Chicken avidin-related proteins show altered biotin-binding and physico-chemical properties as compared with avidin

Olli H. LAITINEN*[1], Vesa P. HYTÖNEN*, Mervi K. AHLROTH*, Olli T. PENTIKÄINEN†, Ciara GALLAGHER*, Henri R. NORDLUND*, Vladimir OVOD‡, Ari T. MARTTILA*, Eevaleena PORKKA*, Sanna HEINO*[2], Mark S. JOHNSON†, Kari J. AIRENNE*[1] and Markku S. KULOMAA*[3]

*Department of Biological and Environmental Science, University of Jyväskylä, P.O. Box 35, FIN-40351 Jyväskylä, Finland, †Department of Biochemistry and Pharmacy, Åbo Akademi University, P.O. Box 66, FIN-20521 Turku, Finland, and ‡Institute of Medical Technology, University of Tampere, FIN-33014 Tampere, Finland

Chicken avidin and bacterial streptavidin are proteins familiar from their use in various (strept)avidin–biotin technological applications. Avidin binds the vitamin biotin with the highest affinity known for non-covalent interactions found in nature. The gene encoding avidin (*AVD*) has homologues in chicken, named avidin-related genes (*AVR*s). In the present study we used the *AVR* genes to produce recombinant AVR proteins (AVRs 1, 2, 3, 4/5, 6 and 7) in insect cell cultures and characterized their biotin-binding affinity and biochemical properties. Amino acid sequence analysis and molecular modelling were also used to predict and explain the properties of the AVRs. We found that the AVR proteins are very similar to avidin, both structurally and functionally. Despite the numerous amino acid substitutions in the subunit interface regions, the AVRs form extremely stable tetramers similar to those of avidin. Differences were found in some physico-chemical properties of the AVRs as compared with avidin, including lowered pI, increased glycosylation and, most notably, reversible biotin binding for two AVRs (AVR1 and AVR2). Molecular modelling showed how the replacement Lys^{111}→isoleucine in AVR2 alters the shape of the biotin-binding pocket and thus results in reversible binding. Both modelling and biochemical analyses showed that disulphide bonds can form and link monomers in AVR4/5, a property not found in avidin. These, together with the other properties of the AVRs described in the present paper, may offer advantages over avidin and streptavidin, making the AVRs applicable for improved avidin–biotin technological applications.

Key words: avidin–biotin technology, biotin-binding protein, molecular evolution, molecular modelling, structure–function relationship.

INTRODUCTION

Chicken avidin and bacterial streptavidin are known for their extraordinarily high affinity for a water-soluble vitamin, biotin [1]. Owing to their high affinity and specificity for biotin, avidin and streptavidin have been utilized in numerous applications in life sciences, including purification, labelling and targeting of various materials. The methodology has been collectively referred to as (strept)avidin–biotin technology [1,2].

Avidin and streptavidin are tetrameric proteins consisting of four identical subunits. Their three-dimensional (3-D) structures have been solved by X-ray crystallography, and their tertiary and quaternary structures show astonishing similarity, despite relatively low amino acid sequence identity [3–6]. Most of the essential biotin-binding residues are conserved, and the affinity for biotin is nearly identical in both avidin and streptavidin. In addition to high biotin-binding capacity, avidin and streptavidin tetramers show remarkable stability at high temperatures, as well as under strongly denaturing conditions. The stability increases even more upon biotin binding [7–9].

There are, however, prominent differences in some biochemical properties of avidin and streptavidin. Streptavidin is not glycosylated and lacks cysteine residues capable of forming disulphide bridges, whereas avidin has a carbohydrate side chain and one intramolecular disulphide bond. Furthermore, the pI of streptavidin is acidic (pI ≈ 6) in contrast with basic in avidin (pI ≈ 10.5) [10]. From the higher-order structural point of view, the interfaces between the subunits are built somewhat differently in these two proteins. The striking similarity in some properties with simultaneous disparity in others, coupled with a long evolutionary distance between them, make avidin and streptavidin an ideal model system for studying the evolution of ligand-binding proteins.

The avidin gene (*AVD*) in chicken has homologues, called avidin-related genes (*AVR*s) [11–13]. The number of *AVR* genes seems to vary between individuals [14], but seven different genes, *AVR1–AVR7*, have been cloned and sequenced. Two of the genes, *AVR4* and *AVR5*, are 100 % identical in their coding sequence, exhibiting only a single nucleotide difference in their 5′-flanking region, whereas the others are 94–99 % identical to each other. The identity between *AVD* and the different *AVR*s ranges from 91 to 95 % [12,13]. mRNAs for *AVR1*, *AVR2* and *AVR3* have been detected in chicken under inflammatory conditions ([15]; P. Lappalainen, T. Kunnas, E.-L. Punnonen and M. S. Kulomaa, unpublished work), but it is not known whether they are expressed as proteins. The putative avidin-

Abbreviations used: *AVR*, avidin-related gene; *AVD*, avidin gene; AVR, avidin-related protein; 3-D, three-dimensional; RT-PCR, reverse-transcription PCR; Endo H_f, recombinant endoglycosidase H fused ($_f$) to maltose-binding protein; PNGase F, peptide:N-glycosidase F; IgG–AP, IgG–alkaline phosphatase conjugate; L3 (etc.), loops designated by their numeral orders; β3 (etc.), β-sheets designated by their numeral orders; K_d, dissociation constant; k_{diss}, dissociation rate constant; k_{ass}, association rate constant; CPK, Corey–Pauling–Koltun.

[1] Present address: A. I. Virtanen Institute, Department of Molecular Medicine, University of Kuopio, P.O. Box 1627, Kuopio, Finland.

[2] Present address: National Public Health Institute, Department of Molecular Medicine, University of Helsinki, P.O. Box 104, FIN-00251 Helsinki, Finland.

[3] To whom correspondence should be addressed (e-mail markku.kulomaa@csc.fi).

related proteins (AVRs), as deduced from their nucleotide sequences, are 74–81 % identical with avidin and 85–100 % identical with each other.

In our previous studies on the structure–function relationship of avidin, we employed a variety of different approaches to examine the extent to which amino acid residues of avidin can be mutated without affecting biotin binding and stability. One successful strategy has been to use an evolutionary approach. For example, we have engineered avidin according to the putative AVR sequences to lower the pI of avidin down to 4.7, as well as to remove its N-glycosylation site [16,17]. Also, the avidin-like domain of sea-urchin fibropellins was used as a model to dissociate the avidin (and streptavidin) tetramer into stable dimers [18]. A structural approach was used in designing mutations to produce a monomeric avidin [19].

The successful use of AVR sequences as models to improve the properties of avidin encouraged us to perform a detailed analysis of the sequence variability in terms of modelled 3-D structures for the putative AVR proteins. The AVRs seem to be well conserved, with most of the biotin-binding residues preserved. However, they exhibit a number of amino acid replacements that may affect their physico-chemical characteristics as compared with avidin. The structural and functional properties of the AVR proteins are of considerable interest to new avidin–biotin technology.

EXPERIMENTAL

Sequence comparison and molecular modelling

The EMBL database accession numbers for the *AVR* genes used to deduce the corresponding amino acid sequences are as follows: *AVR1*, Z21611; *AVR2*, Z21554; *AVR3*, Z21612; *AVR4*, Z22883, *AVR6*, AJ237658; *AVR7*, AJ237659. *AVR4* represents also the identical *AVR5*, and we therefore use here the term *AVR4/5*. The *AVR* cDNAs were translated with the GCG software package program Map (Genetic Computer Group, Madison, WI, U.S.A.). The theoretical pI for the AVRs and avidin was determined by using the GCG-package program Peptidesort. The comparison of the sequences of avidin and AVR1–AVR7 was performed using Malign [20,21] in the BODIL Modelling Environment (J. Lehtonen, V.-V. Rantanen, D.-J. Still and M. S. Johnson, unpublished work). Model structures of AVR1–AVR7 were made using the rapid homology-modelling program HOMODGE in BODIL on the basis of the X-ray structure of chicken avidin in complex with biotin {Research Collaboratory for Structural Bioinformatics PDB (Protein Data Bank) code: 1avd; [6]}. HOMODGE reduces errors in models of closely related proteins by using, as much as possible, atomic coordinates from the known structure to build the conserved main-chain and side-chain structures. Where the side chains differ, HOMODGE uses a rotamer library to select the side-chain conformation for optimal contacts.

Production and purification of recombinant AVR proteins

The *AVR1–4/5* genes [11,12] were cloned into the *Eco*RI–*Hin*dIII sites of pGEM4 and were *in vitro* transcribed and spliced using Ribomax and RNA Splicing System kits (Promega) according to the manufacturer's instructions. The cDNAs were then produced by reverse-transcription (RT)-PCR and further amplified by PCR using the oligonucleotide primers: AK33 (5′-CTGCTAGATCTATGGTGCACGCAACCTCCCC-3′) and AK44 (5′-GTTGCAAGCTTTGCGGGGCCATCCT-3′) containing *Bgl*II and *Hin*dIII restriction sites respectively. After cutting with *Bgl*II and *Hin*dIII, the AVR1–4/5 cDNAs were cloned into *Bam*HI and *Hin*dIII sites of pFASTBAC. For *AVR6* and *AVR7*, cDNAs were produced by subcloning the corresponding genes [13] into pDsRed1 (ClonTech) where the red-fluorescent-protein-encoding region had been removed, and by subsequent transfection of the constructs into NIH/3T3 cells. Total RNA was extracted from the cells using the SV Total RNA Isolation System (Promega). The AVR6 and AVR7 cDNAs were produced by RT-PCR (RobusT RT-PCR Kit; Finnzymes, Espoo, Finland) using the oligonucleotide primers AK33 and AK44 to produce *Bgl*II and *Hin*dIII restriction sites to the 5′- and 3′-ends of the cDNAs respectively. The synthesized cDNAs were cloned into *Bam*HI/*Hin*dIII-digested pFASTBAC1, the cloning vector for the Bac-To-Bac Baculovirus Expression System (Gibco BRL, Life Technologies, Gaithersburg, MD, U.S.A.). The nucleotide sequences of the cDNAs were confirmed by sequencing. The virus vectors for producing the AVR proteins were constructed and amplified according to the Bac-To-Bac system instructions. Recombinant AVR proteins were produced in *Spodoptera frugiperda* (Sf9) insect cells as previously reported [22]. Proteins were purified from the cells using affinity chromatography on a 2-iminobiotin (AVR1, AVR3, AVR4/5–7) and/or biotin–agarose column (AVR1 and AVR2) as previously described [18]. AVR2 was eluted from biotin–agarose with 1 M acetic acid and the eluted fractions were immediately neutralized with NaOH. The elution of AVR1 was achieved using 1 M HCl, followed by neutralization with Tris (1 g/ml).

Biotin-binding analyses

The biotin-binding characteristics of the AVR proteins were studied as previously described using the IAsys optical biosensor (Thermo Labsystems, Helsinki, Finland) [16,18]. Affinities were measured for 2-iminobiotin or biotin at 20 °C using a stirring power of 100 %. Briefly, the protein in biotin-free buffer was allowed to bind to the biotin-coated surface of the IAsys cuvette. As a negative control, the protein was saturated with an excess of biotin prior to addition into the cuvette. Excess of biotin was also maintained during analysis. Under these conditions, binding to the biotin surface was regarded as unspecific and used as a measure of baseline variation. Steady-state equilibrium in the experimental cuvette was considered to be attained after no additional binding was detected as compared with the negative control.

In order to study the reversibility of biotin binding, avidin and the AVRs were allowed to bind to a biotin-coated cuvette. After measuring the maximal binding (A) as described above, biotin-saturated (0.17 mg/ml) buffer was injected into the cuvette. The amount of protein remaining bound (B) was measured after no further dissociation was seen as compared with the negative control. The dissociation was measured for at least 20 min. The percentage reversibility was calculated as follows:

$$Reversibility\ (\%) = 100 \times (A - B)/A$$

$A - B$ is the amount of liberated proteins after addition of biotin.

Structural analyses

The heat-stability of the AVRs was studied by using an SDS/PAGE-based method [23]. In this method, each protein sample was divided into two fractions, one of which was kept biotin-free and the other was supplemented with an excess of biotin. The samples were then mixed with denaturing SDS/PAGE sample buffer (containing β-mercaptoethanol) and incubated at a given temperature for 20 min. After heat treatment, the samples were subjected to SDS/PAGE and visualized by

Table 1 Amino acid substitutions in the interface regions between AVRs and avidin

Amino acids, which in avidin participate in intersubunit interactions but are substituted in AVRs. Amino acid substitutions: H50L (etc.), $His^{50} \rightarrow$ leucine (etc.). Abbreviation: h-phobic, hydrophobic.

Amino acid …. Change and type of change …	H50L Basic/ h-phobic	N54H Polar/ basic	N57A Polar/ neutral	R59A Basic/ neutral	N69H Polar/basic	K71N Basic/polar	T80V Neutral/ h-phobic	M96K h-phobic basic	S101L polar/ h-phobic	I117N/Y h-phobic/ polar
Occurrence in AVRs	All	All	All	All	All	All	All but 4/5	All but 4/5	All but 4/5	AVR2, AVR3, N; the rest, Y
Side/main chain	M*/O†	S‡	S	S	S	S	M/O	S	M/O	S
Hydrogen bond/VdW§	2¶ H	7 H	1 H	2 H	3 H	1 H	2 H	VdW	1 H	VdW
Interface contacts	1/4	1/4	1/4	1/4	1/4	1/4	1/4	1/3 and 1/4	1/4	1/3

* Main-chain contact.
† Oxygen atom of main chain.
‡ Side-chain contact.
§ Van der Waals contacts.
¶ Number of hydrogen bonds.

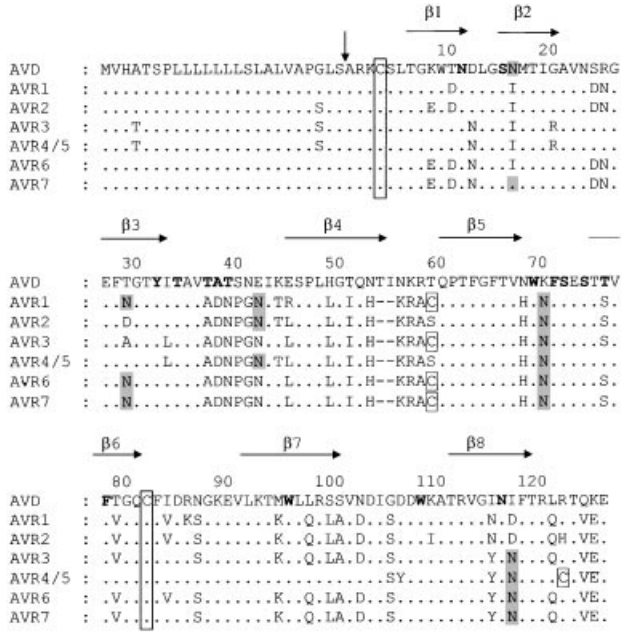

Figure 1 Multiple sequence alignment of avidin (AVD) and the AVRs

Dots indicate identical amino acids in AVRs as compared with avidin, and the two amino acid deletion in AVRs is indicated by dashes. Horizontal arrows designate the β-sheets of avidin, and the vertical arrow indicates the cleavage site of the signal peptide in avidin. Biotin-binding residues are in **bold** in the avidin sequence, and the N-glycosylation site of avidin as well as the potential N-glycosylation sites of the AVRs are highlighted with grey. Cysteine residues are boxed.

staining with Coomassie Brilliant Blue. The relative proportions of tetrameric and monomeric forms of the AVR proteins (with avidin as the control) were detected. A similar assay, with β-mercaptoethanol omitted from the sample buffer, was performed to examine the presence of intersubunit disulphide bridges. Sensitivity to proteinase K was studied in both the absence and presence of biotin as described in [18]. The pI of each AVR was determined by isoelectric focusing as previously reported [16]. The glycosylation patterns of the AVR proteins were studied by treating the proteins with recombinant endoglycosidase H fused ($_f$) to maltose-binding protein (Endo H_f) and PNGase F (peptide: N-glycosidase F) as described in the manufacturer's (New England Biolabs) instructions.

Immunological analyses

In order to investigate whether the immunological properties of AVR proteins differ from those of avidin, an indirect ELISA analysis was utilized [24]. The wells of a 96-well plate were coated with avidin or AVR proteins (1 μg/ml) in 50 mM sodium carbonate buffer at 37 °C for 2 h, followed by washing in PBS/Tween and blocking with 1 % BSA in PBS. ELISAs were performed using two monoclonal anti-avidin antibodies produced in one of our institutions (FIT Biotech), and a polyclonal rabbit anti-avidin antibody (produced at the Laboratory Animal Center,University of Oulu, Oulu, Finland) as primary antibodies. Goat Anti-Mouse IgG–alkaline phosphatase (Bio-Rad) or goat anti-rabbit IgG–alkaline phosphatase (Bio-Rad) was used as secondary antibodies, respectively. The used signal molecule was *p*-nitrophenyl phosphate (1 mg/ml) (Sigma). A_{405} values were measured with an automated ELISA reader.

RESULTS

Sequence comparison and molecular modelling

The putative AVRs were compared with avidin to identify differences that might affect their biotin binding, structural or other biochemical properties. The biotin-binding residues are well conserved in all AVR proteins, showing only three amino acid changes (out of the 16 possible) (Figure 1). Three sequential residues involved in biotin binding (38–40; Thr-Ala-Thr in avidin) are replaced by Ala-Asp-Asn in all AVR proteins. With the exception of AVR7, the AVR proteins lack the glycosylation site (Asn^{17}) present in avidin. Instead, they all exhibit other Asn-Xaa-Ser/Thr sequences, which could be glycosylated during protein maturation along the secretion route. There is one common Asn-Xaa-Ser site in the L5 loop of all the AVR proteins. Additional Asn-Xaa-Ser/Thr sequences are distributed as follows: two in AVR1 (in strand β3 and in L3), one in AVR2 (L3), one in AVR3 (β8), two in AVR4/5 (L3 and β8), and two in AVR6 and AVR7 (β3 and β8). The potential N-glycosylation sites of the AVRs are located on the surface of the tetramers and, therefore, they are expected to be glycosylated.

The theoretical pI values for AVR3 and AVR4/5 are basic, similar to the pI of avidin (avidin, 10.4; AVR3, 10.2; and AVR4/5, 10.0). On the other hand, AVR1, AVR6 and AVR7 are neutral (pI 7.3), and AVR2 is calculated to be acidic with a pI of 4.7. All AVR proteins, except for AVR2, have three cysteine residues (Figure 1). In AVR1, AVR3, AVR6 and AVR7, the third 'extra' cysteine replaces Thr^{60} of avidin located at the

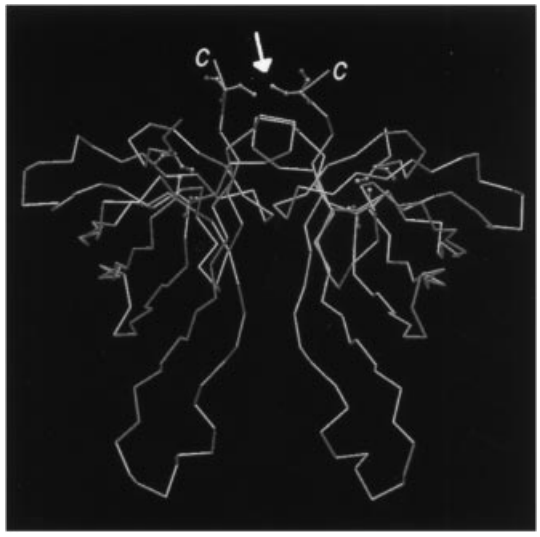

Figure 3 Model structure for two monomers of AVR4/5

Compared with the avidin structure, AVR4/5 has an additional cysteine residue in each monomer located at the C-terminus (arrow). This dimer model, based on the avidin structure, would place these two cysteine residues in an ideal location to form a disulphide bond between the two monomers, but this extra disulphide bond would not prevent the formation of a tetramer similar to that seen for avidin.

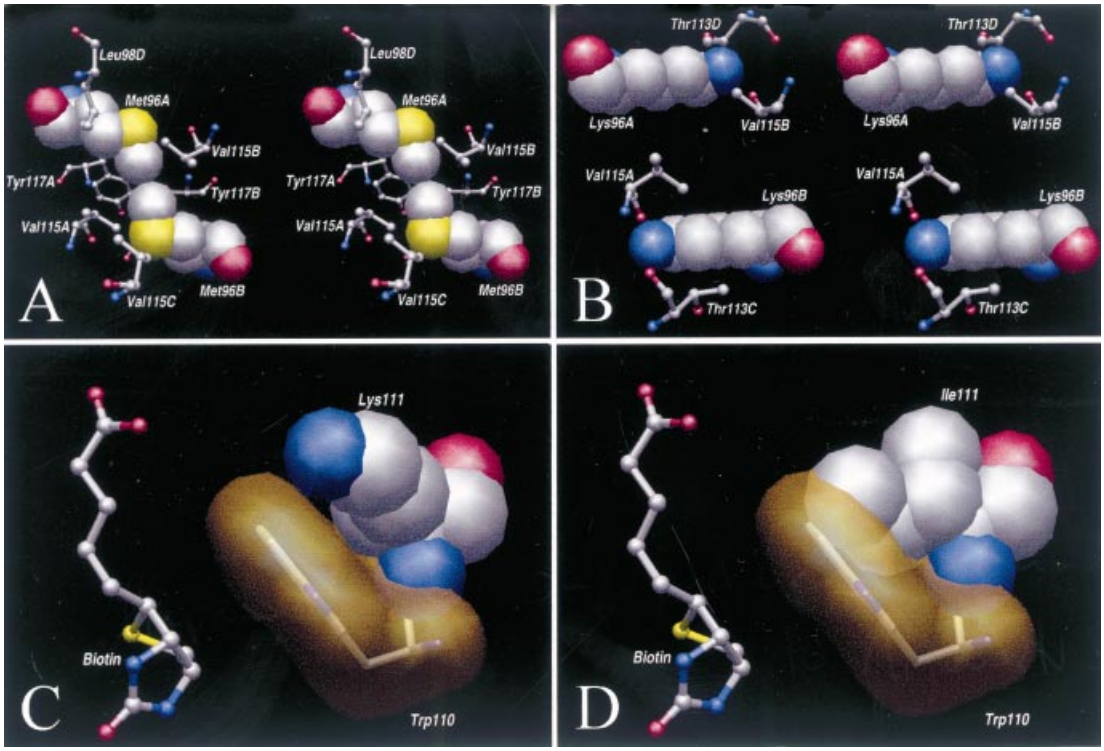

Figure 2 Re-organization of the subunit interface where sequence differences at position 96 are located (in stereo) and the effect of sequence differences at position 111 on the orientation of the side chain of Trp^{110}

The amino acid at position 96 is shown as a CPK (Corey–Pauling–Koltun) model and the surrounding amino acids are shown as ball-and-stick representations. (**A**) Met^{96} (avidin and AVR4/5); (**B**) Lys^{96} (AVR1–3, AVR6 and AVR7); (**C**) Lys^{111} (avidin, AVR1, AVR3–7); (**D**) Ile^{111} (AVR2) as CPK models. Trp^{110} is drawn as sticks along with the solvent-accessible (transparent) surface; ball-and-stick models represent the ligand biotin. Note in (**D**) the severe overlap between Ile^{111} and Trp^{110} that would force the reorientation of the tryptophan side chain, leading to weaker interactions with biotin and reduced binding affinity.

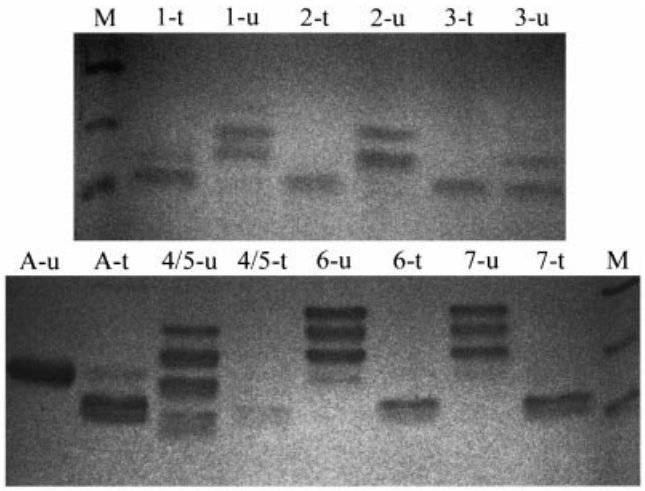

Figure 4 Glycosylation of avidin and AVRs in baculovirus-infected insect cells

Samples were either treated or non-treated with Endo H_f glycosidase followed by SDS/15%-(w/v)-PAGE and staining with Coomassie Brilliant Blue. Avidin is denoted by 'A' and the different AVRs by their corresponding numbers. 'M' indicates the low-molecular-mass markers (31, 21.5 and 14.4 kDa; Bio-Rad). Abbreviations: -u, not treated with enzyme; -t, treated with enzyme.

beginning of β5. In AVR4/5, the 'extra'cysteine residue is close to the C-terminus of the protein where Arg^{124} is found in avidin.

The interface regions between different subunits show a variable number of amino acid substitutions between the AVR proteins and avidin (see [5,6] for the interface structures of avidin). The interface between subunits 1 and 2 (as well as between 3 and 4) is perfectly conserved in all of the AVRs. In contrast, two out of three interface residues between subunits 1 and 3 (2 and 4) show changes in all of the AVR proteins, except for AVR4/5, which shows only one substitution (Ile^{117}→tyrosine). Interestingly, the tightly interacting 1–4 (2–3) subunit interface shows several amino acid substitutions in all of the AVRs. In avidin, a total of 22 residues confer intersubunit contacts within this interface. The AVR proteins have nine substitutions in this region, seven of which are found in all AVRs and two in all AVRs but AVR4/5 (Table 1).

Most of the sequence differences in the subunit interface of avidin and AVR1–AVR7 would not interfere with tetramer formation. The most critical sequence difference is located at position 96, where methionine in avidin and AVR4/5 is replaced by lysine in AVR1–AVR3 and AVR6–AVR7. In avidin and AVR4/5, Met^{96} interacts with Met^{96} from a second monomer of the tetramer (Figure 2A). When Met^{96} is replaced by the positively charged lysine residue, as seen in AVR1–AVR3 and AVR6–AVR7, it would be logical to expect that charge repulsion interferes with the formation of the tetramer. However, this does not seem to occur, but, instead, Lys^{96} from monomer 1 can form hydrogen bonds with the side-chain hydroxy group of Thr^{113} from monomer 4 and with the main-chain oxygen atom of Val^{115} from monomer 3 (Figure 2B).

As compared with avidin, the biotin-binding pocket is highly conserved in AVR1–AVR7, where only minor changes in the sequences – and therefore in the shape of the binding pocket – are observed. These changes are limited to the area surrounding the most flexible part of the biotin molecule, and thus biotin can easily compensate for these changes. The only exception occurs at position 111. In avidin, AVR1 and AVR3–AVR7, lysine is found at this position, whereas, in AVR2, isoleucine is present. In the avidin crystal structure, the hydrophobic part of the lysine side chain interacts with the indole ring of Trp^{110} keeping its position fixed (Figure 2C). The replacement of Lys^{111} by the bulky isoleucine side chain would force the side chain of Trp^{110} either to bend towards the binding pocket (effectively blocking biotin entry to the site) or away from the binding pocket (allowing the entry and exit of biotin into/from the site) (Figure 2D). Trp^{110} is one of the most critical amino acids in the binding pocket, because it conforms exactly to the shape of biotin where the ligand is most rigid. The N-linked glycosylation site on L5, although close to the biotin-binding site, is solvent-exposed. The asparagine that is glycosylated is not in the vicinity of the ligand, whereas the serine side chain is hydrogen-bonded to the carboxylate group of biotin. When biotin is absent (for example, during post-translational glycosylation) the serine side chain should be accessible to the glycosylating enzyme.

All AVRs, with the exception of AVR2, contain a third single cysteine residue (in addition to the two present in avidin) located on the surface of the proteins. Since AVR2 does not have the extra cysteine residue, there is no reason to believe that AVR2 could form tetramers in a way different from that by which

Table 2 Optical biosensor data for biotin and 2-iminobiotin binding for avidin and AVRs

Shown below are the biotin-binding properties of AVRs compared with those of avidin (AVD), determined using the IAsys biosensor. The values for avidin are from Laitinen et al. [19] and Marttila et al. [17]. Reversibility of binding was determined by using a biotin cuvette. ND, not determined.

	Value						
Parameter	AVD	AVR1	AVR2	AVR3	AVR4/5	AVR6	AVR7
K_d (M)‡	$(1.7 \pm 1.3) \times 10^{-8}$*	$(1.0 \pm 0.8) \times 10^{-7}$†	$(1.7 \pm 1.5) \times 10^{-6}$†	$\ll 10^{-8}$†	$(5.7 \pm 2.5) \times 10^{-7}$*	$\ll 10^{-8}$†	$(4.5 \pm 6.2)\ 10^{-9}$†
K_d (M)§	$(2 \pm 0.9) \times 10^{-8}$*	$(4.4 \pm 1.9) \times 10^{-8}$†	$(5.2 \pm 1.7) \times 10^{-8}$†	ND	$(9.1 \pm 3.6) \times 10^{-8}$*	ND	ND
k_{ass} ($M^{-1} \cdot s^{-1}$)	$(1.6 \pm 0.7) \times 10^4$	$(5.5 \pm 1.5) \times 10^4$	$(1.8 \pm 0.1) \times 10^4$	ND	$(1.8 \pm 0.2) \times 10^5$	ND	$(1.8 \pm 0.2) \times 10^5$
k_{diss} (s^{-1})	$(3.1 \pm 1.4) \times 10^{-4}$	$(2.4 \pm 0.8) \times 10^{-3}$	$(4.6 \pm 1.1) \times 10^{-3}$	ND	$(1.7 \pm 0.7) \times 10^{-3}$	ND	ND
Reversibility (%)	1 ± 1	18 ± 10	94 ± 3	3 ± 2	2 ± 2	5 ± 3	3 ± 3

* Measured in a 2-iminobiotin cuvette.
† Measured in a biotin cuvette.
‡ Dissociation constants were calculated from the equilibrium-response data.
§ The dissociation constants were calculated directly from the binding curves.

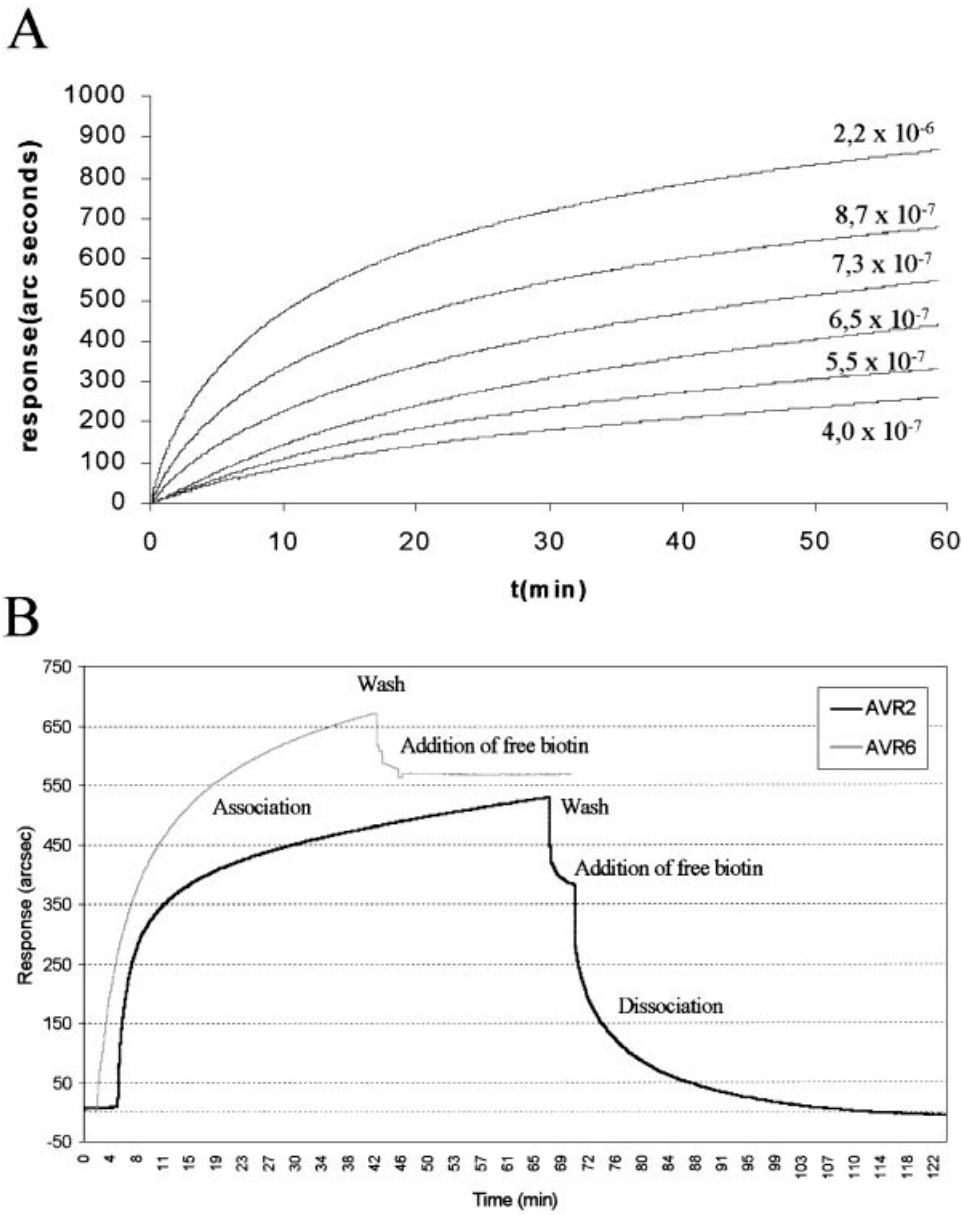

Figure 5 (A) Binding curves for 2-iminobiotin binding by AVR4/5 at different concentrations, and (B) an example of an IAsys reversibility experiment

(**A**) The first 500 s were used to calculate k_{ass}. The equilibrium state was obtained after measurement for about 1 h, depending on the concentration. The maximal binding (R_{max}) for the highest concentration was approx. 1000 arc-s. Note: 2,2 = 2.2 (etc.). (**B**) The protein was allowed to bind on to a biotin–aminosilane cuvette. After reaching equilibrium, the cuvette was washed and dissociation was measured. A surplus of free biotin was added and the measurement continued to achieve steady state. The cuvette was washed again. The values measured after binding and dissociation, and after biotin treatment and second washing step, were used to define the reversibility of the biotin binding. In this particular experiment, the reversibility was practically 0% for AVR6 and 100% for AVR2.

avidin does. The model structure for AVR4/5 (Figure 3) places the additional cysteine residue at position 124 (avidin numbering), close to the C-terminus. It is logical to expect two monomers to optimize their domain–domain interactions by forming a disulphide-bond between these unpaired cysteine residues (arrow in Figure 3), which, in the model structure of AVR4/5, are located near each other. The presence of an intermonomer disulphide bond at this location is not expected to interfere with tetramer formation. However, in AVR1, AVR3, AVR6 and AVR7, the extra cysteine residue is located in L4 at position 60 (avidin numbering), where it is impossible to form an extra disulphide bond between subunits without severely disturbing tetramer formation.

Protein expression and purification

The AVR proteins were successfully produced in Sf9 insect cells. However, all AVRs showed different patterns on SDS/PAGE as compared with avidin [22] (Figure 4). One-step 2-iminobiotin (AVR1, AVRs 3–7) – and/or biotin (AVRs 1 and 2) – agarose affinity chromatography yielded highly homogenous proteins showing no contaminants as judged by SDS/PAGE analysis.

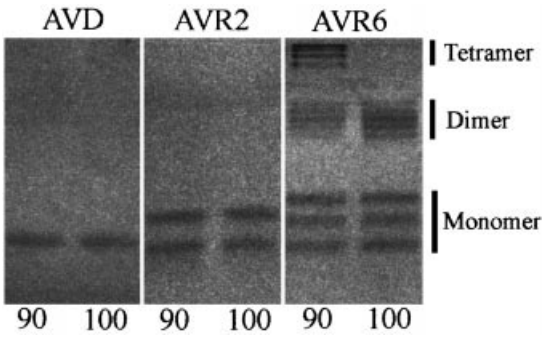

Figure 6 Non-reducing SDS/PAGE of avidin (AVD), AVR2 and AVR6 at 90 and 100 °C

The tetrameric, dimeric and monomeric forms are indicated with bars on the right side of the Figure.

Biotin-binding properties

In the reversibility assay, AVRs 3–7 showed totally irreversible biotin binding, similarly to avidin. In contrast, AVR1 exhibited 18 and AVR2 93 % reversibility following addition of free biotin (Table 2). The actual dissociation constants (K_d) were calculated for the AVRs according to the k_{ass} (association) and k_{diss} (dissociation) rate constants measured in a separate assay at different protein concentrations (varying between 10 μM and 5 nM) (Table 2 and Figure 5). The binding seemed to follow second-order or even more complex kinetics. However, the first 1000 s of the binding fitted rather well to the biphasic curve. All measured k_{diss} values followed the first-order kinetics. Because of the extremely high binding, affinities of the AVRs for biotin could not be determined except for AVR1, AVR2 and AVR7. Binding to 2-iminobiotin was determined only for AVR4/5, and it was similar to that of avidin. The lack of binding for the other AVRs was surprising, since most of them were purified using 2-iminobiotin–agarose. This may be due to the relatively short linker between 2-iminobiotin and the activated group of the IAsys cuvette that did not allow the 2-iminobiotin to penetrate deep enough into the biotin-binding pocket of the other AVRs.

Structural analyses

All AVRs exhibited heat-stability comparable with avidin in a reducing SDS/PAGE assay (results not shown). In the absence of biotin, AVR tetramers began to dissociate into monomers at around 60 °C. In the presence of biotin, a portion of AVRs remained tetrameric even upon boiling. Non-reducing SDS/PAGE showed that AVR1 and AVRs 3–7 have a tendency to form dimers. Under reducing conditions, only tetramers and monomers could be discerned. This result suggests that the dimers are cross-linked via disulphide bridges (stable enough to hold the dimers together under non-reducing conditions) and are slowly degraded into monomers by heat (Figure 6 and Table 3). The shift from a fully tetrameric state into a dimeric and monomeric state was detected upon raising the temperature gradually. AVR2 is an exception, since it disintegrated directly into monomers, similarly to native avidin.

The AVRs also showed remarkable resistance against proteolysis. In the presence of biotin, all AVRs remained intact after 16 h of proteinase K treatment. Proteolysis occurred, albeit slowly, in the absence of biotin. Taken together, AVRs showed stability similar to, or even greater than, that of avidin (results not shown).

The multiple bands observed for the AVRs in SDS/PAGE were found to be differently glycosylated forms. Treatment with Endo H_f (Figure 5) and PNGase F (results not shown) eliminated the higher-molecular-mass bands. The number of putative glycosylation sites seems to correlate well with the observed glycosylation patterns. Preliminary deglycosylation results using non-denaturing conditions suggest that N-glycan is attached to the glycosylation site in L5 in AVRs 1,2,4/5 (not shown). Isoelectric focusing showed that the pI values for the AVRs were approximately the same as expected from theoretical calculations, namely $\approx$ 7 for AVRs 1, 6 and 7, $\approx$ 5 for AVR2, and $\approx$ 10 for AVRs 3, and 4/5.

Table 3 Comparative thermostability of avidin and AVRs under non-reducing conditions

Shown below are the relative proportions of different oligomeric forms of AVRs and avidin (AVD) in SDS/PAGE under non-reducing conditions, boiled 20 min before loading on to the gel. Note the presence of dimeric forms in the case of AVR1 and AVR3–7.

	Relative proportion of oligomeric forms (%)						
Form	AVD	AVR1	AVR2	AVR3	AVR4/5	AVR6	AVR7
Tetrameric	–	20	–	–	–	–	–
Dimeric	–	40	–	50	50	50	20
Monomeric	100	40	100	50	50	50	80

Table 4 Monoclonal anti-avidin antibody and polyclonal rabbit anti-avidin antibody ELISAs for avidin (AVD) and AVRs

The values are means of two independent measurements of A_{405} nm, after 60 min colour reaction.

	A_{405}						
Antibody	AVR1	AVR2	AVR3	AVR4/5	AVR6	AVR7	AVD
Monoclonal anti-avidin							
Clone 1	0	0	0	0	0	0	0.66
Clone 2	0	0	0	0	0	0	0.64
Polyclonal anti-avidin	0.22	0.02	0.16	0.40	0.38	0.35	0.82

Immunological analyses

Polyclonal rabbit anti-avidin antibody recognized AVRs 4/5–7 more weakly than it did avidin. Furthermore, the recognition of AVRs 1 and 3 was even more diminished, and AVR2 was not recognized at all. The two monoclonal anti-avidin antibodies tested did not recognize any of the AVRs (Table 4).

DISCUSSION

The motivation for the present study was basically bipartite. First, we wanted to produce recombinant AVR proteins to examine their biochemical and functional properties, such as biotin binding, stability and immunological properties. Advantageous properties might be utilized to improve the current (strept)avidin–biotin technologies. Secondly, we wanted to reveal in detail the structural basis for the differences between avidin and the AVRs, which was addressed by careful sequence and modelling analyses.

Despite the relatively high primary sequence conservation [25–27] (Figure 1), there are interesting amino acid substitutions that make the physico-chemical properties of AVRs different from those of avidin. For example, the present study showed that the pI of some AVRs is neutral or even acidic, in contrast with the basic pI of avidin, and their glycosylation patterns also differ. Furthermore, differences in biotin binding were observed, regardless of the fact that almost all of the amino acid residues important for biotin binding in avidin [5] and streptavidin [3] are conserved in the AVRs. In previous studies, we have used the AVR sequences as models for lowering the pI of avidin [16] and to produce non-glycosylated avidin [17]. All of the pI-mutants, as well as the non-glycosylated avidins, bind biotin with high affinity and form stable tetramers. Therefore differences in pI or the lack of glycosylation at positions corresponding to residue 17 in avidin were not expected to affect biotin binding and tetramer formation of the AVRs. Consequently, the observed differences must be based on as-yet-unknown features that, while affecting biotin binding, do not impede tetramer stability.

Three biotin-bonding residues in loop L3 are altered in AVRs when compared with avidin. Two of these hydrogen-bonding interactions could be maintained despite having differing side chains at these positions in the AVRs. However, the hydrogen bond to biotin from the side chain of Thr^{38} in avidin [5] is most likely lost when the threonine residue is replaced by alanine. By comparison, streptavidin shows only one hydrogen bond to biotin in loop L3 [3,4]. Therefore the differences in L3, including the presence of additional carbohydrate on some of the AVRs, may not be detrimental for biotin binding. The greatest functional differences were observed for AVR2, where binding to 2-iminobiotin was totally abolished and binding to biotin was virtually reversible. This result is in good agreement with the modelling results that suggest that the conversion of Lys^{111} into isoleucine would alter the shape of the binding pocket.

The most interesting glycosylation site found in AVRs is located in L5. Residues of this loop, including the hydrophobic residues Trp^{70} and Phe^{72}, as well as two hydrogen-bond-forming serine residues, contact biotin. However, AVR4/5 also has a glycosylation site located in L5, and still exhibits 2-iminobiotin binding as high as that shown by avidin. Therefore it seems that, in itself, glycosylation at this site does not affect biotin binding. The contribution of glycosylation to the reduced biotin binding seen in the other AVRs cannot be ruled out. Additional carbohydrate may also explain why none of the AVRs were recognized by the monoclonal anti-avidins antibodies: the epitope(s) could be masked by the sugar moieties in the AVRs.

Another structurally interesting feature of the AVRs (except for AVR2) is the third cysteine residue. The non-reducing-SDS/PAGE results, as well as molecular modelling, suggest that this extra cysteine residue forms an intermonomeric disulphide bridge, leading to SDS/PAGE patterns consistent with dimer formation. Cross-linking of monomers is most likely to occur in AVR4/5. In contrast, other cysteine-bearing AVRs either form bigger aggregates or their quaternary structures are different from that of avidin. Their inability to bind 2-iminobiotin in IAsys analyses may reflect the presence of such structural differences.

A $His^{127} \rightarrow$ aspartate mutation introduced into the subunit interface of streptavidin prevented tetramer formation through charge repulsion [28]. Mutation of all of the residues at the 1–3 (and 2–4) interface of avidin into alanine residues also showed that the stability of the resultant mutant was reduced [19]. However, the $Met^{96} \rightarrow$ lysine substitution does not appear to interfere with tetramer formation in the AVRs, but may function to stabilize the tetramer (Figure 2). At the large 1–4 (and 2–3) interface, the substitution of Asn^{54} and Asn^{69} by histidine residues is especially interesting [5,13]. The two asparagine residues participate in a network of ten hydrogen bonds in avidin [5] at both ends of the 1–4 (2–3) dimer. Thus these two residues are involved in networks of 20 hydrogen bonds between subunits 1 and 4 (and 2 and 3) [18]. However, histidine is also capable of forming hydrogen bonds and may functionally substitute for asparagine.

There are several possible explanations for the stability of the AVRs. Substitutions at the 1–4 (2–3) interface can be complementary; while an amino acid change in one subunit decreases the interface affinity, the coincident mutation in the other subunit may restore it. Bogan and Thorn [29] have introduced a 'hot spot' model for interface contacts, where they propose that the free energy of interface binding is not evenly distributed, but is concentrated at hot spots consisting of small subsets of residues. A hot spot contains energetically important key amino acids (commonly hydrophobic) surrounded by energetically unimportant residues (called 'O-ring' residues), whose effect is to exclude bulk solvent from the hot spot. Thus the energetically less important residues may be mutated frequently without causing major effects on interface affinity and stability. Furthermore, water molecules often play an essential role in protein–protein interactions [30]. The conversion of bulky charged or polar residues into smaller ones may result in a situation where a water molecule functions to bridge the hydrogen bond previously formed directly between the bulky residue and its counterpart. Therefore the subunit interfaces of the AVRs can tolerate mutations within the O-ring zone, and some substitutions may allow water molecules in the interface without seriously affecting interface affinity.

Avidin is generally thought to act as an antimicrobial agent by depriving invading micro-organisms of biotin. It is therefore possible that the different AVRs exist to broaden the range of host defence. Recently, Zerega and co-workers [31] found that avidin is expressed in skeletal muscle and growth plate hypertrophic cartilage of the developing chicken embryo. This finding, together with their *in vitro* results suggest that by interfering with fatty acid metabolism avidin assists the terminal differentiation of chondrocytes and myoblasts. Similarly, the fibropellins expressed during ontogenesis of sea urchins contain a domain similar to that of avidin [32,33]. Fibropellins may form dimers or higher-order oligomeric structures and thereby promote protein–protein interactions during embryogenesis. The AVRs in chicken may offer comparable oligomeric scaffolds for embryonic development and/or host-defence functions.

In conclusion, the recombinant AVRs are functional and show interesting physico-chemical properties. These new biotin binders may provide advantages over avidin and streptavidin in several applications [34]. For example, AVR2 could be used in affinity-purification protocols that require mild elution conditions. Nonetheless, further analyses are required to solve the quaternary structures, precise glycosylation compositions and immunological properties of the AVRs in detail.

We thank Mrs Irene Helkala, Mrs Pirjo Käpylä, Mr Paavo Niutanen and Mr Jarno Hörhä for first-class technical assistance. This study has been supported by grants from the Academy of Finland.

REFERENCES

1 Wilchek, M. and Bayer, E. A. (1999) Foreword and introduction to the book (Strept)avidin–Biotin System. Biomol. Eng. **16**, 1–4

2 Bayer, E. A. and Wilchek, M. (1990) Application of avidin–biotin technology to affinity-based separations. J. Chromatogr. **510**, 3–11

3 Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. and Salemme, F. R. (1989) Structural origins of high-affinity biotin binding to streptavidin. Science **243**, 85–88
4 Hendrikson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A. and Phizackerley, R. P. (1989) Crystal structure of core streptavidin determined from multiwavelength anomolous diffraction of synchrotron radiation. Proc. Natl. Acad. Sci. U.S.A. **86**, 2190–2194
5 Livnah, O., Bayer, E. A., Wilcheck, M. and Sussman, J. L. (1993) Three-dimensional structures of avidin and the avidin–biotin complex. Proc. Natl. Acad. Sci. U.S.A. **90**, 5076–5080
6 Pugliese, L., Coda, A., Malcovati, M. and 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. J. Mol. Biol. **231**, 698–710
7 Green, N. M. (1990) Avidin and streptavidin. Methods Enzymol. **184**, 51–67
8 Wang, C., Eufemi, M., Turano, C. and Giartosio, A. (1996) Influence of the carbohydrate moiety on the stability of glycoproteins. Biochemistry **35**, 7299–7307
9 Gonzalez, M., Argarana, C. E. and Fidelio, G. D. (1999) Extremely high thermal stability of streptavidin and avidin upon biotin binding. Biomol. Eng. **16**, 67–72
10 Green, N. M. (1975) Avidin. Adv. Prot. Chem. **29**, 85–133
11 Keinänen, R. A., Laukkanen, M. L. and Kulomaa, M. S. (1988) Molecular cloning of three structurally related genes for chicken avidin. J. Steroid Biochem. **30**, 17–21
12 Keinänen, R. A., Wallen, M. J., Kristo, P. A., Laukkanen, M. O., Toimela, T. A., Helenius, M. A. and Kulomaa, M. S. (1994) Molecular cloning and nucleotide sequence of chicken avidin-related genes 1–5. Eur. J. Biochem. **220**, 615–621
13 Ahlroth, M. K., Kola, E. H., Ewald, D., Masabanda, J., Sazanov, A., Fries, R. and Kulomaa, M. S. (2000) Characterization and chromosomal localization of the chicken avidin gene family. Anim. Genet. **31**, 367–375
14 Ahlroth, M., Alhroth, P. and Kulomaa, M. S. (2001) Copy-number fluctuation by unequal crossing-over in the chicken avidin gene family. Biochim. Biophys. Res. Commun. **26**, 400–406
15 Kunnas, T. A., Wallen, M. J. and Kulomaa, M. S. (1993) Induction of chicken avidin and related mRNAs after bacterial infection. Biochim. Biophys. Acta **1216**, 441–445
16 Marttila, A. T., Airenne, K. J., Laitinen, O. H., Kulik, T., Bayer, E. A., Wilchek, M. and Kulomaa, M. S. (1998) Engineering of chicken avidin: a progressive series of reduced charge mutants. FEBS Lett. **441**, 313–317
17 Marttila, A. T., Laitinen, O. H., Airenne, K. J., Kulik, T., Bayer, E. A., Wilchek, M. and Kulomaa, M. S. (2000) Recombinant NeutraLite avidin: a non-glycosylated, acidic mutant of chicken avidin that exhibits high affinity for biotin and low non-specific binding properties. FEBS Lett. **467**, 31–36
18 Laitinen, O. H., Airenne, K. J., Marttila, A. T., Kulik, T., Porkka, E., Bayer, E. A., Wilchek, M. and 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. **461**, 52–58
19 Laitinen, O. H., Marttila, A. T., Airenne, K. J., Kulik, T., Livnah, O., Bayer, E. A., Wilchek, M. and Kulomaa, M. S. (2001) Biotin induces tetramerization of a recombinant monomeric avidin. A model for protein–protein interactions. J. Biol. Chem. **276**, 8219–8224

Received 5 October 2001/17 January 2002; accepted 22 February 2002

20 Johnson, M. S. and Overington, J. P. (1993) A structural basis for sequence comparisons. An evaluation of scoring methodologies. J. Mol. Biol. **233**, 716–738
21 Johnson, M. S., May, A. C., Rodionov, M. A. and Overington, J. P. (1996) Discrimination of common protein folds: application of protein structure to sequence/structure comparisons. Methods Enzymol. **266**, 575–598
22 Airenne, K. J., Oker-Blom, C., Marjomäki, V. S., Bayer, E. A., Wilchek, M. and Kulomaa, M. S. (1997) Production of biologically active recombinant avidin in baculovirus-infected insect cells. Protein Expression Purif. **9**, 100–108
23 Bayer, E. A., Ehrlich-Rogozinski, S. and Wilchek, M. (1996) Sodium dodecyl sulfate–polyacrylamide gel electrophoretic method for assessing the quaternary state and comparative thermostability of avidin and streptavidin. Electrophoresis **17**, 1319–1324
24 Engvall, E. and Perlmann, P. (1972) Enzyme-linked immunosorbent assay, ELISA. III. Quantitation of specific antibodies by enzyme-labelled anti-immunoglobulin in antigen coated tubes. J. Immunol. **109**, 129–135
25 Delange, R. J. and Huang, T. S. (1971) Egg white avidin. III. Sequence of the 78-residue middle cyanogen bromide peptide: complete amino acid sequence of the protein subunit. J. Biol. Chem. **246**, 698–709
26 Gope, M. L., Keinänen, R. A., Kristo, P. A., Conneely, O. M., Beattle, W. G., Zarucki-Schulz, T., O'Malley, B. W. and Kulomaa, M. S. (1987) Molecular cloning of the chicken avidin cDNA. Nucleic Acids Res. **15**, 3595–3606
27 Ahlroth, M. K., Grapputo, A., Laitinen, O. H. and Kulomaa, M. S. (2001) Sequence features and evolutionary mechanisms in the chicken avidin gene family. Biochem. Biophys. Res. Commun. **285**, 734–741
28 Sano, T., Vajda, S., Smith, C. L. and Cantor, C. R. (1997) Engineering subunit association of multisubunit proteins: A dimeric streptavidin. Proc. Natl. Acad. Sci. U.S.A. **94**, 6153–6158
29 Bogan, A. A. and Thorn, K. S. (1998) Anatomy of hot spots in protein interfaces. J. Mol. Biol. **280**, 1–9
30 Janin, J. (1999) Wet and dry interfaces: the role of solvent in protein–protein and protein–DNA recognition. Structure Fold. Des. **7**, R277–R279
31 Zerega, B., Camardella, L., Cermelli, S., Sala, R., Cancedda, R. and Descalzi Cancedda, F. (2001) Avidin expression during chick chondrocyte and myoblast development *in vitro* and *in vivo*: regulation of cell proliferation. J. Cell Sci. **114**, 1473–1482
32 Bisgrove, B. W. and Raff, R. A. (1993) The SpEGF III gene encodes a member of the fibropellins: EGF repeat-containing proteins that form the apical lamina of the sea urchin embryo. Dev. Biol. **157**, 526–538
33 Bisgrove, B. W., Andrews, M. E. and Raff, R. A. (1995) Evolution of the fibropellin gene family and patterns of fibropellin gene expression in sea urchin phylogeny. J. Mol. Evol. **41**, 34–45
34 Chinol, M., Casalini, P., Maggiolo, M., Canevari, S., Omodeo, E. S., Caliceti, P., Veronese, F. M., Cremonesi, M., Chiolerio, F., Nardone, E. et al. (1998) Biochemical modifications of avidin improve pharmacokinetics and biodistribution, and reduce immunogenicity. Br. J. Cancer **78**, 189–197