
Unique observations	36425 (5172)
I/sigma	12.3 (3.2)
R _{factor} (%) ^[b]	7.8 (48.2)
Completeness	99.8 (99.9)
Redundancy	4.9 (5.0)
Refinement	
Space group	
Unit cell:	P2 ₁
a, b, c (Å)	41.8, 79.3, 71.7
α, β, γ (°)	90, 98.7, 90
Monomers (asymmetric unit)	4
Resolution (Å)	25 – 1.9
R _{work} (%) ^[c]	20.3
R _{free} (%) ^[c]	25.1
Protein atoms	3556
Heterogen atoms	64
Solvent atoms	172
R.m.s.d:	
Bond lengths (Å)	0.013
Bond angles (°)	1.6
[a] The numbers in parenthesis refer to the highest resolution bin. [b] Observed R-factor from XDS. ^[33] [c] From Refmac 5 (TLS & restrain refinement). ^[34]	

The biotin-binding site

All four subunits of cpAvd4→3(N118M) had biotin bound at the ligand-binding site. However, the electron density around the ligand of subunit C (BTN-C) was clearly weaker (data not shown), and the B-factors for the ligand were higher in comparison to biotin of the other subunits, indicating the higher thermal motion and conformational flexibility of BTN-C. Phe72 at the biotin-binding site of subunit C of cpAvd4→3(N118M) also exhibited a conformation different from that seen in the other subunits, further indicating the lower occupancy of BTN-C at subunit C. In comparison to wt avidin (PDB entry 2AVI)^[1] and streptavidin (PDB entry 1MK5),^[30] the overall biotin-binding mode of each subunit of cpAvd4→3(N118M), as well as the conformations of the bound biotin, were well conserved (Figure 6).

The hydrophobic side chain of Met118 of cpAvd4→3(N118M) did, as expected, lead to different interactions with biotin in comparison to wt avidin. The hydrogen bonds seen in the avidin-biotin complex structure,^[1] respectively between the side chains of Asn12 and Asn118 of the protein and the oxygen (O2') and

nitrogen (N3') atom of the ureido ring of biotin (naming according to DeTitta *et al.*),^[32] were substituted either partially or completely by weak (multi)polar and van der Waals interactions in the cpAvd4→3(N118M) structure (Figure 7). In the ligand-binding sites of subunits A-C of cpAvd4→3(N118M) (see, e.g., subunit B in Figure 7(c)), a weak hydrogen bond between the O2' atom of biotin and the side chain nitrogen atom of Asn12 may exist as in the wt avidin structure (Figure 7(a)), but in subunit D the hydrogen bond was missing (Figure 7(d)). Moreover, the hydrophobic side chain of Met118 could not form a strong hydrogen bond with the N3' atom of biotin in any of the subunits of cpAvd4→3(N118M). This interaction is known to be important for biotin binding in wt avidin, where the Asn118 side chain is hydrogen-bonded to N3'.^[1] In addition, the lack of the L3,4 loop residues Ala39 and Thr40 of wt avidin at the biotin-binding site of cpAvd4→3(N118M) made the site clearly more accessible to solvent (see above; Figure 5), while the biotin-binding mode was nonetheless very similar to that in both streptavidin and wt avidin (Figure 6). Some rearrangement of water molecules due to the deletion of the L3,4 loop and the artificial C-terminus was, however, detectable in the X-ray structure of cpAvd4→3(N118M). For example, in chains C and D only, a water molecule was seen between the C-terminal threonine residue and the carboxylic acid group of biotin in the cpAvd4→3(N118M) structure.

Thermal stability of the novel proteins

The thermostability of cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 (Table 5) was analyzed using an SDS-PAGE assay.^[35] Tetramers of cpAvd4→3 without biotin were found to breakdown into subunits at 55 °C and, in the presence of biotin, at 75 °C. Similarly, cpAvd4→3(N118M) lost its quaternary structure at 50 °C without biotin and at 75 °C with biotin. DcAvd2, in turn, was found to breakdown into subunits at less than 60 °C and with biotin the breakdown temperature was found to be between 75 and 80 °C. By comparison, wt avidin breaks down in the presence and absence of biotin at 95 and 60 °C, respectively (Table 5).

Table 5. Thermal stability of the proteins.^[a]

Protein	T _r (°C) -BTN	T _r (°C) +BTN
Avidin	60	95
cpAvd4→3	55	75
cpAvd4→3(N118M)	50	75
dcAvd2	55	80
[a] Transition midpoint temperatures (T _r) of the oligomeric disassembly measured by an SDS-PAGE-based method are shown in the presence (+BTN) and absence (-BTN) of biotin.		

Conclusion

Since the cloning of the genes, numerous modified (strept)avidins have been reported in which the pI, glycosylation, stability and ligand-binding of these proteins has been altered, as recently reviewed by Kulomaa and associates.^[7] However, two relatively recent innovations have had a major impact on the engineering of (strept)avidin beyond the original products. First, the creation of circularly permuted forms has made it possible to create mutant proteins possessing new N- and C- termini at artificial locations. A

circular permutation was originally observed in nature, as reviewed by Grishin,^[36] but was also applied to streptavidin^[12] and avidin.^[9] Second, the introduction of the dual chain methodology^[9] based on circular permutation has made it possible to produce rational designs generating chimeric avidins and, more recently, streptavidins.^[13]

In the present study, our working hypothesis was that removal of the L3,4 loop located between β -strands 3 and 4 of wt avidin would alter its ligand-binding preferences, i.e., reduce the affinity for biotin and increase for HABA. This hypothesis was made based on results from Ellison and colleagues^[14] and comparisons of several crystallographic structures of (strept)avidin. Furthermore, we sought to apply the recently introduced dual chain technology to generate chimeric avidins which would have a wt-like ligand-binding site and a binding site lacking the L3,4 loop.

Our topologically novel cpAvd4 \rightarrow 3 construct was created as follows: i) six amino acids, residues 39-44 of the original L3,4 loop of wt avidin, were deleted; ii) the original N- and C-termini of wt avidin were connected with a Ser-Gly-Gly-Ser (SGGS) linker and iii) novel N- and C-termini were introduced – the translation of cpAvd4 \rightarrow 3 starts before the β 4-strand and stops after the β 3-strand. The protein was successfully produced in *E. coli*, and spectroscopic analyses showed that the affinity for HABA was increased in comparison to the wt protein, albeit only a little. In order to improve the HABA-binding properties of cpAvd4 \rightarrow 3, we introduced, based on the X-ray structures of avidin-biotin and avidin-HABA complexes, a point mutation to the bottom of the ligand-binding site. This mutation changed the polar side chain of asparagine (Asn118) to the hydrophobic side chain of methionine, thus directly modulating the physicochemical properties of the binding site. In the X-ray structure of cpAvd4 \rightarrow 3(N118M) (Figure 7), it can be seen that, depending on the conformation of Met118, either one or two hydrogen bonds are eliminated. This is in good agreement with our thermodynamic measurements (see below).

Loss of the hydrogen bond between the ureido nitrogen and streptavidin reduces the biotin-binding free energy by approximately 4.3 kcal/mol.^[30] Although the free energy of a hydrogen bond can vary over -0.2 to -40 kcal/mol, as calculated and reviewed by Stainer,^[37] the hydrogen bonds in avidin-biotin binding are thought to represent strong bonds with a free energy of about four kcal/mol. Due to the lack of the amino acid residues Ala39 and Thr40, which form hydrogen bonds with the carboxylic tail of biotin,^[38] cpAvd4 \rightarrow 3 has two hydrogen bonds less than wt avidin. The observed decrease in free energy ($\Delta\Delta G = 8.2$ kcal/mol) can be satisfactorily explained by the loss of the bonds. Our thermodynamic measurements revealed that the difference in Gibb's free energy ($\Delta\Delta G$) between wt avidin and Avd(N118M) was 7.2 kcal/mol, indicating that the N118M mutation affected more than one hydrogen bond, or alternatively, the bond is co-operative as in streptavidin.^[30] Indeed, the bond between biotin and the Asn12 side chain was found to be disturbed due to mutation of N118M (Figure 7). These results could suggest that cpAvd4 \rightarrow 3(N118M) has significantly lower affinity for biotin in comparison to either Avd(N118M) or cpAvd4 \rightarrow 3. The ITC experiments revealed that the biotin binding of cpAvd4 \rightarrow 3(N118M) was weakened 21% and 13% in terms of ΔG when compared to Avd(N118M) and cpAvd4 \rightarrow 3, respectively. The ITC experiments also revealed that cpAvd4 \rightarrow 3(N118M) had increased affinity for HABA in comparison to wt avidin, Avd(N118M) or cpAvd4 \rightarrow 3. Although we were not able to convert the ligand-binding preference of avidin to favour HABA over its natural ligand biotin, the relative binding affinity was modified over one billion-fold. The relative dissociation constant

$K_d(\text{BTN})/K_d(\text{HABA}) = 1.64 \times 10^{-10}$ for wt avidin was changed to $K_d(\text{BTN})/K_d(\text{HABA}) = 0.23$ for cpAvd4 \rightarrow 3(N118M).

Thermodynamic ITC measurements showed that the endothermic binding of HABA to wt avidin was changed dramatically to an exothermic binding reaction in the case of cpAvd4 \rightarrow 3 and cpAvd4 \rightarrow 3(N118M). However, the net entropy term ($T\Delta S$) for both decreased at the same time, which is thermodynamically the reason why even higher affinities were not observed. It has been reported that the binding of HABA to avidin leads to displacement of five water molecules from the binding site.^[19] This displacement is favourable in terms of entropy, and although the binding of HABA is stabilized by hydrogen bonds, binding does not take place without energy contributions from the surrounding solution; this can be detected as a positive net enthalpy in ITC measurements.

These previous reasons may also clarify the observed, enhanced HABA-binding and the reduced biotin-binding properties of cpAvd4 \rightarrow 3. The avidin-biotin complex, one of the tightest non-covalent ligand-protein complexes, is known to be well optimized, having a large number of hydrogen bonds and van der Waals interactions.^[1] Since cpAvd4 \rightarrow 3 lacks the L3,4 loop, two hydrogen bonds stabilizing the interactions between wt avidin and the valeryl moiety of biotin are lost, this at least partially explains the decreased affinity of cpAvd4 \rightarrow 3 for biotin (Figure 6). On the other hand, the deletion of the L3,4 loop 'opens the lid' of the ligand-binding pocket and leaves the pocket more exposed to solvent. Interestingly, deletion of the L3,4 loop or the N118M mutation alone does not cause a clear change in the entropy term when compared to wt avidin, whereas the point mutation N118M and loop deletion together affect the entropy term such that it favours biotin binding. Favourable entropy is usually observed when water molecules are liberated upon burial of a hydrophobic surface.^[39] In the case of avidin-biotin complex, ligand binding stabilizes the structure of avidin, reducing the number of internal degrees of freedom in the system, which leads to the unfavourable contribution of the free energy (negative ΔS).^[40] In the case of cpAvd4 \rightarrow 3(N118M), the entropy term of biotin-binding changed to a positive sign, thus favouring binding, which is totally opposite to any other cases of avidins or streptavidins.^[30] The effect of the positive entropy on binding at room temperature is over 5 kcal/mol when compared to biotin binding of wt avidin and cpAvd4 \rightarrow 3(N118M).

We successfully created a novel subunit fusion form of avidin, dcAvd2, by fusing cpAvd4 \rightarrow 3 to the C-terminus of cpAvd5 \rightarrow 4. Judging from our ITC measurements, the fusion of cpAvd4 \rightarrow 3 to cpAvd5 \rightarrow 4 did not seem to affect the biotin-binding properties of the subunit; cpAvd4 \rightarrow 3 of dcAvd2 binds biotin in a manner similar to the "non-fused" form. Importantly, analysis of dcAvd2 using ITC indicated a HABA-binding enthalpy of -2.3 kcal/mol, which agrees with the theoretical value of -1.6 kcal/mol obtained by summing the separately determined enthalpies for the wt and cpAvd4 \rightarrow 3 proteins (Figure 4).

What is the added value of the "next-generation" avidins presented here? There is a significant demand for advanced tools in biotechnology and nanotechnology, especially as more and more experimental systems are miniaturized on nanometer scale. Furthermore, strept(avidin) with novel properties, especially multiple binding sites which can be specifically and separately targeted for binding, would facilitate the development of more complicated assay systems or materials than those now available.

One demand in both biotechnology and nanotechnology is for new binders for new ligands. The biotin-binding site of strept(avidin) is well known and fairly flexible physically – thus offering excellent starting material which can be modified to bind

new or existing ligands with good specificity and affinity. Although the avidin forms presented here do not yet completely fulfil these requirements with regard to biotin- and HABA-binding, the conceptual steps made here bring us closer to more sophisticated protein engineering for that purpose.

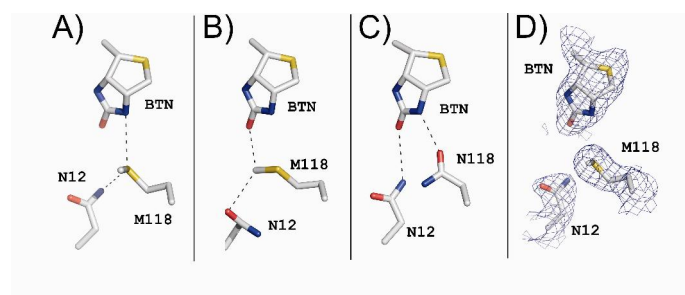


Figure 7. Mutation N118M. The putative hydrogen bonding of amino acid residues N12 and N118 with biotin (BTN; in A, C-E) or HABA (in B) are illustrated with dotted lines. A, avidin-biotin complex (chain A; PDB code 2AVI); B, avidin-HABA complex^[19]; C, subunit II of cpAvd4→3(N118M); and D, subunit IV of cpAvd4→3(N118M). Only the side-chain atoms of residues N12 and N118, or a portion of the BTN and HABA structures near these residues, are shown as sticks (C, grey; N, blue; O, red; S, yellow). E, 2Fo-Fc electron density map (blue) calculated in the absence of biotin and contoured at 1 σ around the residues shown in D is drawn.

Experimental Section

D-biotin was obtained from Sigma Chemicals. Chicken avidin, a generous gift from Belovo S. A. (Bastogne, Belgium), was used as control protein in all analyses. Low molecular weight marker was obtained from New England BioLabs or Fermentas. The concentration of the avidin solution was calculated from the absorbance at 280 nm using an extinction coefficient of 24280 M⁻¹ cm⁻¹ per binding site.

The nucleotide and amino acid sequences of the studied proteins are available at EMBL (<http://www.ebi.ac.uk/embl/>) with the following access numbers: cpAvd4→3, [AM779410](#); cpAvd4→3(N118M), [AM779411](#) (named in the database cp43N76M); dcAvd2, [AM779412](#); and dcAvd, [AJ616762](#).

Design and mutagenesis of the expression constructs

In order to construct a DNA encoding for circularly-permuted cpAvd4→3, PCR reactions were performed using oligonucleotides 5.4N2 and 3.4C1 in the first reaction, oligonucleotides 3.4N1 and 5.4C1 in the second (Supplementary data, Table S1) and the cDNA of wt avidin as template. The resulting DNA fragments were connected together using another round of PCR, using primers overlapping the PCR products from the first and second PCR reactions, together with a DNA sequence encoding the polypeptide linker connecting the original termini. The resulting DNA (cpAvd4→3) was cloned into a pGemTeasy vector containing attL recombination sites and a DNA sequence encoding for the ompA signal peptide^[41]. Mutagenesis of cpAvd4→3 was carried out with the QuikChange mutagenesis kit (Stratagene, La Jolla, CA, USA). In order to create the expression construct for dcAvd2, cpAvd4→3 was digested with the restriction enzymes BamHI and HindIII and subcloned into a pGemTeasy vector already containing the coding sequence for cpAvd5→4.^[9]

Production and purification of the expressed recombinant proteins

In order to achieve bacterial secretion for cpAvd4→3 and dcAvd2, the signal peptide from the OmpA-protein was used at the N-terminus of the expressed proteins as previously described.^[41] For expression, the DNA constructs were transferred into a pBVboostFG plasmid^[42] using the Gateway LR-cloning reaction (Invitrogen). *E. coli* BL21-AI cells (Invitrogen) were used for protein production. The expression constructs were validated using DNA sequencing (ABI PRISM 3100 Genetic Analyzer, Applied Biosystems).

Protein purification was accomplished in a single step using either 2-iminobiotin- or biotin-Sepharose 4 Fast flow (Affiland) affinity chromatography columns as previously described.^[43] Protein concentration was determined by UV/VIS spectrophotometer (ND-1000 NanoDrop Spectrophotometer) by measuring the absorbance at 280 nm and using an extinction coefficient of 23615 M⁻¹ cm⁻¹ for cpAvd4→3 and 48320 M⁻¹ cm⁻¹ for dcAvd2.

Isothermal titration calorimetry

High-sensitivity VP-ITC titration calorimetry (Microcal Inc., Northampton, MA) was used to measure heat release or uptake, which is mostly associated with ligand binding but may also reflect conformational changes. All protein samples were dialyzed against a buffer containing 50 mM sodium phosphate, 100 mM NaCl (pH 7.0) or 0.1 mM potassium acetate (pH 5.0); solid biotin and HABA were dissolved in the same dialysis buffer as protein. In order to eliminate air bubbles and long equilibration periods, prior to VP-ITC analysis, the sample solutions were degassed and preheated near the measuring temperature (25 °C). Biotin (0.1 – 0.4 mM) or HABA (1.0 – 1.2 mM) was injected into the calorimeter cell containing protein samples (0.01 – 0.1 mM) at 4-5 minute intervals. The titration cell was stirred continuously at 440-510 rpm. Both ligand-to-buffer and buffer-to-protein titrations were made in order to detect and correct the heat effect of dilution.

The heat of binding after each injection was calculated by integrating the area of the measured peak using the ORIGIN software suite (v 7.0 SR4, Originlab Corporation) tuned to the ITC instrument. All data except those related to dcAvd2 were analyzed with the “One Set of Sites” fitting procedure. The “Two Sets of Sites” procedure, designed for cases where two different kinds of binding sites exist, was used for dcAvd2. The heat (Q) evolving from each titration can be represented by equation [1]:

$$\Delta Q = V_0 \Delta H [M]_t K_a [L] / (1 + K_a [L]), \quad [1]$$

where V_0 is the volume of the calorimeter cell, ΔH is the enthalpy of binding per mole of ligand, $[M]_t$ is the total macromolecule concentration, K_a is the association constant, and $[L]$ the free ligand concentration. The values of ΔH and K_a can be determined directly from the chromatogram for a titration, whereas the change in Gibb's free energy term ΔG and its entropic component $T\Delta S$ can be calculated from equations [2] and [3]:

$$\Delta G = - RT \ln K_a \quad [2]$$

$$\Delta G = \Delta H - T\Delta S, \quad [3]$$

where R is the gas constant and T is the temperature in kelvins.

Biotin dissociation assays

The biotin-binding properties of avidin, cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 were studied by measuring the dissociation of *d*-[8,9-³H]biotin and fluorescent biotin from these proteins. The dissociation of *d*-[8,9-³H]biotin (Amersham) was determined at different temperatures as described by Klumb and colleagues^[44] The dissociation of fluorescent biotin was detected by measuring the reversal of quenching of the biotin-coupled fluorescent probe ArcDia™ BF560 (ArcDia Ltd., Turku, Finland) after addition of a 100-fold molar excess of free biotin; analyses were made in a 50 mM sodium phosphate (pH 7.0) buffer containing 650 mM NaCl at two different temperatures (25 ± 1 °C and 50 ± 1 °C) and using a PerkinElmer LS55 luminometer as previously described.^[41]

The fluorescence data were interpreted using the “single phase dissociation model” as described elsewhere.^[41] The dissociation rate constants (k_{diss}) were determined using the equation [4]:

$$-k_{\text{diss}}t = \ln(B/B_0), \quad [4]$$

where B_0 is the maximum binding (100 %) measured (the difference between the fluorescence of the free dye and that of the protein-dye complex) and B is the amount of complex measured as a function of time. The first 500 seconds of measurements were omitted from the analyses to abolish the “fast initial phase” effect, which is characteristic of the avidin-BF560-biotin interaction.^[11] The release of the fluorescent biotin was determined after one hour of measurements. In order to determine the dissociation rate for the cpAvd4→3 subunit of dcAvd2, half of the four biotin-binding sites were hypothesized to behave as in wt avidin and a ligand dissociation model was fitted to the subtracted data as previously described for other dual chain avidins.^[11]

Size exclusion chromatography

The oligomeric state of the proteins was assayed at +4 °C using the Superose 6 10/300 GL (Tricorn) gel filtration column and the ÄKTA Purifier 10 Chromatography system as previously described.^[27] A buffer containing 50 mM sodium phosphate buffer (pH 7.0) and 650 mM NaCl was used as the running buffer and gel filtration was performed with a flow rate of 0.5 ml/min. The high concentration of NaCl reduces the binding of the basic avidin protein ($pI \approx 10.5$)^[29] to the chromatography column. The gel filtration MW standards were bovine thyroglobulin, bovine gamma-globulin, chicken ovalbumin, horse myoglobin and vitamin B₁₂ (Gel Filtration Standard, Bio-Rad).

SDS-PAGE-based thermostability assay

Protein samples with biotin, HABA or both ligands were acetylated *in vitro* and the temperature-dependent dissociation of the subunits was monitored. Samples were heated to a given temperature between 25 °C and 80 °C for 20 minutes. The oligomeric state of the protein was analyzed by SDS-PAGE in the presence of 2-mercaptoethanol and the gels were stained with Coomassie blue.^[35]

Disulfide bridge reduction

Disulfide reduction was obtained using the protocol of Scigelova and colleagues.^[45] Briefly, the protein sample with an excess

amount of DL-dithiothreitol (Sigma-Aldrich, St. Louis, MO) was incubated at 70 °C for 5 min, after which the sample tube was placed on ice and diluted with an appropriate solvent for mass spectrometry.

Mass spectrometry

All mass spectrometric experiments were performed on a 4.7-T Bruker APEX-Qe hybrid Fourier transform ion cyclotron resonance (FT-ICR) instrument (Bruker Daltonics, Billerica, MA, USA), having a mass-selective quadrupole interface. Ions were produced in an external Apollo II dual ion-funnel electrospray ionization (ESI) source. For broadband spectra (m/z 400-4000), the quadrupole was operated in an RF-only mode and ions were accumulated in the second hexapole (collision cell) for 50-100 ms. Ions were then transmitted through a high voltage ion optics region, prior to “Sidekick” trapping inside the Infinity ICR cell. Ions were excited by a conventional “RF-chirp” and detected in a direct broadband mode. A total of eight co-added, 512-kWord time-domain transients were acquired, zero-filled twice, followed by magnitude calculation, fast Fourier transform and external frequency-to- m/z calibration with respect to the ions of an ES Tuning Mix (Agilent Technologies, Santa Clara, CA). All data were processed using Bruker XMASS 7.0.8 software.

Determination of HABA affinity by fluorescence spectroscopy

The HABA-binding affinity of Avd(N118M), cpAvd4→3, cpAvd4→3(N118M) and dcAvd2 was measured using fluorescence spectroscopy. The protein samples (50 nM) in neutral 50 mM phosphate buffer containing 650 mM NaCl were titrated with increasing concentrations of 4'-hydroxyazobenzene-2-carboxylic acid (HABA). The samples were excited using a wavelength of 280 nm, emission was measured at 350 nm, and the decrease in intrinsic fluorescence of avidin as a result of HABA binding was monitored. However, since HABA absorbs light also when not bound to protein, there is an error component which cannot be fully subtracted from the data. Thus only the apparent dissociation constants could be measured using this method.

The affinity of HABA to different avidin forms was determined using UV/VIS spectrometer. The assay is based for the colour change of the HABA due to complexation with avidin^[16]. The concentration of the avidin-HABA complex was determined by using extinction coefficient of 35500 cm⁻¹ at 500 nm wavelength.

X-ray structure determination

The Classics™ (Nextal Biotechnology) screen, the vapour diffusion method and sitting drops on 96-well plates (Corning Inc.) were used to search for suitable conditions for crystallization of cpAvd4→3(N118M). Using a 10:1 (v/v) ratio of protein and ligand, a protein solution of cpAvd4→3(N118M) (1.7 mg/ml) containing 50 mM sodium acetate (pH 4) and 100 mM sodium chloride was mixed with a D-biotin (Sigma) solution (1 mg/ml) containing 5 mM Tris (pH 8.8) and 8 mM CHES (pH 9.5). After optimization, crystals formed at 22 °C in conditions where 0.6–1.2 μl of the sample solution and 0.5–0.7 μl of a well solution containing 0.18–0.20 M magnesium chloride, 0.09–0.10 M Tris-HCl (pH 8.5) and 27–30% (w/v) PEG 4000 were used. For data collection, two “cubic”-looking crystals with dimensions less than 0.2 x 0.2 x 0.2 mm were used. The first data set (2.6 Å resolution) was collected from a flash-frozen crystal using MPD as a cryoprotectant (0.7 μl

of 100% MPD was added to the crystallization drop immediately prior to freezing) at the MAX-lab beam line I911_2 (Lund, Sweden) at 100 K using a MarCCD detector. The second, higher resolution (1.9 Å) data set was later collected at the ESRF beam line ID-29 (Grenoble, France) using a 100 K liquid nitrogen stream (Oxford Cryosystem) and an ADSC detector. Sodium formate was used as cryoprotectant (1 µl of 4 M sodium formate was added to the crystallization drop immediately prior to freezing).

Both data sets were processed with programs from the XDS program package.^[33] The 2.6 Å data set was used to solve the structure of cpAvd4→3(N118M) by molecular replacement with the program Phaser^[46] from the CCP4i suite^[47, 48]; four copies of the polypeptide chain A of a known avidin structure (PDB entry 2AVI)^[1] were used as search models. After molecular replacement, the initial 2.6 Å X-ray model of cpAvd4→3(N118M) was extended to higher resolution and completed using the 1.9 Å diffraction data: 1) the initial model was selected as input for automatic model building with ARP/wARP,^[49] 2) the structure was refined with Refmac5^[34] (TLS & restrained refinement)^[50] and 3) modified and rebuilt with Coot.^[51] Solvent atoms and other non-protein atoms were added to the model either by an automatic procedure in ARP/wARP^[52] or in Coot, and the final structure was analyzed using the inbuilt tools of Coot. The data collection and structure determination statistics for the 1.9 Å structure of cpAvd4→3(N118M) are summarized in Table 4, and the coordinates and structure factors of it have been deposited in the Protein Data Bank with entry codes 2JGS and 2JGSSF, respectively.

Miscellaneous methods

The solvent accessible surface areas were calculated with the program Areaimol^[53] of the CCP4i suite.^[47, 48] Figures 5-7 were created with the PyMol Molecular Graphics System.^[54] CorelDRAW X3 was used to edit Figures 5-7. The Supplementary figure S3 was generated using VMD 1.8.6.^[55]

Acknowledgements

This work was supported by the ISB (National Graduate School in Informational and Structural Biology), the Academy of Finland, the Sigrid Jusélius Foundation, and the Foundation of Åbo Akademi (Center of Excellence in Cell Stress). We would like to thank the staff at the MAX-lab beam line I911 and ESRF beam line ID-29 for excellent support. We acknowledge support by the European Community - Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme. We thank Ulla Kiiskinen for excellent technical assistance and Jarkko Valjakka for valuable discussions.

Keywords: avidin, protein engineering, circular permutation, thermodynamics, X-ray structure

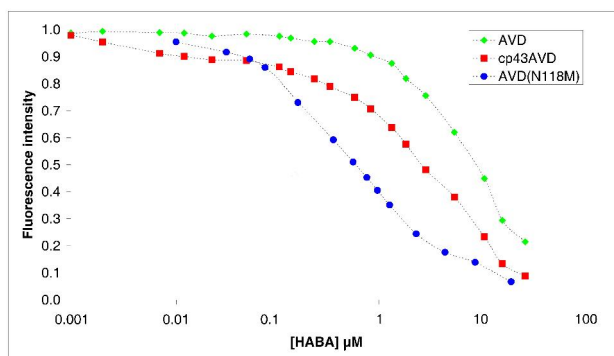
- [1] Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L. *Proc.Natl.Acad.Sci.U.S.A.* 1993, **90**, 297-5080.
- [2] Chaiet, L.; Wolf, F. J. *Arch.Biochem.Biophys.* 1964, **106**, 1-5.
- [3] Nordlund, H. R.; Hytonen, V. P.; Laitinen, O. H.; Kulomaa, M. S. *J.Biol.Chem.* 2005, **280**, 13250-13255.
- [4] Helppolainen, S. H.; Nurminen, K. P.; Maatta, J. A.; Halling, K. K.; Slotte, J. P.; Huhtala, T.; Liimatainen, T.; Yla-Herttuala, S.; Airene, K. J.; Narvanen, A.; Janis, J.; Vainiotalo, P.; Valjakka, J.; Kulomaa, M. S.; Nordlund, H. R. *Biochem.J.* 2007, **405**, 397-405.
- [5] Wilchek, M.; Bayer, E. *Meth.Enzymol.* 1990, **184**, 29-51.
- [6] Wilchek, M.; Bayer, E. A. *Biomol.Eng.* 1999, **16**, 1-4.
- [7] Laitinen, O. H.; Hytonen, V. P.; Nordlund, H. R.; Kulomaa, M. S. *Cell Mol.Life Sci.* 2006, **63**, 2992-3017.
- [8] Laitinen, O. H.; Nordlund, H. R.; Hytonen, V. P.; Kulomaa, M. S. *Trends Biotechnol.* 2007, **25**, 269-277.
- [9] Nordlund, H. R.; Laitinen, O. H.; Hytönen, V. P.; Uotila, S. T.; Porkka, E.; Kulomaa, M. S. *J.Biol.Chem.* 2004, **279**, 36715-36719.
- [10] Nordlund, H. R.; Hytonen, V. P.; Horha, J.; Maatta, J. A.; White, D. J.; Halling, K.; Porkka, E. J.; Slotte, J. P.; Laitinen, O. H.; Kulomaa, M. S. *Biochem.J.* 2005, **392**, 485-491.
- [11] 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. *Proteins.* 2005, **61**, 597-607.
- [12] Chu, V.; Freitag, S.; Le Trong, I.; Stenkamp, R. E.; Stayton, P. S. *Protein Sci.* 1998, **7**, 848-859.
- [13] Aslan, F. M.; Yu, Y.; Mohr, S. C.; Cantor, C. R. *Proc.Natl.Acad.Sci.U.S.A.* 2005, **102**, 8507-8512.
- [14] Ellison, D.; Hinton, J.; Hubbard, S. J.; Beynon, R. J. *Protein Sci.* 1995, **4**, 1337-1345.
- [15] Weber, P. C.; Pantoliano, M. W.; Simons, D. M.; Salemme, F. R. *J.Am.Chem.Soc.* 1994, **116**, 2717-2724.
- [16] Green, N. M. *Biochem.J.* 1965, **94**, 23C-24C.
- [17] Green, N. M. *Meth.Enzymol.* 1970, **18**, 418-424.
- [18] Hendrickson, W. A.; Pahler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P. *Proc.Natl.Acad.Sci.U.S.A.* 1989, **86**, 2190-2194.
- [19] Livnah, O.; Bayer, A.; Wilchek, M.; Sussman, J. L. *FEBS.* 1993, **328**, 165-168.
- [20] Keinänen, R. A.; Wallen, M. J.; Kristo, P. A.; Laukkanen, M. O.; Toimela, T. A.; Helenius, M. A.; Kulomaa, M. S. *Eur.J.Biochem.* 1994, **220**, 615-621.
- [21] Hytönen, 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. *J.Biol.Chem.* 2004, **279**, 9337-9343.
- [22] Repo, S.; Paldanius, T. A.; Hytonen, V. P.; Nyholm, T. K.; Halling, K. K.; Huuskonen, J.; Pentikainen, O. T.; Rissanen, K.; Slotte, J. P.; Airene, T. T.; Salminen, T. A.; Kulomaa, M. S.; Johnson, M. S. *Chem.Biol.* 2006, **13**, 1029-1039.
- [23] Hinds, B. J.; Chopra, N.; Rantell, T.; Andrews, R.; Gavalas, V.; Bachas, L. G. *Science.* 2004, **303**, 62-65.
- [24] Keren, K.; Berman, R. S.; Buchstab, E.; Sivan, U.; Braun, E. *Science.* 2003, **302**, 1380-1382.
- [25] Paganelli, G.; Bartolomei, M.; Grana, C.; Ferrari, M.; Rocca, P.; Chinol, M. *Neuro.Res.* 2006, **28**, 518-522.
- [26] Walsh, C. T. In *Posttranslational Modification of Proteins: Expanding Nature's Inventory* Murdzek, J.; Roberts and Company Publishers:Greenwood village, Colorado USA, 2005; pp 121-149.
- [27] Hytönen, V. P.; Laitinen, O. H.; Grapputo, A.; Kettunen, A.; Savolainen, J.; Kalkkinen, N.; Marttila, A. T.; Nordlund, H. R.; Nyholm, T. K.; Paganelli, G.; Kulomaa, M. S. *Biochem.J.* 2003, **372**, 219-225.
- [28] 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. *J.Biol.Chem.* 2005, **280**, 10228-10233.
- [29] Green, N. M. *Adv.Protein Chem.* 1975, **295**, 85-133.
- [30] Hyre, D. E.; Le Trong, I.; Merritt, E. A.; Eccleston, J. F.; Green, N. M.; Stenkamp, R. E.; Stayton, P. S. *Protein Sci.* 2006, **15**, 459-467.
- [31] Pugliese, L.; Coda, A.; Malcovati, M.; Bolognesi, M. *J.Mol.Biol.* 1993, **231**, 698-710.
- [32] DeTitta, G. T.; Edmonds, J. W.; Stallings, W.; Donohue, J. *J.Am.Chem.Soc.* 1976, **98**, 1920-1926.
- [33] Kabsch, W. *J. Appl. Cryst.* 1993, **26**, 795-800.
- [34] Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.D Biol.Crystallogr.* 1997, **53**, 240-255.
- [35] Bayer, E. A.; Ehrlich-Rogozinski, S.; Wilchek, M. *Electrophoresis.* 1996, **17**, 1370, 1319-1324.
- [36] Grishin, N. V. *J.Struct.Biol.* 2001, **134**, 167-185.
- [37] Steiner, T. *Angew.Chem.Int.Ed Engl.* 2002, **41**, 49-76.

- [38] Pazy, Y.; Kulik, T.; Bayer, E. A.; Wilchek, M.; Livnah, O. *J.Biol.Chem.* 2002, **277**, 30892-30900.
- [39] Smithrud, D. B.; Wyman, T. B.; Diederich, F. *J.Am.Chem.Soc.* 1991, **113**, 5420-5426.
- [40] Williams, D. H.; Stephens, E.; O'Brien, D. P.; Zhou, M. *Angew.Chem.Int.Ed Engl.* 2004, **43**, 6596-6616.
- [41] Hytönen, V. P.; Laitinen, O. H.; Airene, T. T.; Kidron, H.; Meltola, N. J.; Porkka, E.; Hörhä, J.; Paldanius, T.; Määttä, J. A.; Nordlund, H. R.; Johnson, M. S.; Salminen, T. A.; Airene, K. J.; Ylä-Herttuala, S.; Kulomaa, M. S. *Biochem.J.* 2004, **384**, 385-390.
- [42] Laitinen, O. H.; Airene, K. J.; Hytonen, V. P.; Peltomaa, E.; Mahonen, A. J.; Wirth, T.; Lind, M. M.; Makela, K. A.; Toivanen, P. I.; Schenkwein, D.; Heikura, T.; Nordlund, H. R.; Kulomaa, M. S.; Ylä-Herttuala, S. *Nucleic Acids Res.* 2005, **33**, e42 1-10.
- [43] Laitinen, O. H.; Marttila, A. T.; Airene, K. J.; Kulik, T.; Livnah, O.; Bayer, E. A.; Wilchek, M.; Kulomaa, M. S. *J.Biol.Chem.* 2001, **276**, 8219-8224.
- [44] Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry.* 1998, **37**, 7657-7663.
- [45] Scigelova, M.; Green P. S.; Giannokopoulos, A. E.; Rodger, A.; Crout, D. H. G.; Derrick, P. J. *Eur. J. Mass Spectrom.* 2001, **7**, 29-34.
- [46] McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. *Acta Crystallogr.D Biol.Crystallogr.* 2005, **61**, 458-464.
- [47] Collaborative Computational Project, Number 4. *Acta Crystallogr.D Biol.Crystallogr.* 1994, **50**, 760-763.
- [48] Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. *Acta Crystallogr.D Biol.Crystallogr.* 2003, **59**, 1131-1137.
- [49] Perrakis, A.; Morris, R.; Lamzin, V. S. *Nat.Struct.Biol.* 1999, **6**, 458-463.
- [50] Winn, M. D.; Isupov, M. N.; Murshudov, G. N. *Acta Crystallogr.D Biol.Crystallogr.* 2001, **57**, 122-133.
- [51] Emsley, P.; Cowtan, K. *Acta Crystallogr.D Biol.Crystallogr.* 2004, **60**, 2126-2132.
- [52] Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr. A.* 1991, **47**, 110-119.
- [53] Lee, B.; Richards, F. M. *J.Mol.Biol.* 1971, **55**, 379-400.
- [54] DeLano, W. L. *Curr.Opin.Struct.Biol.* 2002, **12**, 14-20.
- [55] Humphrey, W.; Dalke, A.; Schulten, K. *J.Mol.Graph.* 1996, **14**, 33-8, 27-8.
-

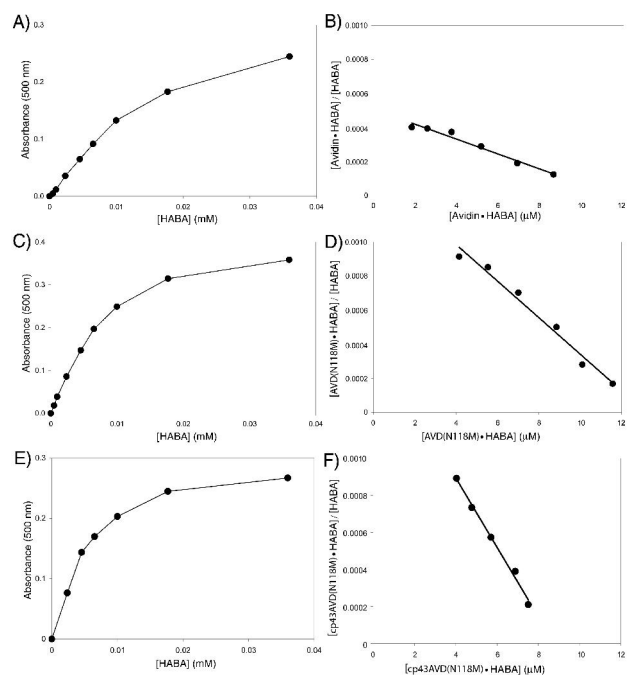
Received: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

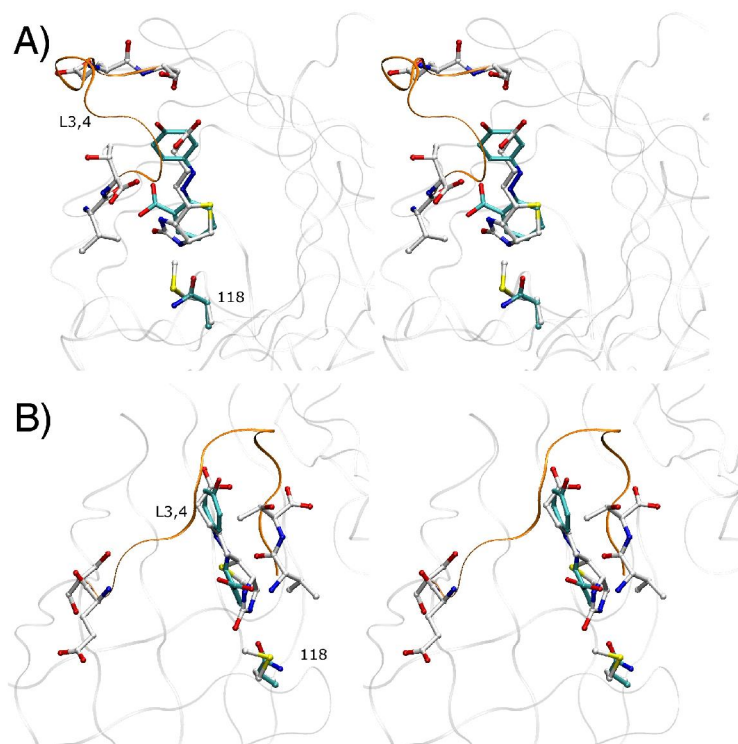
Supplementary data



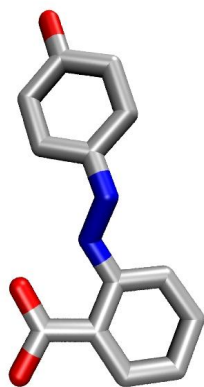
Supplementary Figure S1. Quenching of the intrinsic tryptophan fluorescence of different avidins by HABA. Protein sample in 50 mM Na-PO pH 7.0 containing 650 mM NaCl was titrated with HABA. The measured fluorescence intensity (excitation wavelength 280 nm, emission measured at 350 nm) is plotted normalized by the intensity of the free protein. The decrease in intensity caused by the absorbance of unbound HABA is not subtracted from the data. X-axis is logarithmic and data points are connected by dotted line. Measurement is carried out with PerkinElmer LS55 instrument at 25°C and the sample is continuously mixed with magnetic stirrer.



Supplementary Figure S2. The affinity of HABA to different avidin forms was assayed using UV/VIS spectrometer. A, The absorbance at 500 nm wavelength measured for avidin in the presence of various concentrations of HABA. B, Scatchard analysis for the binding reaction shown in (A) and a linear fit to the data. C, Absorbance of Avd(N118M)-HABA mixture at 500 nm wavelength plotted over HABA concentration. D, Scatchard analysis for the binding reaction shown in (C) and a linear fit to the data. E, Absorbance of cp43AVD(N118M)-HABA mixture at 500 nm wavelength plotted over HABA concentration. F, Scatchard analysis for the binding reaction shown in (E) and a linear fit to the data. The measured points are connected by lines in (A), (C), and (E). Extinction coefficient value of 35500 cm^{-1} at 500 nm wavelength was used for the avidin-HABA-complexes.



Supplementary Figure S3. Ligand binding to cpAvd4→3 as compared to avidin-biotin and avidin-HABA. The structure of cpAvd4→3(N118M) is superimposed with avidin-biotin^[1] (PDB 2AVI) and avidin-HABA^[19] complexes. The residues of cpAvd4→3 at the ends of the L3,4 are shown as white sticks whereas the rest of the protein is shown as gray ribbon. The L3,4 of wt avidin is shown as orange ribbon model representing the closed state of the loop (biotin bound). The biotin bound to cpAvd4→3 is shown as white sticks. The side chain of the residue 118 is also shown for cpAvd4→3 (methionine, white sticks) and for wt avidin (asparagine, cyan sticks). The HABA of avidin-HABA complex^[19] is shown as cyan sticks. A, Stereo view from the top of the binding site. B, Stereo view from the side of the binding site. The figure was generated using VMD 1.8.6^[55].



Supplementary Figure S4. Structure of HABA. The coordinates of HABA were obtained from avidin-HABA complex^[19]. The figure was generated using VMD 1.8.6^[55].

Table S1. Oligonucleotides used in polymerase chain reactions (PCR)	
Oligonucleotide	Sequence
5.4N2	5'-ATATAGGATCCAAGAGTCACCACTGCAT-3'
3.4C1	5'-CTACATCACAGCCGTAACAT-3'
3.4N1	5'-GGAGGCTCCGGAGGCTCCGCCAGAAAGTGCTCGCTG-3'
5.4C1	5'-GGAGCCTCCGGAGCCTCCCTCCTTCTGTGTGCGCAG-3'