BIOLOGICAL RESEARCH REPORTS FROM THE UNIVERSITY OF JYVÄSKYLÄ

68

Kari Airenne

# Production of Recombinant Avidins in *Escherichia coli* and Insect Cells

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 1998

BIOLOGICAL RESEARCH REPORTS FROM THE UNIVERSITY OF JYVÄSKYLÄ 68

Kari Airenne

# Production of Recombinant Avidins in *Escherichia coli* and Insect Cells

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa (S212)
kesäkuun 27. päivänä 1998 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Natural Sciences of the University of Jyväskylä,
in Auditorium S212, on June 27, 1998 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 1998

# Production of Recombinant Avidins in *Escherichia coli* and Insect Cells

BIOLOGICAL RESEARCH REPORTS FROM THE UNIVERSITY OF JYVÄSKYLÄ 68

Kari Airenne

# Production of Recombinant Avidins in *Escherichia coli* and Insect Cells

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 1998

Editors
Jukka Särkkä
Department of Biological and Environmental Science, University of Jyväskylä
Kaarina Nieminen
Publishing Unit, University Library of Jyväskylä

URN:ISBN:978-951-39-8340-6
ISBN 978-951-39-8340-6 (PDF)
ISSN 0356-1062

ISBN 951-39-0267-6
ISSN 0356-1062



Jyväskylä University Printing House,
Jyväskylä and ER-Paino, Lievestuore 1998

*To Paula, Teemu and the loving memory of my parents*

## SELOSTUS OSUUDESTANI TÄSSÄ VÄITÖSKIRJASSA ESITETTYYN JULKAISUMATERIAALIIN

Arvostaen työni ohjaajalta ja kollegoiltani saamaani metodista ja käytännön apua, olen päävastuussa jokaisen osajulkaisun materiaalin tuottamisesta ja sen kirjallisesta julkaisusta seuraavin poikkeuksin:

1. Osatyön yksi elektronimikroskooppisen osuuden suoritti pääosin Eeva-Liisa Punnonen.

2. Osatyön kolme elektronimikroskooppisen osuuden suoritti pääosin Varpu Marjomäki.

3. Osatyön neljä heveiinin karakterisoinnin suorittivat pääosin Harri Alenius, Jari Mikkola, Timo Palosuo ja Nisse Kalkkinen. Biosensorimittauksista vastasi Olli Laitinen.

Jyväskylässä 8.5.1998,

Kari Airenne

# ABSTRACT

Airenne, Kari
Production of recombinant avidins in *Escherichia coli* and insect cells
Jyväskylä: University of Jyväskylä, 1998, 96p.
(Biological Research Reports from the University of Jyväskylä,
ISSN 0356-1062; 68) ISBN 951-39-0267-6
Yhteenveto: Avidiinin ja avidiini-yhdistelmäproteiinien (fuusioproteiinit) tuotto bakteeri- (*E. coli*) ja perhossoluissa (Sf9)
Diss.

Avidin (Avd) is a basic glycoprotein which constitutes about 0.05% of the protein in the chicken egg-white. Due to the extraordinarily strong binding characteristics between Avd and biotin, Avd is widely used tool in numerous applications of (strept)avidin-biotin technology. In order to further expand this technology and to be able to remove the unwanted intrinsic properties of Avd (the high pI and the sugar moieties) by genetic engineering, this study focused on the establishment of a production system for recombinant Avd (re-Avd) and Avd fusion proteins.

Our initial attempts to produce re-Avd in *E. coli* resulted in low amounts of protein. However, by adapting the first ten codons of re-Avd to the so-called optimal *E. coli* codon usage, we were able to increase the production of re-Avd at 37°C to over one milligram per liter of culture. However, at 37°C most of the re-Avd was in insoluble form. At 25°C the protein was mostly soluble, but the overall yield of re-Avd was significantly reduced. The production of Avd as N- or C-terminally linked to glutathione S-transferase (GST) followed the same pattern, but significantly larger amounts of fusion proteins were synthesized at both temperatures. The position of the Avd-tag in the fusion protein has no affect on either the biological activity of re-Avd or that of GST. Biologically active re-Avd and its fusion proteins were found to be easily purified in a single step.

The baculovirus expression vector system proved to be more efficient tool for the synthesis of avidins (Avd and its derivatives) in milligram amounts than *E. coli*. The avidins were obtained in both secreted and intracellular forms from insect cells in a single step. The purified avidins were glycosylated and assembled mainly in tetramers. Like native avidin, the recombinant tetramer also exhibited a high level of thermostability, and was further stabilized upon binding biotin.

Key words: Avidin-biotin technology; affinity chromatography; affinity tag; baculovirus, BEVS; fusion protein; recombinant avidin.

*K. Airenne, University of Jyväskylä, Department of Biological and Environmental Science, P.O.Box 35, FIN-40351 Jyväskylä, Finland*

# CONTENTS

**List of original publications**

This thesis is based on the following publications, which will be referred to in the text by their Roman numerals:

I Airenne, K. J., Sarkkinen, P., Punnonen, E-L. & Kulomaa, M. S. 1994: Production of recombinant avidin in *Escherichia coli*. Gene 144: 75-80.

II Airenne, K. J. & Kulomaa, M. S. 1995: Rapid purification of recombinant proteins fused to chicken avidin. Gene 167: 63-68.

III 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. Prot. Express. Purif. 9: 100-108.

IV Airenne, K. J., Laitinen, O. H., Alenius, H., Mikkola, J., Palosuo, T, Kalkkinen, N., Arif, S. A. M., Yeang, H. Y. & Kulomaa, M. S. 1998: Avidin provides an efficient tag for fusion protein production in baculovirus infected insect cells (manuscript submitted to Biotechniques).

In addition, some unpublished data are presented.

**Abbreviations**

| | |
|---|---|
| AcMNPV | *Autographa californica* multiple nuclear polyhedrosis virus |
| ATP | adenosine triphosphate |
| Avd | chicken egg-white avidin |
| Avidins | chicken egg-white avidin and its derivatives |
| *Avr(s)* | avidin relate gene(s) |
| ß-gal | ß-galactosidase |
| BEVS | baculovirus expression vector system |
| BmNPV | *Bombyx mori* nuclear polyhedrosis virus |
| BSA | bovine serum albumin |
| BV | budded virus |
| *cor* | the basic 6.9 kD gene of AcMNPV |
| cGMP | cyclic guanosine monophosphate |
| 3-D | three-dimensional |
| *E. coli* | *Escherichia coli* |
| ELISA | enzyme-linked immunoadsorbent assay |
| Endo H | Endoglycosidase H |
| Endo $H_f$ | recombinant protein fusion of Endoglycosidase H and MBP |
| EPA | barley endoproteinase A |
| ER | endoplasmic reticulum |
| F-68 | Pluronic F-68 detergent |
| FBS | fetal calf serum |
| FITC | fluorescein isothiocyanate |
| GEP | general export pathway |
| gp64 | the major baculovirus virion glycoprotein |
| GSP | general secretory pathway |
| GST | glutathione-*S*-transferase (GST) from *Schistosoma japonicum* |
| GV | granulosis virus |
| HPLC | high pressure liquid chromatography |
| IAA | indole acrylic acid |
| *ie*1 | ie1 gene of AcMNPV |
| IgG3 | hinge region of mouse immunoglobin G3 |
| IPTG | isopropyl-ß-D-thiogalactopyranoside |
| $K_d$ | dissociation constant |
| kDa | kilodalton |
| *lac* | lactose promoter |
| lacI | lactose repressor gene |
| *lacZ* | ß-galactosidase activity abolished |
| *late* | late basic protein promoter of AcMNPV |
| LB | Luria-Bertani |
| *lpp* | *E. coli* Lpp outer membrane lipoprotein promoter |
| M | molar concentration |
| $M_r$ | molecular weight |
| MBP | maltose binding protein |

| | |
|---|---|
| MCS | multiple cloning site |
| MNPV | multiple nuclear polyhedrosis virus |
| MOI | multiplicity of infection |
| NHS | *N*-hydroxysuccinimidyl |
| NPV | nuclear polyhedrosis virus |
| OVs | occluded viruses |
| *p10* | AcMNPV p10 gene |
| PCB | photocleavable biotin derivative |
| PCR | polymerase chain reaction |
| PGs | prostaglandins |
| pI | isoelectric point |
| PIBs | polyhedral inclusion bodies |
| pKa | negative logarithm dissociation constant (acid-base) |
| *Polh* | polyhedrin promoter |
| PEST | peptide regions rich in the amino acid residues proline, aspartic acid, glutamic acid, serine and threonine. |
| $P_L$ and $P_R$ | bacteriofage λ promoters |
| PNGase F | N-glycosidase F |
| PR | progesterone receptor |
| PrA or B | progesterone receptor A or B |
| RBS | ribosome binding site |
| re- | recombinant |
| SDS | sodium dodecyl sulfate |
| PAGE | polyacrylamide gel electrophoresis |
| Parvo | the nuclear localization signal of canine parvovirus capsid protein |
| RFs | release factors |
| Sf | *Spodoptera frugiperda* |
| SNPV | single nuclear polyhedrosis virus |
| T7 | bacteriofage T7 RNA polymerase |
| TEV | tobacco etch virus |
| *trp* | tryptophan promoter |
| *vp39* | the major capsid protein gene of AcMNPV |
| X-Gal | 5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside |
| X-ray | roentgen crystallography |

# 1 INTRODUCTION

It was discovered in the beginning of this century that raw chicken egg had toxic effects for different animals as the sole source of protein in the diet (Steinitz 1898, Bateman 1916, Boas 1927). This toxicity was shown to be due to a protein constituent present in egg-white which was named avidin after its unique ability to bind biotin (vitamin H) (Eakin et al. 1941). These early investigations in the field of vitamins and nutrition was well summarized by György in 1954. With time interest in the nutritional effects of avidin was displaced by interest in the protein itself, and especially its biotin-binding properties. By the end of the 1970s the protein was purified to homogeneity and its molecular structure as well as physical and chemical properties were well characterized (Green 1975).

After the development of basic recombinant techniques, the full-length cDNA for avidin was cloned by Gope and coworkers in 1987. The cloning of the avidin gene, however, proved to be a more difficult task. After extensive work, the gene was cloned and sequenced by Wallén and coworkers in 1995. At that time it was also clear that at least five avidin-related genes (*avr1-avr5*) were present in the chicken genome (Keinänen et al. 1994). The 3-D structure of avidin had been determined two years earlier by two independent groups (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994) after a long succession of difficulties in growing the acceptable avidin crystals of sufficient size (Green & Toms 1970, Pinn et al. 1982).

Most of the work on avidin since the mid 1970s has concentrated on a refining variety of applications based on avidin-biotin interaction and has gradually led to what is now known as the avidin-biotin technology. The whole technology was based on the exceptionally high affinity and stability of the avidin-biotin complex, and on the fact that biotin can readily be attached to most binders and probes without destroying the biological function of these substrates. The methodology has become very popular and is widely used in

the detection and purification of different target molecules. It can be considered as a chemist's version of an antibody-antigen labeling procedure (Wilchek & Bayer 1990).

Despite the fact that avidin is considerably more abundant and cheaper than its bacterial relative streptavidin (Chaiet & Wolf 1964), the positive charge (pI ~ 10.5) and oligosacharide side chain of the molecule may result in the "nonspecific" binding of avidin to extraneous material. This has reduced the use of avidin in some applications. In order to create avidin more suitable for certain applications and to further expand the applications of the avidin-biotin technology, we have established a production system for recombinant avidin in *E. coli* (I) and insect cells (III) in the present study. In addition, we have developed expression vectors for the production of desired proteins as N- or C-terminally linked with recombinant avidin in these systems (II, IV). This was carried out to demonstrate the ease with which such hybrid proteins could be detected and purified using the biotin-binding activity of avidin.

# 2 REVIEW OF THE LITERATURE

## 2.1 Avidin

### 2.1.1 Distribution, function and natural gene expression

Small amounts (< 0.05% of total protein) of avidin are found in the eggs of avian, amphibian and reptilian species. No avidin activity has been observed in mammals (Hertz & Sebrel 1942, Jones & Briggs 1962, Elo 1980, Korpela et al. 1981a,b, Stevens 1991). However, close bacterial relatives of avidin, streptavidin or streptavidin v1 and 2, are found in certain strains of *Streptomyces* (Chaiet & Wolf 1964, Bayer et al. 1995B). Despite a profound knowledge of the properties of avidin little is known about its biological function. Since the growth of a number of micro organisms is inhibited by it, and its occurrence after tissue damage has been clearly demonstrated, avidin most likely has a role as a factor in host defense factor, although reproductive, enzymatic, antibiotic and remedial roles have also been proposed (Elo et al. 1980b, Elo & Korpela 1984, Korpela, 1984, Wilchek et al. 1988, Tuohimaa et al. 1989, Wilchek & Bayer 1989, White et al. 1992, Vetter et al. 1994). Avidin and streptavidin have also recently been shown to be insecticidal (Morgan et al. 1993).

Avidin was long regarded as an chicken oviduct (estrogen pretreated) and progesterone-specific protein (Hertz et al. 1943, O'Malley 1967, Kohler et al. 1968, Korenman & O'Malley 1968, O'Malley et al. 1969, Korpela et al 1981a, Kunnas et al. 1992). It was, however, later shown that avidin can also be induced in several other chicken tissues in a progestin-independent manner by various treatments causing an inflammatory reaction such as tissue trauma, pinching or heat injury, or the introduction of toxic drugs, bacteria or viruses (Elo et al. 1975, Heinonen & Tuohimaa 1976, Heinonen et al. 1978, Heinonen &

Tuohimaa 1978, Elo et al. 1979a,b, Elo 1980, Elo et al. 1980a,b, Elo et al. 1981, Nordback et al. 1981, Korpela et al. 1982, Nordback et al. 1982a,b, Korpela et al. 1983, Elo & Korpela 1984, Niemelä 1986). These studies were summarized by Tuohimaa et al. (1989). The two mechanisms leading to avidin synthesis in chicken tissues are shown schematically in Fig. 1.

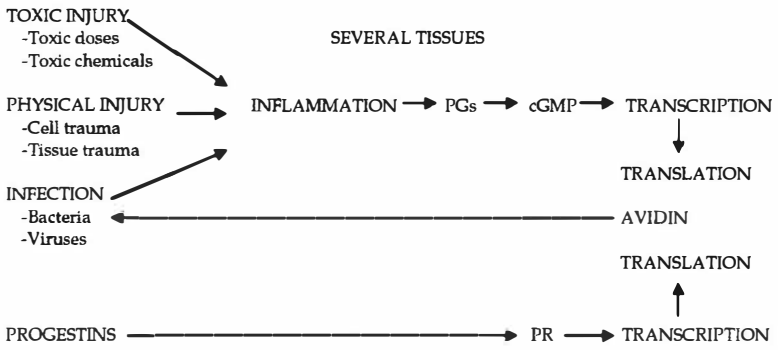



FIGURE 1 A schematic presentation of two mechanisms leading to avidin synthesis in chicken tissues. Avidin can be induced in most tissues by inflammation caused by systemic or local injurious treatments and by mediators of inflammatory reaction such as prostaglandins (PGs) and cyclic guanosine monophosphate (cGMP). A new RNA (transcription) and a new protein synthesis (translation) are needed for the expression of avidin. There is closed negative feedback control between biotin-requiring bacteria and avidin production, owing to the antimicrobial activity of the latter. The other avidin induction mechanism is progestin-dependent and is present only in the differentiated oviduct. The figure is taken from Tuohimaa et al. (1989).
→ = Stimulatory action; ← = Inhibitory action; Progesterone receptor (PR).

The biotin-binding properties of inflammation-associated avidin seem to be similar to those of the progesterone-dependent form, but at least in some cases the antigenic structure of the latter is slightly altered (Korpela et al. 1982). Whether the progestin-dependent and inflammation-associated avidin products are products of the same gene, with a multifactorial regulation, or products of structurally related and differentially regulated genes, remains to be studied (Tuohimaa et al. 1989). The discovery of avidin-related genes (Keinänen et al. 1994, Wallen et al. 1995), of which the transcripts to *avr2* and *avr3* have been detected (Kunnas et al. 1993), raises the possibility that the chicken genome may have more than just one functional gene for avidin.

### 2.1.2 General properties

Chicken egg-white avidin and bacterial streptavidin (from *Streptomyces avidinii*) are both tetrameric proteins which bind biotin with similar association

constants ($K_d$ ~ $10^{-15}$ M), the strongest protein-ligand interaction known in nature (Chaiet & Wolf 1964, Green 1975). The two proteins differ, however, since avidin is a strongly basic (pI >10) disulphidebridged glycoprotein, whereas neutral streptavidin is unglycosylated and is devoid of cystein amino acid residues. X-ray crystallographic analyses have shown clearly that the fold of both proteins is generally conserved. Moreover, most of their binding-site residues are also conserved, despite the fact that there is only about 35% overall sequence identity between the proteins (Bayer & Wilchek 1990).

The avidin tetramer ($M_r$ ~63 kDa) contains four identical subunits ($M_r$ ~15.7 kDa), each of which consists of a single polypeptide chain bearing 128 amino acids and possessing one biotin-binding site (Green 1975). The content of threonine and tryptophan is high and each subunit has an intrachain disulfide bond between amino acid residues 4 and 83 (DeLange & Huang 1971). Heterogeneity (either isoleucine or threonine) was originally observed by protein sequencing at residue 34. In the revealed avidin cDNA and gene structure the residue was isoleucine. As distinct from the original results by DeLange & Huang (1971), the residue 53, originally found to be glutamic acid, is glutamine according to avidin gene (Wallen et al. 1995). Of the ten asparagine residues in each subunit, only Asn17 is glycosylated. According to Bruch at al. (1982), carbohydrate accounts for about 10% of the molecular weight and exhibits extensive glycan microheterogeneity. Avidin contains at least three distinct structural types of oligosaccharide, similar in composition and size. The oligosaccharide side chain is not essential for biotin-binding (Hiller et al. 1987) and can be removed enzymatically or by mixing avidin with a microbial culture to avoid unwanted interactions in some avidin-biotin applications (Bayer et al. 1995A).

The isoelectric point of avidin is within the range 10-10.5 (according to computer prediction 10.43), making avidin a very basic glycoprotein (Green 1975). It is readily soluble in water and salt solutions, and it is stable over a wide range of pH and temperature. The remarkable stability of avidin is increased further when biotin is bound (Green 1963, Wei & Wright 1964, Donovan & Ross 1973, Ross et al. 1986, Durance & Wong 1992). The avidin-biotin complex is even resistant to proteolysis by the enzymes of the digestive tract and neither trypsin nor pronase inactivate free avidin. The biotin-binding of avidin can, however, be blocked by oxidation on the part of any of several tryptophan residues or by the dinitrophenylation of what appears to be a single lysine residue (Green 1975).

### 2.1.3 Structure and biotin-binding

The three-dimensional structure of avidin (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994) as well as that of streptavidin (Hendrickson et al. 1989, Weber et al. 1989) has been determined. The overall fold of the avidin monomer closely resembles that of streptavidin; it is constructed of eight antiparallel $\beta$-strands, which form a classical up-and-down $\beta$-barrel. The

biotin-binding site is positioned near one end of the monomer (Fig. 2). An array of polar and aromatic residues are involved in the tight binding in this area of $\beta$-barrel.

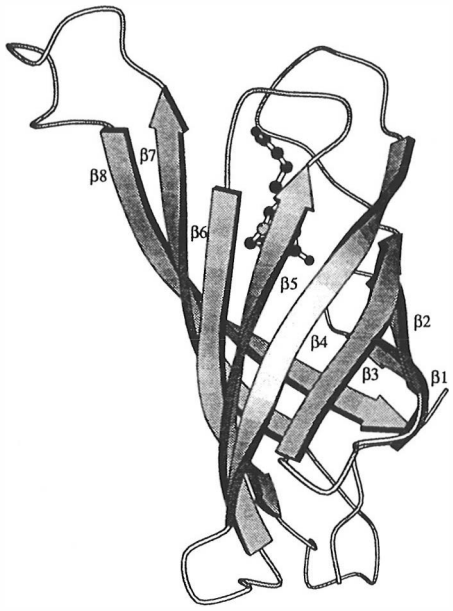



FIGURE 2 A MOLSCRIPT ribbon diagram of the avidin-biotin monomer, with eight strands of the β-barrel labeled. The biotin molecule is shown in a ball and stick model. The figure is taken from Livnah et al. (1993).

There are three regions of monomer-monomer interaction in avidin. These interactions contribute to the rigidity of the quaternary structure and also provide part of the framework for the tight binding of biotin (Livnah et al. 1993). The intersubunit contacts made by tryptophan residue 110 of avidin and 120 of streptavidin with biotin are in special importance. The biotin-binding activity of streptavidin was reduced substantially, to approximately $K_d$ $10^{-8}$ M, when tryptophan 120 was mutated to phenylalanine (Pugliese et al. 1993, Sano & Cantor 1995).

The conserved residues of avidin and streptavidin are usually confined to short homologous stretches which form six relatively defined domains. Interestingly, of the 17 residues in avidin that interact directly with biotin, 12-14 are also conserved in all avidin related genes (Keinänen et al. 1994). Indeed, proteins encoded by the *avr1*, *avr2* and *avr4* genes have biotin-binding activity when expressed in insect cells (unpublished results).

Already in the early 1960s it was shown that tryptophans are apparently involved in the biotin-binding of avidin (Green 1975). The involvement of the majority of the tryptophans, especially those homologous to streptavidin, was

confirmed by Gitlin et al. (1988a,b), who showed that modification of tryptophans by 2-hydroxy-5-nitrobenzyl bromide abolished biotin-binding in both avidin and streptavidin. Fluorescence studies indicated that, in addition to tryptophan residues 70 and 110 in avidin or residues 92, 108 and 120 in streptavidin, an additional tryptophan residue is also involved in the active binding sites of both proteins (Kurzban et al. 1989). The importance of these tryptophan residues was further shown by site-directed mutagenesis studies of streptavidin (Chilkoti et al. 1995, Sano & Cantor 1995). Chemical modification studies have also suggested that the single tyrosine in avidin (residue 33) and its homologue in streptavidin (residue 43) is important in biotin-binding (Gitlin et al. 1989, Gitlin et al. 1990, Hiller 1991). In addition, it has been shown that lysine residues probably comprise part of the biotin-binding site of avidin (Gitlin et al. 1987).

The presence of tryptophan residues in the biotin-binding site of both avidin (residues 70, 97 and 110) and streptavidin (residues 79, 92, 108 and 120) have been confirmed by X-ray crystallographic studies. (Hendrickson et al. 1989, Weber et al. 1989, Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994, Morag et al. 1996). In addition to tryptophan residues 70 and 97, phenylalanine 72 and 79 (substituent of tryptophan residue 92 in streptavidin) from one momomer, and tryptophan 110, which is provided by the adjacent symmetry-related monomer, are also all involved in the formation of the hydrophobic pocket in which the biotin molecule resides in avidin (Fig. 3A). In streptavidin the hydrophobic pocket is provided by tryptophan residues 79, 92, 108 from one monomer, and tryptophan 120 from the adjacent monomer (Livnah et al. 1994). In addition to hydrophobic interactions, the heteroatoms in the ureido ring of biotin exhibit five crucial hydrogen-bond interactions with the side chains of polar amino acid residues. In avidin the hydrophilic interactions are made by the asparagine 12, serine 16, tyrosine 33 (one of the most important bonds, the equivalent of tyrosine 43 in streptavidin), threonine 35, and asparagine 118 residues. In addition, the biotin sulfur may interact with threonine 77, and the two carboxylate oxygens of the valeryl moiety of biotin form five hydrogen bonds with threonine 38 and 40, alanine 39, and serine 73 and 75 (Fig. 3C). In streptavidin, a network of hydrogen bonds, similar to that in the binding site of avidin, is formed with the biotin rings (Livnah et al. 1994). The ureido ring oxygen forms three hydrogen bonds with asparagine 23, serine 27, and tyrosine 43. Each of the ring nitrogens forms one hydrogen bond with serine 45 and asparagine 128, and the biotin sulfur may form a hydrogen bond with threonine 90. Carboxylate oxygens of the biotin valeryl moiety form a total of only two hydrogen bonds i.e., one with the main chain N•••H of asparagine 49 (substituted for alanine 39 in avidin) and the other with serine 88 (the equivalent of serine 75 in avidin). These results are in good agreement with the chemical modification data.

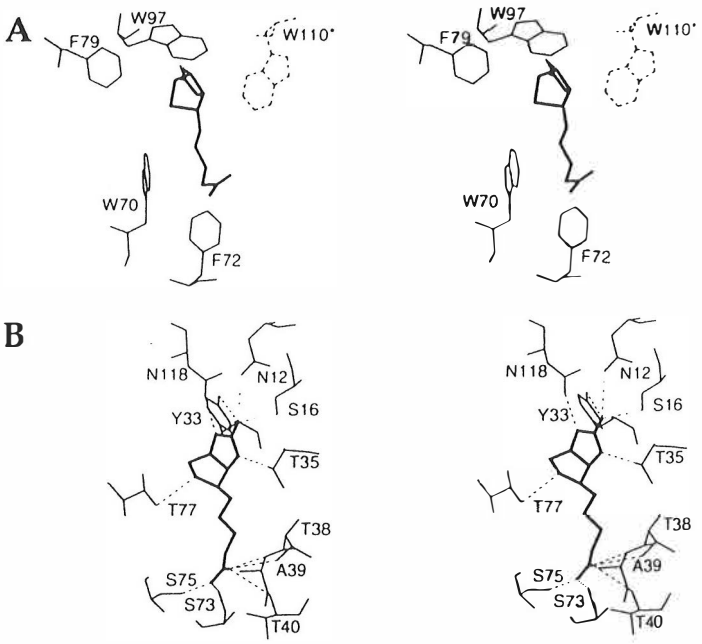
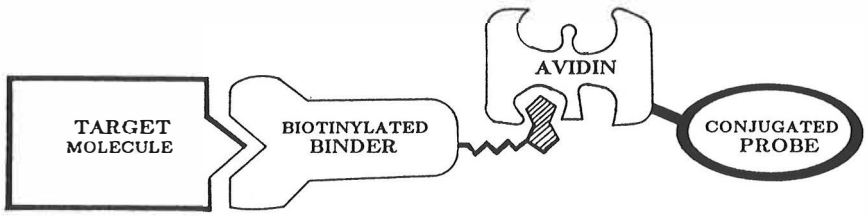



FIGURE 3 Binding sites of avidin. (A) Hydrophobic residues in the binding site of avidin. These include Trp-70, Phe-72, Phe-79, and Trp-97 from one monomer, and Trp-110 (dashed lines), which is provided by the adjacent symmetry-related monomer. (B) Hydrophilic interactions in the binding site of avidin. The ureido oxygen of the biotin molecule forms three hydrogen bonds with the side chains of Asn-12, Ser-16, and Tyr-33 of avidin forming a tetrahedral oxyanion. In addition, each of the two ureido nitrogens participates in a single hydrogen-bond interaction with Thr-35 and Asn-118. The biotin sulfur may interact with Thr-77. The two carboxylate oxygens of the valeryl moiety of biotin form five hydrogen bonds. One interacts with the main chain N•••H of Ala-39 and Thr-40 as well as the side chain of Thr-38 (these residues are part of the loop that is stabilized when biotin is bound), and the other forms hydrogen bonds with the side chains of Ser-73 and Ser-75. The figure was reconstructed from Livnah et al. (1993).

### 2.1.4 Avidin-biotin technology

The extraordinarily high affinity ($K_d \sim 10^{-15}$ M) that characterizes the complex formed between the vitamin biotin and avidin has provided the basis for establishing a new technology, (strept)avidin-biotin technology, which has broad applications in virtually all fields of biology, medicine and biotechnology (Fig. 4). The major areas of application are localization, isolation, diagnostics and affinity-based targeting studies (for a review, see Wilchek & Bayer 1989, Bayer & Wilchek 1990, Wilchek & Bayer 1990).

There are some major advantages in this system which make it very popular. (1) The exceptionally high affinity and stability of the avidin-biotin




| TARGET : BINDER | |
|---|---|
| Antigens | Antibodies |
| Antibodies | Antigens |
| Lectins | Glycoconjugates |
| Glycoconjugates | Lectins |
| Enzymes | Substrates, cofactors, inhibitors, etc. |
| Receptors | Hormones, effectors, toxins, etc. |
| Transport proteins | Vitamins, amino acids, sugars, etc. |
| Hydrophobic sites | Lipids, fatty acids |
| Membranes | Liposomes |
| Nucleic acids, genes | DNA/RNA probes |
| Phages, viruses, bacteria, subcellular organelles, cells, tissues, whole organisms | All of the above |

AVIDIN-BIOTIN COMPLEX

| PROBES |
|---|
| Enzymes |
| Radiolabels |
| Fluorescent agents |
| Chemiluminescent agents |
| Chromophores |
| Heavy metals |
| Colloidal gold |
| Ferritin |
| Hemocyanin |
| Phages |
| Macromolecular carriers |
| Liposomes |
| Solid supports |

| APPLICATIONS | |
|---|---|
| 1. Affinity chromatography: isolation studies | 10. Bioaffinity sensor |
| 2. Affinity cytochemistry: localization studies | 11. Gene probe |
| a. light microscopy | 12. Crosslinking agent |
| b. fluorescence microscopy | 13. Affinity targeting |
| c. electron microscopy | 14. Affinity perturbation |
| 3. Histochemistry | 15. Drug delivery |
| 4. Pathological probe | 16. Fusogenic agent |
| 5. Diagnostics | 17. Immobilizing agent |
| 6. Signal amplification | 18. Selective retrieval |
| 7. Immunoassay | 19. Selective elimination |
| 8. Hybridoma technology | 20. Flow cytometry |
| 9. Blotting technology | 21. Cytological probe |

FIGURE 4 Generalized scheme illustrating the essentials of (strept)avidin-biotin technology. A biologically active target molecule is recognized by a biotinylated binder that is subsequently recognized by (strept)avidin conjugated to an appropriate probe. A list of target-binder pairs, the various probes that have been used, and broad spectrum of applications of (strept)avidin-biotin technology are also presented. The figure is taken from Wilchek & Bayer (1990).

complex ensures the desired conjunction of binder and probe. (2) Biotin can be easily attached to most binders and probes, without destroying their biological activity. (3) The tetrameric structure of avidin leads to the amplification of the signal and (4) the system is amenable to double-labeling. (5) The system is extremely versatile, because a given molecular target can interact with a single type of biotinylated binder that can then be analyzed by various means, using different avidin-conjugated probes. (6) A wide spectrum of different biotinylating reagents, biotinylated and avidin-containing probes are available from a variety of commercial sources (Wilchek & Bayer 1990).

Due to the fact that avidin is highly alkaline and glycosylated, in some systems avidin undergoes undesired interaction with negatively charged macromolecules such as nucleic acids and acid proteins, which leads to background problems. Similarly, lectins or other sugar-binding materials may create unwanted background problems. For this reason streptavidin has replaced avidin in many applications, despite the fact that it is much more expensive reagent.

## 2.2 Heterologous protein production

Due to the rapid advances made during last two decades in recombinant DNA and gene expression techniques, it is now possible to produce many proteins which exist naturally in very low amounts and/or are otherwise difficult to purify to homogeneity with traditional methods. The recombinant hosts currently available include a number of different cell systems. The most commonly used of these are bacteria (*Escherichia coli, Bacillus subtilis, Staphylococcus aureus* etc.), yeast (*Saccharomyces cerevisae, Pichia pastoris* etc.), insect cells (*Spodoptera frugiperda* Sf9 and Sf21, *Trichoplusia ni*, High-Five™ etc.), mammalian cells (Chinese hamster ovary cells, CHO; Baby hamster kidney cells, BHK etc.), filamentous fungi (*Aspergillus niger, Aspergillus oryzae, Trichoderma reesei* etc.), plant cells (green algae, protoplast, tobacco BY-2 cells etc.), and animals (mouse, rat, rabbit, pig, sheep, goat, cow etc.). The choice of the host is of great importance and will largely depend on the nature of protein to be produced. Because a universal expression system that works for all heterologous proteins is not yet available, one should select a production system as near to the origin of the target protein as possible. All of the various posttranslational modifications (e.g. phosphorylation, N-linked and O-linked glycosylation, signal peptide cleavage, proteolytic processing, palmitylation, myristylation, farnesylation, carboxyl methylation, disulfide formation and multimeric assembly) which may be necessary for the production of a functional protein should then occur (for a review, see Goeddel 1990a, Jänne et al. 1994, Foster et al. 1995, Frommer & Ninnemann 1995, Cleland & Craik 1996). However, the use of a natural expression system is not always possible and then, in addition to cell type *per se*, the practical considerations (need of expertise, cost, efficiency etc.) may favor one system rather than another. The by far most commonly used production host for protein of bacterial origin or protein not heavily postranslationaly modificated, continues to be *E. coli.* On the other hand, for more complicated proteins the recent advances in baculovirus expression systems have made insect cells an attractive alternative as an eukaryotic expression host.

### 2.2.1 Synthesis in *Escherichia coli*

For many reasons *E. coli* is still the prefered host for recombinant protein production, not least because it is a very well-known and easy-to-grow organism. Large amounts (up to 1g per liter of culture) of desired polypeptides can be obtained inexpensively for research and commercial purposes via this host. In addition, both a variety of vectors bearing strong regulatable promoters and techniques easy to adopt by any laboratory are available for *E. coli*. The two major drawbacks of the system are that eukaryotic posttranscriptional and posttranslational modifications can not be properly done in these cells and the capacity of *E. coli* to secrete proteins is limited (Gold 1990).

#### 2.2.1.1 DNA elements behind efficient gene expression

An efficient gene expression plasmid vector should supply the following elements: (1) a well-characterized origin of replication and selection marker for plasmid propagation and maintenance, (2) a strong and regulatable promoter, (3) a ribosome binding site and, usually, translation initiation (ATG) codon, and (4) translation and transcription termination sites. In addition to the host and gene itself, each of these elements is vital for efficient gene expression and is described in more detailed below. For practical reasons an efficient vector should also have a multiple cloning site (MCS) in all reading frames with a large number of restriction enzyme sites from which to choose. In the case of a fusion protein vector, the site for the efficient cleavage of the target protein from the fusion partner may also have practical importance.

***Replicon.*** The copy-number of the plasmid effects the gene dosage in the cell. Therefore, the higher the copy number, the great the number of transcripts available for translation and, at least in theory, the greater the amount of protein capable of being produced. In order to maximise the copy-number, most vectors for *E. coli* expression are based on pBR322 or pUC18/19 plasmids, oll of which bear the Col E1 replicon or its derivative for the relaxed replication of the plasmid (Old & Primrose 1989). The copy-number of the plasmid is controlled also by the growth rate of the cells (Seo & Bailey 1985) and by the segregational instability of the plasmid (Klotsky & Schwartz 1987), which has to be taken into account to achieve the best expression levels. In some cases, when the gene for toxic protein is to be expressed, the high copy-number of the plasmid may have lethal effects on the host. To avoid this, plasmids with controllable copy-numbers have been constructed (Kishimoto et al. 1986, Chew & Tacon 1990).

***Selection.*** The selection element plays an important role not only in the DNA cloning work but also in protein production. It helps to avoid the segregational instability of the plasmid. In addition to selection by an antibiotic resistance gene, the use of a suicidal mechanism for those cells that segregate the plasmid, the incorporation of a partition locus into the plasmid, and the

immobilization of cells in organic supports all diminish the effects of segregational instability as selection elements (Balbas & Boliviar 1990).

***Promoter.*** A large number of promoters for *E. coli* have been analyzed (Harley & Reynolds 1987) and a consensus sequence which consists of the -35 region (5'-TTGACA-) and the -10 region or Pribnow box (5'-TATAAT) can be formulated on the basis of these studies for an efficient promoter (Pribnow 1975, Dickson et al. 1975). These regions are part of the core promoter and they are separated by a spacer the optimal length of which is (or near) 17 base pairs (Hawley & McLure 1983). The DNA sequences flanking the core promoter can also have an important influence on expression (Gralla 1990, Déthiollaz et al. 1996).

There are several very well characterized promoters in routine use for recombinant protein production in *E. coli*, including promoters for tryptophan (*trp*) biosynthesis, lactose (*lac*) catabolism, an outer membrane protein (*lpp*), repression of bacteriofage λ ($P_L$ and $P_R$) and bacteriophage T7 RNA polymerase (T7)(Foster et al. 1995). However, none of these shows absolute identity with the consensus sequence (Old & Primrose 1989). An important feature of these promoters is that they can be regulated.

The *lac* promoter makes the expression of the gene dependent upon lactose or a lactose analogue such as isopropyl β-D-thiogalactosidase (IPTG) in the culture medium. In the absence of lactose the repressor (lacI) binds the operator site and the gene is not expressed (Gralla 1990). To ensure the tight regulation of the *lac* operator, a gene for the repressor can be expressed from the same plasmid as the protein encoding gene or, an *E. coli* strain expressing the repressor can be used. A derivative of the *lac* promoter (*lac*UV5) has been constructed in which the catabolite repression of the *lac* promoter by glucose is destroyed (Old & Primrose 1989).

The *trp* promoter is controlled by tryptophan. When tryptophan is present, it binds to the TrpR repressor, and this complex inhibits gene expression. Expression from the *trp* promoter is induced by starvation for tryptophan or by the addition of its analog, indole acrylic acid (IAA) (Yanofsky et al. 1981, Yansura & Henner 1990). A hybrid promoter (*tac*) consisting of the -10 region of the lacUV5 promoter and the -35 region of the *trp* promoter has also been constructed in order to obtain an even stronger promoter than *lac* and *trp* alone (Amman et al. 1983, DeBoer et al. 1983). In the *trc* promoter the spacing of these elements is changed from 16bp to 17bp (Brosius et al. 1985, Mulligan et al. 1985).

The *lpp* promoter which controls the synthesis of Lpp outer membrane lipoprotein of *E. coli* can be used with both constitutive and regulated systems (Nakamura & Innoye 1982, Masui et al. 1984). Bacteriofage λ promoters ($P_L$ and $P_R$) are used for the temperature-sensitive control of gene expression through the cI857 repressor, which is supplied in *E. coli* on a lysogen or plasmid (Shimatake & Rosenberg 1981, Mott 1985).

Although all the promoters mentioned above can be regulated tightly, leakage can occur. This may be a problem if a highly toxic protein is produced; however, it can be circumvented by using the bacteriophage T7 operator

system created by Studier et al. (1990). The system relies on the fact that the RNA polymerase of bacteriophage T7 is very selective for specific promoters that are rarely encountered in DNA unrelated to T7 DNA. By placing the gene of interest under the control of the T7 promoter, one can obtain a very tightly controlled gene expression. The RNA polymerase introduced into such a cell will transcribe actively and selectively only the DNA under the control of the T7 promoter.

Recent reports of the arbinose-inducible *araB* promoter, pH-inducible *cadA* promoter, and xanthosine inducible *xapA* promoter further expand the promoter options available for high-level expression in *E. coli* (Weickert et al. 1996).

***Translation.*** The ribosome binding site (RBS) complementary to 16S rRNA (known as Shine-Dalgarno sequence in bacteria), and the initiation codon with surrounding nucleotides play an important role in an efficient expression vector. The RBS not only determines the translation initiation place on the gene, but also regulates the translation efficiency (Sprengart et al. 1996). According to the compilation of Stormo et al. (1982) and Stormo (1986), general rules for an efficient translation initiation can be postulated: (1) the preferential initiation codon is AUG, although GUG, UUG, AUU, and AUA are not uncommon; (2) the RBS sequence has at least four nucleotides taken from the sequence AGGAGG; (3) the spacing between the RBS and AUG is 9±3 nucleotides; (4) the nucleotides located 5′ of the AUG should be As and Us (A at position -3); (5) the sequence GCAU or AAA located after the AUG enhances translation; and (6) the region around the initiation site should be relatively unstructured (Balbas & Bolivar 1990).

***Terminators.*** To avoid undesirable secondary structures, synthesis of unnecessarily long transcripts, a decrease in plasmid copy-number, and promoter occlusion (interference from the transcript of the promoter on the expression of the gene from a neighboring promoter), an efficient transcription termination element at the end of the expressed gene is needed (Old & Primrose 1989). The termination sequence of the *trp* operon, the bacterifage λ terminators and the rrnB ribosome RNA terminator are widely used for this purpose in expression vectors. For especially long messages (mRNA), opposite elements, antitermination elements, have also been used (Podhajska et al. 1985, Hasan & Szybalski 1987, Shatzman & Rosenberg 1987).

In prokaryotes there are two release factors (RFs) that each recognize two of the three stop codons: RF-1 recognizes UAG and UAA, RF-2 UGA and UAA. Early suppression studies already suggested that there may be a larger context than three nucleotides specifying the termination of translation by these factors. Recent studies have confirmed this and shown that the base next to known stop codons has a certain bias in highly expressed genes in *E. coli* (Tate & Brown 1992, Brown et al. 1993, Brown et al. 1994, Poole et al. 1995). Hence instead of using the most efficient triplet codon (UAA) in the vector (or gene) one should use the tetra sequence of UAAU for efficient translation termination in *E. coli*. The use of tandem termination codons after the gene of interest is also

a common practice in ensuring efficient translation termination in the available expression vectors (Itakura et al. 1977).

Much is known about the vector-based DNA elements that are important for efficient gene expression in *E. coli* and many vectors bearing these optimized elements are commercially available. However, the final amount of the desired recombinant protein is also determined, in addition to these DNA elements, which can easily be incorporated into any vector, by the sequence of the gene to be expressed.

***Codon usage.*** Non-random usage of the synonymous codons has been clearly demonstrated in both prokaryotes and eukaryotes, and a high abundance of proteins show a sharp avoidance of codons that are in low usage in the overall gene population (Zhang et al. 1991). Codon usage has been assumed to correlate with the expression rate of a gene (Gouy & Gautier 1982, Grosjean and Fiers 1982). In particular, the codons located in the N-terminal coding region have been demonstrated to play an important role in determining the expression rate in *E. coli.* (Chen & Innouye 1990, Looman et al. 1987, Furlong et al. 1992). Because weakly expressed genes are characterized by the occurrence of codons recognized by rare tRNA species (Ikemura 1981a,b, Bulmer 1987), it has been suggested that optimizing codon usage will increase the level of production of recombinant proteins in general. Indeed, the optimization of codon usage with respect to those codons that are preferentially recognized in *E. coli* has been shown to have a positive effect on the yield of several recombinant proteins (Robinson et al 1984, Thompson & Weber 1993, I).

Recent experiments and compilations of gene expression data have, however, revealed that there is not necessarily correlation between codon usage *per se* and translation efficiency (Ernst & Kawashima 1988, Kurland 1991, Date et al. 1992). Codon usage can also effect translation efficiency through translation accuracy (Kurland & Gallant 1996). In many eukaryotic genes the triplets AGA and AGG, coding for arginine, are used at much higher frequencies than in *E. coli*, and the shortage of $tRNA_{arg4}$ frequently causes ribosomal framesifts of +1 at tandem AGA and AGG codons. This can be avoided by expressing the argU gene encoding the $tRNA_{arg4}$ along with the gene containing AGA/AGG-rich regions (Schenk et al. 1995).

***Secondary structures.*** In addition to codon use *per se*, the above codon optimization procedures may affect translation efficiency (along with vector-based translation initiation elements, i.e. the Shine-Dalgarno sequence and its flanking region) through destabilizing the stem-loop structures of mRNAs. These have been assumed for many years to have an important role in translation initiation for (Iserentant & Fiers 1980, Hall et al. 1982, Ganoza et al. 1987, Gross et al. 1990, Wikström et al. 1992). The secondary structure also affects the half life of mRNA, and the reduction in the degradation of mRNA in the cell may be important for efficient translation initiation by controlling the absolute amount of the desired mRNA (Belasco & Higgins 1988). Two $3' \rightarrow 5'$ - exonucleases, RNAase II and polynucleotide phosphorylase seem to be primarily responsible for the degradation of mRNA. The presence or absence

of ribosomes, the rate of translation and specific sequences in the 5′ or 3′ ends of the DNA (that responds to trans-acting effectors) may also affect mRNA decay as well as the stability determinants within specific genes, the cell growth rate and the utilization of Rnase-deficient or conditionally lethal mutants. Even if general approaches have been suggested for the manipulation of mRNA stability, more work is needed to establish the real impact that this issue has on the final product yield of an *E. coli* plasmid-based expression system (Balbas & Bolivar 1990).

### 2.2.1.2 Enhancement of solubility and yield of recombinant proteins

***Inclusion bodies.*** Many recombinant proteins form large electron-dense particles, called inclusion bodies, when expressed in high amounts in the cytoplasm of *E. coli.* Inclusion bodies consist primarily of aggregated recombinant protein, but it has also been suggested that four subunits of RNA polymerase, some combination of the outer membrane proteins OmpC, OmpF and OmpA, 16S and 23S rRNA, and plasmid DNA are involved (for a review, see Marston 1986, Schein 1989). The formation of inclusion bodies is not restricted to the overproduction of recombinant proteins in *E. coli* (Williams 1982). They can also form due to the overproduction of normal *E. coli* proteins (Cheng 1983, Gribskov & Burgess 1983) as well as in the presence of a normal concentration of homologous proteins into which have been amino acid analogs incorporated (Prouty & Goldberg 1972, Prouty et al. 1975). Formation of inclusion bodies is not unique to *E. coli.* They have also been found in yeast and insect cells (Hayakawa et al. 1987, Rudolph 1996).

The exact mechanism that causes formation inclusion bodies is not known, but several suggestions have been made to account to their formation: (1) foreign proteins are recognized as such in *E. coli*; (2) a high production rate allows insufficient time for the nascent chain to fold properly; (3) high local concentration in the cytoplasm leads to non-specific precipitation due intermolecular interactions; (4) lack of cellular compartmentation causes the protein to be produced in a reducing environment, preventing formation of S-S bonds necessary for proper folding; and (5) lack of mammalian post-translational modifying enzymes and proper chaperons during production leads to decreased solubility of, especially, eukaryotic proteins (Schein 1989). Recently evidence has shown that formation of inclusion bodies occurs through the specific interaction of certain conformations of folding intermediates rather than by non-specific coaggregation (Speed et al. 1996).

Although the production of a protein as an insoluble inclusion body can offer the advantage of easy purification and protease protection, devising an appropriate solubilization and refolding procedure can be difficult, with no guarantee of success in all cases (Marston 1986. Fischer et al. 1993). Hence, in practice, There is great interest in finding ways of avoiding inclusion body formation and increasing the solubility of the expressed proteins.

***Growth temperature.*** One way to present the formation of inclusion bodies is to reduce the growth temperature of the culture (and thus avoid thermally induced plasmid vectors). This method has been used successfully for many proteins. Changing the growth temperature is thought, besides enhancing the productive folding pathway, to facilitate the association of the foreign protein with the necessary chaperons or circumvent the necessity for chaperon association (Schein & Noteborn 1988, Browner et al. 1991, Ghosh et al. 1995). Lower temperature also reduces the growth rate of the cells, which has been shown to decrease the fraction of aggregated recombinant proteins as well as limited induction of gene expression (Takagi et al. 1988, Kopetzki et al. 1989). The use of sarkosyl after bacterial expression has also been shown to increase the fraction of the soluble end product (Frankel et al. 1991, Frangioni & Neel 1993).

***Site-directed mutagenesis.*** A more difficult way to increase the solubility of recombinant proteins is to alter the structure of the protein by site-directed or random mutagenesis (Wetzel et al. 1991, Wetzel 1994). The difficulty in employing this approach arises from the fact that no general rules predicting the solubility of the target protein can be given. However, according to the summary by Schein (1993) the placement of charged residues (especially Arg and Asp) on the surface of proteins may prevent aggregation either by increasing the repelling charge on the protein surface or by stabilizing interactions with the aqueous shell around the proteins. In addition, the replacement of lysine by an arginine or glutamic acid by aspartic acid may increase the thermostability of the protein and thus possibly also increase its solubility at higher temperatures. In addition, to avoid great changes in the structure of the protein, histidine substitutions or insertions must take into account the effective pKa of the residues in their surrounding environment. If the histidine is not expected to be protonated, it should be replaced with 'hydrophobic residues' rather than charged residues. Further, insertion of proline residues may decrease the entropy of the unfolded state, thus increasing thermostability. Strongly hydrophobic regions have also been found to increase insolubility, and hence elimination of these regions may increase the stability and/or solubility of recombinant protein (Smith & Johnson 1988).

***Chaperones and foldases.*** Chaperones are proteins that bind to and stabilize an otherwise unstable conformer of another protein. By controlled binding and release, they facilitate the correct folding and oligimeric assembly of the protein and its transport to a particular subcellular compartment or its disposal by degradation (Hartl 1996). Molecular chaperones prevent incorrect interactions within and between non-native polypeptides, thus typically increasing the yield but not the rate of folding reactions. This distinguishes them from folding catalyst enzymes (disulfide isomerases and peptidyl-prolyl isomerases) which accelerate intrinsically slow steps in the folding of some proteins. They are particularly involved in the rearrangement of disulfide bonds in secretory proteins and *cis-trans* isomerization of peptide bonds preceding proline residues (Hartl 1996). Recent advances in the knowledge of the function of these proteins together with the possibility of co-expressing

these molecules with recombinant protein, may reveal whole new ways to assist folding and thus solubility (export/secretion). This is of particular relevance to proteins, which have thus far been difficult or even impossible to produce in *E. coli*. Indeed, some chaperones (GroES/EL and DnaK, DnaJ) and foldases (DsbA, PPIase, PDI) have already been co-expressed with recombinant proteins with good results (Hockney 1994, Georgiou & Valax 1996, Weickert et al. 1996). Co-expression of binding molecules with specific affinity to the target protein can be an attractive alternative to the engineered expression of chaperones or foldases in enhancing yields of correctly folded protein products (Samuelsson et al. 1996, Weickert et al. 1996).

***Export and secretion.*** The export of recombinant protein into the periplasm of *E. coli* may help to keep it in soluble form in some cases. Unlike the cytoplasm, the periplasmic space is an oxidizing compartment which facilitates the folding and assembly of many proteins, especially those containing disulfide bridges. Export may also help one to achieve a correct protein primary amino acid structure without initiatior fMet or other extra amino acids in the N-terminus. In addition, the possibility for continuous cell culture, simplified purification (the periplasm contains only 4% of bacterial proteins, Nossal & Heppel 1966), and the avoidance of protein degradation (Stader & Silhavy 1990) may be achieved by exporting the recombinant protein into the periplasm or by secretion out of the cell.

Protein export through the inner membrane to the periplasm or outer membrane is achieved in *E. coli* by a unique, apparently universal mechanism. The signal peptide-dependent general export pathway (GEP), comprises the sec gene products and the signal peptidases (Wickner et al. 1991). However, protein secretion through both membranes of gram-negative bacteria is not due to the operation of a simple mechanism but is rather the outcome of multiple systems (Lory 1992, Wandersman 1992, Salmond & Reeves 1993). The most common secretion route, the general secretory pathway (GSP), is a two-step process in which the secreted protein crosses the inner membrane to the periplasm by the GEP and crosses the outer membrane by a separate mechanism (Pugsley 1993). The GEP has been studied to many years and it consists of at least six distinct proteins: the peripheral inner membrane protein SecA, and the integral inner membrane proteins SecD, SecE, SecF, SecG (band1) and SecY, each acting at distinct stages during export. In addition, to be efficiently exported, many but not all proteins must be prevented from premature folding or aggregation by cytoplasmic chaperones such as SecB, DnaK, DnaJ, GrpE or GroES/EL (Nilsson et al 1993, Höhfeld & Hartl 1994, Rapoport et al. 1996, Wild et al. 1996). The leader peptidase is also involved in this process in the periplasmic place (Wickner et al. 1991, Dalbey & Heijne 1992).

The signal sequences of secreted proteins seem to be highly conserved. An *E. coli* derived signal peptide can usually be used to direct a eukaryotic protein to the export pathway of *E. coli* and vice versa (Fraser & Bruce 1978, Talmadge et al. 1980, Müller et al. 1982, Gray et al. 1985, Hall et al. 1990). The signal sequences of *E. coli* genes *ompA*, *ompF* and *lamB*, which encode

membrane proteins, and the signal sequences from the genes of *Bacillus subtilis* are among the most used for the secretion of heterologous proteins in *E. coli* (Ghrayeb et al. 1984, Smith et al. 1988a, Marullo et al. 1989, Shibui et al. 1989). As mentioned earlier, the export of some proteins may be enhanced by the co-over-expression of some factors known to be involved in GSP (Hockney 1994, Georgiou & Valax 1996).

A few *E. coli* proteins are naturally secreted into the culture medium (Müller et al. 1983). However, the leakage of periplasmic proteins may be induced by using *E. coli* with the outer membrane permeabilized by specific mutations or by over-production of some (fusion)proteins (Abrahmsén 1988, Georgiou et al. 1988, Blanchin-Roland & Masson 1989, Hsiung et al. 1989). The fact that the signal peptide is not sufficient to accomplish the transport of all proteins in *E. coli*, diminishes the utility of export/secretion for the enhancement of recombinant protein solubility (Bassford et al. 1979, Moreno et al. 1980, Ito et al. 1981).

***Protein degradation.*** Proteolysis may also be a serious problem when proteins are produced in *E. coli*. Many recombinant proteins are recognized as abnormal by the proteolytic system of the cell and rapidly hydrolyzed (Itakura et al. 1977). In *E. coli*, 24 different endoproteinases and at least 12 exopeptidases, distributed in cytoplasm, the inner membrane, and the periplasm have been identified. Indeed, according to Maurizi (1992) > 3% of the enzymatic activities present in a cell at any given time are proteolytic. This reflects the normal physiological importance of these highly specific proteases in regulating the amount of proteins and in eliminating damaged or abnormal proteins. Of the more than 90% energy-dependent protein degradation occurring in the cytoplasm two ATP dependent proteases, Lon and Clp, are responsible for 70-80% (for review see Maurizi 1992).

There are several ways to avoid unwanted degradation of recombinant proteins. As already stated, the production of recombinant protein in insoluble form usually protects it from degradation, but difficulties arise from the subsequent denaturation/renaturation procedure needed to obtain an active protein molecule (Cheng et al. 1981, Kleid et al. 1981). Also, as discussed earlier, the cellular location (cytoplasm/periplasm/medium) of the protein can effect its stability (Talmadge & Gilbert 1982). In some cases, fusion of extra material to the C- or N-terminus (or to both ends) of a recombinant protein or the production of the protein as a tandem repeat can help to protect it from degradation (Shen 1984, Schulz et al. 1987, Butt et al. 1989, Hammarberg et al. 1989). This especially concerns small proteins. The use of protease-deficient strains of *E. coli* offers yet one more possibility to decrease unwanted protein degradation (Gottesman 1990, Nakano et al. 1994).

In addition to all the factors discussed thus far, the physiology of the host greatly affects the overall efficiency of homo/heterologous gene expression in *E. coli*. Even if the genetic background of a strain sets the limits to the general physiology of the cell, it is the growth conditions (nutrients, temperature, and oxygen availability), which regulate the actual situation in the cell (Meyer et al. 1985). The addition of sorbitol, betaine, ethanol, low molecular weight thiols or

disulfides, glutathione, and NaCl in the growth medium, or use of a rich medium, may increase the solubility of some recombinant proteins (Blackwell & Horgan 1991, Moore et al. 1993, Chopra et al. 1994, Hockney 1994, Georgiou & Valax 1996, Samuelsson et al. 1996). Because the same expression vector can yeild varying amounts of protein in different strains of *E. coli*, it is desirable to have several possible strains available with which to start (Kaytes et al. 1986).

### 2.2.1.3 Fusion proteins

Fusion proteins have a large range of applications in the field of gene expression, which have made them a popular alternative to direct gene expression in *E. coli* as well as other expression systems. (for a review, see Uhlén & Moks 1990, Nygren et al. 1994, Nilsson et al. 1997). Among their many applications is their use as a probe for translational and transcriptional activity (Casadaban et al. 1983). The fusion protein strategy also enables the tracking and localization of the gene product to different compartments of the host cell (Takahara et al. 1988, Moks et al. 1987, Endo et al. 1995, Kain et al. 1995), thus providing a very useful tool in, for example, membrane traffic research (Marjomäki et al. 1998). It is also possible to constitues bifunctional molecules which have the activity of two different proteins (Uhlén et al. 1983, Sano & Cantor 1991a, Dübel et al. 1995). The stability of small proteins in particular can be increased when they are produced as a fusion protein (see previous section and Marston 1986). In addition the problems associated with a heterogeneous N-terminus in some recombinant protein may be avoided by *in vitro* cleavage of the fusion protein (Nygren et al. 1994). Further, the solubility of an otherwise insoluble protein may be increased (Smith & Johnson 1988, LaVallie et al. 1993, Samuelsson et al. 1994, Yasukawa et al. 1995) or, if required, decreased to obtain inclusion bodies (Nagai & Thorgersen 1984, Yansura 1990). Usually the produced fusion protein can also be easily detected using antibodies or specific activity of one end of the protein (II, Sano et al. 1992, Zentgraf et al. 1995, O'Shannessy et al. 1995, Pharmacia Biotech 1996). Last but not least, in order to obtain, for example, efficient biosensors, the orientation of the molecules on the sensor surface of the detector should be unique, a process which may be achieved by fusion strategy (O'Shannessy et al. 1995, Stempfer et al. 1996). There exists no other limits to further applications than the imagination of the scientist.

In addition to all the above-mentioned applications, fusion proteins are widely used to ease the purification of recombinant proteins, and a wide range of possible fusion tags/tails/affinity handles for purification purposes are available, each having a specific affinity for a certain ligand (for a review, see Uhlén & Moks 1990, Uhlén et al. 1992, Nygren et al. 1994). The most widely used of these are the polyhistidine tag (6-10 amino acid)(Smith et al. 1988b), glutathione-S-transferase (GST) from *Schistosoma japonicum* (Smith & Johnson 1988), staphylococcal protein A (Uhlén et al. 1983), and maltose-binding protein (MBP)(diGuan et al. 1988). The specific ligands of these proteins are

$Ni^{2+}$, glutathione, IgG, and starch respectively. The tails can usually be attached to the N- or C-terminus of the desired protein, and when a native gene product is desired a site for specific chemical or enzymatic treatment can be put between the fusion partners. Proteases like factor Xa, enterokinase, and H64A subtilisin, which enable the desired protein to be released from the fusion with its native N-terminus, are very useful for this purpose (Nygren et al. 1994).

Despite the fact that there are many heavily studied fusion partners currently available, none of them works perfectly in every case. Hence an obvious need remains for new partners with enhanced properties.

### 2.2.2 Synthesis in insect cells

Since 1983, when the baculovirus expression vector was introduced (Smith et al. 1983a,b), the baculovirus expression vector system (BEVS), has become a popular choice among eukaryotic expression systens, and its use has increased dramatically. The system has several unique features which account for its popularity over other expression systems. First, the insect cells used in baculovirus expression systems allow expression of a eukaryotic gene in a eukaryotic cell, thereby taking advantage of the pathways in these cells that facilitate the folding, modification, and assembly of the protein product. Second, compared to other higher eukaryotic expression systems, very high expression levels (up to 1 g/1 liter $\cong 10^9$ cells) can be achieved. Third, the system is safe to use. Baculoviruses are essentially non-pathogenic to mammals and plants, and they have a restricted host range, which is often limited to specific invertebrate species. They are not known to infect permissively any non-arthropod hosts, although at least the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) can penetrate some mammalian cells by endocytosis (see below). Indeed, wild-type baculoviruses have been used as efficacious pest-control agents for many years. Fourth, the capsid structure of baculoviruses allows the packaging and expression of very large genes, and the vectors are not dependent on helper viruses. Fifth, multisubunit proteins and particles can be expressed in this system by co-infecting cells with more than one baculovirus or a virus bearing multiple recombinant protein-encoding genes. Sixth, recombinant proteins are usually localized in the same subcellular compartment as the authentic protein when expressed in insect cells, i.e. the nuclear proteins will be transported to the nucleus, membrane proteins to the cell membrane and secreted protein will be secreted into the culture medium. Seventh, recent advances in baculovirus methodology has made it easy to construct recombinant viruses and has reduced the time required for this process to 7-10 days. Finally, eighth, the availability of cell lines capable of growing in a suspension culture (as well as plates) with a serum-free medium renders it relatively easy to scale up this system. Although all genes are not expressed well or properly by baculovirus expression vectors, the reasons listed above make them an attractive choice for those wishing to produce high levels of biologically active eukaryotic proteins (for a review, see Luckow &

Summers 1988, Miller 1988, Kidd & Emery 1993, O'Reilly et al. 1994, Patterson et al. 1995).

### 2.2.2.1 Baculovirus biology

Baculoviruses are a diverse group of viruses having a restricted host range, which often is limited to specific invertebrate species, especially to insects (Gröner 1986). They are not known to infect permissively any non-arthropod hosts although at least the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) can penetrate some mammalian cells by endocytosis. However, neither normal virus gene expression nor viral DNA replication has been observed in these cells (Hofmann et al. 1995). The double-stranded and circular DNA genome (80-200 kbp) of baculoviruses (Summers & Anderson 1972, Burgess 1977) is condensed into a nucleoprotein structure known as the core (Tweeten et al. 1980). The core is located within the flexible rod-shaped capsid, which is 40-50 nm in diameter and 200-320 nm in length (Harrap 1972a) and which can expand to accommodate even very large recombinant molecules (Fraser 1986). The core and capsid are known collectively as the nucleocapsid. Nucleocapsids are made in the nucleus of infected cells and they acquire membrane envelopes either by budding through the plasma membrane of the cell or by a nuclear envelopment process. Membrane-enveloped nucleocapsids are referred to as virus particles or virions (O'Reilly et al. 1994).

Three morphologically distinct groups of baculoviruses can be detected: nuclear polyhedrosis viruses (NPV), granulosis viruses (GV), and non-occluded viruses. In the NPV group, those virions that obtain an envelope by an intracellular envelopment process can be occluded within a paracrystalline protein matrix, forming large (1-5 μm) polyhedral occlusion bodies (also known as occluded viruses, OVs, and polyhedral inclusion bodies, PIBs) containing multiple virions. NPVs are further distinguished on the basis of whether they contain a single nucleocapsid (SNPV) or multiple nucleocapsids (MNPV) per envelope (Miller 1988). In contrast to NPVs, granulosis viruses (GV) have only a single virion embedded in a very small occlusion body. As indicated by the term non-occluded, baculoviruses in this group do not synthesize polyhedrin and so do not have an occluded viral form (O'Reilly et al. 1994).

The two forms of AcMNPV as well as those of other NPVs (the BV and the PIB) play different and important roles during the natural life cycle of the virus (Fig. 5). PIBs are formed during late phase of natural infection (know also as the occluded cycle) by the embedding of virions in the crystalline protein matrix, which is composed mostly of polyhedrin protein (Harrap 1972b). PIBs enable the horizontal infection of a larva by contaminating the plant on which the larva feeds (Granados & Federici 1986). The viruses are protected from environmental factors within the PIBs, but in the alkaline midgut of the larva the crystalline polyhedrin matrix is solubilized (Harrap & Longworth 1974) and the released viruses enter the midgut cells by receptor-mediated fusion

with the membrane of the microvilli (Granados & Williams 1986). During the lytic cycle of infection the cells release BVs from the basolateral area of the mid-gut cells (Keddie et al. 1989). Although it was previously thought that the spread of infection within the insect occurred via hemocytes in the hemocoel, this role has been recently ascribed to cells of the tracheal system (Vialard et al. 1995). Eventually the larva dies and the PIBs that are produced in the very late phase of infection are released into the environment, and the cycle can start again.

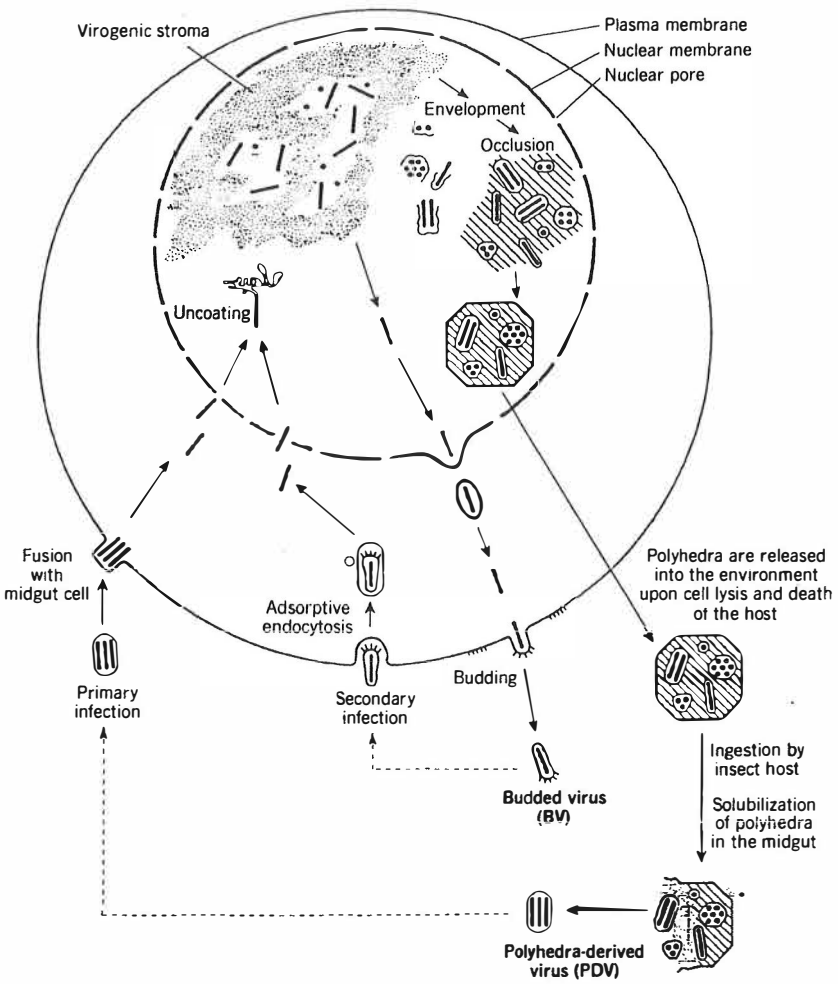



FIGURE 5 Life cycle of nuclear polyhedrosis virus. Polyhedra dissolve in the alkaline environment of the midgut to release nucleocapsids, which are uncoated, as they enter the nucleus. Viral genes are expressed in a temporally coordinated cascade to control DNA replication and the synthesis of structural proteins. Budded viruses spread the infection to other tissues within the host organism. Occluded viruses, produced late in the infection, are released into the environment when the infected cells lyse and the host organism dies. The figure is taken from Cleland & Craik (1996).

In cell culture, the infection cycle of baculoviruses can be considered to occur in three basic phases: early (the first 6 h), late (6-20 to 24 h postinfection), and very late (begins around 20 h postinfection). Biologically these phases correspond to reprogramming the cell for virus replication (BV production and OV production) (O'Reilly et al. 1994).

### 2.2.2.2 Baculovirus expression vectors

The prototype of the family *Baculoviridea* and the most extensively studied NPV type baculovirus is the *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV). Its genome (~134 Kbp) has been recently sequenced (Ayres et al. 1994). The other NPV-type baculovirus commonly used in expression vector construction, is the *Bonbyx mori* (single) nuclear polyhedrosis virus (BmNPV). The genome size of BmNPV is ~130 Kbp (Maeda 1989). The range of transfer plasmids and parent viruses available for the AcMNPV-based system, and the characteristics (growth, expression level) of the cell lines supporting AcMNPV, has made this the virus of choice for the cell culture production of recombinant proteins. However, because the size of the preferred larvae (*Bombyx mori*) for BmNPV infection is approximately 10 times larger than that for AcMNPV, a BmNPV -based expression system can be useful if insect larvae are used for recombinant protein overproduction (O'Reilly et al. 1994). The most commonly used cell lines with AcMNPV are SF-9 and SF-21AE. They both originate from IPLB-SF-21 cells, which were derived from *S. frugiperda* pupal ovarian tissue (Vaughn et al. 1977). A new cell line derived from *Trichoplusia ni* egg cell homogenates (BTI-TN-5B1-4 = High Five™, Invitrogen), however, has become a popular alternative to Sf cell lines due to fact that these cells have been shown to produce up to 28-fold more secreted proteins than any of the other insect lines (Wickham et al. 1992, Davis et al. 1993, Wickham et al. 1995). All these cell lines grow well in both monolayer and suspension culture and a serum-free culture medium can be used if desired.

In order to enable the expression of the recombinant protein in baculovirus systems, the gene for the desired protein is usually placed under a strong polyhedrin promoter (*polh*) of AcMNPV (O'Reilly et al. 1994). The *polh* is normally responsible for the synthesis of polyhedrin which can constitute up to 50% of the total protein of the infected cell. Fortunately, polyhedrin is not essential to virus replication or vertical infection in the cell culture (Smith et al. 1983b). However, the use of the *polh* promoter may be restricted in some cases by the fact that it is activated very late in the infection at a point when the host cell machinery for post-translational modifications is no longer working efficiently/properly. Problems with the *polh* promoter have been encountered, especially with proteins whose biological activity depends on proper glycosylation (Rosa et al. 1996). In such cases, the use of an alternative strong viral promoter which initiate transcription earlier in the infection while the host modification pathways are still functional (and so provide a longer period

for export and post-translational processing), could be useful. Indeed, Chazenbalk & Rapoport (1995) were able to produce a more highly glycosylated and functional extracellular domain of the human thyrotropin receptor under a late basic protein promoter (*late*). However, the expression level was low and factors other than the degree of glycosylation and signal peptide were also thought to play an important role in the expression of this and some other secreted proteins. Other promoters, which will activate earlier than *pohl*, include the promoters for the p10 gene (*p10*), the major capsid protein gene (*vp39*), the basic 6.9 kD protein gene (*cor*), and the viral ie1 gene (ie1). All these promoters are available in a variety of baculovirus transfer plasmids (Miller 1993, O'Reilly et al. 1994, Jarvis et al. 1996, Jones & Morikawa 1996).

### 2.2.2.3 Construction of recombinant baculoviruses

It was on account of the large size of baculovirus genomes (80kbp-200kbp) that the homologous recombination procedure was originally adopted to insert foreign genes into baculovirus genome instead of conventional plasmid cloning techniques (Smith et al. 1983a). In practice, the target gene is subcloned into a transfer vector containing a baculovirus promoter flanked by baculovirus DNA derived from a nonessential locus such as the polyhedrin gene of AcMNPV (most common). The viral DNA (usually AvMNPV) and transfer plasmid are then co-tranfected into insect cells (or more recently into yeast cells) where the recombination event takes place. Typically 0.1% to 1% of the resulting progeny are recombinant, which complicates their identification. Because in the most of the cases the target gene continues to be inserted into the polyhedrin locus, the altered plaque morphology of recombinant viruses can be used for identification. The cells in which the nuclei do not contain an occluded virus, thus contain a recombinant virus. However, the detection of the desired occlusion-minus plaque phenotype against the background of a greater than 99% wild-type parental virus is difficult. Viral identification may be facilitated by the introduction of a *lacZ* cassette along with the foreign gene to be expressed, which enables the recombinant viruses to be selected according to color (O'Reilly et al. 1994). Drug selection may also be used (Godeau et al. 1992).

Recently several techniques have been developed to further facilitate and speed up the construction of recombinant baculoviruses (for a review, see Davies 1994, Jones & Morikava 1996). By introducing a unique restriction site (Bsu361) at the polyhedrin locus, Kitts and coworkers (1990) were able to linearize the double stranded circular genome of AcMNPV. The linerization reduced the background of wild-type viruses and as a result, 10-25% of the progeny virus were recombinant. To obtain an even higher proportion of recombinants (85-99%), Kitts and Possee (1993) further modified the AcMNPV genome to enable Bsu361 digestion to also remove an essential gene (ORF 1629) from the AcMPMV genome. Infective viruses will only be reconstituted by

recombination with the transfer vector also carrying the gene of interest, whereby an intact ORF 1629 will be restored to the genome. The system also enables the color selection (*lacZ*) of recombinant viruses. However, it suffers from the need for time-consuming plaque assays to purify the recombinant virus. Novagen Inc. sells a commercial version of this system in which the target gene is amplified with specific primers using the polymerase chain reaction. This enables the ligase-free coupling of the linearized transfer vector and amplified gene in the mixture, which can then be used directly to transfect insect cells with cut viral DNA. The avoidance of cloning steps in *E. coli* speeds up the construction of the recombinant virus (Bishop et al. 1995).

In order to accelerate up the recombinant virus construction process the genetic material can be introduced into the baculovirus genome outside of insect cells. Patel and coworkers (1991) have reported novel method of propagating the viral genome by homologous recombination in the yeast *Saccharomyces cerevisae,* where the appropriate recombinants can be more easily selected. Viruses are then obtained by transfecting insect cells. This method is rapid (pure recombinant virus within 10-12 days), efficient and it ensures that there is no parental virus background. It also eliminates the need for time-consuming plaque assays, and multiple recombinants can be readily isolated simultaneously. The major disadvantages of this system are the need for experience of yeast culture and the incompatibility of traditional transfer vectors with the system. An even faster approach for generating a recombinant baculovirus uses site-specific transposition with Tn7 to insert foreign genes into bacmid DNA propagated in *E. coli* cells. The recombinant bacmid containing the *E. coli* clones are selected by color (*lacZ*), and the DNA purified from a single white colony is used to transfect insect cells (Lukow et al. 1993). The system has the same advantages as the yeast system but, in addition, it is faster (pure recombinant virus within 7-10 days) and easier to work with for those not familiar with the yeast system. The system is also commercially available (Bac-to-Bac™, Gibco BRL). The poor selection procedure of the original system has recently been overcome by introducing a temperature-sensitive selection methodology combined with the blue/white screening of recombinant viruses (Leusch et al. 1995).

The *in vitro* construction of a recombinant baculovirus genome by traditional cloning techniques was published by Ernst and coworkers in 1994. They introduced the I-*Sce*I meganuclease site into the AcMNPV genome by homologous recombination. The new virus genome, called Ac-omega, can be cut with I-*Sce*I meganuclease and the target gene bearing compatible ends can be ligated straight to the linearized Ac-omega DNA under a polyhedrin promoter. This method is simple and less time consuming than conventional homologous recombination, but owing to the normal background problems encountered with traditional cloning techniques, the need of plaque purification can not be entirely avoided.

#### 2.2.2.4 Optimization of protein production

There is usually no need to modify the regions responsible for transcriptional initiation in baculovirus expression vectors, because they have usually been optimized for efficient initiation. However, some attention should be paid to other factors unrelated to the promoter region if maximal protein production is wanted. All the factors listed below are known to be important to the efficient expression of heterologous genes in baculovirus systems, but the real effect of each of these factors to the overall yield of the desired protein remains to be studied case by case (for a review, see O'Reilly et al. 1994, Jones & Morikawa 1996).

##### 2.2.2.4.1 Nucleic acid modification

***Codon usage.*** Although the codon bias for highly expressed late baculovirus genes has been noted (Rohrmann 1986, Ayres et al. 1994, Ranjan & Hasnain 1994), it is unlikely that it seriously affects the translational efficiency of recombinant proteins in insect cells. An example of this is the *E. coli lacZ* gene which is efficiently expressed by the baculovirus expression system, although the gene has many codons that are underrepresented in late, highly expressed baculovirus genes (Pennock et al. 1984). However, it is preferable to use a UAA stop codon if possible, as it is used by most of the natural AvMNPV genes studied thus far (O'Reilly et al. 1994). As discussed in relation to the bacterial expression system, codon usage may also affect the secondary structure of the mRNA and in some cases unwanted stemloop structures may be avoided by the use of the correct codons. For the same reason consideration should be given to removing the leader region of the target gene if it is long (>30bp) or has a high G+C content (Pendergast et al. 1989).

***Promoter.*** The gene of interest should neither contain an internal TAAG sequence nor additional translation initiation codons (ATGs) upstream of the one intended for the proper initiation purpose. The TAAG sequence is the primary determinant of both late and very late promoter activity in insect cells and is located at the transcriptional start point of all known late and very late transcripts (Possee & Howard 1987, Wilson et al. 1987, Thiem & Miller 1989). TAAG sequences within heterologous genes, either in the sense or antisense direction, may adversely influence the level of heterologous gene expression. TAAG sequences in the sense direction would be expected to be initiated internally in the gene, leading to a rise in truncated gene products. On the other hand, TAAG sequences in the antisense direction (CTTA) could promote the formation of antisense transcripts leading to translational block or destabilization of the transcripts (Ooi & Miller 1990).

***Translation initiation context.*** In addition to obvious need to avoid ATG codons upstream of the proper translational initiation site, the immediate ATG context in the recombinant gene can affect the translation efficiency of the heterologous protein in insect cells. At least two major base requirements

emerge from the derived consensus initiation codon context (AAaATGa) for AcNMPV. In 91% and 70% of cases, respectively, A is found in -3 or -2 position (Ranjan & Hasnain 1994).

***mRNA.*** Most baculovirus expression vectors already contain the sequence for the polyadenylation signal and there is no need replace this with the gene of interest. However, the role of the polyadenylation sequence in the expression level is not clear. Even if it is not usually considered critical, it may have an effect on mRNA stability along with the other regions in the untranslated 5'- and 3'- ends of mRNA (Jackson & Standart 1990, Beelman & Parker 1995). Some specific AU-rich sequences have been shown to govern mRNA stability in the 3' untranslated regions, and eliminating such regions from the recombinant gene is worth of considering (Yost et al. 1990, Zubiaga et al. 1995). Although splicing has been shown to occur in insect cells (Jeang et al. 1987, Iatrou et al. 1989), strong expression of proteins from spliced mRNAs has not been observed. Thus, it is strongly recommended that the gene to be expressed in insect cells does not contain introns (O'Reilly et al. 1994).

#### 2.2.2.4.2 Cell-related factors

***Host cell.*** Not only the cell line *per se* (Wickham et al. 1995) but also the overall physiology of the host cell has an important effect on protein production and quality in baculovirus systems as well as in other systems (Jenkins 1996). For example, the highest levels of expression in *Spodoptera frugiperda* cells require healthy cells in which more than 97% are viable and in exponential growth phase with a doubling time of 16-18 hours (Lukow & Summers 1988). The physiology of the cells can be influenced by many factors and not least by the culture medium, which has to be optimized for each cell line. In addition to physical factors, such as oxygen availability and growth temperature, media supplements such as dextran sulfate or essential nutrients can further improve the yield and quality of recombinant protein (Kling 1996).

***MOI.*** Attention should also be given to the MOI (multiplicity of infection = number of plaque-forming units per cell added at the time of infection) used in cell infection. The relationship of the final protein concentration to the MOI is highly dependent on the growth phase of the cells prior to infection. Product yields are relative steady for MOI values from 0.1 to 100 with cultures infected during the early exponential phase but are strongly dependent on the MOI in the case of cells infected in the late-exponential growth phase. Although the maximum amount of end product has been achieved by infecting cells in the late-exponential phase with an MOI of 100, it is more practical to infect the culture earlier in the exponential phase with a lower MOI and obtain final product concentrations of the same order of magnitude (Licari & Bailey 1991). By using a low MOI ($\cong$ 0.1), the virus stocks can also be amplified without the risk of the accumulation of defective interfering particles which interfere with virus replication (Kool et al. 1991). However, extended passage of viruses in the cell culture, even with a low dilution, results in FP (few polyhedra) and

deletion (in viral *egt* gene) and insertion mutations (Kumar & Miller 1987, O'Reilly & Miller 1989).

***Protein degradation.*** Protein degradation may also be an important factor influencing heterologous protein yields in baculovirus systems. Proteins that are foreign to the cell, including recombinant proteins, mutated proteins, proteins incorporating exogenously supplied amino acid analogs, and proteins resulting from transcripional and translational errors, are usually recognized as abnormal by mammalian and bacterial cell systems (Goldberg & Dice 1974, Licari & Bailey 1991, Goldberg 1995). In eukaryotic systems, these intracellular abnormal proteins are usually degraded by ubiquitin-mediated process in proteasomes (Goldberg 1995, Hilt & Wolf 1996, Hochstrasser 1996). Ubiquitin-like protein has also been found in the genome of AvMNPV, but its function is unknown (Guarino 1990). The proteolytic machinery of insect cells has been not studied as extensively as other systems, and the extent to which the insect degradation system is really similar to known ubiquitin mediated degradation system remains to be studied more carefully. However, Licari and Bailey (1991) have shown that the degradation activity of the baculovirus system declines as the infection proceeds and is insignificant late in the infection when recombinant protein expression is intense. Fortunately, most baculovirus systems are based on a polyhedrin promoter which will be activated very late in the infection, but some degradation of recombinant protein may nonetheless occur, especially if early promoters are used. This may be avoided by modification of the N-terminal amino acid(s) of the recombinant protein, which is (are) known to regulate the half-life of proteins (Varshavsky 1992, Varshavsky 1996).In addition, modification of longer amino acid stretches known to involved in ubiqutin-mediated protein degradation in proteosomes, such as PEST (regions rich in the amino acid residues proline, aspartic acid, glutamic acid, serine and threonine) sequences (Kornitzer et al. 1994) and destruction box (RAALGNISN) sequences (Glotzer et al. 1991, Irniger et al. 1995) can be considered. Producing the target protein as a fusion protein or/and exporting it out of the cytoplasm may also help (see 2.2.1.3). In addition, it may be necessary to supplement the culture post-infection with 0.1% to 0.5% FBS (fetal calf serum) or BSA (bowine serum albumin) for the expression of some recombinant products and/or rAcMNPV to protect the recombinant product or virus from proteolysis (Gibco BRL, Bac-to-Bac™ manual).

***Post-translational modifications.*** In addition to the yield, the quality of recombinant protein is also an important factor in an optimal expression system. Post-translational modifications such as signal peptide cleavage, proteolytic cleavage, N-glycosylation, O-glycosylation, acylation, amidation, phosphorylation, prenylation, and carboxymethylation have been shown to occur in insect cells in ways similar to those in mammalian cells. Albeit the sites of such modifications are usually in identical positions on the produced proteins in insect and mammalian cells, the precise nature of the post-translational modifications of some recombinant proteins may differ from those found in authentic protein (O'Reilly et al. 1994). For example, in addition

to the protein heterogeneity problems encountered in all overexpression systems (Jenkins 1996), glycoproteins produced in the baculovirus system typically lack complex biantennary N-linked oligosaccharide side chains containing penultimate galactose and terminal sialic acid residues (James et al. 1995, Jarvis & Finn 1995). Problems with phosphorylation have also been encountered (Fuchs et al. 1995). As already discussed in the context of baculovirus promoters, this may be due to the failure of the N-glycosylation and phosphorylation systems to keep pace with the very high levels of expression occurring in the last phase of the infection process. As a solution one can try to use a promoter activating earlier in the infection.

***Chaperons.*** The amount of folded recombinant protein and efficiency of post-translational modifications might also be improved by co-expressing chaperons or/and enzymes needed for the proper modification of the desired recombinant protein. Jarvis and Finn (1996) have shown that expression of bovine β1,4-galactosyltransferase from the viral early *ie1* promoter can contribute to the host N-glycosylation pathway and add galactose to at least one N-linked oligosaccharide chain on gp64 (the major baculovirus virion glycoprotein). In addition, overproduction of BiP chaperone by modified cell line Tn5B1-4 (Kling 1996) or co-infection of it with the desired protein (Hsu et al. 1994) has been reported to increase the yield of recombinant protein. An improved glycosylation pattern of baculovirus-expressed recombinant protein has also been reported by Wagner and co-workers (1996) by the co-expression strategy. In addition, Ogonah et. al. (1996) have reported a new *Estigmene acrea*-derived insect cell line, Ea4, which is naturally capable of complex sugar formation. This demonstrates that screening for new potential insect cell lines, different from *Spodoptera frugiperda*, may lead to the discovery of cell lines more suitable as the hosts of proteins whose activity is strongly dependent on correct and complicated post-translational modifications. Finally, the phosphatase inhibitor okadaic acid has been shown to drastically enhance the overall phosphorylation of p53. This suggests that the degree of phosphorylation of p53 and, perhaps, any given phosphoprotein overexpressed in insect cells (Sf21) is mainly determined by high phosphatase activity rather than unfavorable kinase/substrate ratios (Fuchs et al. 1995).

# 3 AIM OF THE STUDY

Chicken egg-white avidin is a widely accepted and used tool in virtually all fields of biology and biotechnology. The protein-ligand interaction between avidin and biotin ($K_d \sim 10^{-15}$ M), which is the strongest known in the nature, forms the basis of a rapid and specific detection/affinity system, the major areas of application of which are in localization, isolation, diagnostics and affinity-based targeting studies. Although avidin is considerably more abundant and cheaper than its bacterial counterpart streptavidin, the sugar side chain and high pI of avidin have been a persistent obstacle to its use in avidin-biotin technology on account of the nonspecific (charge-related) binding of the molecule to extraneous material via these unwanted chemical properties (Wilchek & Bayer 1990). In order to improve the intrinsic properties of avidin by site-directed mutagenesis and further expand its usage and applications in avidin-biotin technology, the aims of this study were to:

1. Establish an efficient system for the overproduction of biologically active recombinant avidin in bacterial and eukaryotic cells.

2. Study the use of avidin as a versatile N- or C-terminal tag/tail to ease the detection and purification of fusion proteins.

3. Create expression plasmids suitable for recombinant protein production as N- or C-terminal fusions with avidin.

4. Develop a rapid and easy purification protocol for recombinant avidin and avidin fusion proteins.

# 4 SUMMARY OF MATERIALS AND METHODS

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

## 4.1 Cell lines and plasmids

The bacterial synthesis of recombinant avidin and avidin fusion proteins was carried out in *E. coli* JM109 cells (I,II). *E. coli* RRI cells were used for plasmid propagation in some cases (I-III). The site specific transposition of gene-encoding recombinant avidin to the bacmid carrying the baculovirus genome was carried out in *E. coli* DH10Bac cells (Bac-to-Bac™ expression system, Gibco BRL) (III, IV). A clonal isolate of the *Spodoptera frugiperda* cell line IPLB-Sf21-AE, Sf9 (ATCC CRL-1711), was used as a host for baculovirus-mediated production of recombinant avidin (III, IV). All bacterial expression constructs were based on the pTrc99A vector from Pharmacia Biotech (I, II). The baculovirus expression vector (pFastBac1) as well as other components for recombinant baculovirus production were obtained from Gibco BRL (Bac-to-Bac™ expression system) (III, IV). The gene fragments amplified by polymerase chain reaction (PCR) were cloned to suitable vectors using standard molecular biology methods (Sambrook et al. 1989). All constructs involving PCR amplification were sequenced. The PCR primers were designed to add sequence modifications as required into the amplified fragments (I-IV).

The bacterial cells were grown in Luria-Bertani medium at the desired temperature with vigorous shaking, and the protein expression was induced with isopropyl-β-D-thiogalactosidase (IPTG), as described in (I, II). The insect cells were maintained in serum-free culture medium (SF 900IISFM, Gibco BRL); however, to achieve efficient purification of recombinant avidin this

medium was replaced by biotin- and detergent-free SF 900IISFM for infection of cells with the recombinant virus (III-IV).

## 4.2 Protein purification

The bacterial recombinant proteins were liberated from the cells by sonication and insoluble and soluble fractions were separated by centrifugation as earlier described (I-IV). The recombinant avidin as well as avidin fusion proteins were purified from the soluble fraction by affinity chromatography using biotin-agarose (B-0519) from Sigma (I, II). The fusion proteins were also affinity purified with glutathione-Sepharose 4B from Pharmacia Biotech (II).

The insect cells were broken by lysis buffer or/and sonication and the recombinant avidin was purified from the soluble fraction by pH-dependent affinity chromatography, using 2-iminobiotin-agarose (B-4507) from Sigma (III, IV).

## 4.3 Protein characterization

The protein samples were analyzed by SDS/PAGE using the standard procedure described by Laemmli (1970). Immunoblotting was performed as described by Towbin et al. (1979). Recombinant proteins were visualized in the blot using avidin antibody from DAKO A/S (I-III) or from Sigma (IV). The secondary antibody was goat anti-rabbit IgG (H+L) alkaline phosphatase conjugate from Bio-Rad. Proteins in polyacrylamide gels were visualized with Coomassie blue staining (I-IV). Enzyme-linked immunoadsorbent assay (ELISA) was performed to quantify the amount of recombinant avidin in (III). The ELISA was performed as described by Joensuu et al. (1991). The Bradford assay was used to quantify the yield of AvHev (IV). Enterokinase-cleaved re-Hevein was analyzed by reverse phase chromatography on a PepRP 5/5 column from Pharmacia Biotech, and the N-terminal amino acid sequencing was performed using an ABI 494 A procise™ Sequencer from Perkin-Elmer Applied Biosystems (IV). The mass spectroscopy of re-Hevein was carried out using a Bruker Biflex II from Bruker-Franzen Analytic and the functional analysis by IgE-ELISA and ELISA inhibition assays (IV).

### 4.3.1 Biotin-binding activity

The biotin-binding activity of recombinant avidins was demonstrated in all cases by affinity purification using biotin- or 2-iminobiotin-agarose (I-IV). Biotin-binding activity was also shown by immunoblotting, using biotinylated alkaline phosphatase from Sigma as a probe (I) and by performing the

thermostability test as described by Bayer et al. (1995)(III, IV). In publication IV affinity sensor technology (IASyS, Affinity sensors) was also used for this purpose.

### 4.3.2 Quaternary state and glycosylation

The recombinant avidin was shown to form a natural-like tetramer by acetylating the sample prior to analysis by SDS/PAGE as described by Bayer et al. (1995)(III, IV). In order to study the glycosylation pattern, the sample was treated with Endo $H_f$ or PNGase F glycosidase and the results were revealed by immunoblotting (III, IV).

## 4.4 Microscopical methods

The effect of the growth temperature on the production of recombinant avidin in bacterial cells was analyzed by electron microscopy. Both conventional plastic sections and cryosections of *E. coli* cells induced for protein production were prepared (I). Cryosections were immunoblotted as described in Punnonen et al. (1992). Electron microscopy was also used to visualize the secretion of recombinant avidin into the culture medium by insect cells (III). The Sf9 cells infected with the recombinant virus as well as mock-infected cells were prepared for cryosectioning and immunoelectron microscopy as described by Marjomäki et al. (1990).

# 5 REVIEW OF THE RESULTS

## 5.1 Production of recombinant avidin in *E. coli* (I)

In order to produce recombinant avidin (re-Avd) in *E. coli*, a fragment encoding the amino acids 1-123 of chicken egg-white avidin was cloned to the pTrc99A expression plasmid (I, Fig. 1). For technical reasons due to the cloning the obtained expression plasmid, pAVEX8, encoded an re-Avd molecule with an extra N-terminal methionine (M) and C-terminal glutamine (Q) residue (M+aminoacids 1-123 from chicken egg-white avidin+Q). Following induction of the JM 109 cells carrying the pAVEX8 vector at +37°C, a distinct band of about 14 kDa was clearly visible in the immunoblot by avidin antibody. The estimated size of 14 kDa agrees well with the computer-predicted size for the re-Avd peptide (13960 Da, including the starting methionine). A rough estimation from the immunoblot of the re-Avd yield was 0.3-0.5 mg/l of culture (I, Fig. 2B).

Due to the low initial yield of re-Avidin with the pAVEX8 vector, a new vector, pAVEX15, was designed. As a result of the codon modification in the pAVEX15, the yield of re-Avd increased to a level in which the re-Avd band was also detectable in the Coomassie-stained SDS-PAGE gel (I, Fig. 2A). The yield was estimated from the blot to be at a minimum 1 mg/l (I, Fig. 2B), which is at least 2-4 times higher than the yield obtained using the pAVEX8 vector.

In order to reveal the mechanism underlying the improved expression from pAVEX15, computer prediction for possible secondary structures within the first 200 nucleotides of pAVEX15 or pAVEX8 encoding re-Avd mRNA was carried out. The prediction was done with the aid of the fold program in the GCG package using the default values (Zuker & Stiegler 1981). Interestingly, the changes in the codons in pAVEX15 seemed to destabilize the predicted

loop structure in the mRNA, including the ribosome-binding site (AGGA) and start codon (AUG), as compared to the corresponding region in pAVEX8 mRNA.

The form (soluble v. insoluble) of the produced bacterial re-Avd was studied by immunoblotting, which clearly showed that re-Avd was accumulating as inclusion bodies inside the bacterial cells at +37°C (Fig. 1). In order to get the protein into soluble form, the effect of the growth of the cells at a lower temperature was examined. Indeed, reduction in the growth temperature to or under +30°C resulted in the accumulation of re-Avd in mostly soluble form. This was also demonstrated by electron microscopy. Following growth at +37°C, inclusion bodies were visible in many cells, which were heavily labeled with avidin antibody (I, Fig. 3B,F).Contrary to this, following growth at +24°C, no inclusion bodies were found, and only the scattered cytoplasmic re-Avd was visible in the immunolabeled sections (I, Fig. 3E,F).

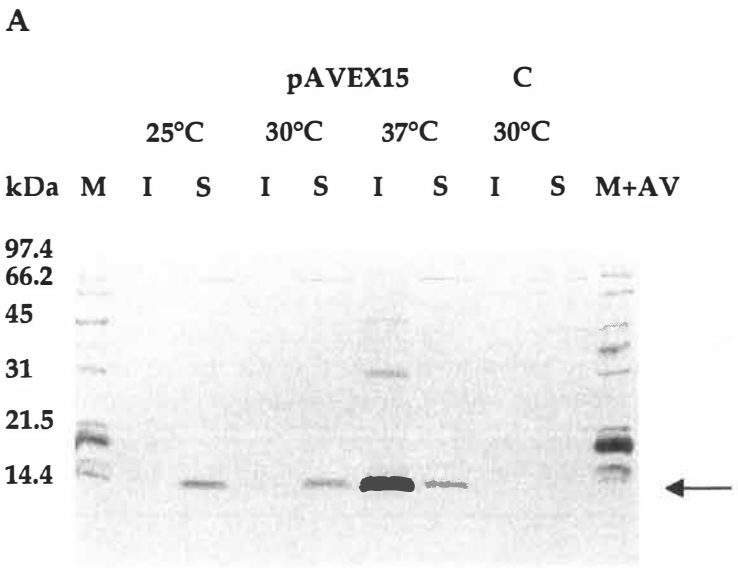



FIGURE 1 Immunoblot showing the effect of growth temperature on the form and yield of produced re-Avd. Lanes: (1) **M**, 10 µl of low range molecular weight standard (Bio-Rad) and 0.15 µg of Avd; (2-7) 30 µl of insoluble fraction (**I**) and soluble fraction (**S**) sample after 4 hours of induction at the indicated temperature; **C**, 30 µl of control samples after 4 hours of induction at 30°C i.e. JM 109 [pTrc99A]; (10) **M+AV**, 10 µl of low range molecular weight standard (Bio-Rad) and 1 µg of Avd; ←, re-Avd. The samples were prepared as described in figure legend 4 of original publication I.

The properties of the re-Avd were studied by immunoblotting and affinity chromatography. Antibody against natural avidin detected the re-Avd band in immunoblotting, which demonstrated that the produced re-Avd was

antigenically consistent with egg-white avidin (I, Fig. 2B). The immunoblotting with biotinylated alkaline phosphatase showed that the re-Avd had biotin-binding activity (I, Fig. 4A). This was further demonstrated by purification of re-Avd in a single step from the soluble fraction by biotin-agarose (I, Fig. 4B). However, it was only possible to purify a small proportion of the soluble re-Avd in the cells by the protocol used (I, Fig. 4C. Note that lane B corresponds to about five to seven times more of the original sample than lane U). By also taking our unpublished result into account, the recovery has been estimated to be at best 40-50% when Sarkosyl detergent has been used in the extraction buffer (1.5%) and a final concentration of 1M NaCl has been used in the soluble fraction before biotin-agarose purification (II).

We also constructed an expression vector for full-length avidin production in *E. coli* cells to study the possible effect of the amino acid modifications (deletion and an extra Q) in the re-Avd on the recovery of active protein (unpublished results). To achieve this, the fragment encoding the mature part of chicken avidin was amplified by PCR from avidin cDNA containing plasmid pGEMAV using primers AK10 (II) and AK25 (I) and the resulted fragment was cloned to the pTrc99A expression plasmid as a NcoI-HindIII fragment. The resulted plasmid, pAVn, encoded protein corresponding to the mature part of egg-white avidin (amino acids 1-128) with an extra M in the N-terminal end of the protein. The coding sequence for re-Avd in the pAVn was adapted for *E. coli* codon usage in the same way as the fragment in the pAVEX15 construct. As seen in Fig. 2, following the synthesis and purification of mature re-Avd, a band corresponding to the expected size of the mature part of avidin (14.3 kDa) was revealed by an avidin antibody from the immunoblot. The band migrated slightly slower than re-Avd but, as expected, as fast as the sugarless band in the avidin control sample. Fig. 2 suggests that the low recovery of re-Avd is not the result of the amino acid modifications made to the re-Avd *per se*, because the recovery of the full-length avidin is not better than that of re-Avd. In both cases the best yield of pure re-Avd was obtained from the samples induced overnight (~20h) at ~+24°C (Fig. 2).

## 5.2 Avidin fusion protein production in *E. coli* (II)

To study whether it is possible to use avidin as an affinity tail in fast and simple purification and in the easy detection of fusion proteins, we produced avidin as N- and C-terminally linked to *Schistosoma japonicum* glutathione *S*-transferase (GST). The expression vector encoding Avd-GST, pAV3X, was constructed from pAVEX15 by cloning into it a special linker region and GST encoding region from the pGEX-3X vector (II, Fig. 1). The GST-Avd vector was based on pGEX-3X into which a fragment corresponding to avidin cDNA region 114-569 was cloned to obtain p3XAV (II, Fig. 1).

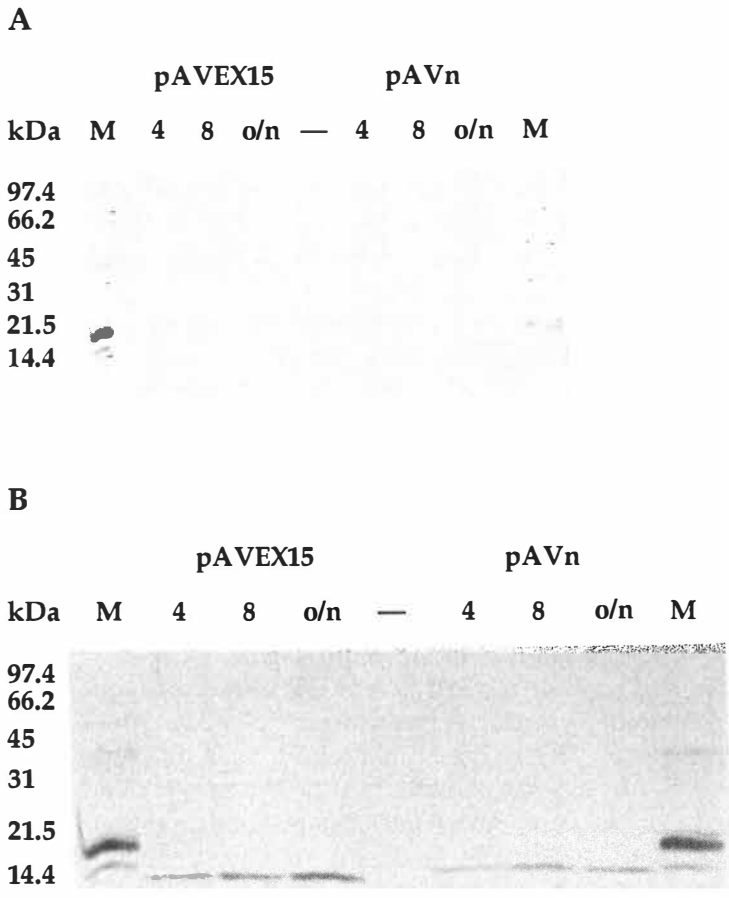



FIGURE 2 Immunoblots of 0.1% SDS-15% polyacrylamide-gels showing the amount of re-Avd left in crude extracts (**A**) and extracts purified (**B**) with biotin-agarose. The samples were prepared as described in Airenne et al. (II). Anti-avidin was used for detection. Lanes: (1) **M**, 10 µl of low range molecular weight standard (Bio-Rad) and 1 µg of Avd; (2-4) Samples after 4 h, 8 h and o/n of induction at 24°C using pAVEX15; (5) **—**, Empty lane; (6-9) Samples after 4 h, 8 h and o/n of induction at 24°C using pAVn. Samples in panels A and B correspond to about 230 µl and 1ml of original culture, respectively.

Both constructs produced a peptide of a size near to that predicted by computer (Fig. 2B). A protein corresponding to the band 41.2 kDa or 40.9 kDa was revealed by avidin antibody from induced cells carrying the vector pAV3X or p3XAV, respectively (In original publication (II) 39.1 kDa was given as the size of the p3XAV encoding peptide, but the revised corrected value is 40.9 kDa). This shows that avidin can be used as an antigenic tail for detecting fusion proteins. Interestingly, Avd-GST seemed to be more stable than GST-Avd. A major degradation product of about 30 kDa was evident in the samples

containing GST-Avd after immunoblotting (II, Fig. 2B) and it was also clearly present in the purified sample of GST-Avd (II, Fig. 4A).

The expression of both proteins was performed at +22°C to increase the solubility of the fusion proteins. At this temperature both fusion proteins were estimated (immunoblotting) to accumulate mostly in soluble form to a level of 5-10 μg/ml of culture (II, Fig. 2B). Significantly higher yields of both fusion proteins were obtained by growing the cells at the optimal bacterial temperature of +37°C, but most of the both fusion proteins were then in insoluble form (unpublished results).

By using a simple affinity purification protocol including biotin-agarose, we were able to get 1-2 mg/l culture (rough estimation from SDS-PAGE; II, Fig. 4A) of highly purified Avd-GST and GST-Avd in a single step which corresponds to approximately 10-20% of the total amount of each protein (II, Fig. 2B). Thus it is possible to produce biologically active avidin as an N- or C-terminal fusion protein in *E. coli*. Further, avidin can be used as an efficient purification tail at both ends of the fusion protein without destroying the natural affinity of avidin for biotin. Interestingly, the best yield of pure fusion proteins was obtained from samples collected after eight hours or even overnight induction, although the overall yield of both proteins did not seem to increase after four hours of induction (II, Fig. 3).

It was also possible to purify Avd-GST with glutathione-Sepharose 4B as well as with biotin-agarose, but purification of GST-Avd resulted in very poor yields (II, Fig. 4B). This suggests that avidin does not destroy the biological activity of GST. In accordance with this, the expression of avidin C-terminally linked to β-galactosidase (Avd-β-gal) also resulted in protein whose presence could be readily identified by following its β-gal activity. Figure 3 shows a photograph of *E. coli* JM 109 cells expressing Avd-β-gal on an LB-plate containing ampicillin and X-Gal. As can be seen from thefigure (Fig. 3) most of the cells are colored, indicating the presence of β-galactosidase activity inside the cells. This together with the GST-fusion results suggests that in fusion proteins avidin can be used as universal tail which does not harm the biological activity of the other part of the fusion protein, at least when the desired protein is linked to the C-terminal end of avidin.

## 5.3 Baculovirus expression system for avidin (III)

Avidin was produced in insect cells using a baculovirus expression system from Gibco BRL (Bac-To-Bac). The PCR amplified fragment, encoding a similarly truncated form of avidin as described in original publication (I) plus the natural signal sequence of avidin, was transferred to the baculovirus (*Autographa californica* nuclear polyhedrosis virus, AcMNPV) genome under a strong polyhedrin promoter in order to get recombinant baculoviruses bearing the re-Avd encoding gene fragment. In order to produce re-Avd, *Spodoptera frugiperda* cells (Sf9) were infected with these viruses.

Interestingly, after immunoblotting and immunostaining of the sample of infected insect cells, three different forms of re-Avd were revealed on the membrane by avidin antibody (III, Fig. 1B). The three intracellular forms of re-Avd were estimated to be 18, 17, or 15.5 kDa in size. The majority of the secreted re-Avd seemed to fit with the 17 kDa form, but the two other forms were also present in the culture medium (III, Fig. 1B). Whether the different species of secreted re-Avd were actually identical to the intracellular forms remains to be studied more carefully. Two of the three different species of insect re-Avd (18 and 17 kDa bands) were shown to represent differentially glycosylated forms of re-Avd by treatment of the purified re-Avd sample with Endo $H_f$ or PNGase F glycosidases (III, Fig. 3). Endo $H_f$ treatment removed the 18 kDa band, but seemed to have little effect on the other two forms of re-Avd. PNGase F treatment removed both of the high molecular-weight forms, leaving only the 15 kDa band. The 15.5 kDa form represented the unglycosylated form of re-Avd, as might be expected from the theoretical size 13,8 kDa (without carbohydrate) calculated by computer for recombinant product of this kind.

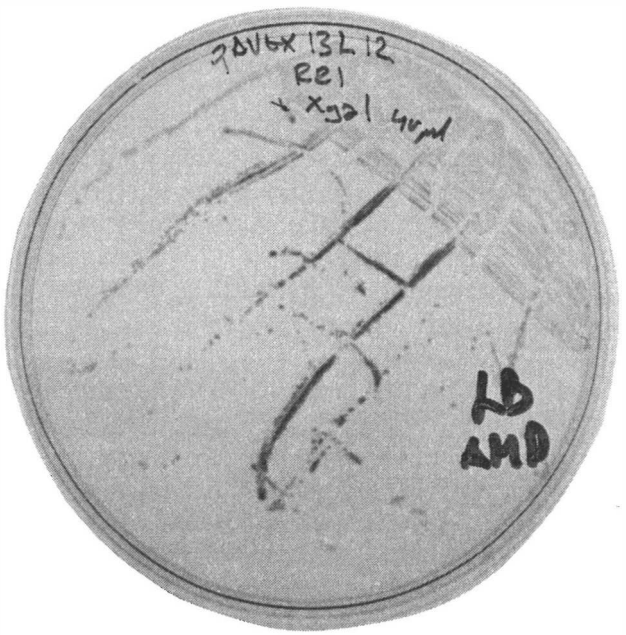



FIGURE 3 Photograph of Avd-β-gal synthesizing *E. coli* RR1 cells. Most of the cells are colored indicating the presence of β-galactosidase activity inside the cells.

The amount of re-Avd from both the cells and the culture medium was quantified by ELISA. Accordingly, approximately $5 \times 10^7$ cells produced ~0.61 mg of re-Avd (≅ 12.2 mg/L/$10^9$ cells), of which about 59% (~0.36 mg) was

secreted into the culture medium. Most (~0.23 mg, 92%) of the intracellular re-Avd was in soluble form after disrupting the cells in lysis buffer (III, Fig. 1B). We were able to increase the initial low yield of highly purified (2-iminobiotin-agarose affinity chromatography) re-Avd from a recovery of 12% (calculated from the intracellular fraction of soluble re-Avd; III, Fig. 1B) to over 90% by excluding the biotin from the culture medium after one day of infection (III, Fig. 1C). We were also able to purify the secreted re-Avd from the culture biotin-free medium (result not shown).

The production of re-Avd by Sf9 cells was also studied by immunoelectron microscopy (III, Fig. 2). Immunolabelling showed high amounts of re-Avd inside the vesicular compartments, which most likely participate in the secretory pathway of infected Sf9 cells (III, Fig. 2B). Large numbers of vesicles were observed close to the surrounding plasma membrane, suggesting that re-Avd is in the process of being secreted from the cells (III, Fig. 2A).

The properties of insect re-Avd were compared to egg-white avidin (Avd) by electrophoretic techniques, ELISA analysis and affinity purification. The thermostabilty test, in which re-Avd or Avd was incubated in different temperatures with or without additional biotin, resulted in a similar pattern of bands for both re-Avd and Avd as shown by immunoblotting (III, Fig. 4). Like Avd, biotin-saturated re-Avd was stabilized at +80°C, suggesting that re-Avd binds biotin in a manner similar to the native protein. The natural-like biotin-binding activity was also demonstrated by an ELISA-like method using biotinylated alkaline phosphatase as a primary antibody (unpublished results). According to the results, the biotin-binding activity of the re-Avd was estimated to be about 90% of that of Avd, which further supports the biological viability of baculovirus-produced re-Avd. In addition, the biotin-binding activity was also demonstrated by purification of re-Avd on a 2-iminobiotin-agarose column (III, Fig. 1A).

The re-Avd was shown to form native-like tetramers by acetylating the sample before the SDS-PAGE analysis. Acetylation inhibited the aggregation of positively charged re-Avd in the SDS-containing sample buffer, enabling the penetration of re-Avd in tetramer form into the gel at the ambient temperature. Interestingly, the re-Avd sample did not seem to contain dimer or oligo-tetramer forms, both of which were observed in the Avd sample (III, Fig. 5). This data indicates that, similar to Avd, re-Avd assembles into a stable tetramer in solution.

## 5.4 Avidin fusion protein production in insect cells (IV)

Hevein is a major cysteine-rich allergen of natural rubber latex which make it (and its mutated forms) an interesting object for detailed molecular studies of the immunological and allergic properties behind latex allergy (Alenius et al. 1996). In order to produce large amount of heveins for X-ray and NMR studies

and to test the usefulness of avidin-tag in BEVS, we produced Avd-hevein in insect cells. As a first step to obtaining AvHev-producing baculoviruses, a PCR-amplified fragment, encoding a recognition site for enterokinase and the hevein cDNA sequence, was cloned into pbacAVs+C vector (IV, Fig. 1). The resulted plasmid was named pbAv+ent+hev. The pbacAVs+C vector was derived from pbAvd$_r$ (III) by introducing a linker sequence C (II) encoding for the hinge region of mouse IgG3 and for a multiple cloning site (MCS) between the *Pst*I and *Hin*dIII sites of pbAvd$_r$. The pbacAVs+C thus encodes for a secretion-compatible form of recombinant avidin (aa $^{-}$24-$^{+}$123 of egg-white avidin and an extra Gln at the C-terminal end) connected C-terminally to the IgG hinge region and MCS (5′ - *Bam*HI, *Sma*I, *Sal*I, *Pst*I, *Hin*dIII - 3′ in frame). The plasmid can be used for the production and secretion of a desired avidin fusion protein in insect cells. All the above constructs were based on the pFastBAC1-donor plasmid and were designed for the construction of recombinant baculovirus genomes by site-specific transposition in *E. coli* using the Bac-TO-Bac Baculovirus Expression System. High titer AvHev baculovirus stock was obtained by transfection of Sf9 cells with the constructed recombinant genome. Automated DNA dideoxynucleotide sequencing was used to confirm all the cloned sequences.

Three days after infection a protein near the expected size of AvHev (20.2 kDa) was detected in the SDS-PAGE-gel (IV, Fig. 2A) and by immunoblotting of samples derived from recombinant baculovirus-infected Sf9 cells (IV, Fig. 2B). The band was absent from cells infected with the wild-type virus. In addition to this major band, other forms of AvHev were also identified (IV, Fig. 2B). The major band was detected by antibodies raised against avidin and hevein, which indicated that the fusion protein contained antigenic determinants for both components. Like re-avidin, AvHev seemed to be secreted efficiently into the culture medium by the avidin signal sequence (IV, Fig. 2B).

The AvHev was purified from the crude cell extract as well as from the culture medium by affinity chromatography on a 2-iminobiotin-agarose column (IV, Fig. 2A,B). In order to enable efficient recovery of the AvHev, biotin and Pluronic F-68 were excluded from the culture medium. As a result (ELISA), about 0.22 mg of highly purified fusion protein was obtained in a single step from 25ml culture (~5 X $10^7$ cells, about 8.8 mg/l). The recovery was efficient from both the intracellular and the extracellular source (IV, Fig. 2B). Most of the intracellular AvHev was soluble after disrupting the cells into lysis buffer (IV, Fig. 2B).

Affinity-purified AvHev separated into two clearly distinct bands in SDS-PAGE (IV, Fig. 3). The upper band was near the theoretical molecular weight of AvHev (20.2 kDa). The lower molecular weight band was close to the theoretical molecular weight (~15.7 kDa) of egg-white avidin and most probably reflects the avidin component as a stable degradation product of AvHev. In addition to the two distinct bands, several adjacent minor bands were detected (IV, Fig. 3). In order to further characterize these bands, the AvHev was deglycosylated with PNGase F glycosidase. The PNGase F treatment resulted only in bands of about 20.2 kDa and 14 kDa (IV, Fig. 3). This

indicates that the majority of the purified AvHev and degradation product (avidin) was glycosylated and that there was heterogeneity in the oligosacharides of the baculovirus-produced Avidin-tail of AvHev. This is in a good agreement with our studies on re-avidin production by BEVS (III).

In order to study whether avidin subunits connected to hevein are still capable of forming tetramers, the quaternary state of AvHev was studied by SDS-PAGE (IV, Fig. 3). Under nondenaturing conditions, most of the AvHev seemed to be in the tetramer or oligotetramer state, and the remainder was the free monomer.

Properties of AvHev were also studied by real-time optical biosensor technology (IASyS). The binding of purified AvHev to the biotin surface of an assay cuvette was both rapid and irreversible at concentrations of $1 \times 10^{7}$ M, thus closely resembling commercial chicken egg-white avidin (IV, Fig. 4). We were unable to dissociate the proteins from the cuvette by 20 mM HCl or even by an excess of free biotin. The AvHev bound to the biotin surface was specifically recognized by Anti-hevein even after the HCl or biotin treatment. This strongly suggests that the biotin-binding property of the avidin-tag (in AvHev) is indeed equal to that of egg-white avidin.

In order to cleave hevein from the avidin moiety, a proteolytic cleavage site for the factor Xa, a commonly used protease, was constructed between the C-terminus of avidin and the N-terminus of hevein. Factor Xa had, however, no effect on the fusion protein. We therefore constructed new viruses which encoded a cleavage site for enterokinase, thrombin, and the tobacco etch virus (TEV) poteases in place of the factor Xa site. It proved possible to cleave AvHev to avidin and hevein by enterokinase (IV, Fig. 3) or by thrombin (unpublished results), but it also seemed to be resistant (besides factor Xa) to TEV protease cleavage (unpublished results).

In reversed phase HPLC the enterokinase-cleaved hevein eluted after 20 min as a single narrow peak with base-line separation. N-Terminal sequencing of the recombinant hevein revealed the sequence DDDKEQCGRQAGGKLCPNNLCCSQ (unpublished results). This indicated that the recombinant hevein contained four extra amino acids in front of the published N-terminus of hevein. This was also confirmed by MALDI-TOF mass spectrometry in which the purified recombinant hevein gave a molecular mass of 5191.8 D (unpublished results). The determined molecular mass was equal to the mass of native hevein (4727.2 D) with disulfide bridges plus the mass of the sequence DDDK.

In ELISA, the binding of IgE in sera from latex-allergic patients and control subjects both to the recombinant and native hevein correlated highly significantly ($r = 0.92$, Pearson's product moment correlation). When the ability of the re-hevein (in solution) to inhibit the binding of hevein-specific IgE to native hevein (on solid-phase) was assessed and compared to that of native hevein, identical inhibition curves were observed (IV, Fig. 5). This indicated that the main IgE-binding epitope(s) of native hevein were present also in the re-hevein, and that in terms of their immunological properties, both preparations appeared to be identical.

# 6 DISCUSSION

Avidin is a widely used tool in numerous applications of the (strept)avidin-biotin technology (Wilchek & Bayer 1990). High amounts of it can be easily purified from chicken egg-white (Heney & Orr 1981) and there is no direct need to produce natural chicken egg-white avidin in any recombinant system. However, the high pI (~10.5) of the molecule and the presence of the sugars in avidin has been a persistent obstacle to its usage in some applications on account of its non-specific binding to extraneous material. It is for this reason that streptavidin, a nonglycosylated and neutrally charged bacterial relative of avidin, has displaced avidin in many (strept)avidin-biotin technology applications, although streptavidin is less abundant and more expensive than its eukaryotic counterpart.

In order to modify the unwanted intrinsic properties of avidin and to create an avidin more suitable for any application of (strept)avidin-biotin technology, we have established a production system for recombinant avidin in *Escherichia coli* (I) and insect cells (III) as reported in this study. Although avidin can be deglycosylated by enzymatic treatment or by incubating it with certain microbial culture (Bayer et al. 1995), and the positive charge can be decreased by chemical treatment (acetylation, succinylation, etc.; Bayer & Wilchek 1990), the recombinant avidin system offers an more direct way to achieve all these by site-directed mutagenesis. In addition, the recombinant system enables the production of avidins impossible by traditional methods. This can further expand the already widely used (strept)avidin-biotin technology. For example, avidin fusion proteins should prove useful in many applications due to the ease with which such hybrid proteins can be detected and purified using their biotin-binding activity (II, IV). The expression system for the avidin gene also enables the more detailed study of the exceptionally high (highest known non-covalent interaction) affinity between avidin and

biotin in detail, which can lead to a better understanding of the general laws behind such extremely high affinity between two molecules.

## 6.1 *E. coli* as a recombinant avidin host (I)

As an easy-to-handle and cheap expression system (Gold 1990), we decided to try first the *E. coli* expression system for recombinant avidin production. Although avidin is an eukaryotic protein, it is not heavily post-translationally modified and contains only one intramolecular disulfide bridge (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994). The sugars are not necessary for the biotin-binding of avidin (Hiller et al. 1987) and the size of the molecule is relatively small (15.5-15.8 kDa per identical subunit, Green 1975). Thus, at least in theory, the structural properties of avidin should not restrict its bacterial production (Goeddel 1990b).

According to the 3D-structure of avidin (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994), it is higly unlikely that the last five amino acids of the subunit are important for the molecule and hence, for cloning technical reasons, they were deleted from the re-Avd (I). The estimated size of 14 kDa for re-Avd corresponded closely to the computer-predicted size for unglycosylated re-Avd containing the starting methionine (13960 Da). Whether the N-terminal end of the re-Avd was correctly processed in *E. coli* was, however, not studied by N-terminal peptide sequencing, and it thus awaits further study. In any case, the bacterial cells produced biologically active re-Avd of the expected size, which we were able to purify in a single step by biotin-agarose affinity chromatography. The natural-like biotin-binding activity was also demonstrated by immunoblotting using biotinylated alkaline phosphatase as a probe. In accordance with the studies by Hiller and coworkers (1987) our results indicate that the oligosacharide side chains of avidin are not required for biotin-binding.

According to rough estimation from immunoblot, our first construct (pAVEX8) produced 0.3-0.5 mg re-Avd per litre of culture at +37°C (I, Fig. 2B). Although the final concentration of the recombinant protein is always protein-dependent, the yield was quite low. According to the literature the overall yield of recombinant protein per l of culture can at best approach to 1 g in the *E. coli* system (Gold 1990).

In order to improve the expression of re-Avd, we adapted the first ten codons of re-Avd to so-called optimal *E. coli* codon usage, i.e. they were replaced by codons which are the most common in highly expressed genes in *E. coli* (Zhang et al. 1991). Synonymous codons were used for the adaptation, so that the only difference between pAVEX8 and pAVEX15 was in codon usage. The resulting plasmid, pAVEX15, produced over 1 mg re-Avd per litre of culture, which was at least 2-4 higher than the yield obtained using pAVEX8 (I, Fig. 2B). At best the yield has been estimated to be ~ 5 mg/l of culture.

However, the estimation from the immunoblot is far from a quantitative measure, and the actual yield remains to be studied more carefully.

The positive effect of the codon modification for re-Avd production by *E. coli* cells is in accordance with reports from other groups (Robinson et al 1984, Thompson & Weber 1993). The importance of optimal codon usage *per se* in determining the expression of recombinant proteins has been contested for years (Ernst & Kawashima 1988, Kurland 1991, Date et al. 1992), because the substitution of rare codons by optimal ones affects not only the rate of elongation but also mRNA stability (Belasco & Higgins 1988), translation initiation frequency (Iserentant & Fiers 1980, Hall et al. 1982, Ganoza et al. 1987, Gross et al. 1990, Wikström et al. 1992), and translation accuracy (Kurland & Gallant 1996). Interestingly, the changes in the codons in pAVEX15 seemed to destabilize the computer-predicted loop structure in the mRNA, including the ribosome-binding site (AGGA) and start codon (AUG), as compared to the corresponding region in pAVEX8 mRNA (unpublished result). Thus, by exposing the essential control elements, the effect of the codon change on the secondary structures of the mRNA might partially explain the better expression from pAVEX15. The effect of codon choice *per se*, the stability of mRNA, all these together, or the effect of other still unknown factor(s) for improved expression cannot, however, be excluded.

The expression of recombinant proteins in *E. coli* is often limited by the formation of inclusion bodies (Marston 1986, Schein 1989). The only major advantage of inclusion body formation is that these cytoplasmic aggregates protect otherwise unstable recombinant proteins from degradation by host proteases (Cheng et al. 1981, Kleid et al. 1981). However, inclusion bodies usually complicate the purification of recombinant protein, because a case dependent denaturation/renaturation procedure has to be set up to obtain active protein (Marston 1986, Fischer et al. 1993). This is laborious, time-consuming and expensive.

Streptavidin has been expressed in insoluble form in *E. coli* and a successful denaturation/renaturation procedure has been established to obtain the protein biologically active (Sano & Cantor 1990). However, we wanted to obtain the re-Avd in soluble form in order to simplify the purification procedure and study of mutated re-Avd molecules in the future. Indeed, we were able to get re-Avd mainly in soluble form by simply lowering the culture growth temperature. The biochemical (5.1, Fig. 1) and microscopic studies (I, Fig. 3) clearly showed that re-Avd was produced mainly as inclusion bodies at +37°C, but was mainly in soluble form after expression at a temperature at or below +30°C. Our results thus strongly support previous reports suggesting the use of a lower growth temperature as an general and easy way to increase the solubility of an otherwise insoluble recombinant protein in *E. coli* (Schein 1988, Schein 1989, Schein 1993).

Although we were able to obtain re-Avd in mainly soluble form in *E. coli* and biologically active re-Avd was easy to purify from this fraction in a single step, the recovery from the soluble fraction was relatively low, being about 40-50% at best (I, Fig. 4C. Note that lane B corresponds to about five to seven

times more original sample than lane U). In addition, the expression of re-Avd in soluble form resulted in significantly lower overall yields of end product than expression at +37°C (I, Fig. 3; 5.1, Fig. 1). As a result, the amount of purified re-Avd remained low, yields at best amounting to about 0.5 mg/l of culture (unpublished results). In order to increase the recovery, and so yield of pure re-Avd, we also attempted the expression of re-Avd in a biotin-deficient strain of *E. coli* (SL 133, Lesley et al. 1993), in a biotin-free culture medium (M9) and in both of these to avoid the saturation of the re-Avd by biotin during expression (unpublished results). The recovery in general seemed to be better when the SL 133 strain was used, but in practice almost same yields of re-Avd were also achieved with the JM109 strain. This was due to the fact that the SL133 cells producing re-Avd did not grow well without either a rich medium or the addition of biotin to the medium in a concentration sufficient to saturate most of the biotin-binding sites of re-Avd. The toxicity of biologically active re-Avd to the cells was also indicated by the fact that the expression of re-Avd in soluble form in the JM109 strain resulted in the reduced growth of cells even when a biotin-rich medium (LB) was used (unpublished results). The use of an extra-rich medium (Terrific) seemed to result in slightly increased amounts of pure re-Avd after affinity chromatography in both cases. The effect was most probably due to the increased growth rate of the cells bearing re-Avd encoding plasmids in this medium (unpublished results).

The low recovery of re-Avd may also be partly explained by improper folding and assembly of the molecule in the cytoplasm of *E. coli*. In addition to the reducing environment present in the cytoplasm of the *E. coli* cells, the C-terminal deletion of the five last amino acids (124-128) could be involved in this. Although the deleted amino acids should not be critical for the biological activity of re-Avd on account of the 3-D structure (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994) of the molecule, we decided to express full-length re-Avd as well to find out if it results in a better recovery and/or higher yield (5.1, Fig. 1). The results suggest that there is no major difference between full length re-Avd and re-Avd in terms of recovery or yield. Thus the deletion does not effect the biotin-binding activity of re-Avd. This is also supported by re-Avd produced by baculovirus (III, Fig. 1C).

As an essential cofactor of biotin-dependent carboxylases, biotin plays an important role in all cells, being an absolute requirement for higher organisms (Dakshinamurti & Chauhan 1990). The detrimental effect of biologically active re-Avd to *E. coli* cells may thus be a hard/impossible obstacle to resolve in efforts to increase the yield and recovery of re-Avd, if re-Avd is produced in the cytoplasmic space in soluble/active form. One way to overcome or reduce this lethal effect might be to produce re-Avd into the periplasmic space of *E. coli*. As a naturally secreted protein, avidin should be recognized by the export machinery of *E. coli*. The periplasmic space should also favor the correct folding of re-Avd by providing an oxidizing environment in which the intramolecular disulfide bond of the re-Avd subunit can form more easily. We have not studied this option intensively, but our initial result indicates that the natural signal sequence of avidin is capable of leading re-Avd to the export

pathway of *E. coli* (unpublished results). By fusing the re-Avd peptide to the signal sequence of a readily exported protein in *E. coli*, an efficient re-Avd production system might be achieved into the periplasmic space of *E. coli* (Ghrayeb et al. 1984, Smith et al. 1988a, Marullo et al. 1989, Shibui et al. 1989).

One way to increase the overall yield and recovery of pure re-Avd might also be its production as inclusion bodies in the cytoplasm or propably into periplasm of *E. coli*. As a result, most of the produced re-Avd would then be inactive, allowing the fast growth of the host cells. This strategy, however, suffers from the fact that the subsequent denaturation/renaturation procedure has to be set up. Although this kind of protocol should be easy to adopt from the strategy designed to make monomer avidin affinity columns (Kohanski & Lane 1990), it might proved to be difficult to find a general scheme which works with all kinds of re-Avd molecules (mutated and fusion proteins) in future. In any case, it would make the purification of these molecules more laborious and time-consuming.

The overall yield of re-Avd might also be improved by producing it as a tandem repeats or as another kind of fusion protein in *E. coli*. This kind of strategy has been used successfully with some recombinant proteins, especially with small polypeptides (Shen 1984, Schulz et al. 1987, Butt et al. 1989, Hammarberg et al. 1989), which are recognized as abnormal by the proteolytic system of the cell (Itakura et al. 1977). Indeed, the results from our avidin fusion protein experiments (II) show, that the overall yield of avidin fusion proteins is considerably better than that of re-Avd in *E. coli*. However, this kind of re-Avd production strategy suffers from the fact that an additional protease digestion step is needed to separate the re-Avd from its fusion partner(s).

In conclusion, a biologically active recombinant form of chicken egg-white avidin can be produced in *E. coli* in a soluble form which can then be easily purified in a single step using affinity chromatography. However, without further improvement in the future, the relatively low production efficiency of re-Avd in *E. coli* reduce the utility of bacterial cells as an attractive host for avidin-related studies.

## 6.2 *E. coli* and avidin fusion proteins (II)

Chicken egg-white avidin meets most of the criteria required of an ideal polypeptide tail for use in fusion proteins. It is small (15.5-15.8 kDa per identical subunit), stable, heat resistant, soluble, and needs no co-factors to be functional (Green 1975). Although many peptide tails have been introduced since the early 1980s (Uhlén & Moks 1990, Cronan 1990), none of them is perfect. Avidin could thus offer an efficient new tool for fusion protein studies due to the ease with which such hybrid proteins can be detected and purified using its biotin-binding activity. In addition, a large number of reagents for (strept)avidin-biotin technology are already commercial available and Avd-

fusion proteins could further expand this already widely used technology (Wilchek & Bayer 1990).

In order to study whether it is possible to use Avd as an affinity tail in fast and simple detection and purification of fusion proteins, we produced Avd as C- and N-terminally linked with *Schistosoma japonicum* glutathione S-transferase (GST) in *E. coli* (II, Fig. 1A). A commercial pGEX-3X expression vector (Pharmacia Biotech) was used to synthesize a fusion protein in which Avd was linked to the C-terminus of GST. The GST-Avd encoding vector was named p3XAV. In order to enable production of GST (or any desired protein) as N-terminally linked with the C-terminus of re-Avd, a new expression vector, pAVEX16C, was constructed. Vector pAVEX16C was based on pAVEX15 (I), to which the IgG3 hinge region (Pack & Plückthun 1992) and multiple cloning site was added behind the re-Avd region (II, Fig. 1B). The hinge region was included in pAVEX16C to facilitate the autonomous folding and activity of both parts of the fusion protein. This particular hinge region was chosen as it meets well the theoretical criteria for a good one (Argos 1990) and it is from a natural source. A protease site between Avd and the desired protein was not included in pAVEX16C owing to the fact that it is easy to include this kind of site in any 5'- primer used to amplify the encoding fragment of the desired protein. This strategy also offers advantages over the pre-existing protease site in the expression vector. By including the protease site in the primer, any available protease can then be used to cleave the desired protein out of the re-Avd component. In addition, this enables the authentic N-terminal end to be created to any desired protein irrespective of the actual cloning site used in the expression vector.

The Avd-GST encoding expression plasmid was named pAV3X. The pAV3X plasmid encoded a slightly larger fusion protein (41.2 kDa) than p3XAV (40.9 kDa), which was also seen by immunoblotting and by SDS-PAGE (II, Fig. 4A). The region between the Avd and GST was mostly responsible for the difference in size between Avd-GST and GST-Avd (II, Fig. 1A). As mentioned above, Avd-GST contained the IgG3 hinge region (10 amino acids), which was not in the GST-Avd fusion protein. The difference in size between these two constructs was slightly compromised by the fact that in Avd-GST the Avd component was actually the re-Avd form (four amino acids smaller than natural Avd; I, Fig. 1), not natural Avd as in GST-Avd.

The yields of both fusion proteins were about 5-10 mg/l culture at room temperature (21-22°C). These yields were estimates from immunoblot (II, Fig. 2B), as our attempts to quantify the yield of Avd-GST or GST-Avd as well as some other Avd-fusion proteins by the ELISA method failed (Joensuu et al.1991). The reason for this is not known, but the antigenic epitopes of avidin may be shielded by the fusion partner, which prevents the Avd-antibody binding to epitopes of Avd.

In accordance with the re-Avd results, both fusion proteins were synthesized to notably (5-10 x) higher amounts at +37°C, but most of the proteins were then accumulated as inclusion bodies (unpublished results). As distinct from re-Avd, the best yield of pure fusion proteins (1-2 mg/l culture,

estimated from SDS-PAGE gel; II, Fig. 4A) was obtained from cultures grown at room temperature. Production at +30°C, and the use of *N*-laurylsarcosine (Frankel et al. 1991, Frangioni & Neel 1993) in the cell lysis buffer, resulted in both fusion proteins being produced in soluble form. However, for unknown reasons the amount of pure Avd-GST or GST-Avd obtained from this sample was lower than from an otherwise identically prepared sample grown at room temperature (unpublished results). Thus, although *N*-laurylsarcosine seemed to efficiently solubilize the Avd fusion protein inclusion bodies, the activity of the Avd-component of the *N*-laurylsarcosine solubilized fusion protein remained low. It may be that the Avd-fusion proteins existed in the inclusion bodies in such a form that a biologically active protein could not be obtained by simply solubilizing them with a lysis buffer containing *N*-laurylsarcosine. Biologically active Avd does anyhow contain a disulfide bridge, which seems to be important for the correct 3D structure and biotin-binding of Avd (unpublished results). *N*-laurylsarcosine was shown not to harm the biotin-binding activity of natural Avd when Avd was purified according to the same protocol as the fusion proteins (unpublished results). This is in accordance with the results of Ross and co-workers (1986), who found that Avd is quite stable against a number of detergents such as SDS, Tween-20 and Triton X-100.

Despite the fact that the expression of both fusion proteins seemed to be toxic to host cells at the lowered temperature, as was the case with re-Avd (cells containing the expression vector grew poorly), the overall yield of these proteins was still considerably higher than the yield of re-Avd. This may be explained in several ways. The Avd-fusion protein encoding gene may be more efficiently transcribed for mRNA than the re-Avd gene. The Avd-fusion protein mRNA may be more stable and it may be less secondary-structured than re-Avd mRNA. All these factors may lead to increased translation efficiency. On the other hand, re-Avd may be more stable in *E. coli* when fused to GST, which would be in accordance with the published results showing increased stability in some proteins when expressed as fusion proteins in *E. coli* (Shen 1984, Schulz et al. 1987, Butt et al. 1989, Hammarberg et al. 1989).

As with re-Avd (I), it proved possible to purify only a relatively small proportion (10-20%, estimation from immunoblot; II, Fig. 2B) of the soluble Avd-GST and GST-Avd with the affinity chromatography protocol used. This was most probably due mainly to the same reasons as discussed in conjunction with re-Avd. To avoid the toxic effect of the re-Avd component of the fusion protein on the cells, biotin had to be present in the culture medium in a concentration sufficient to saturate most of the four re-Avd biotin-binding sites. In addition, several other factors could explain the low overall recovery: (1) it may be that the folding of re-Avd is more difficult when synthesized as fusion protein than alone; (2) the fusion partner may decrease or destroy the biological activity of re-Avd by shielding the re-Avd biotin-binding sites; (3) because the affinity of the Avd-monomer to biotin ($K_d \sim 10^{-7}$ M) is far from what it is in tetramer ($K_d \sim 10^{-15}$ M) form (Kohanski & Lane 1990), any problems in oligomerization of the re-Avd caused by the fusion partner may result in poor yields of purified end product.

According to the purification results, the activity of the Avd component of the fusion protein seemed not to be affected by the orientation of the Avd in the fusion protein (II, Fig. 4A). Avd-GST was purified by biotin-agarose affinity chromatography as well as was GST-Avd. Avd-GST was also purified by glutathione-Sepharose-4B (II, Fig. 4B). On the contrary, barely any GST-Avd was purified by glutathione-Sepharose-4B (II, Fig. 4B), suggesting problems in folding or in the activity of the GST as C-terminally linked to N-terminus of re-Avd. This was interesting because GST-Avd was expressed using the p3XAV construct, which was based on a commercial expression vector designed especially for the production of desired proteins as N-terminally linked to the C-terminus of GST (Smith & Johnson 1988). It may be that in GST-Avd the lack of the linker region between GST and Avd leads to a situation in which these two molecules are oriented wrongly or too close together, and this creates problems in GST activity. In accordance with these results, the activity of β-galactosidase was not effected either when produced as N-terminally linked to the C-terminus of re-Avd (5.2, Fig. 2), suggesting that the desired proteins produced by the pAVEX16C vector are fully functional.

The comparison of the purification results of differentially prepared samples of Avd-GST and GST-Avd was complicated by the fact that the 2-iminobiotin-agarose tended to adhere to the wall of the eppendorf tube during the purification process. It was therefore not used for the purification of Avd-GST or GST-Avd in this study. However, it should be useful in the large scale retrieval of Avd fusion proteins in a reversible manner. At a high pH (>9.5) the free base form of 2-iminobiotin retains its specific high affinity binding to avidin ($K_d$ = 3.5 $10^{-11}$ M, Fudem-Goldin & Orr 1990), whereas at acidic pH values (<4), the salt form of the analog interacts poorly with avidin (Hofmann et al. 1980, Heney & Orr 1981, Orr 1981, Bayer et al. 1986). The disadvantage of using this affinity material is that a high pH must be used during the purification process, a condition that may not be tolerated by all polypeptides.

The use of biotin-agarose is complicated by the fact that the binding of Avd to biotin is in practice irreversible, and to elute out of the column the fusion proteins have to be denatured. However, for certain applications this may even be an advantage by enabling the coupling of the desired fusion proteins to different matrixes in an oriented and stable manner. The use of biotinylation reagents with cleavable spacer arms, i.e. S-S-biotin, which has a chemically cleavable disulfide bond (Shimkus et al. 1985), or a photocleavable biotin derivative (PCB-NHS ester) (Olejnik et al. 1995) remain a further possibility for the reversible purification of Avd fusion proteins. The disadvantages of using these chemicals for purification are that such columns are not reusable and that after elution the biotin will adhere to the avidin destroying its biological activity.

In addition to Avd-GST and GST-Avd we have also produced some other prokaryotic or eukaryotic proteins as fusions to the C-terminal end of re-Avd, such as Avd-β-galactosidase (130.9 kDa), Avd-EPA (Avd-barley endoproteinase A, 41.2 kDa), Avd-PrA and B (Avd-chicken progesterone receptor A and B; 87.2 kDa and 101.2 kDa, respectively), Avd-Parvo (Avd-part

of the nuclear localization signal of the canine parvovirus capsid protein, 23.2 kDa) and Avd-IgG3 (Avd-hinge region of mouse IgG3, 14.9 kDa) (unpublished results). All of these were easily detected by Avd antibody after production in *E. coli* cells, but only in the case of Avd-IgG3 and Avd-Parvo was the purification efficiency at same level as with GST fusions when biotin-agarose affinity chromatography was performed for purification. Purification efficiency seemed to be better the smaller the fusion partner, which is in good accordance with the published results on the purification of GST-fusion proteins (Frangioni & Neel 1993).

The problems encountered in the purification of Avd-β-galactosidase, Avd-PrA and B or Avd-EPA may, however, have also been due to some other factors than merely the molecular size of the fusion partner *per se.* In the case of Avd-β-galactosidase, the tetrameric nature of both re-Avd and β-galactosidase might lead to problems such as unnatural or intermolecular oligomerization of this fusion protein, resulting in aggregation or/and reduced Avd activity. The problems with Avd-PrA and B may be related to the extremely unstable nature of PrA and PrB.

Interestingly, the computer-predicted isoelectric point of the Avd fusion partner (EP-A, 4.34; β-galactosidase, 5.28; PrB, 5.10; PrA, 7.28; IgG3, 9.67; Parvo, 11.32) seemed to correlate well with the purification efficiency of the Avd-fusion protein with biotin-agarose so that the isoelectric point of the GST (6.26) seemed to be the below which under the recovery was poor, with the exception of Avd-PrA. This may reflect problems arising from the highly basic isoelectric point of Avd (10.43) when fused to a highly acidic fusion partner. The oppositely charged molecules may attract each other, resulting in molecules in which the biotin-binding activity of the Avd is blocked or destroyed. The purification efficiency also correlated with the predicted isoelectric point for the whole fusion protein (Avd-Parvo, 11.00; Avd-IgG3, 10.43; Avd-PrA, 7.95; Avd-β-galactosidase, 5.68; Avd-PrB, 5.10; Avd-EPA, 4.99). The limit for efficient (i.e. comparable to that with GST fusions) purification was again the isolectric point of Avd-GST and GST-Avd (7.98), with the exception of the Avd-PrA.

The Avd fusion protein system might be improved in several ways in future. As in the re-Avd case, the production of fusion proteins into the periplasmic space of *E. coli* might lead to the better recovery and high overall yield of pure peptides. The efficiency of the system might also be improved by reducing the isoelectric point of Avd by targeted mutagenesis. In this way the possible charge-related factors discussed above might be avoided. The use of an improved linker region between the Avd component and the desired protein partner of the fusion protein could also improve the system. By creating molecules of in which both the fusion protein partners have more space to fold independently and far enough away from each other might lead to improved results. By developing a general denaturation/renaturation protocol for Avd-fusion proteins, the overall yield could be markedly increased simply by synthesizing them as inclusion bodies at a higher growth temperature (+37°C). This strategy has been successfully used in streptavidin

fusion protein production (Sano et al. 1997). The fact that it makes purification more expensive, inconvenient and complicated, however, reduces the utility of this strategy as an attractive alternative.

## 6.3 Baculovirus expressed recombinant avidin (III)

The relative low efficiency of the *E. coli* production system in terms of an active and pure end product makes it inconvenient for site-directed mutagenesis studies of Avd. Although the yield of the Avd fusion proteins was at a reasonable level in *E. coli*, this system also suffered from a low rate of recovery. These facts combined lead us to search for a better production system for Avd. After on disappointment with the prokaryotic expression system, it was a natural next step to try a eukaryotic expression one. Avd is, after all, a eukaryotic protein, which is naturally secreted from the chicken oviduct into the egg-white. From among all the available eukaryotic expression systems the baculovirus expression vector system (BEVS, O'Reilly et al. 1994) seemed attractive for this purpose due to the recent advances in this field. The new virus construction techniques had speeded up recombinant virus construction, and the primary results of protein production could be achieved in about two weeks (Davies 1994, Jones & Morikava 1996 and Bac-to-Bac Baculovirus expression system, Gibco BRL). This and all the other advantages of the BEVS (see 2.2.2), such as high yields of recombinant proteins, led us finally to choose the BEVS as a next generation production system for Avd.

According to ELISA quantification, the insect cells produced about 12mg of re-Avd per 1 liter of suspension culture ($10^6$ cells/ml or $10^9$ cells/L). Most (59%) of this re-Avd was secreted into the culture medium, suggesting that the natural signal sequence of Avd is efficiently recognized and processed by the insect cell export machinery. Although some portion of the detected re-Avd in the culture medium might be and was due to the lysis of the cells late in the infection, the electron microscopy results also supported the notion that the natural signal sequence of Avd is efficiently recognized by the insect cell export machinery (III, Fig. 2). This is in line with previous results which have shown that signal peptides of mammalian, plant, and yeast origin are all able to direct proteins into the endoplasmic reticulum (ER) and be properly cleaved in baculovirus-infected cells (O'Reilly et al. 1994).

The overall production level of about 12 mg/l of re-Avd was good when compared to published production levels of other secretory glycoproteins. Although the levels of many recombinant proteins produced intracellularly in the BEVS are extraordinarily high (≥ 100 mg), in some cases approaching 1 g/l, the expression level of the secreted glycoproteins is usually much lower (0.5-22 mg/l). This is most probably due to the fact that the integrity of the secretion/processing machinery becomes compromised during the late stages of baculovirus infection (Wickham et al. 1995).

As in the bacterial system, the initial attempts to purify intracellular re-Avd by 2-iminibiotin-agarose resulted in a low recovery (12%) of highly purified re-Avd (III, Fig. 1B). However, the recovery was markedly increased (>90%) simply by excluding the biotin from the culture medium (III, Fig. 1C), indicating that the problems with recovery (activity) were connected with the precence and amount of biotin in the medium. Indeed, the original medium contained enough biotin to saturate the all the free biotin-binding sites of re-Avd.

Contrary to the bacterial system, biotin did not seem to be needed for the efficient progression of the baculovirus infection and the production of re-Avd. This is reasonable, because already within the first 6 h of infection, the normal cellular functions of insect cells decline precipitously. Within 6 to 24 h the cells cease to divide and the lytic baculovirus infection eventually leads to the death of the cell (O'Reilly et al. 1994).

The major intracellular form of re-Avd produced in the baculovirus-infected cells was a polypeptide of about 18 kDa (III, Fig. 1A). In addition, two other minor forms corresponding in size to about 17 kDa and 15.5 kDa were clearly detected by immunoblotting (III, Fig. 1B). The same three forms were also detected in the culture medium, but interestingly the 17 kDa form seemed to be the major polypeptide form in the culture medium (Fig. 1B). Many mammalian proteins have been shown to undergo N-glycosylation when synthesized in Sf9 cells. Thus it was not a surprising result that these different forms represented glycosylated (18 and 17 kDa) and unglycolsylated (15.5) form of re-Avd (III, Fig. 3) (O'Reilly et al. 1994). The 17 kDa form of re-Avd was partially resistant to Endo $H_f$ glycosidase, which removes immature high-mannose-type- ologosaccharides. This is in line with the previous reports that a high mannose N-linked oligosaccharide can be processed to Endo H-resistant form (to trimmed and fucosylated side chains) in insect cells (Jarvis & Summers 1989, Jarvis and Finn 1995). However, it is generally assumed that baculovirus-infected insect cells cannot process N-linked carbohydrates on most glycoproteins to complex forms, unless the glycoprotein is an unusually good substrate for the late processing reactions (Jarvis and Finn 1995). It will therefore be interesting to study the composition of the carbohydrate residues of re-Avd in more detail to find out if they correspond to those of natural Avd. Natural Avd has shown to contain at least three distinct oligosaccharide structural types of similar composition and size. In addition to oligomannosidic and bisected hybrid components, the avidin carbohydrate also contains nonbisected hybrid structures (Bruch et al. 1982).

Due to the extremely high pI (about 10.5) of Avd (Green 1975), it aggregates extensively when mixed at ambient temperatures with anionic detergents, such as sodium dodecyl sulfate (SDS) (Bayer et al. 1996). This has long complicated the study of the quaternary structure and thermostabilty of Avd by means of a simple SDS/PAGE run. Recently, however, acetylated Avd was found to behave in a manner similar to that of streptavidin; at room temperature, both proteins normally migrated mainly as tetramers with a tendency to form oligomers. When samples were boiled, both proteins

migrated mainly as monomers (Bayer et al. 1996). By using this method we showed that similar to the known quaternary structure of native egg-white Avd (Livnah et al. 1993, Pugliese et al. 1993, Pugliese et al. 1994), purified re-Avd was assembled mainly into tetramers, which exhibited a high level of thermostability, and was further stabilized upon binding biotin (III, Fig. 4). These results together with the glycosylation data strongly suggest that the biotin-binding and structural properties of the re-Avd are similar to those of natural egg-white avidin, and that the insect system is appropriate both for future site-directed mutagenesis studies and for the production of avidin fusion proteins. Indeed, over 30 different mutated versions of Avd and over 20 Avd fusion proteins have thus far been successfully produced by BEVS in our laboratory (unpublished results).

The yield of about 12 mg/re-Avd per liter is sufficient for all protein-chemical work with mutagenised re-Avd's, including X-ray studies. In practice, however, most of the baculovirus-produced re-Avd is secreted into the medium, from which, without having expensive concentration units, purification in large scale is more inconvenient than from inside the cells. For some reason re-Avd can not be recovered from the medium by affinity chromatography simply by optimally adjusting the pH and salt concentration as it can from inside the cells (Hofmann et al. 1980, Heney & Orr 1981, Orr 1981, Bayer et al. 1986). Some component of the medium, most probably Fluronic F-68, interferes with the interaction of re-Avd with the 2-iminobiotin column, decreasing the recovery; consequently, it has to be removed for efficient recovery (IV).

The solution to the practical problem arising from the purification of re-Avd from a large volume of medium as well as to providing for an increase in the overall yield of re-Avd might be its production without a signal sequence, i.e. to intracellular space. This would also provide an easy way to obtain unglycosylated forms of mutagenized re-Avd's with reduced non-specific binding properties (see beginning of this section) without the need to modify the N-linked glycosylation site of the polypeptides. We have not, however, been able to produce biologically active avidin in the cytoplasm of insect cells (unpublished results). This is somewhat surprising since biologically active avidin was produced in the cytoplasm of *E. coli* (I). It may be that certain essential factors, such as chaperons or foldases (Hartl 1996), which are needed for the correct folding of avidin, are sequestered in insect cells in the endoplasmic reticulum but are present in the cytoplasm of *E. coli*.

The overall yield of re-Avd might also be increased simply by using another host cell line such as BTI-TN-5B1-4 (High Five™, Invitrogen) cells, which are reported to produced secreted proteins more efficiently than Sf cell lines (Wickham et al. 1992, Davis et al. 1993, Wickham et al. 1995). This strategy, however, suffers from the fact that these cells grow poorly on suspension culture, and, as mentioned, the large-scale purification of re-Avd from culture medium is inconvenient.

## 6.4 Insect cells and multipurpose avidin-tag (IV)

Karp et al. (1996) and Oker-Blom et al. (1996) have recently reported on the successful production of cytoplasmic streptavidin fusion proteins by BEVS. However, according to our results, the streptavidin signal sequence, does not seem to be processed correctly by insect cells, which may complicate its use as a secretable protein-tag in BEVS (unpublished results). This coupled with the success in the synthesis and efficient purification of biologically active re-Avd from insect cells led us to study the value of avidin as a secretable multipurpose fusion protein-tag for BEVS.

As discussed earlier (6.2), avidin meets most of the criteria for an ideal tag. It is relatively small, soluble in water, very stable and efficiently secreted. In addition, it can be considered a versatile multidomain tag (Nilsson et al. 1996), since it can be used both as a specific label and as an isolation handle. In this study, we used avidin successfully as an antigenic domain and purification tag. In addition, AvHev could also have been detected by its biotin-binding activity. A wide variety of biotinylated reagents, such as biotinylated alkaline phosphatase or biotinyl-FITC, are commercially available for this purpose. Similarly, the purification of AvHev could have been accomplished using a biotin-column instead of 2-iminobiotin. This would not, however, have allowed us to release AvHev from the column, since avidin binds to biotin in an irreversible manner ($K_d \sim 10^{-15}$ M). On the other hand, for some applications the irreversible nature of the binding offers a unique advantage in enabling the tight interaction of a desired fusion protein with different matrixes in an oriented and stable manner.

As discussed earlier, the general use of 2-iminobiotin for the purification of avidin fusion proteins might be restricted, since pH 9.5 (or higher) is needed for a tight binding and a relatively low pH (4 or below) is required for the efficient elution of avidin from the column (Bayer et al. 1986). Elution can, however, be performed at all pH values by using biotin as an eluent, but this will destroy the biotin-binding activity of the avidin-tag. This is not a problem if biotin-binding activity is not needed after purification, since biotin does not effect the cleavage of the fusion partner from the avidin-tag or the antigenic properties of avidin. For applications of this kind, biotin or biotin derivatives such as S-S-biotin (Shimkus et al. 1985) or photocleavable biotin (Olejnik et al. 1995), might also offer an additional method of purification.

The results indicated that AvHev formed tetramers, since its migration in SDS-PAGE (IV, Fig. 3) and its association rate with the biotin surface (IASyS) and the stability of the complex were in line to that of the avidin tetramer (IV, Fig. 4). This strongly suggests that the avidin-tag folded and acted as an independent N-terminal domain in the fusion protein. On the other hand, hevein contained eight cysteine residues, all of which are involved in the formation of disulfide bonds. This could easily have led to the formation of a wrong intramolecular disulfide bridge between the cysteine residues of hevein and avidin. Mass-spectrometric analysis suggested, however, that all four

correct disulfide bridges were present in the re-hevein. The observed difference of 8.8 mass units between the determined (5191.8 D) and calculated (5200.6 D) molecular mass supports the idea that four disulfide bridges were formed in the re-hevein. The independent folding of both ends of avidin fusion proteins is also supported by our results from the production of N- and C-terminal avidin fusion proteins in *E. coli* (II, Fig. 4).

Real-time kinetic assays confirmed that AvHev had the strong and stable biotin-binding ability characteristic of tetrameric avidin (IV, Fig. 4). Due to the fact that hevein is the major allergen of latex (Alenius et al. 1996), the irreversible and oriented nature of AvHev binding to biotin thus enables the construction of biosensor cyvettes (IASyS) or ships (BIACore) for the testing and analysis of patients with potential latex allergy. Indeed, one major advantage of using avidin as a fusion-tag is that it can be used to form an irreversible binding between the desired fusion partner and different biotin-surfaces in an oriented manner. The small size of hevein (4.7 kDa) may partially explain the high biotin-binding activity of the Avd-tag seen in this study. The efficiency of the Avd-tag may vary with larger fusion partners, as it does, for example, with the GST-tag (Frangioni & Neel 1993).

The N-terminal sequencing of the purified hevein following cleavage by enterokinase revealed that it had four extra amino acids (DDDK) in the N-terminus. This was somewhat surprising since the enterokinase cleavage site (DDDDK↓) was added between avidin and hevein in order for the protease to cleave hevein in the native form from AvHev. The antigenic properties of the DDDK-hevein were, however, indistinguishable from native hevein (IV, Fig. 5), which makes it (and its mutated forms) an interesting object(s) for detailed studies of the immunological and allergic properties of hevein.

Despite the fact that hevein is considered to be a stable protein, it seemed to be prone to proteolysis in the insect cell system. Only small amounts of hevein were obtained when produced as such (unpublished results). Although the avidin-tag allowed for the efficient production of hevein, a part of the affinity-purified fraction continued to represent avidin as a stable degradation product of AvHev (IV, Fig. 3). Addition of the protease inhibitor mix to the cell extract during the purification process did not yield more intact AvHev. Several reasons can be tentatively proposed for the instability of re-hevein. First, hevein is naturally synthesized as preproprotein that is posttranslationally cleaved into aminoterminal hevein and the carboxyterminal domain. This may protect hevein from "unnatural" proteolytic attack when produced in the laticifer cells of the rubber tree (Lee et al. 1991). Second, the deleterious effect of baculovirus infection on insect cells may lead to proteolytic enzymes having access to the synthesizing route of AvHev, an event which does not occur in healthy rubber tree cells. Third, as mentioned earlier, it is a well-known fact that small proteins, especially, may be vulnerable to proteolytic attack when produced as recombinant forms, and their stability may be increased when produced as fusion proteins (Marston 1986).

In conclusion, our results prove that avidin works well in the baculovirus expression system as a multipurpose secretable tag. The avidin-tag is very

stable and retains its characteristically high biotin-binding activity. It is thus appropriate for use with the large range of commercially available avidin-biotin technology reagents. The ease, with which avidin fusion proteins can be detected, purified and immobilized, provides the basis for the use of this system as an elegant alternative in the production of eukaryotic fusion proteins. Indeed, thus far we have already successfully produced over 20 different avidin fusion protein forms in BEVS (unpublished results).

# 7 CONCLUSIONS

The study presented here is part of a larger project, the aims of which are (1) to understand better the molecular factors behind the extremely high affinity of chicken egg-white avidin to biotin and (2) to create a molecules more appropriate for use in numerous applications of (strept)avidin-biotin technology. The main focus of this study was to establish a recombinant production system for Avd. Another important focus was to study whether it is possible to use Avd as a versatile polypeptide tag in fast and simple detection and/or purification of desired proteins. The main conclusions are:

1. Biologically active re-Avd was produced in soluble form in *E. coli* cells despite the potentially harmful biotin-binding activity of Avd. The biologically active re-Avd was easily purified to high purity in a single step.

2. Biologically active re-Avd was also produced as C- or N-terminally linked to the desired protein in *E. coli* and insect cells. These kind of biologically active fusion proteins were easily detected and purified by the Avd-tag, using commercially available reagents.

3. The pAVEX16C expression plasmid was constructed and used to produce the desired proteins as N-terminally linked to the C-terminus of Avd. The biological activity of the fusion partner was not effected by the Avd-tag. Thus, it is possible to produce novel bifunctional Avd fusion proteins for different applications of avidin-biotin technology in *E. coli* with this vector.

4 The *E. coli* production system for re-Avd's did, however, suffer from the fact that the yield of soluble re-Avd's was relatively low and that it was possible to purify only a small portion of the produced polypeptide by the biotin-binding activity of re-Avd(-tag).

5. Contrary to the *E. coli* expression system, the baculovirus expression vector system (BEVS) was found to be an efficient way of producing biologically active re-Avd in insect cells. Both the yield and the recovery of the produced protein was high in this system. The biotin-binding and structural properties of baculovirus re-Avd were similar to those of natural Avd, and thus the insect system is appropriate for future site-directed mutagenesis studies of Avd.

6 BEVS also provided an efficient means by which to produce biologically active re-Avd fusion proteins. The avidin-tag was very stable, retained its characteristically high biotin-binding activity, and was appropriate for use with the existing large range of commercially available avidin-biotin technology reagents. The ease with which avidin fusion proteins can be detected, purified and immobilized provides the basis for the use of the avidin-tag as an elegant alternative-tag for secretable fusion protein production in BEVS.

*Acknowledgements*

Most of this work was carried out at the Department of Biological and Environmental Science at the University of Jyväskyä, although it was started at the Department of Biomedical Sciences at the University of Tampere. The study was supported by the Academy of Finland, Developmental Foundation of the Province of Central Finland, Ministry of Education, Technology Development Centre of Finland (TEKES), United States-Israel Binational Science Foundation (BSF) and the University of Jyväskylä.

First of all, I wish to express my deepest gratitude to my supervisor, Professor Markku Kulomaa for his inspiring guidance and valuable support throughout this study. Your everlasting optimism has encouraged me to go on so many times. I am also grateful for you for giving me the chance to try out new ideas independently and express freely my personal quirks. I also appreciate your warm friendship during these years.

I am indebted to the official referees of the thesis, Hans Söderlund and Ismo Ulmanen, for their valuable criticism and comments. I am especially grateful to them for their fast review of this study. I also wish to thank Michael Freeman for revising the language of this thesis.

I express my warmest thanks to all the members of the Kulomaa group, both in Jyväskylä and Tampere, for all their help and comments as well as for their contribution in creating a friendly and jovial atmosphere in the laboratory. My special thanks go to Olli Laitinen and Ari Marttila for being outstanding in the field of undergraduate students and to being respectable opponents in our numerous Four Hills events, and to Varpu Marjomäki for sharing our spacious room and putting up with me despite disagreement about which radio channel to switch on. My special thanks also go to Mervi Alhroth, Kati Juuti, Anna Laukkanen and Eevaleena Vierula for their patience during the scientific sessions with Olli and Ari, and to Irene Helkala and Anne Ettala for their excellent technical assistance throughout and their contribution in the war against cockroaches and smaller contaminants in the lab.

I further express my thanks to Harri Alenius, Edward Bayer, Nisse Kalkkinen, Markku Kulomaa, Olli Laitinen, Varpu Marjomäki, Jari Mikkola, Christian Oker-Blom, Timo Palosuo, Eeva-Liisa Punnonen, Paula Sarkkinen, M. Siti Arija, Meir Wilchek, and H. Y. Yeang, the co-authors of the original publications. Special thanks to Christian, who introduced me to the baculovirus technology, and Ed and Meir for their kind help and support on avidin-related matters.

My sincere thanks are also due to the entire personnel of the Department of Biological and Environmental Science at the University of Jyväskylä, for creating an inspiring and friendly atmosphere in which to do research. In particular, I acknowledge the first-rate technical assistance of Arja Mansikkaviita, Paavo Niutanen, Pasi Purhonen, Raija Vassinen, the co-operation I have enjoyed with Ilkka Porali, Ilari Suominen, Maija-Vihinen Ranta and Matti Vuento, and the great help I have had from Maija-Leena

Kiiveri, Anna-Liisa Kotiranta and Marjatta Suhonen with all the copying and red tape involved in the study.

My valued colleagues and dear friends, so many in number that naming them all is impossible, are warmly thanked for their support and the good times shared both at work and during leisure.

I also deeply thank my brother Tomi, grandmother Ida and my godmother "Kaaru" for their support and love throughout my life.

Finally, I wish to express my warmest thanks to my dear wife Paula for her love, support and patience. I also want to thank Teemu, my beloved son, for all the joy he has brought to my life.

This work is dedicated to Paula, Teemu and the loving memory of my parents.

# YHTEENVETO (Résumé in Finnish)

## Avidiinin ja avidiini-yhdistelmäproteiinien (fuusioproteiinit) tuotto bakteeri- (*E. coli*) ja perhossoluissa (Sf9)

Kananmunan valkuainen sisältää pieninä pitoisuuksina (0.05%) avidiininimistä valkuaisainetta. Avidiini on tunnettu sen erikoisen voimakkaasta kyvystä sitoa biotiini vitamiinia. Sidos on yksi voimakkaimmista tunnetuista ei-kovalenttisista sidoksista ($K_a \approx 10^{15}$ $M^{-1}$). Avidiinista onkin tullut laajalti käytetty leimaus- ja puhdistustyökalu monille biotieteiden ja lääketieteen tutkimusaloille. Nykyään puhutaankin avidiini-biotiini -teknologiasta. Avidiinin voimakas emäksisyys sekä molekyylin sisältämät sokeriketjut ovat kuitenkin haitanneet sen käyttöä eräissä sovelluksissa, koska nämä tekijät aiheuttavat taustaa (väärää signaalia) kokeissa. Niinpä streptavidiini, avidiinin neutraali ja sokeriton bakteeriperäinen sukulainen, onkin korvannut avidiinin monissa sovelluksissa, vaikka se onkin huomattavasti kalliimpi vaihtoehto. Lisäksi streptavidiininkin on huomattu aiheuttavan taustaa sitoutumalla epäspesifisesti solujen pintaan, mikä on ongelmallista erityisesti silloin kun halutaan tutkia kokonaisia soluja.

Väitöskirjatyöni tavoite on ollut kehittää tehokas tuottosysteemi rekombinantti-avidiinille (re-Avd), jotta pystyisimme tuottamaan suuria määriä ominaisuuksiltaan muunneltua avidiinia sen rakenteen ja toiminnan tutkimiseen sekä erilaisiin sovelluksiin. Lisäksi tarkoitukseni on ollut selvittää avidiinin käyttökelpoisuus fuusioproteeinien puhdistus kahvana, jonka avulla halutun yhdistelmä proteiinin eristystä ja puhdistusta muista isäntäorganismin proteiineista voitaisiin helpottaa. Taustaton avidiini sekä avidiini-fuusioproteiinit laajentaisivat edelleen jo ennestään laajaa avidiini-biotiini –teknologiaa ja tarjoaisivat uusia välineitä perus- ja soveltavan tutkimuksen käyttöön.

Biologisesti aktiivisen avidiinin tuottaminen suurina määrinä *E. coli*ssa osottautui odotettua vaikeammaksi. Yhdistelmä-DNA -tekniikoilla saimme kuitenkin tuottotason nousemaan kohtuulliseksi (n. 1mg/l) +37°C:ssa, kun muutimme tuotantovektorissa avidiinia koodaavan cDNA:n 10 ensimmäistä kodonia vastaamaan *E. coli*n voimakkaasti ilmenevissä geeneissä käytettyjä kodoneja. Ongelmaksi muodostui kuitenkin se, että +37°C:ssa tuotettu re-Avd ei ollut liukoista vaan kertyi soluihin ns. inkluusiojyväsinä. Tuotantolämpötilan laskeminen +25°C:een tai sen alle mahdollisti re-Avd:n tuottamisen liukoisessa muodossa, mutta tuotetun re-Avd:n määrä putosi tällöin huomattavasti. Avidiini osoittautui kuitenkin toimivaksi puhdistus kahvaksi *E. coli*ssa. Tuottamamme N- ja C-terminaaliset avidiini-fuusioproteiinit voitiin helposti osoittaa ja puhdistaa bakteerin muista proteiineista käyttäen hyväksi avidiini-kahvan biotiininsitomiskykyä. Tuotto-olosuhteiden vaikutus fuusioproteiineihin noudatti samaa kaavaa kuin re-Avd:n kanssa oli havaittu, mutta tuottomäärät olivat huomattavasti korkeammat kummassakin lämpötilassa.

Vaikka biologisesti aktiivisen avidiinin tai avidiini-fuusioproteiinien tuottaminen olikin mahdollista *E. coli*ssa, tuottotasot eivät kuitenkaan vastanneet odotuksia ja tarpeita. Niinpä löytääksemme tehokkaamman tuottosysteemin, kokeilimme seuraavaksi hyönteissoluja, jotka tarjoavat mahdollisuuden rekombinanttiproteiinien tuottoon tumallisissa soluissa baculovirusten avulla. Hyönteissolut osoittautuivatkin erinomaisiksi ominaisuuksiltaan munanvalkuaisen kaltaisen re-Avd:n tuottamiseen suurina määrinä (>12mg/l). Lisäksi ne soveltuivat hyvin tehokkaaseen avidiini-fuusioproteiini -tuottoon. Heveiini on luonnonkumiallergioiden pääallergeeni ja avidiini kahva mahdollisti sen (ja tulevaisuudessa sen mutatoitujen muotojen) tehokkaan tuoton ja puhdistuksen hyönteissoluissa. Näin voimme jatkossa selvittää yhteistyössä allergiatutkijoiden kanssa heveiinin korkean allergeenisyyden syitä molekyylitasolla.

# REFERENCES


Abrahmsén, L. 1988. Different strategies to utilize the staphylococcal protein A gene in the expression of recombinant proteins in bacteria. Academic dissertation. Department of Biochemistry and Biotechnology, Royal Institute of Technology, Stockholm.

Alenius, H., Kalkkinen, N., Reunala, T., Turjanmaa, K. & Palosuo, T. 1996. The main IgE-binding epitope of a major latex allergen, prohevein, is present in its N-terminal 43 amino acid fragment, hevein. J. Immunol. 156: 1618-1625.

Amann, E., Brosius, J. & Ptashne, M. 1983. Vectors bearing a hybrid trp-lac promoter useful for regulated expression of cloned genes in Escherichia coli. Gene 25: 167-178.

Argos, P. 1990. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J. Mol. Biol. 211: 943-958.

Ayres, M. D., Howard, S. T., Kuzio, J., Lopez-Ferber, M. & Possee, R. D. 1984. The complete DNA sequence of Autographa californica nuclear polyhedrosis virus. Virology 202: 586-605.

Balbas, P. & Bolivar, F. 1990. Design and construction of expression plasmid vectors in Esherichia coli. Methods Enzymol. 185: 14-37.

Bassford, P. J. Jr., Silhavy, J. T. & Beckwith, J. R. 1979. Use of gene fusion to study secretion of maltose-binding protein into Escherichia coli periplasm. J. Bacteriol. 139: 19-31.

Bateman, W. G. 1916. The digestibility and utilization of egg proteins. J. Biol. Chem. 26: 263-291.

Bayer, E. A. & Wilchek, M. 1990. Application of avidin-biotin technology to affinity-based separations. J. Chromatograph. 510: 3-11.

Bayer, E. A., Ben-Hur, H., Gitlin, G. & Wilchek, M. 1986. An improved method for the single-step purification of streptavidin. J. Biochem. Biophys. Methods 13: 103-112.

Bayer, E. A., Meester, F. D., Kulik, T. & Wilchek, M. 1995A. Preparation of deglycosylated egg-white avidin. Appl. Biochem. Biotech. 53: 1-9.

Bayer, E. A., Kulik, T., Adar, R. & Wilchek, M. 1995B. Close similarity among streptavidin-like, biotin-binding proteins from streptomyces. Biochim. Biophys. Acta 1263: 60-66.

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 17: 1319-1324.

Belasco, J. G. & Higgins, C. F. 1988. Mechanisms of mRNA decay in bacteria: a perspective. Gene 72: 15-23.

Beelman, C. A. & Parker, R. 1995. Degradation of mRNA in eukaryotes. Cell 81: 179-183.

Bishop, D. H. L., Novy, R. & Mierendorf, R. 1995. The BacVector system: Simplified cloning and protein expression using novel baculovirus vectors. Innovations 4: 1-6.

Blackwell, J. R. & Horgan, R. 1991. A novel strategy for production of a highly expressed recombinant protein in an active form. Febs Lett. 295: 10-12.

Blanchin-Roland, S. & Masson, J-M. 1989. Protein secretion controlled by a synthetic gene in Escherichia coli. Protein Eng. 2: 473-480.

Boas, M. A. 1927. The effect of desiccation upon the nutritive properties of egg-white. Biochem. J. 21: 712-724.

Brosius, J., Erfle, M. & Storella, J. 1985. Spacing of the -10 and -35 regions in the tac promoter. J. Biol. Chem. 260: 3539-3541.

Brown, C. M, Dalphin, M. E., Stockwell, P. A. & Tate, W. P. 1993. The translational termination signal database. Nucleic Acids Res. 21: 3119-3123.

Brown, C. M, Stockwell, P. A., Dalphin, M. E. & Tate, W. P. 1994. The translational termination signal database (TransTerm) now also includes initiation contexts. Nucleic Acids Res. 22: 3620-3624.

Browner, M. F., Rasor, P., Tugendreich, S. & Fletterick, R. J. 1991. Temperature-sensitive production of rabbit muscle glycogen phosphorylase in Escherichia coli. Protein Eng. 4: 351-357.

Bruch, R. C. & White, H. B. III 1982. Compositional and structural heterogeneity of avidin glycopeptides. Biochemistry 21: 5334-5341.

Bulmer, M. 1987. Coevolution of codon usage and transfer RNA abundance. Science 325: 728-730.

Burgess, S. 1977. Molecular weights of lepidopteran baculoviruses DNAs: Derivation by electron microscopy. J. Gen. Virol. 37: 501-510.

Butt, T. R., Jonnalagadda, S., Monia, B. P., Sternberg, E. J., Marsh, J. A., Stadel, J. M., Ecker, D. J. & Crooke, S. T. 1989. Ubiquitin fusion augments the yield of cloned gene products in Escherichia coli. Proc. Natl. Acad. Sci. USA 86: 2540-2544.

Carpousis, A. J., Stefano, J. E. & Gralla, J. D. 1982. 5' nucleotide heterogeneity and altered initiation of transcription at mutant lac promoters. J. Mol. Biol. 157: 619-633.

Casadaban, M. J., Martinez-Arias, A., Shapira, S. K. & Chou, J. 1983. β-Galactosidase gene fusions for analyzing gene expression in Escherichia coli and yeast. Methods Enzymol. 100: 293-308.

Chaiet, L. & Wolf, F. J. 1964. The properties of streptavidin, a biotin-binding protein produced by streptomycetes. Arch. Biochem. Biophys. 106: 1-5.

Chazenbalk, G. D. & Rapoport, B. 1995. Expression of the extracellular domain of the thyrotropin receptor in the baculovirus system using a promoter active earlier than the polyhedrin promoter. J. Biol. Chem. 270: 1543-1549.

Chen, G-F. T. & Inouye, M. 1990. Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the Escherichia coli genes. Nucleic Acid Res. 18: 1465-1473.

Cheng, Y-S. E. 1983. Increased cell buoyant densities of protein overproducing Escherichia coli cells. Biochem. Biophys. Res. Comm. 111: 104-111.

Cheng, Y-H. E., Kwoh, D. Y., Kwoh, T. J., Soltvedt, B. C. & Zipser, D. 1981. Stabilization of a degradable protein by its overexpression in Escherichia coli. Gene 14: 121-130.

Chew, L. C. & Tacon, W. C. 1990. Simultaneous regulation of plasmid replication and heterologous gene expression in Esherichia coli. J. Biotechnol. 13: 47-60.

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, 120. Proc. Natl. Acad. Sci. USA 92: 1754-1758.

Chopra, A. K., Brasier, A. R., Das, M., Xu, X-J. & Peterson, J. W. 1994. Improved synthesis of salmonella typhimurium enterotoxin using gene fusion expression systems. Gene 144: 81-85.

Cleland, J. L. & Craik, C. S. 1996 (ed.). Protein engineering. - 101-247 p. Wiley-Liss, New York.

Cronan, J. E., Jr. 1990. Biotination of proteins in vivo. J. Biol. Chem. 256: 10327-10333.

Dakshinamurti, K. & Chauhan, J. 1990. Nonavidin biotin-binding proteins. Methods Enzymol. 184: 93-102.

Dalbey, R. E. & von Heijne, G. 1992. Signal peptidases in prokaryotes and eukaryotes - a new protease family. Trends Biochem. Sci. 17: 474-478.

Date, T., Tanihara, K., Yamamoto, S., Nomura, N. & Matsukage, A. 1992. Two regions in human DNA polymerase β mRNA suppress translation in Escherichia Coli. Nucleic Acids Res. 20: 4859-4864.

Davies, A. H. 1994. Current methods for manipulating baculoviruses. Biotechnology 12: 47-50.

Davis, T. R., Wickham, T. J., McKenna, K. A., Granados, R. R., Shuler, M. L. & Wood, H. A. 1993. Comparative recombinant protein production of eight insect cell lines. In Vitro Cell. Dev. Biol. 29A: 388-390.

Delange, R. J. & 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.

DeBoer, H. A., Comstock, L. J. & Vasser, M. 1983. The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc. Natl. Acad. Sci. USA 80: 21-25.

Déthiollaz, S., Eichenberger, P. & Geiselmann, J. 1996. Influence of DNA geometry on transcriptional activation in Escherichia coli. Embo J. 15: 5449-5458.

Dickson, R. C. & Abelson, J. 1975. Genetic regulation: the lac control region. Science 187: 27-35.

DiGuan, C., Li, P., Riggs, P. D. & Inouye, H. 1988. Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. Gene 67: 21-30.

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 12: 512-517.

Durance, T. D. & Wong, N. S. 1992. Kinetics of thermal inactivation of avidin. Food Res. Int. 25: 89-92.

Dübel, S., Breitling, F., Kontermann, R., Schmidt, T., Skerra, A. & Little, M. 1995. Bifunctional and multimeric complexes of streptavidin fused to single chain antibodies (scFv). J. Immunol. Methods 178: 201-209.

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. 140: 535-543.

Elo, H. A. 1980. Occurrence of avidin-like biotin-binding capacity in various vertebrate tissues and its induction by tissue injury. Comp. Biochem. Physiol. 67B: 221-224.

Elo, H. A. & Korpela, J. 1984. The occurrence and production of avidin: a new conception of the high-affinity biotin-binding protein. Comp. Biochem. Physiol. 72B: 15-20.

Elo, H., Tuohimaa, P. & Jänne, O. 1975. Cumulative superinduction of avidin in the chick oviduct by tissue damage and actimycin D. Mol. Cell Endocrinol. 2: 203-211.

Elo, H. A., Kulomaa, M. S. & Tuohimaa, P. J. 1979a. Avidin induction by tissue injury and inflammation in male and female chickens. Comp. Biochem. Physiol. 62B: 237-240.

Elo, H. A., Kulomaa, M. S. & Tuohimaa, P. J. 1979b. Progesterone-independent avidin induction in chick tissues caused by tissue injury and inflammation. Acta Endocrinol. 90: 743-752.

Elo, H. A., Jänne, O. & Tuohimaa, P. J. 1980a. Avidin induction in the chick oviduct by progesterone and non-hormonal treatments. J. Steroid. Biochem. 12: 279-281.

Elo, H. A., Räisänen, S. & Tuohimaa, P. J. 1980b. Induction of an antimicrobial biotin-binding egg-white protein (avidin) in chick tissues in septic Escherichia coli infection. Experientia 36: 312-313.

Elo, H. A., Kulomaa, M. S. & Niemelä, A. O. 1981. Kinetics of avidin induction by tissue injury and inflammation in chicks and the dependence of induction on protein and RNA synthesis. Comp. Biochem. Physiol. 68A: 323-330.

Endo, T., Nakayama, Y. & Nakai, M. 1995. Avidin fusion protein as a tool to generate a stable translocation intermediate spanning the mitochondrial membranes. J. Biochem. 118: 753-759.

Ernst, J. F & Kawashima, E. 1988. Variations in codon usage are not correlated with heterologous gene expression in saccharomyces cerevisiae and Escherichia coli. J. Biotecnol. 7: 1-10.

Ernst, W. J., Grabherr, R. M. & Katinger, W. D. 1994. Direct cloning into the Autographa californica nuclear polyhedrosis virus for generation of recombinant baculoviruses. Nucleic Acids Res. 14: 2855-2856.

Fischer, B., Summer, I. & Goodenough, P. 1993. Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in Escherichia coli as inclusion bodies. Biotech. Bioeng. 41: 3-13.

Foster, K. A., Frackman, S. & Jolly, J. F. 1995. Production of enzymes as fine chemicals. In: Reed, G. & Nagodawithana T. W. (ed.), Biotechnology: 73-120. VHC Publishers Inc., New York.

Frangioni, J. V. & Neel, B. G. 1993. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal. Biochem. 210: 179-187.

Frankel, S., Sohn, R. & Leinwand, L. 1991. The use of sarkosyl in generating soluble protein after bacterial expression. Proc. Natl. Acad. Sci. USA 88: 1192-1196.

Fraser, M. J. 1986. Ultrastructural observations of virion maturation in Autographa californica nuclear polyhedrosis virus infected spodoptera frugiperda cell cultures. J. Ultrastruct. Mol. Struct. Res. 95: 189-195.

Fraser, T. H. & Bruce, B. J. 1978. Chicken ovalbumin is synthesized and secreted by Escherichia coli. Proc. Natl. Acad. Sci. USA 75: 5936-5940.

Frommer, W. B. & Ninnemann, O. 1995. Heterologous expression of genes in bacterial, fungal, animal, and plant cells. Annu. Rev. Plant Physiol. Plant. Mol. Biol. 46: 419-444.

Fuchs, B., Hecker, D. & Scheidtmann, K. H. 1995. Phosphorylation studies on rat p53 using the baculovirus expression system. Eur. J. Biochem. 228: 625-639.

Fudem-Golding, B. & Orr, G. A. 1990. 2-iminobiotin-containing reagent and affinity columns. Methods Enzymol. 184: 167-173.

Furlong, J., Meighan, M., Conner, J., Murray, J. & Clements, J. B. 1992. Methods for improved protein expression using pET vectors. Nucleic Acids Res. 20: 4668.

Ganoza, M. C., Kofoid, E. C., Marliére, P. & Louis, B. G. 1987. Potential secondary structure at translation-initiation sites. Nucleic Acid Res. 15: 345-360.

Georgiou, G. & Valax, P. 1996. Expression of correctly folded proteins in Escherichia coli. Curr. Opin. Biotech. 7: 190-197.

Georgiou, G., Shuler, M. L. & Wilson, D. B. 1988. Release of periplasmic enzymes and other physiological effects of β-lactamase overproduction in Escherichia coli. Biotech. Bioeng. 32: 741-748.

Ghosh, S., Basu, S., Strum, J. C., Basu, S. & Bell, R. M. 1995. Identification of conditions that facilitate the expression of GST fusions as soluble, full-length proteins. Anal. Biochem. 225: 376-378.

Ghrayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y. & Inouye, M. 1984. Secretion cloning vectors in Escherichia coli. EMBO J. 3: 2437-2442.

Gitlin, G., Bayer, E. A. & Wilchek, M. 1987. Studies on the biotin-binding site of avidin. Biochem. J. 242: 923-926.

Gitlin, G., Bayer, E. A. & Wilchek, M. 1988a. Studies on the biotin-binding site of avidin. Biochem. J. 250: 291-294.

Gitlin, G., Bayer, E. A. & Wilchek, M. 1988b. Studies on the biotin-binding site of streptavidin. Biochem. J. 256: 279-282.

Gitlin, G., Khait, I., Bayer, E. A., Wilchek, M. & Muszkat, K. A. 1989. Studies on the biotin-binding sites of avidin and streptavidin. Biochem. J. 259: 493-498.

Gitlin, G., Bayer, E. A. & Wilchek, M. 1990. Studies on the biotin-binding sites of avidin and streptavidin. Biochem. J. 269: 527-530.

Glotzer, M., Murray, A. W. & Kirschner, M. W. 1991. Cyclin is degraded by the ubiquitin pathway. Science 349: 132-138.

Godeau, F., Saucier, C. & Kourilsky, P. 1992. replication inhibition by nucleoside analogues of a recombinant Autographa californica multicapsid nuclear polyhedrosis virus harboring the herpes thymidine kinase gene driven by the IE-1(0) promoter: A new way to select recombinant baculoviruses. Nucleic Acids Res. 20: 6239-6246.

Goeddel, D. V. 1990a (ed.). Gene expression technology. Methods Entzymol. 185.

Goeddel, D. V. 1990b. Systems for heterologous gene expression. Methods Enzymol. 185: 3-7.

Gold, L. 1990. Expression of heterologous proteins in Esherichia coli. Methods Entzymol. 185: 11-14.

Goldberg, A. L. 1985. Functions of the proteasome: The lysis at the end of the tunnel. Science 268: 522-523.

Goldberg, A. L. & Dice, J. F. 1974. Intracellular protein degradation in mammalian and bacterial cells. Annu. Rev. Biochem. 43: 835-869.

Gope, M. L., Keinänen, 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. 15: 3595-3606.

Gottesman, S. 1990. Minimizing proteolysis in Escherichia coli: Genetic solutions. Methods Enzymol. 185: 119-129.

Gouy, M. & Gautier, C. 1982. Codon usage in bacteria: Correlation with gene expressivity. Nucleic Acids Res. 10: 7055-7074.

Gralla, J. D. 1990. Promoter recognition and mRNA initiation by Esherichia coli $E\sigma^{70}$. Methods Enzymol. 185: 37-54.

Granados, R. R. & Federici, B. A. 1986. The biology of baculoviruses, Vol. 1. p. 89-127 CRC Press, Boca Raton.

Granados, R. R. & Williams K. A. 1986. In vivo infection and replication of baculoviruses. In: Granados, R. R. & Federici, B. A. (ed.), The biology of baculoviruses, Vol. 1: 89-108. CRC Press, Boca Raton.

Gray, G. L., Baldridge, J. S., McKeown, K. S., Heyneker, H. L. & Chang, C. N. 1985. Periplasmic production of correctly processed human growth hormone in Escherichia coli: Natural and bacterial signal sequences are interchangeable. Gene 39: 247-254.

Green, N. M. 1963. Avidin. 4. Stability at extremes of pH and dissociation into sub-units by guanide hydrochloride. Biochem. J. 89: 609-620.

Green, N. M. 1975. Avidin. Avd. Protein Chem. 29: 85-133.

Green, N. M. & Toms, E. J. 1970. Purification and crystallization of avidin. Biochem. J. 118: 67-70.

Gribskov, M. & Burgess, R. R. 1983. Overexpression and purification of the sigma subunit of Escherichia coli RNA polymerase. Gene 26: 109-118.

Grosjean, H. & Fiers, W. 1982. Preferential codon usage in prokaryotic genes: The optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene 18: 199-209.

Gross, G., Mielke, C., Hollatz, I., Blöcker, H. & Frank, R. 1990. RNA primary sequence or secondary structure in the translational initiation region controls expression of two variant interferon-β genes in Escherichia coli. J. Biol. Chem. 265: 17627-17636.

Gröner, A. 1986. Specificity and safety of baculoviruses. In: Granados, R. R. & Federici, B. A. (ed.), The biology of baculoviruses, Vol 1: 177-202. CRC Press, Boca Raton.

Guarino, L. A. 1990. Identification of a viral gene encoding a ubiquitin-like protein. Proc. Natl. Acad. Sci. USA 87: 409-413.

György, P. 1954. In: Sebrell, W. H., Jr & Harris, R. S. (Ed.), "The vitamins": Vol. 1: 525-618. Academic Press, New York.

Hall, M. N., Gabay, J., Débarbouillé, M., & Schwartz, M. 1982. A role for mRNA secondary structure in the control of translation initiation. Nature 295: 616-618.

Hall, J., Hazlewood, G. P., Surani, M. A., Hirst, B. H. & Gilbert, H. J. 1990. Eukaryotic and prokaryotic signal peptides direct secretion of a bacterial endoglucanase by mammalian cells. J. Biol. Chem. 265: 19996-19999.

Hammarberg, B., Nygren, P-Å., Holmgren, E., Elmblad, A., Tally, M., Hellman, U., Moks, T. & Uhlén, M. 1989. Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II. Proc. Natl. Acad. Sci. USA 86: 4367-4371.

Harley, C. B. & Reynolds, R. P. 1987. Analysis of E. coli promoter sequences. Nucleic Acids Res. 15: 2343-2361.

Harrap. K. A. 1972a. The structure of nuclear polyhedrosis viruses. II. The virus particle. Virology 50: 124-132.

Harrap, K. A. 1972b. The structure of nuclear polyhedrosis viruses. I. The inclusion body. Virology 50: 114-123.

Harrap, K. A. & Longworth, J. F. 1974. An evaluation of purification methods for baculoviruses. J. Invertebr. Pathol. 24: 55-62.

Hartl, F. U. 1996. Molecular chaperones in cellular protein folding. Nature 381: 571-580.

Hasan, N. & Szybalski, W. 1987. Control of cloned gene expression by promoter inversion in vivo: Construction of improved vectors with a multiple cloning site and the $P_{tac}$ promoter. Gene 56: 145-151.

Hawley, D. K. & McLure, W. R. 1983. Compilation and analysis of Escherichia coli promoter DNA sequences. Nucleic Acids Res. 11: 2237-2255.

Hayakawa, T., Toibana, A., Marumoto, R., Nakahama, K., Kikuchi, M., Fujimoto, K. & Ikehara, M. 1987. Expression of human lysozyme in an insoluble form in yeast. Gene 56: 53-59.

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

Heinonen, P. K. & Tuohimaa, P. 1976. Induction of avidin in the chick oviduct by tissue damage. Effect of promethazine chloride, $CaCl_2$ and hydrocortisone on local induction. Mol. Cell. Endocrinol. 5: 209-214.

Heinonen, P. K. & Tuohimaa, P. 1978. The effects of 6-hydroxydopamine, atropin and propranolol on avidin synthesis by the chick oviduct. Med. Biol. 56: 161-162.

Hendrickson, W. A., Pähler, A., Smith, J. L., Satow, Y., Merritt, E. A. & Phizackerley, R. P. 1989. Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. Proc. Natl. Acad. Sci. USA 86: 2190-2194.

Heney, G. & Orr, G. A. 1981. The purification of avidin and its derivatives on 2-iminobiotin-6-aminohexyl-sepharose 4B. Anal. Biochem. 114: 92-96.

Hertz, R., Fraps, R. M. & Sebrell, W. H. 1943. Induction of avidin formation in the avian oviduct by stilbestrol plus progesterone. Proc. Soc. Exp. Biol. Med. 52: 142-144.

Hertz, R. & Sebrell, W. H. 1942. Occurrence of avidin in the oviduct and secretions of the genital tract of several species. Science 96: 257.

Hiller, Y., Gershoni, J. M., Bayer, E. A. & Wilchek, M. 1987. Biotin-binding to avidin. Biochem. J. 248: 167-171.

Hiller, Y, Bayer, E. A. & Wilchek, M. 1991. Studies on the biotin-binding site of avidin. Biochem. J. 278: 573-585.

Hilt, W. & Wolf, D. H. 1996. Proteasomes: Destruction as a programme. Trends Biol. Sci. 21: 96-102.

Hochstrasser, M. 1996. Protein degradation or regulation: Ub the judge. Cell 84: 813-815.

Hockney, R. C. 1994. Recent developments in heterologous protein production in Escherichia coli. Trends Biotech. 12: 456-463.

Hofmann, K., Wood, S. W., Brinton, C. C., Montibeller, J. A. & Finn, F. M. 1980. Iminobiotin affinity columns and their application to retrieval of streptavidin. Proc. Natl. Acad. Sci. USA 77: 4666-4668.

Hofmann, C., Sandig, V., Jennings, G., Rudolph, M., Schlag, P. & Strauss, M. 1995. Efficient gene transfer into human hepatocytes by baculovirus vectors. Proc. Natl. Acad. Sci. USA 92: 10099-10103.

Hsiung, H. M., Cantrell, A., Luirink, J., oudega, B., Veros, A. J. & Becker, G. W. 1989. Use of bacteriocin release protein in E. coli for excretion of human growth hormone into the culture medium. Biotechnology 7: 267-271.

Hsu, T-A., Eiden, J. J., Bourgarel, P., Meo, T. & Betenbaugh, M. J. 1994a. Effects of co-expressing chaperone BiP on functional antibody production in the baculovirus system. Prot. Express. Purif. 5: 595-603.

Höhfeld, J. & Hartl, F. U. 1994. Post-translational protein import and folding. Curr. Opin. Cell Biol. 6: 499-509.

Iatrou, K., Meidinger, R. G. & Goldsmith, M. R. 1989. Recombinant baculoviruses as vectors for identifying proteins encoded by intron-

containing members of complex multigene families. Proc. Natl. acad. Sci. USA 86: 9129-9133.

Ikemura, T. 1981a. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes. J. Mol. Biol. 146: 1-21.

Ikemura, T. 1981b. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the E. coli translational system. J. Mol. Biol. 151: 389-409.

Irniger, S., Piatti, S., Michaelis, C. & Nasmyth, K. 1995. Genes involved in sister chromatid separation are needed for B-type cyclin proteolysis in budding yeast. Cell 81: 269-277.

Iserentant, D. & Fiers, W. 1980. secondary structure of mRNA and efficiency of translation initiation. Gene 9: 1-12.

Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F. & Boyer, H. W. 1977. Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin. Science 198: 1056-1063.

Ito, K., Bassford, P. H. Jr., Beckwith, J. 1981. Protein localization in E. coli: Is there a common step in the secretion of periplasmic and outer-membrane proteins ? Cell 24: 707-717.

Jackson, R. J. & Standart, N. 1990. Do the poly(A) tail and 3' untranslated region control mRNA translation?. Cell 62: 15-24.

James, D. C., Freedman, R. B., Hoare, M., Ogonah, O. W., Rooney, B. C., Larionov, O. A., Dobrovolsky, V. N., Lagutin, O. V. & Jenkins, N. 1995. N-glycosylation of recombinant human interferon-γ produced in different animal expression systems. Biotechnology 13: 592-596.

Jänne, J., Hyttinen, J-M., Peura, T., Tolvanen, M., Alhonen, L., Sinervirta, R. & Halmekytö, M. 1994. Transgenic bioreactors. Int. J. Biochem. 26: 859-870.

Jarvis, D. L. & Finn, E. E. 1995. Biochemical analysis of the N-glycosylation pathway in baculovirus-infected lepidopteran insect cells. Virology 212: 500-511.

Jarvis, D. L. & Finn, E. E. 1996. Modifying the insect cell N-glycosylation pathway with immediate early baculovirus expression vectors. Na Biotechnology 14: 1288-1292.

Jarvis, D. L. & Summers, M. D. 1989. Glycosylation and secretion of human tissue plasminogen activator in recombinant baculovirus-infected insect cells. Mol. Cell. Biol. 9: 214-223.

Jarvis, D. L., Weinkauf, C. & Guarino, L. A. 1996. Immediate early baculovirus vectors for foreign gene expression in transformed or infected insect cells. Prot. Express. Purif. 8: 191-203.

Jeang, K-T., Holmgren-Konig, M. & Khoury, G. 1987. A baculovirus vector can express intron-containing genes. J. Virol. 61: 1761-1764.

Jenkins, N. 1996. Role of physiology in the determination of protein heterogeneity. Curr. Opin. Biotech. 7: 205-209.

Joensuu, T., Tuohimaa, P. & Vilja, P. 1991. Avidin and ovalbumin induction by progesterone in chicken oviduct detected by sensitive immunoenzymometric assays. J. Endocrinol. 130: 191-197.

Jones, P. D. & Briggs, M. H. 1962. The distribution of avidin. Life Sciences 11: 621-623.

Jones, I. & Morikawa, Y. 1996. Baculovirus vectors for expression in insect cells. Curr. Opin. Biotech. 7: 512-516.

Kain, S. R., Adams, M., Kondepudi, A., Yang, T-T., Ward, W. W. & Kitts, P. 1995. Green fluorescent protein as a reporter of gene expression and protein localization. Biotechniques 19: 650-655.

Karp, M., Lindqvist, C., Nissinen, R., Wahlbeck, S., Åkerman, K. & Oker-Blom C. 1996. Identification of biotinylated molecules using a baculovirus-expressed luciferase-streptavidin fusion protein. Biotechniques 20: 452-459.

Kaytes, P. S., Theriault, N. Y., Poorman, R. A., Murakami, K. & Tomich, C-S. C. 1986. High-level expression of human renin in Escherichia coli. J. Biotechnol. 4: 205-218.

Keddie, B. A., Aponte, L. E. & Volkman, L. E. 1989. The pathway of infection of Autographa californica nuclear polyhedrosis virus in insect host. Science 243: 1728-1730.

Keinänen, R. A., Wallén, M. J., Kristo, P. A., Laukkanen, M. O., Toimela, T.A., Helenius, M.A. & Kulomaa, M. S. 1994. Molecular cloning and nucleotide sequence of chicken avidin-related genes 1-5. Eur. J. Biochem. 220: 615-621.

Kidd, M. & Emery, V. C. 1993. The use of baculoviruses as expression vectors. Appl. Biochem. Biotech. 42: 137-159.

Kishimoto, F., Gomi, H., Kanaoka, M., Nakatani, T., Ito, A., Katoh, T., Agui, H., Sumida, S. & Ogino, S. 1986. Direct expression of urogastrone gene in Escherichia coli. Gene 45: 311-316.

Kitts, P. A., Ayres, M. D. & Possee, R. D. 1990. Lineriazation of baculovirus DNA enhances the recovery of recombinant virus expression vectors. Nucl. Acids Res. 18: 5667-5672.

Kitts, P. A. & Possee, R. D. 1993. A method for producing recombinant baculovirus expression vectors at high frequency. Biotechniques 14: 810-817.

Kleid, D. G., Yansura, D., Small, B., Dowbenko, D., Moore, D. M., Grubman, M. J., McKercher, P. D., Morgan, D. O., Robertson, B. H. & Bachrach, H. L. 1981. Cloned viral protein vaccine for foot-and-mouth disease: Responses in cattle and swine. Science 214: 1125-1129.

Kling, J. 1996. "Complementary" advances in baculovirus recombinant protein expression systems. Na Biotechnology 14: 648.

Klotsky, R-A. & Schwartz, I. 1987. Measurement of cat expression from growth-rate-regulated promoters employing β-lactamase activity as an indicator of plasmid copy number. Gene 55: 141-146.

Kohanski, R. A. & Lane, M. D. 1990. Monovalent avidin affinity columns. Methods. Enzymol. 184: 194-200.

Kohler, P. O., Grimley, P. M. & O'Malley, B. W. 1968. Protein synthesis: differential stimulation of cell-specific proteins in epithelial cells of chick oviduct. Science 160: 86-87.

Kool, M., Voncken, J. W., van Lier, F. L. J., Tramper, J. & Vlak, J. M. 1991. Detection and analysis of Autographa californica nuclear polyhedrosis virus mutants with defective interfering properties. Virology 183: 739-746.

Kopetzki, E., Schumacher, G. & Buckel, P. 1989. Control of formation of active soluble or inactive insoluble baker's yeast α-glucosidase PI in Escherichia coli by induction and growth conditions. Mol. Gen. Genet. 216: 149-155.

Korenman, S. G. & O'Malley, B. W. 1968. Progesterone action: regulation of avidin biosynthesis by hen oviduct in vivo and in vitro. Endocrinology 83: 11-17.

Kornitzer, D., Raboy, B., Kulka, R. G. & Fink, G. R. 1994. Regulated degradation of the transcription factor Gcn4. Embo J. 13: 6021-6030.

Korpela, J. K. 1984. Avidin, a high affinity biotin-binding protein, as a tool and subject of biological research. Med. Biol. 62: 5-26.

Korpela, J. K., Elo, H. A. & Tuohimaa, P. J. 1981a. Avidin induction by estrogen and progesterone in the immature oviduct of chicken, Japanese Quail, duck, and Gull. Gen. Comp. Endocrinol. 44: 230-232.

Korpela, J. K., Kulomaa, M. S., Elo, H. A. & Tuohimaa, P. J. 1981b. Biotin-binding proteins in eggs of oviparous vertebrates. Experientia 37: 1065-1066.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. 1982. Induction of avidin in chickens infected with the acute leukemia virus ok 10. Int. J. Cancer 30: 461-464.

Korpela, J., Kulomaa, M., Tuohimaa, P. & Vaheri, A. 1983. Avidin is induced in chicken embryo fibroblasts by viral transformation and cell damage. Embo J. 2: 1715-1719.

Kumar, S. & Miller, L. K. 1987. Effects of serial passage of Autographa californica nuclear polyhedrosis virus in cell culture. Virus Res. 7: 335-349.

Kunnas, T. A., Joensuu, T. K., Viitala, K. K., Sopanen, P., Tuohimaa, P. & Kulomaa, M. S. 1992. Induction of the avidin messenger ribonucleic acid in the chick oviduct by progesterone and other steroids. Endocrinology 130: 3421-3426.

Kunnas, T. A., Wallén, M. J. & Kulomaa, M. S. 1993. Induction of chicken avidin and related mRNAs after bacterial infection. Biochim. Biophys. Acta 1216: 441-445.

Kurland, C. G. 1991. Codon bias and gene expression. Febs Lett. 285: 165-169.

Kurland, C. & Gallant, J. 1996. Errors of heterologous protein expression. Curr. Opin. Biotech. 7: 489-493.

Kurzban, G. P., Gitlin, G., Bayer, E. A., Wilchek, M. & Horowitz, P. M. 1989. Shielding of tryptophan residues of avidin by the binding of biotin. Biochemistry 28: 8537-8542.

Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. & McCoy, J. M. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cells. Biotechnology 11: 187-193.

Lee, H-I., Broekaert, W. F. & Raikhel, N. V. 1991. Co- and post-translational processing of the hevein preproprotein of latex of the rubber tree (Hevea brasiliensis). J. Biol. Chem. 266: 15944-15948.

Lesley, S. A., Patten, P. A. & Schultz, P. G. 1993. A genetic approach to the generation of antibodies with enhanced catalytic activities. Proc. Natl. Acad. Sci. USA 90: 1160-1165.

Leusch, M. S., Lee, S. C. & Olins, P. O. 1995. A novel host-vector system for direct selection of recombinant baculoviruses (bacmids) in Escherichia coli. Gene 160: 191-194.

Licari, P. & Bailey, J. E. 1991. Factors influencing recombinant protein yields in an insect cell-baculovirus expression system: Multiplicity of infection and intracellular protein degradation. Biotech. Bioeng. 37: 238-246.

Livnah, O., Bayer, E. A., Wilchek, M. & Sussman, J. L. 1993. Three-dimensional structures of avidin and the avidin-biotin complex. Proc. Natl. Acad. Sci. USA 90: 5076-5080.

Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., deBoer, H. A. & Van Knippenberg, P. H. 1987. Influence of the codon following the AUG initiation codon on the expression of a modified lacZ gene in Escherichia coli. Embo J. 6: 2489-2492.

Lory, S. 1992. Determinants of extracellular protein secretion in gram-negative bacteria. J. Biol. Chem. 174: 3423-3428.

Luckow, V. A. & Summers, M. D. 1988. Trends in the development of baculovirus expression vectors. Biotechnology 6: 47-55.

Luckov, V. A., Lee, S. C., Barry, G. F. & Olins, P. O. 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. J. Virol. 67: 4566-4579.

Maeda, S. 1989. Gene transfer vectors of a baculovirus, Bombyx mori, nuclear polyhedrosis virus, and their use for expression of foreign genes in insect cells. In: Mitsuhashi, J. (ed.), Invertebrate cell system applications: 167-181. CRC Press, Boca Raton.

Marjomäki, V. S., Huovila, A. J., Surkka, M. A., Jokinen, I. & Salminen, A. 1990. Lysosomal trafficking in rat cardiac myocytes. J. Histochem. Cytochem 38: 1155-1164.

Marjomäki, V., Juuti, K., Airenne, K. J., Punnonen, E-L., Olkkonen, V. M., Laukkanen, A., Gruenberg, J. & Kulomaa, M. 1998. Targeting of avidin fusion proteins to early or late endosomes. Manuscript.

Marston, F. A. O. 1986. The purification of eukaryotic polypeptides synthesized in Escherichia coli. Biochem. J. 240: 1-12.

Marullo, S., Delavier-Klutchko, C., Guillet, J-G., Charbit, A., Strosberg, A. D. & Emorine, L. J. 1989. Expression of human β1 and β2 adrenergic receptors in E. coli as a new tool for ligand screening. Biotechnology 7: 923-927.

Masui, Y., Mizuno, T. & Inouye, M. 1984. Novel high-level expression cloning vehicles: 104-fold amplification of Escherichia coli minor protein. Biotechnology 2: 81-85.

Maurizi, M. R. 1992. Proteases and protein degradation in Escherichia coli. Experientia 48: 178-201.

Meyer, H-P., Käppeli, O. & Fiechter, A. 1985. Growth control in microbial cultures. Annu. Rev. Microbiol. 39: 299-319.

Miller, L. K. 1988. Baculoviruses as gene expression vectors. Ann. Rev. Microbiol. 42: 177-199.

Miller, L. K. 1993. Baculoviruses: High-level expression in insect cells. Curr. Opin. Gen. Dev. 3: 97-101.

Moks, T., Abrahmsén, L., Holmgren, E., Bilich, M., Olsson, A., Uhlén, M., Pohl, G., Sterky, C., Hultberg, H., Josephson, S., Holmgren, A., Jörnvall, H. & Nilsson, B. 1987. Expression of human insulin-like growth factor I in bacteria: Use of optimized gene fusion vectors to facilitate protein purification. Biochemistry 26: 5239-5244.

Moore, J. T., Uppal, A., Maley, F. & Maley G. F. 1993. Overcoming inclusion body formation an a high-level expression system. Prot. Express. Purif. 4: 160-163.

Morag, E., Bayer, E. A. & Wilchek, M. 1996. Reversibility of biotin-binding by selective modification of tyrosine in avidin. Biochem. J. 316: 193-199.

Morgan, T. D., Oppert, B., Czapla, T. H. & Kramer, K. J. 1993. Avidin and streptavidin is insecticidal and growth inhibiting dietary proteins. Entomol. Exp. Appl. 69: 97-108.

Moreno, F., Fowler, A. V., Hall, M., Silhavy, T. J., Zabin, I. & Schwartz, M. 1980. A signal sequence is not sufficient to lead β-galactosidase out of the cytoplasm. Nature 286: 356-359.

Mott, J. E., Grant, R. A., Ho, Y-S. & Platt, T. 1985. Maximizing gene expression from plasmid vectors containing the λ $P_L$ promoter: Strategies for overproducing transcription termination factor ρ. Proc. Natl. Acad. Sci. USA 82: 88-92.

Müller, M., Ibrahimi, I., Chang, C. N., Walter, P. & Blobel, G. 1982. A bacterial secretory protein requires signal recognition particle for translocation across mammalian endoplasmic reticulum. J. Biol. Chem. 257: 11860-11863.

Müller, D., Hughes, C. & Goebel, W. 1983. Relationship between plasmid and chromosomal hemolysin determinants of Escherichia coli. J. Bacteriol. 153: 846-851.

Mulligan, M .E., Brosius, J. & McClure, W. R. 1985. Characterization in vitro of the effect of spacer length on the activity of Escherichia coli RNA polymerase at the TAC promoter. J. Biol. Chem. 260: 3529-3538.

Nagai, K. & Thøgersen, H. C. 1984. Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in Escherichia coli. Nature 309: 810-812.

Nakamura, K. & Innoye, M. 1982. Construction of versatile expression cloning vehicles using the lipoprotein gene of Escherichia coli. EMBO J. 6: 771-775.

Nakano, H., Yamazaki, T., Ikeda, M., Masai, H., Miyatake, S-I. & Saito, T. 1994. Purification of glutathione S-transferase fusion proteins as a non-degraded form by using a protease-negative E. coli strain, AD202. Nucleic Acids Res. 22: 543-544.

Niemelä, A., Kulomaa, M. S. & Tuohimaa, P. 1986. Avidin induction by theophylline in vivo and in vitro. Endocrinology 118: 2319-2323.

Nilsson, I., Gafvelin, N. & von Heijne, G. 1993. Different sec-requirements for signal peptide cleavage and protein translocation in a model E. coli protein. Febs lett. 318: 7-10.

Nilsson, J., Larsson, M., Ståhl, S., Nygren, P-Å. & Uhlén, M. 1996. Multiple affinity domains for the detection, purification and immobilization of recombinant proteins. J. Mol. Recog. 9: 585-594.

Nilsson, J., Ståhl, S., Lundeberg, J., Uhlén, M. & Nygren, P-Å. 1997. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. Prot. Express. Purif. 11: 1-16.

Nordback, I., Wigham, T., Kulomaa, M. S. & Tuohimaa, P. 1981. Progesterone-independent avidin in chick oviduct fibroblast culture. Endocrinology 109: 596-601.

Nordback, I., Joensuu, T. & Tuohimaa, P. 1982a. Effects of glucocorticoids and disodium cromoglycate on avidin production in chick tissues. J. Endocrinol. 92: 283-291.

Nordback, I., Kulomaa, M., Joensuu, T. & Tuohimaa, P. 1982b. Characterization of progesterone-independent avidin production of chicken tissues in culture. Comp. Biochem. Physiol. 71A: 389-393.

Nossal, N. G. & Heppel, L. A. 1966. The release of enzymes by osmotic shock from Escherichia coli in exponential phase. J. Biol. Chem. 241: 3055-3062.

Nygren, P-Å., Ståhl, S. & Uhlén, M. 1994. Engineering proteins to facilitate bioprocessing. Trends Biotech. 12: 184-188.

Ogonah, O. W., Freedman, R. B., Jenkins, N., Patel, K. & Rooney, B. C. 1996. Isolation and characterization of an insect cell line able to perform complex N-linked glycosylation on recombinant proteins. Biotechnology 14: 197-202.

Oker-Blom, C., Orellana, A. & Keinänen, K. 1996. Highly efficient production of GFP and its derivatives in insect cells for visual in vitro applications. FEBS Lett. 389: 238-243.

Old, R. W. & Primrose, S. B. 1989. Expression in E. coli of cloned DNA molecules. In: Old, R. W. & Primrose, S. B. (ed.), Principles of gene manipulation: 143-167. Blackwell Scientific Publications, Oxford.

Olejnik, J., Sonar, S., Krzymañska-Olejnik, E. & Rothschild, K. J. 1995. Photocleavable biotin derivatives: A versatile approach for the isolation of biomolecules. Proc. Natl. Acad. Sci. USA 92: 7590-7594.

O'Malley, B. W. 1967. In vitro hormonal induction of a specific protein (avidin) in chick oviduct. Biochemistry 6: 2546-2551.

O'Malley, B. W., McGuire, W. L., Kohler, P. O. & Korenman, S. G. 1969. Studies on the mechanism of steroid hormone regulation of synthesis of specific proteins. Recent Prog. Horm. Res. 25: 105-153.

Ooi, B. G. & Miller, L. K. 1990. Transcription of the baculovirus polyhedrin gene reduces the levels of an antisense transcript initiated downstream. J. Virol. 64: 3126-3129.

O'Reilly, D. R. & Miller, L. K. 1989. A baculovirus blocks insect molting by producing ecdysteroid UDP-glucosyl transferase. Science 245: 1110-1112.

O'Reilly, D. R., Miller, L. K. & Luckov, V. A. 1994. Baculovirus expression vectors. A laboratory manual. Oxford University Press, New York.

Orr, G. A. 1981. The use of the 2-iminobiotin-avidin interaction for the selective retrieval of labeled plasma membrane components. J. Biol. Chem. 256: 761-766.

O'Shannessy, D. J., O'Donnell, K. C., Martin, J. & Brigham-Burke, M. 1995. Detection and quantitation of hexa-histidine-tagged recombinant proteins on western blots and by a surface plasmon resonance biosensor technique. Anal. Biochem. 229: 119-124.

Pack, P. & Plückthun, A. 1992. Miniantibodies: Use of amphipathic helices to produce functional, flexibly linked dimeric Fv fragments with high avidity in Escherichia coli. Biochemistry 31: 1579-1584.

Patel, G., Nasmyth, K. & Jones, N. 1991. A new method for the isolation of recombinant baculovirus. Nucleic Acids Res. 20: 97-104.

Patterson, R. M., Selkirk, J. K. & Merrick B. A. 1995. Baculovirus and insect cell gene expression: Review of baculovirus biotechnology. Environ. Health Persp. 103: 756-759.

Pendergast, A. M., Clark, R., Kawasaki, E. S., McCormick, F. P. & Witte, O. N. 1989. Baculovirus expression of functional P210 BCR-ABL oncogene product. Oncogene 4: 759-766.

Pennock, G. D., Shoemaker, C. & Miller, L. K. 1984. Strong and regulated expression of Escherichia coli β-galactosidase in insect cells with a baculovirus vector. Mol. Cell. Biol. 4: 399-406.

Pharmacia Biotech 1996. GST gene fusion system. Second edition, revision 2. Pharmacia Biotech, Inc., Molecular Biology Reagents Division, USA.

Pinn, F., Pähler, A., Saenger, W., Petsko, G. A. & Green, M. N. 1982. Crystallization and preliminary X-ray investigation of avidin. Eur. J. Biochem. 123: 545-456.

Podhajska, A. J., Hasan, N. & Szybalski, W. 1985. Control of cloned gene expression by promoter inversion in vivo: Construction of the heat-pulse-activated att-nutL-p-att-N module. Gene 40: 163-168.

Poole, E. S., Brown, C. E. & Tate, W. P. 1995. The identity of the base following the stop codon determines the efficiency of in vivo translation termination in Esherichia coli. Embo J. 14: 151-158.

Possee, R. D. & Howard, S. C. 1987. Analysis of the polyhedrin gene promoter of the Autographa californica nuclear polyhedrosis virus. Nucleiv Acids Res. 15: 10233-10248.

Pribnow, D. 1975. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc. Natl. Acad. Sci. USA 72: 784-788.

Prouty, W. F. & Goldberg, A. L. 1972. Fate of abnormal proteins in E. coli accumulation in intracellular granules before catabolism. Nat. New Biol. 240: 147-150.

Prouty, W. F., Karnovsky, M. J. & Goldberg, A. L. 1975. Degradation of abnormal proteins in Escherichia coli. J. Biol. Chem. 250: 1112-1122.

Punnonen, E-L., Autio, S., Marjomäki, V. S. & Reunanen, H. 1992. Autophagy, cathepsin L transport, and acidification in cultured rat fibroblasts. J. Histochem. Cytochem. 40: 1579-1587.

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. J. Mol. Biol. 231: 698-710.

Pugliese, L., Malcovati, M., Coda, A. & Bolognesi, M. 1994. Crystal structure of apo-avidin from hen egg-white. J. Mol. Biol. 235: 42-46.

Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. 57: 50-108.

Rapoport, T. A., Jungnickel, B. & Kutay, U. 1996. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. Annu. Rev. Biochem. 65: 271-303.

Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Millican, A., Eaton, M. & Humphreys, G. 1984. Codon usage can affect efficiency of translation of genes in Escherichia coli. Nucleic Acids Res. 17: 6663-6671.

Ranjan, A. & Hasnain, S. E. 1994. Influence of codon usage and translation initiation codon context in the AcNPV-based expression system: Computer analysis using homologous and heterologous genes. Virus Genes 9: 149-152.

Rohrmann, G. F. 1986. Polyhedrin structure. J. Gen. Virol. 67: 1499-1513.

Rosa, D., Campagnoli, S., Moretto, C., Guenzi, E., Cousens, L., Chin, M., Dong, C., Weiner, A. J., Lau, J. Y. N., Choo. Q-L, Chien, D., Pileri, P., Houghton, M. & Abrignani, S. 1996. A quantitative test to estimate neutralizing antibodies to the hepatitis C virus: Cytofluorimetric assessment of envelope glycoprotein 2 binding to target cells. Proc. Natl. Acad. Sci. USA 93: 1759-1763.

Ross, S. E., Carson, S. D. & Fink, L. M. 1986. Effects of detergents on avidin-biotin interaction. Biotechniques 4: 350-354.

Rudolph, R. 1996. Successful protein folding on an industrial scale. In: Cleland, J. L. & Craik, C. S. (ed.), Protein Engineering: 283-298. Wiley-Liss, New York.

Salmond, G. P. C. & Reeves, P. J. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. Trends Biochem. Sci. 18: 7-12.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989. Molecular Cloning. A laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Samuelsson, E., Moks, T., Nilsson, B. & Uhlén, M. 1994. Enhanced in vitro refolding of insulin-like growth factor I using a solubilizing fusion partner. Biochemistry 33: 4207-4211.

Samuelsson, E., Jonasson, P., Viklund, F., Nilsson, B. & Uhlén, M. 1996. Affinity-assisted in vivo folding of a secreted human peptide hormone in Escherichia coli. Biotechnology 14: 751-755.

Sano, T. & Cantor, C. R. 1990. Expression of a cloned streptavidin gene in Escherichia coli. Proc. Natl. Acad. Sci. USA 87: 142-146.

Sano, T. & Cantor, C. R. 1991a. A streptavidin-protein A chimera that allows one-step production of a variety of specific antibody conjugates. Biotechnology 9: 1378-1381.

Sano, T. & Cantor, C. R. 1991b. Expression vectors for streptavidin-containing chimeric proteins. Biochem. Biophys. Res. Com. 176: 571-577.

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

Sano, T., Glazer, A. N. & Cantor, C. R. 1992. A streptavidin-metallothionein chimera that allows specific labeling of biological materials with many different heavy metal ions. Proc. Natl. Acad. Sci. USA 89: 1534-1538.

Sano, T., Smith, C. L. & Cantor, C. R. 1997. Expression and purification of recombinant streptavidin-containing chimeric proteins. Methods Mol. Biol. 63: 119-128.

Schein, C. H. 1989. Production of soluble recombinant proteins in bacteria. Biotechnology 7: 1141-1149.

Schein, C. H. 1993. Solubility and secretability. Curr. Opin. Biotech. 4: 456-461.

Schein, C. H. & Noteborn, M. H. M. 1988. Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperature. Biotechnology 6: 291-294.

Schenk, P. M., Baumann, S., Mattes, R. & Steinbiß, H-H. 1995: Improved high-level expression system for eukaryotic genes in Escherichia coli using T7 RNA polymerase and rare $^{arg}$tRNAs. Biotechniques 19: 196-200.

Schulz, M-F., Buell, G., Schmid, E., Movva, R. & Selzer, G. 1987. Increased expression in Escherichia coli of a synthetic gene encoding human somatomedin C after gene duplication and fusion. J. Bacteriol. 169: 5385-5392.

Seo, J-H. & Bailey, J. E. 1985. Effects of recombinant plasmid content on growth properties and cloned gene product formation in Esherichia coli. Biotechnol. Bioeng. 27: 1668-1674.

Shatzman, A. R. & Rosenberg, M. 1987. Expression, identification, and characterization of recombinant gene products in Escherichia coli. Methods Enzymol. 152: 661-673.

Shen, S-H. 1984. Multiple joined genes prevent product degradation in Escherichia coli. Proc. Natl. Acad. Sci. USA 81: 4627-4631.

Shibui, T., Matsui, R., Uchida-Kamizono, M., Okazaki, H., Kondo, J., Nagahari, K., Nakanishi, S. & Teranishi, Y. 1989. Periplasmic production of human

pancreatic prokallikrein in Escherichia coli. Appl. Microbiol. Biotechnol. 31: 253-258.

Shimatake, H. & Rosenberg, M. 1981. Purified λ regulatory protein cII positively activates promoters for lysogenic development. Nature 292: 128-133.

Shimkus, M., Levy, J. & Herman, T. 1985. A chemically cleavable biotinylated nucleotide: Usefulness in the recovery of protein-DNA complexes from avidin affinity columns. Proc. Natl. Acad. Sci. USA 82: 2593-2597.

Smith, D. B. & Johnson, K. S. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 67: 31-40.

Smith, G. E., Summers, M. D. & Fraser, M. J. 1983a. Production of human beta interferon in insect cells infected with a baculovirus expression vector. Mol. Cell. Biol. 3: 2156-2165.

Smith, G. E., Fraser, M. J. & Summers, M. D. 1983b. Molecular engineering of the Autographa californica nuclear polyhedrosis virus genome: Deletion mutations within the polyhedrin gene. J. Virol. 46: 584-593.

Smith, H., deJong, A., Bron, S. & Venema, G. 1988a. Characterization of signal-sequence-coding regions selected from the Bacillus subtilis chromosome. Gene 70: 351-361.

Smith, M. N., Furman, T. C., Ingolia, T. D. & Pidgeon, C. 1988b. Chelating peptide-immobilized metal ion affinity chromatography. J. Biol. Chem. 263: 7211-7215.

Speed, M. A., Wang, D. I. C. & King, J. 1996. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body formation. Biotechnology 14: 1283-1287.

Sprengart, M. L., Fuchs, E. & Porter, A. G. 1996. The downstream box: An efficient and independent translation initiation signal in Escherichia coli. Embo J. 15: 665-674.

Stader, J. A. & Silhavy, T. J. 1990. Engineering Escherichia coli to secrete heterologous gene products. Methods Enzymol. 185: 166-187.

Steinitz, F. (1898). Ueber das verhalten phosphorhaltiger eiweisskörper im stoffwechsel. Arch. Ges. Physiol. 72: 75-104.

Stempfer, G., Höll-Neugebauer, B., Kopetzki, E. & Rudolph, R. 1996. A fusion protein designed for noncovalent immobilization: Stability, enzymatic activity, and use in an enzyme reactor. Nature Biotechnology 14: 481-484.

Stevens, L. 1991. Egg-white proteins. Comp. Biochem. Physiol. 100B: 1-9.

Stormo, G. D., Schneider, T. D. & Gold L. M. 1982. Characterization of translational initiation sites in E. coli. Nucleic Acids Res. 10: 2971-2996.

Stormo G. D. 1986. Translation initiation. In: Reznikoff, W. & Gold, L. (ed.), Maximizing gene expression: 195-. Butterworths, Boston.

Studier, F. W., Rosenberg, A. H, Dunn, J. J. & Dubendorff, J. W. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. 185: 60-89.

Summers, M. D. & Anderson, D. L. 1972. Granulosis virus deoxyribonucleic acid: a closed, double-stranded molecule. J. Viroil. 33: 710-713.

Takahara, M., Sagai, H., Inouye, S. & Inouye, M. 1988. Secretion of human superoxide dismutase in Escherichia coli. Biotechnology 6: 195-198.

Takagi, H., Morinaga, Y., Tsuchiya, M., Ikemura, H. & Inouye, M. 1988. Control of folding of proteins secreted by a high expression secretion vector, pIN-III-ompA: 16-fold increase in production of active subtilisin E in Ecsherichia coli. Biotechnology 6: 948-950.

Talmadge, K. & Gilbert, W. 1982. Cellular location affects protein stability in Escherichia coli. Proc. Natl. Acad. Sci. USA 79: 1830-1833.

Talmadge, K., Stahl, S. & Gilbert, W. 1980. Eukaryotic signal sequence transports insulin antigen in Escherichia coli. Proc. Natl. Acad. Sci. USA 77: 3369-3373.

Tate, W. P. & Brown, C. M. 1992. Translational termination: "Stop" for protein synthesis or "pause" for regulation of gene expression. Biochemistry 31: 2443-2450.

Thiem, S. M. & Miller, L. K. 1989. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus Autographa californica nuclear polyhedrosis virus. J. Virol. 63: 2008-2018.

Thompson, L. D. & Weber, P. C. 1993. Construction and expression of a synthetic streptavidin-encoding gene in Escherichia coli. Gene 136: 243-246.

Towbin, H., Staehelin, T. & Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA 76: 4350-4354.

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

Tweeten, K . A., Bulla, L .A. & Consigli, R. A. 1980. Characterization of an extremely basic protein derived from granulosis virus nucleocapsids. J. Virology 33: 866-876.

Uhlén, M. & Moks, T. 1990. Gene fusions for purpose of expression: An introduction. Methods Enzymol. 185: 129-143.

Uhlén, M., Nilsson, B., Guss, B., Lindberg, M., Gatenbeck, S. & Philipson, L. 1983. Gene fusion vectors based on the gene for staphylococcal protein A. Gene 23: 369-378.

Uhlén, M., Forsberg, G., Moks, T., Hartmanis, M. & Nilsson, B. 1992. Fusion proteins in biotechnology. Curr. Opin. Biotech. 3: 363-369.

Varshavsky, A. 1992. The N-end rule. Cell 69: 725-735.

Varshavsky, A. 1996. The N-end rule: Functions, mysteries, uses. Proc. Natl. Acad. Sci. USA 93: 12142-12149.

Vaughn, J. L., Goodwin, R. H., Tompkins, G. J. & McCawley, P. 1977. The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae). In Vitro 13: 213-217.

Vetter, S., Bayer, E. A. & Wilchek, M. 1994. Avidin can be forced to adopt catalytic activity. J. Am. Chem. Soc. 116: 9369-9370.

Vialard, J. E., Arif, B. M. & Richardson, C. D. 1995: Introduction to the molecular biology of baculoviruses. In: Richardson, C. D. (ed.), Baculovirus expression protocols: 1-24. Humana Press, Totowa.

Wagner, R., Liedtke, S., Kretzschmar, E., Geyer, H., Geyer, R. & Klenk H-D. 1996. Elongation of the N-glycans of fowl plaque virus hemagglutinin expressed in Spodoptera frugiperda (Sf9) cells by coexpression of human β1,2-N-acetylglucosaminyltransferase I. Glycobiology 6: 165-175.

Wallén, M. J., Laukkanen, M. O. & Kulomaa, M. S. 1995. Cloning and sequencing of the chicken egg-white avidin-encoding gene and its relationship with the avidin-related genes Avr1 - Avr5. Gene 161: 205-209.

Wandersman, C. 1992. Secretion across the bacterial outer membrane. Trends Gen. 8: 317-322.

Weber, P. C., Ohlendorf, D. H., Wendoloski, J. J. & Salemme, F. R. 1989. Structural origins of high-affinity biotin-binding to streptavidin. Science 243: 85-88.

Wei, R-D. & Wright, L. D. 1964. Heat stability of avidin and avidin-biotin complex and influence of ionic strength on affinity of avidin for biotin. Proc. Soc. Exp. Biol. Med. 117: 341-344.

Wetzel, R. 1994. Mutations and off-pathway aggregation of proteins. Trends Biotech. 12: 193-198.

Wetzel, R., Perry, L. J. & Veilleux, C. 1991. Mutations in human interferon gamma affecting inclusion body formation identified by a general immunochemical screen. Biotechnology 9: 731-737.

White, H. B. III, Orth, W. H. III, Schreiber, R. W. Jr. & Whitehead, C. C. 1992. Availability of avidin-bound biotin to the chicken embryo. Arch. Biochem. Biophys. 298: 80-83.

Wickham, T. J., Davis, T., Granados, R. R., Shuler, M. L. & Wood, H. A. 1992. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. Biotechnol. Prog. 8: 391-396.

Wickham, J. T., Nemerow, G. R., Wood, H. A. & Shuler, M. L. 1995. Comparison of different cell lines for the production of recombinant baculovirus proteins. In: Richardson, C. D. (ed.), Baculovirus expression protocols: 385-395. Humana Press, Totowa.

Wickner, W., Driessen, A. J. M. & Hartl, F-U. 1991. The enzymology of protein translocation across the Escherichia coli plasma membrane. Annu. Rev. Biochem. 60: 101-124.

Wilchek, M. & Bayer, E. A. 1989. Avidin-biotin technology ten years on: has it lived up to its expectations?. Trends Biochem Sci. 14: 408-412.

Wilchek, M. & Bayer, E. A. (ed.) 1990. Avidin-biotin technology. Methods in Entzymol. 184.

Wilchek, M., Gitlin, G., Bayer, E. A. & Viswanatha, T. 1988. Enzymatic properties of avidin. Biochemistry 27: 3099.

Wild, J., Rossmeissl, P., Walter, W. A. & Gross, C. A. 1996. Involvement of the DnaK-DnaJ-GrpE chaperone team in protein secretion in Escherichia coli. J. Bacteriol. 178: 3608-3613.

Wilson, M. E., Mainprize, T. H., Friesen, P. D. & Miller, L. K. 1987. Location, transcription, and sequence of a baculovirus gene encoding a small arginine-rich polypeptide. J. Virol. 61: 661-666.

Wikström, P. M., Lind, L. K., Berg, D. E. & Björk, G. R. 1992. Importance of mRNA folding and start codon accessibility in the expression of genes in a ribosomal protein operon of Escherichia coli. J. Mol. Biol. 224: 949-966.

Williams, D. C., Van Frank, R. M., Muth, W. L. & Burnett, J. P. 1982. Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins. Science 215: 687-689.

Yanofsky, C., Platt, T., Crawford, I. P., Nichols, B. P., Christie, G. E., Horowitz, H., VanCleemput, M. & Wu, A. M. 1981. The complete nucleotide sequence of the tryptophan operon of Escherichia coli. Nucleic Acids Res. 9: 6647-6668.

Yansura, D. G. 1990. Expression as trpE fusion. Methods Enzymol. 185: 161-166.

Yansura, D. G. & Henner, D. J. 1990. Use of Esherichia coli trp promoter for direct expression of proteins. Methods Enzymol. 185: 54-60.

Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T. & Ishii, S. 1995. Increase of solubility of foreign proteins in Escherichia coli by coproduction of bacterial thioredoxin. J. Biol. Chem. 270: 25328-25331.

Yost, H. J., Petersen, R. B. & Lindquist, S. 1990. RNA metabolism: strategies for regulation in the heat shock response. Trends Gen. 6: 223-227.

Zentgraf, H., Frey, M., Schwinn, S., Tessmer, C., Willemann, B., Samstag, Y. & Velhagen, I. 1995. Detection of histidine-tagged fusion proteins by using a high-specific mouse monoclonal anti-histidine tag antibody. Nucleic Acids Res. 23: 3347-3348.

Zhang, S., Zubay, G. & Goldman, E. 1991. Low-usage codon in Escherichia coli, yeast, fruit fly and primates. Gene 105: 61-72.

Zubiaga, A. M., Belasco, J. G. & Greenberg, M. E. 1995. The nonamer UUAUUUAUU is the key AU-rich sequence motif that mediates mRNA degradation. Mol. Cell. Biol. 15: 2219-2230.

# ORIGINAL PAPERS

## I

**Production of recombinant avidin in *Escherichia coli***

by

Airenne, K. J., Sarkkinen, P., Punnonen, E-L. & Kulomaa, M. S.

Gene 144: 75-80, 1994

https://doi.org/10.1016/0378-1119(94)90206-2

*Gene*, 144 (1994) 75–80
 0378-1119/94/$07.00



GENE 07910

# Production of recombinant avidin in *Escherichia coli*

(Recombinant DNA; gene expression; codon choice; PCR amplification; affinity chromatography; biotin-binding protein)

Kari J. Airenne, Paula Sarkkinen, Eeva-Liisa Punnonen and Markku S. Kulomaa

*Department of Biology, University of Jyväskylä, FIN-40100, Jyväskylä, Finland*

Received by J.K.C. Knowles: 31 August 1993; Revised/Accepted: 3 December/23 December 1993; Received at publishers: 21 February 1994

SUMMARY

A recombinant avidin (re-Avd), containing amino acids (aa) 1–123 of the native chicken egg-white Avd, was produced in *Escherichia coli*. When cells were grown at 37°C production was over 1 μg/ml, due to altering the codon preference of the first ten codons. The re-Avd was recovered as a soluble protein from cells grown at 25 or 30°C, whereas at 37°C it was mostly insoluble in inclusion bodies. Our results indicated that, despite the potentially harmful biotin-binding activity of Avd, it is possible to produce biologically active Avd in *E. coli* which then can easily be purified by affinity chromatography on a biotin column in a single step.

## INTRODUCTION

Avidin (Avd), a basic glycoprotein (66–69 kDa) in the chicken egg-white, is composed of four identical subunits (15.5–15.8 kDa) each of which can bind one biotin molecule (Green, 1975). The sequence of the 128-aa peptide has been determined (Delange and Huang, 1971). Each subunit has an oligosaccharide side chain which is not essential for biotin binding (Hiller et al., 1987). Avd is soluble in water and salt solutions and stable over a wide range of pH values and temperatures, particularly as a complex with biotin. Its extraordinarily strong and specific affinity for biotin ($K_d = 10^{-15}$ M) has made Avd a frequently used protein in a number of applications, including isolation of biotinylated macromolecules by affinity chromatography and labelling of biologically active materials for immunochemical localisation studies (for a review, see Wilchek and Bayer, 1990). Despite a profound knowledge of the properties of Avd little is known about its biological function. Since the growth of a number of micro organisms is inhibited by Avd, and its occurrence after tissue damage has been clearly de-

*Correspondence to:* Dr. M.S. Kulomaa, Department of Biology, University of Jyväskylä, Vapaudenkatu 4, FIN-40100 Jyväskylä, Finland. Tel. (358-41) 602-272; Fax (358-41) 602-221; e-mail: kulomaa@csc.fi

Abbreviations: aa, amino acid(s); Ap, ampicillin; APA, 0.1 M $NaHCO_3$/1 mM $MgCl_2$ pH 9.8; Avd, avidin; *avd*, gene encoding avd; BCIP, 5-bromo-4-chloro-3-indolyl phosphate in DMF (15 mg/ml); bp, base pair(s); cDNA, DNA complementary to mRNA; CLAP, chymostatin/leupeptin/antipain/pepstatin (10 mg/ml) in DMSO; DMF, dimethylformamide; DMSO, dimethylsulfoxide; dNTP, deoxyribonucleoside triphosphate; DTT, dithiothreitol; ET, 2 mM EDTA/30 mM Tris pH 8; ETCD, ET plus CLAP 1 μl to 1 ml/5 mM DTT; ETT, ET plus 0.5% Triton X-100; GCG package, sequence analysis software package (Genetics Computer Group, Madison, WI, USA); IPTG, isopropyl-β-D-thiogalactopyranoside; $K_d$, dissociation constant; *lacI*$^q$, gene encoding *lac* repressor; LB, Luria-Bertani (medium); NBT, nitro blue tetrazolium (30 mg/ml) in 70% DMF; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; o/n, overnight; PA, polyacrylamide; PAGE, PA-gel electrophoresis; PBS, 1.86 mM $NaH_2PO_4$/12.64 mM $Na_2HPO_4$/150 mM NaCl pH 7.5; PCR, polymerase chain reaction; Pfu, *Pyrococcus furiosus*; re-, recombinant; RT, room temperature; SDS, sodium dodecyl sulfate; SDS sample buffer, 0.125 M Tris pH 6.8/4% SDS/20% glycerol/0.004% bromophenol blue/10% 2-mercaptoethanol; TBE, 90 mM Tris·borate/2 mM EDTA pH 8; TBS, 0.5 M NaCl/20 mM Tris pH 7.5; TBSM, TBS plus 5% non-fat milk powder; TBST, TBS plus 0.2% Tween 20; TBSTM, TBST plus 5% non-fat milk powder; TE, 1 mM EDTA/10 mM Tris pH 8; *Ptrc*, hybrid *trp-lac* promoter; [ ], denotes plasmid-carrier state.

*SSDI* 0378-1119(94)00121-8

monstrated, it most likely has a role as a host defence factor (Korpela, 1984). The full-length cDNA encoding chicken egg-white Avd was cloned and characterized as reported by Gope et al. (1987).

Our purpose is to construct expression vectors for the production in bacterial and eukaryotic cells of fusion proteins containing Avd tail. Streptavidin, a close bacterial relative of Avd, has been used successfully for this purpose in *E. coli* (Sano and Cantor, 1990; 1991). As the first step toward attaining our objective, we describe here the construction of an expression vector to be used in bacterial cells. We also demonstrate that active re-Avd can be produced in *E. coli* and purified easily by affinity chromatography in a single step. A short description of an attempt to produce Avd in *E. coli* has previously been published (Chandra and Gray, 1990), but according to our best knowledge this is the first extensive report on this topic.

## EXPERIMENTAL AND DISCUSSION

### (a) Design of primers for PCR

Primers complementary to the 5′ and 3′ end of the coding sequence for mature Avd were synthesised for a PCR amplification (Fig. 1B). The AK18 primer contained an additional sequence with a *Nco*I site, five extra nt and a start codon (ATG) in front of the coding sequence to ensure subcloning of the amplified product and a proper translational reading frame. The AK20 primer was used to incorporate a restriction site for *Pst*I and *Hin*dIII, as well as a stop codon (TGA) at the end of the amplified sequence. The region was designed for the addition of a desired sequence immediately following the Avd coding sequence in order to produce a protein fusion at the C-terminal end of Avd. Avd encoded by the PCR product thus contained an extra Met residue at the N-terminal and an extra Gln residue at the C-terminal end of the molecule. The last five aa of native Avd were also excluded from the product. According to molecular modelling of the Avd molecule these modifications should not dramatically effect the tertiary structure of Avd (Livnah et al., 1993; Pugliese et al., 1993). The amplified region corresponded to the nt 116–485 of the Avd-encoding cDNA.

The attempts to produce biologically active Avd in *E. coli* cells using the DNA fragment amplified with the AK18 and AK20 primers resulted in a low concentration of re-Avd. It could be detected only by an immunoblot analysis (Western blotting) using an Avd antibody. Therefore, a new 5′ primer, AK25, was designed to mutate the first ten codons of Avd to those widely used by highly

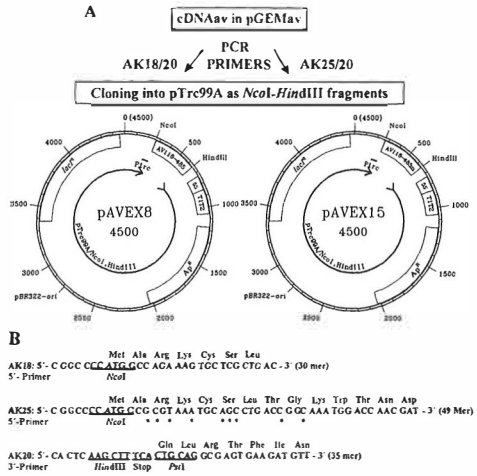



Fig. 1. Construction of recombinant plasmids pAVEX8 and pAVEX15 (**A**). Sequences of oligo primers employed to introduce the required sequence modifications into *avd* cDNA during PCR amplification (**B**). The oligos were synthesised by an Applied Biosystems 381A DNA synthesizer and further treated according to Barnes (1987). The nt, which were changed for adapting codons for *E. coli* usage, are indicated by asterisks. **Methods:** PCR amplification of mature part of Avd from Avd-encoding cDNA in pGEMav. The PCR reactions were carried out in a total volume of 100 µl containing 4.2 ng plasmid DNA/1 µmol of each primer/200 µM each dNTP (Promega, Madison, WI, USA)/Pfu DNA polymerase buffer No. 1 (Stratagene, La Jolla, CA, USA)/2.5 units Pfu DNA polymerase. Prior to the addition of Pfu DNA polymerase, the samples were heated at 94°C for 5 min to denature the plasmid DNA, cooled at 62°C for 2 min and transferred on ice. After the addition of polymerase, the samples were overlaid with 100 µl of paraffin oil, heated at 72°C for 3 min and subjected to 25 cycles of amplification with denaturation at 94°C for 1 min, annealing at 62°C for 2 min and extension at 72°C for 3 min using a DNA thermal cycler (MJ Research). After the final cycle, incubation at 72°C was continued for 10 min to extend incomplete products. The aqueous phase was extracted twice with chloroform, precipitated with ethanol and resuspended in TE buffer (pH 8.0). After digestion with *Nco*I + *Hin*dIII (Promega) the PCR products were applied to a 1.5% preparative agarose gel. The DNA was electroeluted from the gel slices to 0.5 × TBE buffer. The NaCl concentration of the buffer was set to 0.2 M and the fragments were purified using elutip-d columns as described by the manufacturer (Schleicher&Schuell). The purified fragments amplified by the primer pairs AK18/AK20 and AK25/AK20, were cloned into the *Nco*I + *Hin*dIII-digested pTrc99A vector (Pharmacia) to construct the recombinant plasmids pAVEX8 and pAVEX15, respectively. The re-plasmids were introduced into competent *E. coli* cells of strains JM109 and RR1 for preparing mg amounts of the plasmids.

expressed genes in *E. coli* (Zhang et al., 1991). As with the AK18 primer, a *Nco*I site and five additional nt were also incorporated in front of the coding sequence.

### (b) Construction of pAVEX8 and pAVEX15 expression vectors

The PCR products (396 bp each) from the amplification reactions using the AK18/AK20 and AK25/AK20

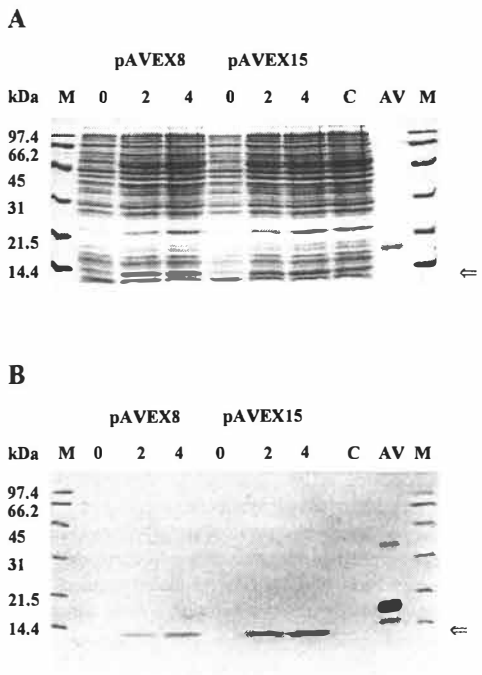



Fig. 2. The effect of codon change to production level of re-Avd. (A) Coomassie-stained 0.1% SDS-15% PA-gel of *E. coli* JM109 cells synthesizing re-Avd from pAVEX8 and pAVEX15. Lanes: (1, 10) **M**, 10 µl of low-range molecular weight standard (Bio-Rad); (2–4) samples after 0, 2 and 4 h of induction using pAVEX8; (5–7) samples after 0, 2 and 4 h of induction using pAVEX15; (8) **C**, sample after 4 h of induction using pTrc99A (negative control); (9) AV, 1 µg of Avd; The arrow points to the re-Avd bands. **Methods:** Production of re-Avd crude cell lysates. *E. coli* JM109 cells transformed with pAVEX8 or pAVEX15 plasmids were streaked out on LB plates containing 50 µg Ap/ml and the plates were incubated overnight at 37°C. Next day, single colonies were inoculated in 2 ml of LB medium supplemented with 200 µg Ap/ml and the cultures were grown overnight at 37°C with vigorous shaking. The overnight cultures were centrifuged at 1500 × *g* for 10 min and resuspended in 2 ml of fresh LB medium. Fresh cultures were diluted 1:50 in 15 ml of LB medium containing 200 µg Ap/ml and grown with vigorous shaking at 37°C until an $A_{595\ nm}$ of approx. 0.2 was reached. IPTG was added to a final concentration of 0.1 mM and the cultures were grown for an additional 4 h with an 1-ml aliquot taken every hour. Immediately prior to the addition of the inducer, a 1-ml aliquot, which served as a pre-induction sample, was also taken. Cells from the 1-ml aliquot were collected by centrifugation at 15000 × *g* for 2 min (RT), resuspended in 400 µl of SDS sample buffer diluted 1:2 in $H_2O$, heated at 96°C for 5 min and stored at −20°C overnight to prepare crude lysate samples for PAGE. The samples were centrifuged at 15000 × *g* for 2 min before loading 40 µl to the gel wells. (B) Western blot of gel shown in panel A. Immunoblotting was performed by electroblotting proteins from 0.1% SDS-15% PAGE onto nitrocellulose membrane (Trans-Blot®, Bio-Rad). The blot was blocked by incubating in TBSM overnight at 4°C. Next day the blot was washed in TBST for 5 min and a rabbit Avd antibody (Dako A/S) diluted 1:6000 in TBSTM was added onto the blot. After 1 to 3 h of incubation the blot was washed 4 times for 5 min in TBST buffer and the Goat Anti-Rabbit IgG (H + L) Alkaline Phosphatase conjugate (Bio-Rad) diluted 1:2000 in TBSTM was added onto the blot. After 1 to 3 h of incubation the blot was washed four times in TBST for 5 min and once for 5 min in APA and the complexes were detected by adding 4 ml of APA, including 40 µl NBT and BCIP. The color was allowed to develop for 30 s to 2.5 min before being stopped by washing with several changes of water.

primer pairs were purified and subcloned as *Nco*I-*Hin*dIII fragments into the pTrc99A vector to obtain the recombinant plasmids pAVEX8 and pAVEX15, respectively (Fig. 1A). The pTrc99A vector (Pharmacia LKB) has a strong IPTG-inducible *trc* promoter and *lacI*$^q$ gene allowing the tight regulation of the gene expression and it has been used in wide range of *E. coli* hosts. The nt sequences of the constructs were confirmed by dideoxynucleotide sequencing of the plasmids with an automated DNA sequencer (ALF, Pharmacia).

**(c) Expressing re-*avd* from pAVEX8**

To synthesize re-Avd the *E. coli* JM109 cells carrying a target plasmid were grown at 37°C in LB medium supplemented with Ap. A polypeptide of the expected size (14 kDa) was barely detectable in a Coomassie blue-stained SDS-PA-gel from a sample corresponding to cells from 100 µl of the original culture (Fig. 2A). A distinct band of the 14-kDa polypeptide was however clearly revealed after immunolabeling the blotted gel (Fig. 2B). The amount of re-Avd was not quantified, but a rough estimation from the Western blotting was 0.3–0.5 µg of re-Avd/ml of culture.

The low level of re-Avd synthesis from pAVEX8 could be due to instability of the expression plasmid and/or toxicity of the translation product for the host cells. The results of the expression studies, however, indicated that the cells harbouring the expression plasmid grew well at 37°C under selection pressure. The amount of plasmid DNA seemed also to increase along the growth of the culture according to the plasmid minipreps. A better explanation may be the known fact that small polypeptides are often degraded by host proteases (Marston, 1986). Indeed, some preliminary results from the production of Avd as a fusion protein have supported this possibility.

**(d) Expressing re-*avd* from pAVEX15 – the effect of codon choice**

*Avd* is a eukaryotic gene which could explain the low production of re-Avd from the pAVEX8 plasmid, since codon choice has been assumed to correlate with the expression rate of a gene (Gouy and Gautier, 1982; Grosjean and Fiers, 1982). In particular, the codons located in the N-terminal coding region have been demon-

strated to play an important role in determining the expression rate in *E. coli* (Chen and Inouye, 1990).

Another plasmid, pAVEX15, was therefore constructed where six of the first ten codons for mature Avd were adapted to the so-called optimal *E. coli* codon usage, i.e., they were changed to codons which are the most common in highly expressed genes in *E. coli* (Zhang et al., 1991). The pAVEX15 plasmid was introduced into *E. coli* JM109 cells and the level of expression was analysed. The change in the codons increased the yield of re-Avd to over 1 μg/ml of the original culture (Fig. 2A,B) which is approx. 2–4-times higher than the yield obtained using pAVEX8 as an expression plasmid .

The secondary structure of mRNA has also been assumed to influence translational initiation frequency in *E. coli* (Hall et al., 1982; Gross et al., 1990; Date et al., 1992; Wikström et al., 1992). Prediction of the possible secondary structure was thus studied within the first 200 nt of both mRNAs for the re-Avd. The prediction was carried out with the aid of the fold program in the GCG package using the default values (Zuker and Stiegler, 1981). Interestingly, the changes in the codons in pAVEX15 seemed to destabilize the predicted loop structure in the mRNA including ribosome-binding site (AGGA) and start codon (AUG) as compared to the corresponding region in pAVEX8 mRNA. Thus, by exposing the essential control elements the effect of the codon change on the secondary structures of the mRNA might partially explain the better expression from pAVEX15. The effect of codon choice per se, stability of mRNA or other, still unknown factors for improved expression cannot, however, be excluded.

**(e) The solubility of the re-Avd**

The production of biologically active recombinant proteins is often hampered by the formation of insoluble inclusion bodies (Marston, 1986). To study the form of the produced re-Avd, *E. coli* JM109[pAVEX15] were grown in LB medium at 37°C. The results revealed that re-Avd was accumulated mainly in insoluble form, i.e., as inclusion bodies, when the cells were grown at 37°C (data not shown).

The positive effect of a reduced growth temperature on the production of soluble recombinant proteins has been indicated by several authors(Schein, 1989; Browner et al., 1991). The effect of the growth temperature on re-Avd production was also studied by electron microscopy. Following growth at 37°C, inclusion bodies were visible

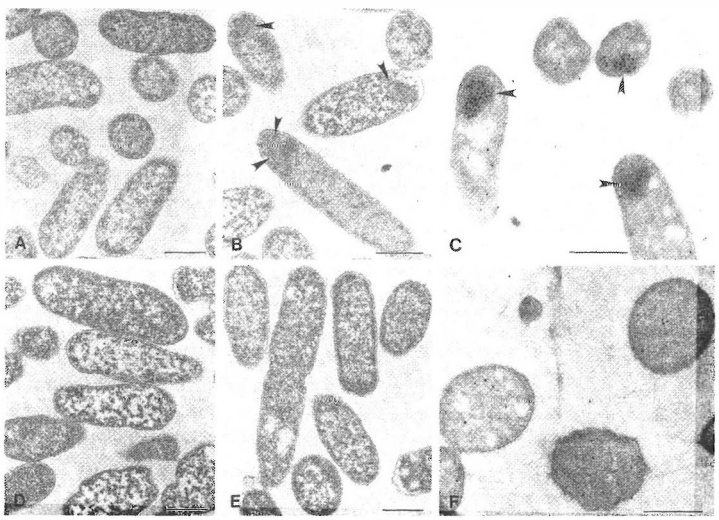



Fig. 3. Electron micrographs of *E. coli* JM109 cells after production of re-Avd at 37°C (**A, B, C**) and 24°C (**D, E, F**). Both conventional plastic sections (**A, B, D, E**) and cryosections, labelled with rabbit anti-Avd and protein A-gold (**C, F**), are shown. **A**, Plastic section of *E. coli* JM109[pTrc99A] grown at 37°C; **B**, Plastic section of *E. coli* JM109[pAVEX15] grown at 37°C; **C**, Cryosection of *E. coli* JM109[pAVEX15] grown at 37°C, labelled with anti-Avd and protein A-gold; **D**, Plastic section of *E. coli* JM109[pTrc99A] grown at 24°C; **E**, Plastic section of *E. coli* JM109[pAVEX15] grown at 24°C; **F**, Cryosection of *E. coli* JM109[pAVEX15] grown at 24°C, labelled with anti-Avd and protein A-gold. Arrowheads indicate inclusion bodies in cells after *avd* expression at 37°C. Bars, 500 nm. **Methods:** The cells were grown and re-Avd production induced as described in Fig. 2. For plastic embedding, the cells were fixed in 2.5% glutaraldehyde in PBS at RT for 1 h, postfixed in 1% $OsO_4$ at 4°C for 1 h, dehydrated in ethanol and propylene oxide, and embedded in Epon. For cryosectioning, the cells were fixed in 4% paraformaldehyde/0.1% glutaraldehyde in PBS at RT for 1 h. Cryosections were prepared and immunolabelled as described in Punnonen et al. (1992).

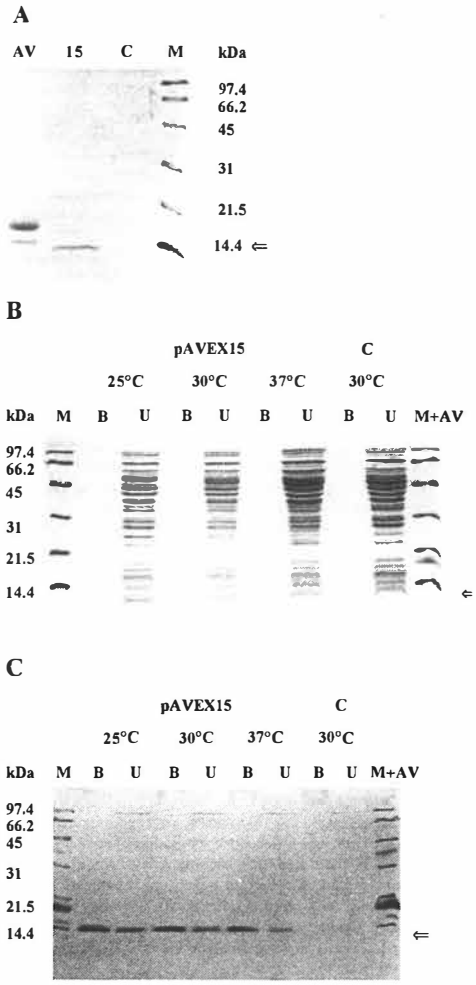



Fig. 4. Characterization of the biotin-binding activity of the produced re-Avd. (Panel **A**) Lanes: (1) **AV**, 5 µg of Avd; (2) **15**, sample corresponding to 100 µl of original culture after 4 h of induction at 37°C using JM109[pAVEX15]; (3) **C**, control sample corresponding to 100 µl of original culture after 4 h of induction at 37°C using JM109[pTrc99A]; (4) **M**, 10 µl of low-range molecular weight standard (Bio-Rad). **Methods:** (**A**) Probing with biotin-labelled alkaline phosphatase (Sigma, P8024). Immunoblotting was performed by electrophoretic transfer of proteins from 0.1% SDS-15% PAGE onto nitrocellulose membrane. The blot was incubated in TBST for 2 h and washed for 5 min in TBST before addition of a fresh solution containing biotin-labelled alkaline phosphatase in TBST (1 µg/ml). After 1 h of incubation, the blot was treated as described in Fig. 2. The color was allowed to develop for 15 min before being stopped by washing with several changes of water. (Panel **B**) Affinity purification of re-Avd. **M**, 10 µl of low-range molecular weight standard (Bio-Rad) and 0.15 µg of Avd; **B**, bound re-Avd; **U**, unbound re-Avd (see below); **M+AV**, 10 µl of low range molecular weight standard (Bio-Rad) and 1 µg of Avd. **Methods:** The *E. coli* JM109 cells harbouring pAVEX15 expression plasmid were grown at a desired temperature and induced as described in Fig. 2, except that the culture volume was 50 ml. After 4 h of induction cells from 4 ml of culture were collected by centrifugation at 1500 × *g* and 4°C for 10 min. The pellets were washed with 5 ml of cold ET buffer, centrifuged as above, frozen with liquid nitrogen and stored at −70°C. The frozen pellets were melted on ice and resuspended in 400 µl of cold ETCD buffer. The cells were lysed by sonication using a Vibra cell™ microtip sonicator (Sonics&Materials). The sonication was performed on ice in two 30-s bursts using a 50% duty cycle and 1-min break between bursts. The lysates were then centrifuged at 15 000 × *g* for 10 min at 4°C. The pellets were resuspended in 400 µl of SDS sample buffer diluted 1:2 in $H_2O$, heated at 96°C for 5 min and stored at −20°C until used in blotting. These samples were referred to as the insoluble fractions. The supernatants were recentrifuged as described and transferred to new tubes. Aliquots of 75 µl were mixed with 75 µl of SDS sample buffer, heated at 96°C for 5 min and stored at −20°C until used in blotting. These samples were referred to as the soluble fractions. The rest of the soluble fractions were mixed with 50 µl of 50% biotin-agarose beads (Sigma, B0519) at RT for 30 min with low speed rotation. Biotin-agarose beads were equilibrated in ET buffer and washed twice in the same buffer before being used for affinity purification. The beads were collected by a short centrifugation at 15 000 × *g* and washed five times with 1 ml of ETT buffer. The supernatants were mixed with an equal volume of SDS sample buffer and saved for analysis of the amount of unbound re-Avd (**U**) in 0.1% SDS-15% PAGE. The beads were resuspended in 100 µl of SDS sample buffer containing 6 **M** urea and heated at 96°C for 5 min to dissociate the bound re-Avd (**B**) from agarose. The samples were then analysed in 0.1% SDS-15% PAGE. (Panel **C**) Western blot of gel shown in panel **B**, which was performed as described in Fig. 2. Lanes: **M**, 10 µl of low range molecular weight standard (Bio-Rad); pAVEX15, samples of bound (**B**, 40 µl/lane) and unbound (**U**, 30 µl/lane) fractions after 4 h of induction at indicated temperatures; **C**, JM109[pTrc99A] control samples after 4 h of induction at 30°C; **M+AV**, 10 µl of low range molecular weight standard (Bio-Rad) and 1 µg of Avd; Arrows mark the re-Avd bands.

in many cells. The inclusions were heavily labelled with anti-Avd in immunolabelled cryosections, but a small amount of Avd was also found scattered in the cytoplasm (Fig. 3A,B,C). Following growth at 24°C, no inclusions were found, and only the scattered cytoplasmic Avd was visible in immunolabelled sections (Fig. 3D,E,F). Microscopy thus supported the biochemical results on the effect of growth temperature: the cells produce mainly insoluble re-Avd at 37°C and soluble re-Avd at 24°C.

### (f) Activity of re-Avd

The crude cell lysates were analysed by 0.1% SDS-15% PAGE and Western blotting with biotinylated alkaline phosphatase to determine whether the re-Avd could bind biotin. The result showed that the re-Avd was capable of binding biotin (Fig. 4A). The biotin-binding capacity of re-Avd was also demonstrated by affinity purification of the re-Avd in a single step from soluble fraction with biotin agarose (Fig. 4B). However, we were able to purify only a fraction of the total amount of re-Avd by the used protocol, which thus requires further development.

**(g) Conclusions**

(*1*) The coding region of the mature Avd was amplified by PCR and subcloned into the pTrc99A cloning vector to create expression plasmids pAVEX8 and pAVEX15. Using the recombinant plasmids, we were able to produce biologically active re-Avd in *E. coli* cells.

(*2*) The adaptation of the first ten codons coding the N-terminal part of mature Avd to the so-called optimal *E. coli* codon usage in pAVEX15 significantly enhanced the yield of the produced re-Avd.

(*3*) Lowering the growth temperature from 37°C to 25°C affected the production of re-Avd. At 37°C most of the synthesised re-Avd was in an insoluble form as compared to 25°C where most of it was soluble. However, at the same time, the overall yield of re-Avd decreased significantly.

(*4*) The biologically active re-Avd could be easily purified in a single step using biotin agarose.

(*5*) According to the results, despite the potentially harmful biotin-binding activity of Avd, it should be possible to synthesize chimeric proteins consisting of re-Avd and the target protein in *E. coli*, what would expand the applications of the avidin-biotin system.

ACKNOWLEDGEMENTS

We wish to thank Irene Helkala, Arja Mansikkaviita, Paavo Niutanen and Raija Vassinen for excellent technical assistance, Sanna Myöhänen (University of Kuopio, Department of Biochemistry and Biotechnology, Kuopio, Finland) for the sequencing of the plasmids and Matti Vuento for critical comments on the manuscript. This work was supported by the Technology Development Centre of Finland (TEKES) and by the Developmental Foundation of the Province of Central Finland.

REFERENCES

Barnes, W.M.: Sequencing DNA with dideoxyribonucleotides as chain terminators: hints and strategies for big projects. Methods Enzymol. 152 (1987) 538–556.

Browner, M.F., Rasor, P., Tugendreich, S. and Fletterick, J.R.: Temperature sensitive production of rabbit muscle glycogen phosphorylase in *Escherichia coli*. Protein Eng. 4 (1991) 351–357.

Chandra, G. and Ray, J.G.: Cloning and expression of avidin in *Escherichia coli*. Methods Enzymol. 184 (1990) 70–79.

Chen, G.-F.T. and Inouye, M.: Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucleic Acids Res. 6 (1990) 1465–1473.

Date, T., Tanihara, K., Yamamoto, S., Nomura, N. and Matsukage, A.: Two regions in human DNA polymerase β mRNA suppress translation in *Escherichia coli*. Nucleic Acids Res. 20 (1992) 4859–4864.

DeLange, R.J. and Huang, T.-S.: 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 (1971) 698–709.

Gope, M.L., Keinänen, R.A., Kristo, P.A., Conneely, O.M., Beattie, W.G., Zarucki-Schulz, T., O'Malley, B.W. and Kulomaa, M.S.: Molecular cloning of the chicken avidin cDNA. Nucleic Acids Res. 15 (1987) 3595–3606.

Gouy, M. and Gautier, C.: Codon usage in bacteria: correlation with gene expressivity. Nucleic Acids Res. 10 (1982) 7055–7074.

Green, N.M.: Avidin. Adv. Prot. Chem. 29 (1975) 85–133.

Grosjean, H. and Fiers, W.: Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene 18 (1982) 199–209.

Gross, G., Mielke, C., Hollatz, I., Blöcker, H. and Frank, R.: RNA primary sequence or secondary structure in the translational initiation region controls expression of two variant interferon-b genes in *Escherichia coli*. J. Biol. Chem. 265 (1990) 17627–17636.

Hall, M.N., Gabay, J., Debarbouille, M. and Schwartz, M.: A role for mRNA secondary structure in the control of translation initiation. Nature 295 (1982) 616–618.

Hiller, Y., Gershoni, J.M., Bayer, E.A. and Wilchek, M.: Biotin binding to avidin. Biochem. J. 248 (1987) 167–171.

Korpela, J.: Avidin, a high affinity biotin binding protein, as a tool and subject of biological research. Med. Biol. 62 (1984) 5–26.

Marston, F.A.O.: The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. Biochem. J. 240 (1986) 1–12.

Livnah, O., Bayer, E.A., Wilchek, M. and Sussman, J.L.: Three-dimensional structures of avidin and the avidin-biotin complex. Proc. Natl. Acad. Sci. USA 90 (1993) 5076–5080.

Pugliese, L., Coda, A., Malcovati, M. and 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. 231 (1993) 698–710.

Punnonen, E.-L., Autio, S., Marjomäki, V.S. and Reunanen, H.: Autophagy, cathepsin L transport, and acidification in cultured rat fibroblasts. J. Histochem. Cytochem. 40 (1992) 1579–1587.

Sano, T. and Cantor, C.R.: Expression of a cloned streptavidin gene in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 87 (1990) 142–146.

Sano, T. and Cantor, C.R.: Expression vectors for streptavidin-containing chimeric proteins. Biochem. Biophys. Res. Commun. 176 (1991) 571–577.

Schein, C.H.: Production of soluble recombinant proteins in bacteria. BioTechnology 7 (1989) 1141–1149.

Wikström, P.K., Lind, L.K., Berg, D.E. and Björk, G.R.: Importance of mRNA folding and start codon accessibility in the expression of genes in a ribosomal protein operon of *Escherichia coli*. J. Mol. Biol. 224 (1992) 949–966.

Wilchek, M. and Bayer, E.A.: Applications to avidin-biotin technology: literature survey. Methods Enzymol. 184 (1990) 14–45.

Zhang, S., Zubay, G. and Goldman, E.: Low usage codons in *Escherichia coli*, yeast, fruit fly and primates. Gene 105 (1991) 61–72.

Zuker, M. and Stiegler, P.: Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. 9 (1981) 133–148.

II

**Rapid purification of recombinant proteins fused to chicken avidin**

by

Airenne, K. J. & Kulomaa, M. S.

Gene 167: 63-68, 1995

https://doi.org/10.1016/0378-1119(95)00631-1

*Gene*, 167 (1995) 63–68
 0378-1119/95/$09.50



GENE 09337

# Rapid purification of recombinant proteins fused to chicken avidin

(Recombinant DNA; gene expression; PCR amplification; affinity chromatography; biotin-agarose; biotin-binding protein; fusion protein)

Kari J. Airenne and Markku S. Kulomaa

*Department of Biological and Environmental Science, University of Jyväskylä, FIN-40100 Jyväskylä, Finland*

Received by J.K.C. Knowles: 1 May 1995; Revised/Accepted: 4 August/11 August 1995; Received at publishers: 7 September 1995

SUMMARY

A novel expression vector (pAVEX16C) has been constructed that directs the synthesis of desired polypetides as fusions with the C terminus of chicken egg-white avidin (Avd). With this and a commercial *GST* gene (encoding glutathione *S*-transferase) fusion vector (pGEX-3X, Pharmacia), we produced Avd as fusions C- and N-terminally linked to GST in *Escherichia coli*. By using the Avd tail and a simple affinity purification protocol, including biotin-agarose, we were able to obtain 1–2 μg/ml of highly purified Avd::GST and GST::Avd from crude bacterial lysates. The produced proteins were, to a great extent, in soluble fraction when the cells were grown at 22°C and disrupted with a detergent, *N*-laurylsarcosine. The fusion proteins could also be affinity-purified with the GST tail using glutathione-Sepharose 4B, but the yield of GST::Avd was significantly lower than when using the Avd tail. Our results therefore indicate that it is possible to produce, in *E. coli*, biologically active fusion proteins consisting of Avd C- or N-terminally linked with the desired protein which then can easily be purified by a simple affinity chromatography procedure.

## INTRODUCTION

Purification of recombinant (re-)protein is often simplified by producing it as a fusion protein with a polypeptide tail, such as a $His_6$-tag (Hochuli, 1990), glutathione *S*-transferase (GST) (Smith and Johnson, 1988) or protein A (Nilsson and Abrahmsén, 1990). The use of a fusion partner may also offer additional advantages over facilitated protein recovery. For instance, the stability, solubility and localization of the recombinant protein in the cell can be influenced by the use of a gene fusion strategy (Uhlén and Moks, 1990). Indeed, gene fusion has become

*Correspondence to:* Dr. M.S. Kulomaa, Department of Biological and Environmental Science, University of Jyväskylä, Vapaudenkatu 4, FIN-40100 Jyväskylä, Finland. Tel. (358-41) 602-272; Fax (358-41) 602-221; e-mail: kulomaa@csc.fi

Abbreviations: *A*, absorbance (200 μl sample in ELISA plate); aa, amino acid(s); Ab, antibody(ies); Ap, ampicillin; APA, 0.1 M $NaHCO_3$/1 mM $MgCl_2$ pH 9.8; Avd, avidin; *Avd*, gene encoding Avd; BCIP, 5-bromo-4-chloro-3-indolyl phosphate in DMF (15 mg/ml); bp, base pair(s); cDNA, DNA complementary to mRNA; CLAP, chymostatin/leupeptin/antipain/pepstatin (10 mg/ml) in DMSO; DMF, dimethylformamide; DMSO, dimethylsulfoxide; dNTP, deoxyribonucleoside triphosphate; DTT, dithiothreitol; ELISA, enzyme-linked immunosorbent assay; ET, 2 mM EDTA/30 mM Tris pH 8; ETCDS, ET+CLAP 1 μl to 1 ml/5 mM DTT/1.5% *N*-laurylsarcosine; ETN, ET+1 M NaCl; GST, glutathione *S*-transferase; *GST*, gene encoding GST; IPTG, isopropyl-β-D-thiogalactopyranoside; $lacI^q$, gene encoding *lac* repressor; LB, Luria-Bertani (medium); LSC, linker sequence C; MCS, multiple cloning site(s); NBT, nitro blue tetrazolium (30 mg/ml) in 70% DMF; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; o/n, overnight; PA, polyacrylamide; PAGE, PA-gel electrophoresis; PBS, 140 mM NaCl/2.7 mM KCl/10 mM $Na_2HPO_4$/1.8 mM $KH_2PO_4$ pH 7.3; PBSDS, PBS plus 5 mM DTT/1.5% *N*-laurylsarcosine; PCR, polymerase chain reaction; *Pfu*, *Pyrococcus furiosus*; *Ptac* and *Ptrc*, hybrid *trp-lac* promoters; re-, recombinant; SDS, sodium dodecyl sulfate; SDS sample buffer, 0.125 M Tris pH 6.8/4% SDS/20% glycerol/0.004% bromophenol blue/10% 2-mercaptoethanol; TBE, 90 mM Tris·borate/2 mM EDTA pH 8; TBS, 0.5 M NaCl/20 mM Tris pH 7.5; TBSM, TBS+5% non-fat milk powder; TBST, TBS+0.2% Tween-20; TBSTM, TBST+5% non-fat milk powder; TE, 1 mM EDTA/10 mM Tris pH 8; [ ], denotes plasmid-carrier state; ::, novel junction (fusion or insertion).

*SSDI* 0378-1119(95)00631-1

an important and widely used tool in molecular biology and biotechnology.

An ideal polypeptide tail for protein fusions should be small, stable, heat resistant and soluble. It should also be able to fold independently and should need no co-factors to be functional. Chicken egg-white Avidin (Avd) meets most of these criteria. It is a glycoprotein (66–69 kDa) consisting of four small identical subunits (15.5–15.8 kDa) each of which can bind one biotin molecule. Avd is soluble in water and salt solutions and stable over a wide range of pH values and temperatures, particularly as a complex with biotin (Green, 1975). The sequence (128 aa) of the subunit (Delange and Huang, 1971) and the 3-D structure of Avd has been determined (Livnah et al., 1993; Pugliese et al., 1993; 1994). The full-length cDNA (Gope et al., 1987) and gene (Wallén et al., 1995) encoding chicken egg-white Avd has been cloned and characterized. It has been shown that the oligosaccharide side chain of each subunit is not essential for biotin binding (Hiller et al., 1987). Avd exhibits one of the highest affinities ($K_d$ $10^{-15}$ M) known between a small ligand and a protein (Green, 1975). It is difficult to disrupt the tetramer structure of Avd, but by denaturating the tetramer on agarose column, the $K_d$ of a single subunit has been determined as approx. $10^{-7}$ M (Kohanski and Lane, 1990). Because of this exceptionally strong interaction, the Avd-biotin system has became a powerful tool in different fields of the biosciences (for a review, see Wilchek and Bayer, 1990). In the protein purification field, streptavidin (Sano and Cantor, 1991; Sano et al., 1992; Karp et al., 1995) and a peptide which becomes biotinylated (Cronan, 1990; PinPoint™ Xa protein purification system, Promega) have been used as a polypeptide tail to enhance protein detection and purification after expression in *E. coli* and insect cells.

We have recently described the production of recombinant Avd (re-Avd) in *E. coli* (Airenne et al., 1994). Here we wanted to study whether it is possible to use Avd as an affinity tail in fast and simple purification of fusion proteins. The question was studied by producing Avd as N- or C-terminally linked with *Schistosoma japonicum* glutathione *S*-transferase (GST; EC 2.5.1.18) in *E. coli*. Avd-fusion proteins could further expand the already widely used Avd-biotin technology due to the ease with which such hybrid proteins could be detected and purified using the biotin-binding activity.

## EXPERIMENTAL AND DISCUSSION

### (a) Construction of expression plasmids pAVEX16C and pAV3X

The pAVEX16C vector was made by inserting a linker sequence C (LSC; Fig. 1B) between the *Pst*I and *Hin*dIII sites of pAVEX15 (Airenne et al., 1994). With this linker a hinge region of mouse IgG3 (Pack and Plückthun, 1992) and MCS was inserted into the pAVEX15 after the re-*Avd* coding sequence to facilitate the production of the desired fusions with re-Avd. The codons in the linker sequence were adapted to the so-called optimal *E. coli* codon usage (Zhang et al., 1991) to enhance protein production. The pAVEX16C vector has a strong IPTG inducible *trc* promoter and *lacI*$^q$ gene allowing tight regulation of the gene expression.

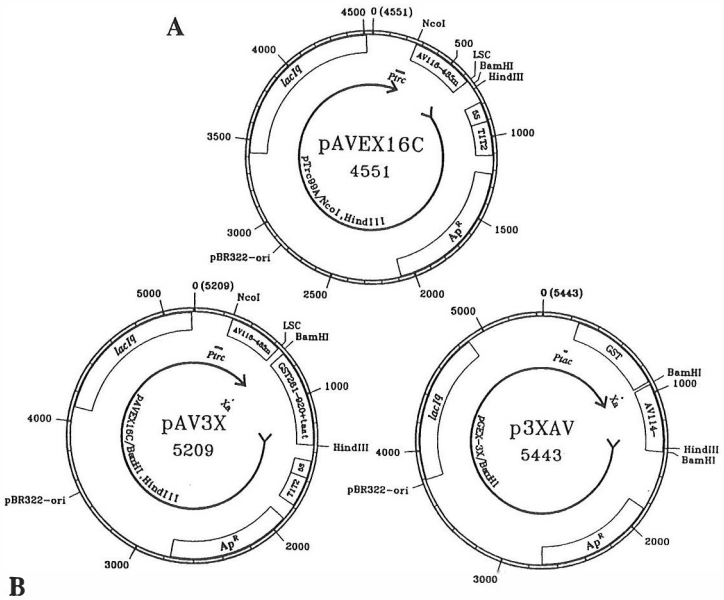
The *Schistosoma japonicum* GST gene contained in pGEX-3X (nt 261–920, Pharmacia LKB) was amplified by PCR using an AK38/AK39 primer pair (Fig. 1B) and subcloned as a *Bam*HI-*Hin*dIII fragment into the pAVEX16C vector (Fig. 1A) to obtain the expression plasmid pAV3X (Fig. 1A), which encodes the Avd::GST fusion protein. The AK38 primer incorporated a *Bam*HI cloning site and factor Xa protease recognition site at the ATG start codon of GST. The AK39 primer incorporated a stop codon (TAAT), as well as a restriction site for *Hin*dIII at the end of the amplified sequence. The stop codon was selected because highly expressed genes are often terminated in *E. coli* by TAAT (Tate and Brown, 1992). The nt of the cloned sequences were confirmed by dideoxynucleotide sequencing of the plasmids with an automated DNA sequencer (ALF, Pharmacia).

### (b) Construction of expression vector p3XAV

In order to construct the GST::Avd producing vector p3XAV (Fig. 1A), the AK10/AK11 primer pair (Fig. 1B) was used in the PCR to amplify the region including the aa 1–128 of Avd from pGEMav (Airenne et al., 1994). The AK11 contained an additional sequence with five extra nt and a *Bam*HI site in front of the Avd coding sequence for subcloning the amplified sequence into the *Bam*HI site of the pGEX-3X. To ensure cloning, an extra Pro was also added between GST and Avd with this primer. The AK10 sequence is taken from the pGEM3 vector (nt 250–269, Promega) in order to obtain several possible cloning sites from the MCS of pGEM3 at the 3′ end of the amplified *Avd* sequence. The pGEX-3X vector is designed for the inducible (*tac* promoter), regulatable (*lacI*$^q$ gene) and high-level intracellular expression of genes or gene fragments as fusions with GST. It also contains the Xa protease recognition site for cleaving the GST part from the fusion product (Smith and Johnson, 1988). The nt of the cloned sequence was confirmed by dideoxynucleotide sequencing of the plasmid with an automated DNA sequencer (ALF, Pharmacia).

### (c) Production of Avd::GST (pAV3X) and GST::Avd (p3XAV)

The cultures of *E. coli* JM109 harbouring pAV3X or p3XAV plasmid were grown at 21–22°C. When the




**AK10:** 5' - GTTATCCCCTGATTCTGTGG - 3' (20 mer)
3' - Primer

Pro Ala Arg Lys Cys Ser Leu
**AK11:** 5' - T TTT TGG ATC CCT GCC AGA AAG TGC TCG CTG - 3' (31 mer)
5' - Primer *Bam*HI

Ser Pro Ile Leu Gly Tyr
**AK38:** 5' - TC ACA GGA TCC *ATT GAA GGC CGT* TCC CCT ATA CTA GGT TAT T - 3' (42 mer)
5' Primer *Bam*HI Xa

Gln Ile Arg Phe Trp Arg
**AK39:** 5' - CCC ACA AGC TTA *TTA* CAG ATC CGA TTT TGG AGG AT - 3' (35 mer)
3' Primer *Hin*dIII Stop

**LSC:** 5' - A CCG AAA CCG AGC ACC CCG CCG GGC AGC AGC GGA TCC CCC GGG
3' - AC GTT GGC TTT GGC TCG TGG GGC GGC CCG TCG TCG CCT AGG GGG CCC
*Pst*I IgG3 hinge region (*E. coli* codons) *Bam*HI *Sma*I

GTC GAC CTG CAG A - 3'
CAG CTG GAC GTC TTC GA - 5'
*Sal*I *Pst*I *Hin*dIII

Fig. 1. Maps of plasmids pAV3X and p3XAV (**A**), sequences of the oligo primers used in PCR amplification, and the sequence of LSC (**B**). Oligos were either (AK10/AK11) synthesised by Cyclone DNA synthesizer (Milligen/Bioresearch) and further treated according to Barnes (1987), or (AK38/39) were purchased from KEBO Lab OY (Helsinki, Finland). **Methods:** Construction of plasmids pAV3X and p3XAV by PCR amplification of GST and Avd was performed basically as described by Airenne et al. (1994), with the exception that Avd was amplified by *Taq* polymerase instead of *Pfu* polymerase. The purified fragments amplified by the primer pairs AK10/AK11 or AK38/AK39 were cloned into the *Bam*HI-digested pGEX-3X (Pharmacia) or *Bam*HI + *Hin*dIII-digested pAVEX16C to construct the re-plasmids p3XAV andpAV3X, respectively. The re-plasmids were introduced into competent *E. coli* cells of strains JM109 and RR1 for preparing mg amounts of the plasmids.

cultures reached the mid-exponential phase they were induced with IPTG and samples were taken at different time points. Polypeptides of the expected sizes (41.2 and 39.1 kDa) were detected in a Coomassie-blue-stained SDS-PA gel. The bands, which were of the expected size, were verified as representing Avd::GST and GST::Avd by an immunoblot analysis using Ab raised against Avd (Fig. 2A,B). The amount of the fusion proteins (estimated from Western blotting) was approx. 5–10 μg/ml for both proteins. The ELISA method for Avd (Joensuu et al., 1991) needs modification, since it did not work with the Avd-fusion proteins. The reason is not known, but the binding of the Ab to Avd may be prevented by a fusion partner.

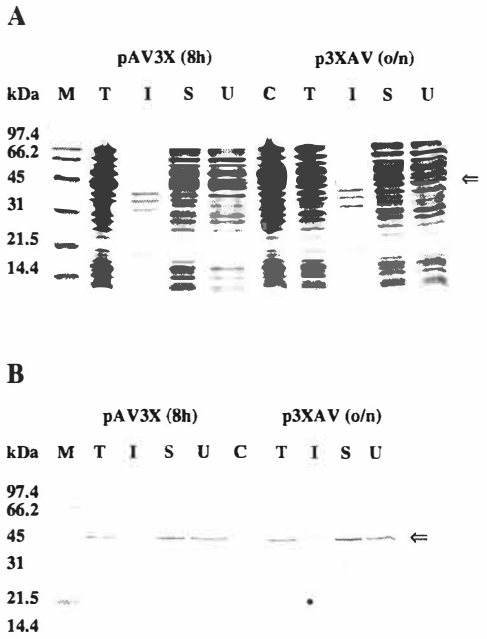



Fig. 2. Production of Avd::GST and GST::Avd from pAV3X and p3XAV, respectively. (**A**) Coomassie-stained 0.1% SDS-15% PA gel of *E. coli* JM109 cells synthesizing Avd::GST and GST::Avd. Samples are prepared according to biotin-agarose purification protocol. Lanes: (1) **M**, 10 µl of low range molecular weight standard (Bio-Rad); (2–5) **T,I,S,U** samples after 8 h of induction using pAV3X; (6) C, JM109 host cells after 8 h of induction (negative control); (7–10) **T,I,S,U** samples after o/n of induction using p3XAV. The arrow points to the Avd::GST or GST::Avd bands. All samples correspond to 50 µl of original culture. **Methods:** Sample preparation: *E. coli* JM109 cells transformed with pAV3X or p3XAV plasmid were streaked out on LB plates contalning 50 µg Ap/ml and the plates were incubated o/n at 37°C. Next day, single colonies were inoculated in 5 ml of LB medium supplemented with 200 µg Ap/ml and the cultures were grown with vigorous shaking o/n at 37°C. The cultures were then centrifuged at 1500 × *g* for 10 min and resuspended in 5 ml of fresh LB medium. Fresh cultures were diluted 1:50 in 65 ml of LB medium containing 200 µg Ap/ml and grown with vigorous shaking at 21–22°C until an $A_{595\ nm}$ of approx. 0.5 was reached. IPTG was added to a final concentration of 0.25 mM and the cultures were grown for an additional periods of time. After 4 h, 8 h or o/n of induction, cells from 5 ml of culture were collected by centrifugation at 2500 × *g* and 4°C for 10 min. The pellets were washed with 5 ml of cold ET buffer, centrifuged as above, frozen with liquid nitrogen and stored at −70°C. The frozen pellets were melted on ice and resuspended in 500 µl of cold ETCDS buffer (biotin-agarose purification) or PBSDS (glutathione-Sepharose 4B purification). Aliquots (50 µl) were mixed with an equal volume of SDS sample buffer, heated at 97°C for 5 min and stored at −20°C. These samples were referred to as the total cell samples (**T**). The cells were lysed by sonication using a Vibra cell™ microtip sonicator (Sonics&Materials). The sonication was performed on ice in two 20-s bursts using a 50% duty cycle, power set 3 and 1-min break between bursts. The lysates were then centrifuged at 15000 × *g* and 4°C for 10 min. The pellets were resuspended in 900 µl of SDS sample buffer diluted 1:2 in $H_2O$, heated at 97°C for 5 min and stored at −20°C. These samples were referred to as the insoluble fractions (**I**). The supernatants were recentrifuged as described and transferred to new tubes. Aliquots of 50 µl were mixed with 50 µl of SDS sample buffer, heated at 97°C for 5 min and stored at −20°C. These samples were referred to as the soluble fractions (**S**). The unbound samples (**U**) were prepared as described in the legend to Fig. 4. (**B**) Western blot of gel shown in panel **A**. Lane 1 supplemented with 1 µg of Avd. Immunoblotting was performed as described by Airenne et al. (1994).

By growing the cells at 37°C the overall yield of fusion proteins was significantly increased, but most of the proteins accumulated as inclusion bodies. Using the lowered growth temperature and the *N*-laurylsarcosine in the cell lysis buffer (ETCDS) we were, however, able to get most of the produced proteins in soluble form (Fig. 2B). The positive effect of the sarkosyl detergent on the solubility of partially or completely insoluble recombinant proteins has also been reported by Frankel et al. (1991) and Frangioni and Neel (1993). The detergent does not harm the biotin-binding activity of Avd (results not shown) as is also the case with SDS, Tween-20 and Triton X-100 (Ross et al., 1986). Interestingly, the Avd::GST seemed to be more stable than the GST::Avd. After four hours of induction, the Western blot showed a major degradation product of approx. 30 kDa from the GST::Avd (Fig. 2B).

**(d) Purification of the fusion proteins**

To test the biotin-binding activity of the produced fusion proteins, the cells were lysed by sonication and the clarified supernatants were mixed with biotin-agarose. Biotin-agarose was used in this study to simplify the analytical process but for preparative purposes it should be replaced by 2-iminobiotin for pH-dependent affinity isolation of Avd fusion proteins (Heney and Orr, 1981). The results showed that the Avd::GST, as well as the GST::Avd were able to bind biotin. By using a simple purification method, we were able to get 1–2 µg/ml (estimation from SDS-PA gel, Fig. 4A) of highly purified Avd::GST and GST::Avd in a single step corresponding to approx. 10–20% of the total amount of each protein (Fig. 2B). The best yield of pure fusion proteins was obtained from samples collected after eight hours or even overnight induction, although the overall yield of both proteins did not seem to increase after four hours (Fig. 3). The delay in the activity of the produced proteins may indicate the need of more time for proper folding in the decreased temperature.

As in the case of re-Avd (Airenne et al., 1994), only a relatively small fraction of the soluble Avd::GST and GST::Avd could be purified by the affinity protocol (10–20%, Fig. 2B). This could be explained by improper

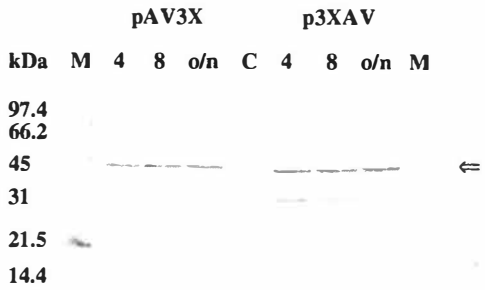



Fig. 3. Western blot of 0.1% SDS-15% PA gel showing the total cell samples (T) from different time points. The samples are made as described in Fig. 2. Lanes: (1) **M**, 10 µl of low range molecular weight standard (Bio-Rad) and 1 µg of Avd; (2–4) Samples after 4 h, 8 h and o/n of induction using pAV3X; (5) **C**, JM109 cells after 8 h of induction (negative control); (6–9) Samples after 4 h, 8 h and o/n of induction using p3XAV. The arrow points to the Avd::GST or GST::Avd bands. All samples correspond to 50 µl of original culture.

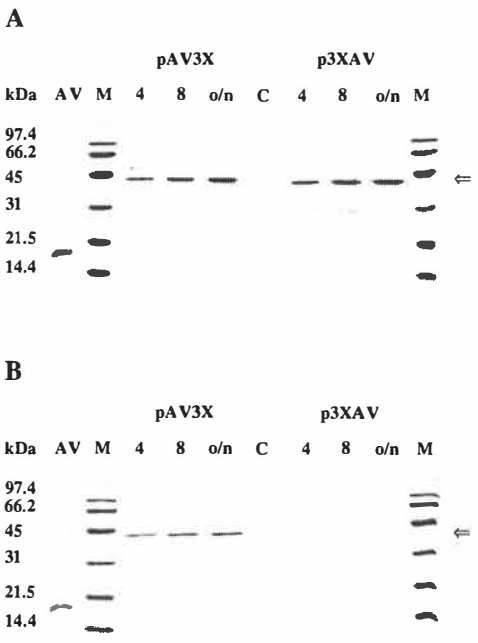



Fig. 4. Purification of Avd::GST and GST::Avd with biotin-agarose (**A**) or glutathione-Sepharose 4B (**B**). Lanes: (1) **AV**, 1 µg of Avd; (2,10) **M**, 10 µl of low range molecular weight standard (Bio-Rad); (3–5) Purified Avd::GST samples corresponding to 1 ml of original culture after 4 h, 8 h and o/n of induction using JM109[pAV3X]; (6) **C**, JM109 cells after 8 h of induction (negative control); (7–9) Purified GST::Avd samples corresponding to 1 ml of original culture after 4 h, 8 h and o/n of induction using JM109[p3XAV]. **Methods:** Affinity purification: The *E. coli* JM109 cells harbouring the re-plasmids pAV3X or p3XAV were treated as described in the legend to Fig. 2. To 400 µl of the soluble fraction (**S**) was added 100 µl 5 M NaCl (for biotin-agarose purification) or 100 µl 10% Triton X-100 (for glutathione-Sepharose 4B purification) and the samples were mixed with 50 µl of 50% biotin-agarose beads (Sigma, B0519) or glutathione-Sepharose 4B (Pharmacia LKB) with low speed rotation at room temperature for 30 min. The biotin-agarose beads were equilibrated in ETN buffer and washed twice in the same buffer before they were used for affinity purification. Glutathione-Sepharose 4B beds were equilibrated in PBS buffer before use. The beads were collected by a short centrifugation at 15 000 × *g* and washed five times with 1 ml of TBST buffer. An aliquot of 75 µl of supernatant was mixed with an equal volume of SDS sample buffer, heated at 97°C for 5 min and saved for analysis of the amount of unbound (**U**) Avd::GST and GST::Avd (Fig. 2B). The beads were resuspended in 100 µl of SDS sample buffer containing 6 M urea and heated at 97°C for 5 min to dissociate the bound Avd::GST or GST::Avd from agarose. The samples were then analysed using 0.1% SDS-15% PAGE.

folding or oligomerization of Avd which could partially be due to the absence of the natural oligosaccharide side chains of Avd in *E. coli*. A further reason for the reduced yield may be saturation of Avd by biotin present in the cytoplasm of *E. coli*. However, the purification protocol was simple. By using the methods of denaturation and renaturation, as has been used with streptavidin fusion proteins (Sano and Cantor, 1991; Sano et al., 1992), recovery might be increased.

The produced fusion proteins were also purified from clarified crude cell lysates with glutathione-Sepharose 4B. It proved possible to purify the Avd::GST with this matrix as well as with biotin-agarose but purification of the GST::Avd resulted in very poor yields (Fig. 4B). The explanation for this may be the hinge sequence (LSC) in Avd::GST which is not present in GST::Avd. The lack of this sequence did not however effect the purification done using biotin-agarose (Fig. 4A).

We have also produced some other prokaryotic or eukaryotic proteins as fusions to the C-terminal end of Avd, such as Avd-β-galactosidase, Avd-EP-A (barley endopoteinase A), Avd-PrA and B (Chicken progesterone receptor A and B) and Avd-IgG3 (hinge region of mouse IgG3)(unpublished results). The yield of the purified fusion proteins has varied depending on the protein, but our results suggest that the smaller the fusion partner the better the yield in affinity purification using Avd tail and biotin-agarose (results not shown).

Even if the purification of Avd-fusion proteins with Avd tail in *E. coli* did not seem to be very efficient in terms of recovery, our system may be useful when mg-amounts of highly purified end product from 1 liter of culture are sufficient. It can also be used in mapping of antigenic epitopes (Vihinen-Ranta et al., unpublished results) or simply as an antigenic tag for detecting

produced proteins (I. Porali, S. Marttila, K. Laukkanen, D. Ho and A. Mikkonen, unpublished results).

**(e) Conclusions**

(*1*) We constructed re-plasmids pAV3X and p3XAV by which we were able to produce biologically active Avd as C- and N-terminally linked to GST in *E. coli* cells. The pAV3X is based on the pAVEX16C expression vector which can also be used for the production of other Avd-fusion proteins in *E. coli*.

(*2*) Purification of the biologically active fusion proteins was easily accomplished in a single step using biotin agarose. We were able to get 1–2 μg of highly purified Avd::GST and GST::Avd from 1 ml of the original cell culture.

(*3*) It was also possible to purify The Avd::GST with GST tail using glutathione-Sepharose 4B, but purification of the GST::Avd with this matrix resulted in low yields.

(*4*) According to the results, it should be possible to produce in *E. coli* and purify to homogeneity other biologically active chimeric proteins with the re-Avd handle.

ACKNOWLEDGEMENTS

We wish to thank Irene Helkala, Arja Mansikkaviita and Paavo Niutanen for first-class technical assistance, Sanna Myöhänen (University of Kuopio, Department of Biochemistry and Biotechnology, Kuopio, Finland) for the sequencing of the plasmids and Professor Hans Söderlund (VTT, Technical Research Center of Finland, Espoo, Finland) for critical review of the manuscript. This work was supported by grants from the Developmental Foundation of the Province of Central Finland, from the Ministry of Education and from the Academy of Finland.

REFERENCES

Airenne, K.J., Sarkkinen, P., Punnonen, E.-L. and Kulomaa, M.S.: Production of recombinant avidin in *Escherichia coli*. Gene 144 (1994) 75–80.

Barnes, W.M.: Sequencing DNA with dideoxyribonucleotides as chain terminators: hints and strategies for big projects. Methods Enzymol. 152 (1987) 538–556.

Cronan Jr., J.E.: Biotination of proteins in vivo. J. Biol. Chem. 265 (1990) 10327–10333.

DeLange, R.J. and Huang, T-S.: 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 (1971) 698–709.

Frangioni, J.V. and Neel, B.G.: Solubilization and purification of enzymatically active glutathione *S*-transferase (pGEX) fusion proteins. Anal. Biochem. 210 (1993) 179–187.

Frankel, S., Sohn, R. and Leinwand, L.: The use of sarkosyl in generating soluble protein after bacterial expression. Proc. Natl. Acad. Sci. USA 88 (1991) 1192–1196.

Gope, M.L., Keinänen, R.A., Kristo, P.A., Conneely, O.M., Beattie, W.G., Zarucki-Schulz, T., O'Malley, B.W. and Kulomaa, M.S.: Molecular cloning of the chicken avidin cDNA. Nucleic Acids Res. 15 (1987) 3595–3606.

Green, N.M.: Avidin. Adv. Prot. Chem. 29 (1975) 85–133.

Heery, D.M., Gannon, F. and Powell, R.: A simple method for subcloning DNA fragments from gel slices. Trends Genet. 6 (1990) 173.

Heney, G. and Orr, G.A: The purification of avidin and its derivatives on 2-iminobiotin-6-aminohexyl-sepharose 4B. Anal. Biochem. 114 (1981) 92–96.

Hiller, Y., Gershoni, J.M., Bayer, E.A. and Wilchek, M.: Biotin binding to avidin. Biochem. J. 248 (1987) 167–171.

Hochuli, E.: Purification of recombinant proteins with metal chelate adsorbent. In: Setlow, J.K. (Ed.), Genetic Engineering, Principle and Methods, Vol. 12. Plenum Press, New York, NY, 1990, pp. 87–98.

Joensuu, T., Tuohimaa, P. and Vilja, P.: Avidin and ovalbumin induction by progesterone in chicken oviduct detected by sensitive immunoenzymometric assays. J. Endocrinol. 130 (1991) 191–197

Karp, K., Linqvist, C., Nissinen, R., Wahlbeck, S., Åkerman, K. and Oker-Blom, C.: Identification of biotinylated molecules using a Baculovirus expressed luciferase-streptavidin fusion protein. BioTechniques 13 (1995) in press.

Kohanski, R.A. and Lane, M.D.: Monovalent avidin affinity columns. Methods Enzymol. 184 (1990) 194–200.

Livnah, O., Bayer, E.A., Wilchek, M. and Sussman, J.L.: Three-dimensional structures of avidin and the avidin-biotin complex. Proc. Natl. Acad. Sci. USA 90 (1993) 5076–5080.

Nilsson, B. and Abrahmsén, L.: Fusions to staphylococcal protein A. Methods Enzymol. 185 (1990) 144–161.

Pack, P. and Plückthun, A.: Miniantibodies: use of amphipathic helices to produce functional, flexibly linked dimeric $F_v$ fragments with high avidity in *Escherichia coli*. Biochemistry 31 (1992) 1579–1584.

Pugliese, L., Coda, A., Malcovati, M. and 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. 231 (1993) 698–710.

Pugliese, L., Malcovati, M., Coda, A. and Bolognesi, M.: Crystal structure of apo-avidin from hen egg-white. J. Mol. Biol. 235 (1994) 42–46.

Ross, S.E., Carson, S.D. and Fink, L.M.: Effects of detergents on avidin-biotin interaction. BioTechniques 4 (1986) 350–354.

Sano, T. and Cantor, C.R.: Expression vectors for streptavidin-containing chimeric proteins. Biochem. Biophys. Res. Commun. 176 (1991) 571–577.

Sano, T., Glazer, A.N. and Cantor, C.R.: A streptavidin-metallothionein chimera that allows specific labeling of biological materials with many different heavy metal ions. Proc. Natl. Acad. Sci. USA 89 (1992) 1534–1538.

Smith, D.B. and Johnson, K.S.: Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene 67 (1988) 31–40.

Tate, W.P. and Brown, C.M.: Translational termination: "Stop" for protein synthesis or "pause" for regulation of gene expression. Biochemistry 31 (1992) 2443–2450.

Uhlén, M. and Moks, T.: Gene fusions for purpose of expression: an Introduction. Methods Enzymol. 185 (1990) 129–143.

Wallén, M.J., Laukkanen, M.O. and Kulomaa, M.S.: Cloning and sequencing of the chicken egg-white avidin-encoding gene and its relationship with the avidin-related genes *Avr1-Avr5*. Gene 161 (1995) 205–209.

Wilchek, M. and Bayer, E.A.: Applications to avidin-biotin technology: literature survey. Methods Enzymol. 184 (1990) 14–45.

Zhang, S., Zubay, G. and Goldman, E.: Low usage codons in *Escherichia coli*, yeast, fruit fly and primates. Gene 105 (1991) 61–72.

# III

**Production of biologically active recombinant avidin in baculovirus-infected insect cells**

by

Airenne, K. J., Oker-Blom, C., Marjomäki, V. S., Bayer, E. A., Wilchek, M. & Kulomaa, M. S.

Prot. Express. Purif. 9: 100-108, 1997

https://doi.org/10.1006/prep.1996.0660

PROTEIN EXPRESSION AND PURIFICATION 9, 100–108 (1997)
ARTICLE NO. PT960660

# Production of Biologically Active Recombinant Avidin in Baculovirus-Infected Insect Cells

Kari J. Airenne,* Christian Oker-Blom,† Varpu S. Marjomäki,* Edward A. Bayer,‡ Meir Wilchek,‡ and Markku S. Kulomaa*,[1]

*Department of Biological and Environmental Science, University of Jyväskylä, FIN-40100, Jyväskylä, Finland; †VTT Biotechnology and Food Research, FIN-02044 VTT, Finland; and ‡Department of Membrane Research and Biophysics, The Weizmann Institute of Science, Rehovot 76100, Israel

Received May 9, 1996, and in revised form July 24, 1996

**An efficient lepidopteran insect cell system was established for the expression of a recombinant form of chicken egg-white avidin. The gene product was obtained in both secreted and intracellular forms, and biologically active recombinant avidin was isolated using affinity chromatography on an iminobiotin–agarose column. Similar to the known quaternary structure of the native egg-white protein, the purified recombinant protein was glycosylated and assembled mainly into tetramers. Like native avidin, the recombinant tetramer also exhibited a high level of thermostability, and was further stabilized upon binding biotin. The biotin-binding and structural properties of the recombinant avidin are thus similar to those of the natural egg-white protein, and the insect system is appropriate both for future site-directed mutagenesis studies and for the production of avidin fusion proteins.**


Chicken egg-white avidin and bacterial streptavidin (from *Streptomyces avidinii*) are tetrameric proteins which bind biotin with similar affinity constants ($K_a \sim 10^{15}$ $M^{-1}$), the strongest protein–ligand interaction known in nature (1,2). The two proteins differ, however, since avidin is a strongly basic (p$I$ > 10) glycoprotein, whereas streptavidin is neutral and unglycosylated. X-ray analyses have shown clearly that the fold of both proteins is generally conserved; moreover, most of their binding-site residues are also conserved, even though there is only about 35% overall sequence identity in both proteins.

The avidin tetramer contains four identical subunits ($M_r$ 15,500), each of which consists of a single polypeptide chain bearing 128 amino acids and possessing one biotin-binding site (3). Of the 10 asparagine residues of each subunit, only Asn17 is glycosylated. According to Bruch *et al.* (4), the carbohydrate accounts for about 10% of the molecular weight and exhibits extensive glycan microheterogeneity. Avidin contains at least three distinct structural types of oligosaccharide, similar in composition and size. The oligosaccharide side chain is not essential for biotin binding (5). The full-length cDNA and gene which encode for chicken egg-white avidin have been cloned (6,7), and the 3D-structure has been determined (8,9).

Despite extensive knowledge about the properties of avidin, little is known about its biological function. Since the growth of a number of microorganisms is inhibited by avidin, and its occurrence after tissue damage has been clearly demonstrated, it most likely has a role as a host defense factor (10). The remarkably high specific affinity of avidin for biotin, together with the facile incorporation of the biotin moiety into various binders and probes, has served to promote the widely used avidin–biotin technology (11).

Streptavidin was recently cloned and expressed in *Escherichia coli* and *Bacillus* systems (12,13). We have previously developed an expression system for recombinant avidin (re-avidin) in *E. coli* (14). The production of biologically active re-avidin, however, was not very efficient in this system. In addition, it is not possible to study the biological role of the glycosylation of avidin in *E. coli.* Many eukaryotic proteins have been produced successfully using a baculovirus expression system. High levels of expression can be achieved by replacing the polyhedrin gene of the *Autographa californica* nuclear polyhedrosis virus (AcNPV) with a gene of interest. Insect cells provide the eukaryotic environment with the capacity for most of the posttranslational

[1] To whom reprint requests should be addressed.



1046-5928/97 $25.00

modifications (signal cleavage, proteolytic cleavage, N-glycosylation, O-glycosylation, acylation, amidation, phosphorylation, prenylation, and carboxymethylation) required for biological activity of a eukaryotic protein (15,16).

In the present study, we describe the efficient expression of biologically active re-avidin in baculovirus-infected lepidopteran insect cells. The protein was purified to homogeneity, and its properties were compared to those of the native egg-white protein. Re-avidin is secreted into the culture medium; it binds biotin and is glycosylated similarly to egg-white avidin. The baculovirus system thus offers an efficient method of avidin production for site-directed mutagenesis and studies of the structure–function relationship. It should also provide a tool for the production of avidin fusion proteins (17) for different applications in biotechnology and bioscience.

## MATERIALS AND METHODS

### *Materials*

Egg-white avidin and 2-iminobiotin–agarose were purchased from Sigma Chemical Co. (St. Louis, MO). Rabbit avidin antibody and its horseradish peroxidase (HRP) conjugate were obtained from DAKO A/S (Glostrup, Denmark). The Bac-To-Bac baculovirus expression system was obtained from GIBCO BRL (Gaithersburg, MD). Goat anti-rabbit IgG (H + L) alkaline phosphatase conjugate and the low-range SDS–PAGE molecular weight standards were purchased from Bio-Rad Laboratories (Hercules, CA). Endo $H_f$ and PNGase F glycosidases were products of New England Biolabs (Beverly, MA). All other reagents were of the purest analytical grade available.

### *Construction of Recombinant Baculoviruses*

The recombinant viruses were generated using the Bac-To-Bac baculovirus expression system. In order to produce the desired recombinant pFastBAC1 donor plasmid (pbAvd$_r$), the region corresponding to nucleotides (nt) 33–501 of avidin cDNA was amplified from pGEMAV (14) by PCR amplification using 5′-**CTGCTAGATCTATGGTGCACGCAACCTCCCC**-3′ (avidin cDNA 33-63, in bold; *Bgl*II site, underlined) as the forward primer and 5′-**CACTCAAGCTTTCACTGCAGGCGAGTGAAGATGTT**-3′ (avidin cDNA 467-501, in bold; stop codon site between *Pst*I and *Hin*dIII sites, underlined) as the reverse primer. The forward primer was purchased from KEBO Lab (Espoo, Finland). The reversed primer was synthesized by an Applied Biosystems 381A DNA synthesizer and treated further according to Barnes (18). The PCR was performed as described by Airenne *et al.* (14). After digestion with *Bgl*II + *Hin*dIII (Promega, Madison, WI), the PCR product was applied to a 1.5% preparative agarose gel. The fragment, which was of the expected size, was recovered from the gel as described by Heery *et al.* (19) and was purified further with the Magic DNA clean-up system (Promega). The purified fragment was cloned into the *Bam*HI + *Hin*dIII-digested pFastBAC1. The resulting recombinant donor plasmid was named pbAvd$_r$. The re-plasmid was introduced into competent *E. coli* cells of strain JM109 for the production of milligram amounts of the plasmid. The nucleotide sequence was confirmed by dideoxynucleotide sequencing with an automated DNA sequencer (ALF, Pharmacia Biotech). The recombinant baculovirus was generated using the Bac-To-Bac system according to the procedures described by the manufacturer (Gibco BRL). The primary virus stock was amplified for a larger scale production of re-avidin and the titer of a virus stock was determined using a plaque assay procedure (16).

### *Production of Re-Avidin in Insect Cells*

In order to produce re-avidin, approximately $5 \times 10^7$ Sf9 cells (ATCC CRL 1711) in SF-900 II SFM serum-free culture medium (GIBCO BRL), supplemented with 0.5 × PSN antibiotic mixture (GIBCO BRL), were seeded to a final volume of 50 ml in a 250-ml Erlenmeyer flask. Recombinant viruses were added to give an m.o.i. of 10 PFU/cell. After 3 days of incubation at 27°C in a shaker at 125 rpm, the cells were pelleted by centrifugation (1000*g*, 22°C, 5 min). The pellet was washed twice with 20 ml of PBS (140 mM NaCl/2.7 mM KCl/10 mM $Na_2HPO_4$/1.8 mM $KH_2PO_4$, pH 7.3), centrifuged as above, and frozen under liquid nitrogen and stored at −70°C.

For biotin-free expression, approximately $2.5 \times 10^7$ Sf9 cells in SF-900 II SFM were seeded to a final volume of 50 ml in a 250-ml Erlenmeyer flask. Recombinant viruses were added to give an m.o.i. of 0.1 PFU/cell. After 1 day of incubation at 27°C as above, the cells were pelleted by centrifugation (100*g*, 22°C, 5 min). The cells were then washed twice with PBS A (137 mM NaCl/3 mM KCl/2 mM $KH_2PO_4$/8 mM $Na_2HPO_4$, pH 7.4) and resuspended in 50 ml of biotin-free SF-900 II SFM, and the infection process was continued for another 2 days. The cell pellet was collected as described above. The medium was made biotin-free using SoftLink soft release avidin resin (Promega) according to the manufacturer's instructions.

### *Purification of Re-Avidin*

All steps were performed on ice unless otherwise stated. The cell pellet was resuspended in 5 ml of lysis buffer (10 mM Tris, pH 7.5/130 mM NaCl/1% Triton X-100/10 mM NaF/10 mM NaPi, pH 7.5/10 mM NaPPi, pH 7.5) supplemented with 2 μl/ml of protease-inhibitor cocktail CLAP (chymostatin/leupeptin/antipain/pep-

statin, 10 mg/ml each in DMSO). After 45 min of incubation, an aliquot of 75 $\mu$l (referred to as the total cell protein sample, T) was taken, and the rest of the lysate was centrifuged (15,000*g*, 4°C, 10 min). The pellet was then resuspended in 2.5 ml of lysis buffer and an aliquot of 75 $\mu$l (referred to as the insoluble fraction of re-Avidin, I) was taken. The supernatant was mixed with an equal volume of binding buffer (50 mM $Na_2CO_3$, pH 11/1 M NaCl). The pH of the mixture was adjusted to 11 with 1 M NaOH and NaCl was added to a final concentration of 1 M. The crude re-avidin mixture was then applied to a 2-iminobiotin–agarose column (1 ml) which had previously been equilibrated with binding buffer. A 75-$\mu$l aliquot was taken before (soluble fraction, S) and after (unbound fraction, U) applying the mixture to the column. The re-avidin (rAv) was eluted from the column with elution buffer (50 mM ammonium acetate, pH 4/0.1 M NaCl) in 250-$\mu$l fractions and stored at −20°C pending further analysis.

For purification after biotin-free expression, the cells were treated as above except that (i) the pellet was lysed for 30 min in 4 ml of HILLO1 buffer (50 mM Tris-HCl, pH 8/1% Triton X-100/2 mM EDTA/150 mM NaCl + 2 $\mu$l/ml CLAP) and (ii) the lysate was sonicated using a Vibra Cell microtip sonicator (Sonics & Materials, Inc., Danbury, CT). The sonication was performed on ice in two 20-s bursts using a 50% duty cycle, power set at 2, and with 1-min break between bursts.

*Protein Analysis*

Protein samples were diluted 1:2 in sample buffer (0.125 M Tris-HCl, pH 6.8/4% SDS/20% glycerol/0.004% bromophenol blue/10% 2-mercaptoethanol) and analyzed by SDS–PAGE (20). For immunoblotting, proteins were transferred onto a nitrocellulose membrane (Trans-Blot; Bio-Rad) and the molecular weight standards were marked onto blots with a pencil according to bands revealed by Ponceau S red staining. Unspecific binding to the membrane was blocked by incubating overnight at 4°C in TBS (0.5 M NaCl/20 mM Tris-HCl, pH 7.5) supplemented with 5% nonfat milk powder. The blot was then washed in TBST (TBS plus 0.2% Tween 20) for 5 min and reacted with rabbit avidin antibody diluted 1:6000 in TBSTM (TBST plus 5% nonfat milk powder) for 1 to 3 h. The blot was washed four times in TBST buffer for 5 min and then reacted with goat anti-rabbit IgG (H + L) alkaline phosphatase conjugate diluted 1:2000 in TBSTM. After 1 to 3 h of incubation, the blot was washed four times in TBST for 5 min and once for 5 min in APA (0.1 M $NaHCO_3$/1 mM $MgCl_2$, pH 9.8). Alkaline phosphatase was detected by adding 4 ml of APA containing 40 $\mu$l of nitro blue tetrazolium solution (30 mg/ml in 70% dimethylformamide) and 40 $\mu$l of 5-bromo-4-chloro-3-indolyl phosphate (15 mg/ml of dimethylformamide). The color was usually allowed to develop for 3–5 min before being terminated by washing several times with water.

*ELISA for Re-Avidin*

The concentration of re-avidin was determined by an enzyme-linked immunoadsorbent assay (ELISA) (21). A microtiter plate (Maxisorp; Nunc, Roskilde, Denmark) was coated at 4°C overnight with avidin antibodies diluted 1:6000 in 60 mM $Na_2CO_3$–$NaHCO_3$, pH 9.6. The plate was then washed three times with PBS A containing 0.05% Tween 20 (PBST), blocked for 30 min at 37°C with PBSTB (PBST containing 1% BSA), and then washed three times with PBST. Samples were added, and the plate was incubated at 37°C for 60 min. The plate was then washed as before and incubated for 60 min at 22°C with HRP-conjugated avidin (diluted 1:1000 with PBSTB). After incubation, the plate was washed as above and incubated for 5 min at 22°C with substrate solution (0.4 mg of 1,2-phenylenediamine-dihydrochloride/1.5 $\mu$l $H_2O_2$ in 1 ml of 50 mM citric acid/0.1 M $Na_2HPO_4$, pH 5.0). The reaction was terminated with 2 M $H_2SO_4$, and color was measured spectrophotometrically at $A_{492}$.

*Thermostability of Re-Avidin*

Purified re-avidin, in the presence or absence of an excess of biotin, was diluted in sample buffer, treated for 20 min at 22, 80, or 97°C and subjected to SDS–PAGE. Commercial egg-white avidin was used as a control (22). The samples were then analyzed by immunoblotting. The stability of avidin increases at elevated temperatures when saturated with biotin (1,23,24).

*Quaternary State of Re-Avidin*

Purified re-avidin was acetylated prior to analysis by SDS–PAGE as described by Bayer *et al.* (22) to study its quaternary structure. Briefly, samples (0.5 $\mu$g in 10 $\mu$l) of re-avidin or egg-white avidin were mixed with 5 $\mu$l of 50 mM sodium bicarbonate, pH 8, and then with 5 $\mu$l of acetyl *N*-hydroxysuccinimide ester (5 mg/ml of dimethylformamide). The samples were briefly mixed by Vortex, and the reaction was allowed to continue for 10 min, after which the samples were diluted 1:2 in gel buffer and analyzed by SDS–PAGE.

*Glycosylation Analysis*

The glycosylation pattern of re-avidin was studied by treating 0.5 $\mu$g of purified re-avidin in elution buffer with 1000 U of Endo $H_f$ or PNGase F glycosidase. The enzyme treatments were performed according to the manufacturer's instructions (New England Biolabs).

*Immunoelectron Microscopic Studies*

The Sf9 cells infected with the recombinant virus as well as mock-infected cells were prepared for cryosec-

tioning and immunoelectron microscopy as described by Marjomäki *et al.* (25). Briefly, cells ($6 \times 10^5$) were fixed in 1 ml of 8% paraformaldehyde in 250 mM Hepes, pH 7.4, with simultaneous centrifugation (1000*g*, 22°C, 10 min). Fixation was continued at 4°C for 24 h without disturbing the cell pellet. Small blocks of the cell pellet were infused with 2.1 M sucrose for 30 min and then mounted on metal specimen stubs. Thin frozen sections were labeled using avidin antibodies, followed by treatment with 10-nm protein A–gold particles. The preparation of protein A–gold particles and the immunogold labeling procedure were as described by Marjomäki *et al.* (25).

## RESULTS AND DISCUSSION

### *Construction of the Donor Vector (pbAvdr) and Recombinant Baculoviruses*

The Bac-To-Bac Baculovirus Expression System allows rapid and efficient generation of re-baculovirus-DNA by site-specific transposition in *E. coli* (26). As a first step to obtain re-avidin-producing baculoviruses, the PCR-amplified fragment coding for re-avidin was inserted into the multiple cloning site of the pFastBAC1 donor vector under the control of the AcNPV polyhedrin promoter. The cloned fragment encoded the C-terminally truncated, modified form of avidin, previously described by Airenne *et al.* (14), wherein an extra Gln residue was included at the C-terminal end of the molecule and the last five amino acids of native avidin were excluded. However, the fragment cloned in this work contained the native signal sequence of avidin (nt 44–115 of avidin cDNA) at the N-terminal end. Recombinant bacmids were then constructed by transposing a mini-Tn7 element from a donor plasmid (pbAvdr) to the mini-*att*Tn7 at tachment site on the bacmid. The resulted bacmids were transfected into Sf9 cells.

### *Synthesis and Purification of Re-Avidin*

Insect cells (Sf9), infected with recombinant virus, synthesized sufficient re-avidin for easy visualization by immunoblot analysis (Fig. 1). In addition to a polypeptide of the expected size, two other forms of re-avidin were identified by avidin antibody. These data demonstrate that the recombinant polypeptides produced are antigenically consistent with egg-white avidin.

The intracellular concentration of re-avidin appeared to reach a maximum by ~72 h. Cultures were therefore routinely harvested 3 days after infection. At that time, protein was also detectable in the culture medium (Fig. 1, Sec). The amount of re-avidin from both the cells and the culture medium was quantified by ELISA (21). Accordingly, $5 \times 10^7$ Sf9 cells produced ~0.61 mg of re-avidin, of which about 59% (~0.36 mg) was secreted into the extracellular culture medium. Most (~0.23 mg, 92%) of the intracellular re-avidin was in soluble form after disrupting the cells in lysis buffer.

Attempts to purify re-avidin by affinity chromatography on 2-iminobiotin–agarose (27) yielded only about 27 $\mu$g (12%) of highly purified protein from the solubilized intracellular material. This indicated that only a small portion of re-avidin was active in binding biotin. Attempts to purify re-avidin from the culture medium failed, suggesting that the secreted re-avidin was saturated with biotin. Indeed, the amount of free biotin in the culture medium was sufficient to block all of the free biotin-binding sites of the re-avidin produced.

In order to improve the recovery of the re-avidin, we excluded biotin from the SF-900 II SFM culture medium after 1 day of infection (Fig. 1C). As a consequence, we were able to increase the recovery of the solubilized intracellular re-avidin fraction remarkably. The improved recovery is shown in Fig. 1. In Fig. 1B, no major difference could be seen between lanes containing identical sample of re-avidin before (S) and after passing through a 2-iminobiotin column (U). On the contrary, hardly any re-avidin could be detected by anti-avidin in lane U of Fig. 1C. This and the ELISA results suggest that the recovery is over 90%. Because the major difference between Figs. 1B and 1C was in the lanes S and U, only these lanes are shown in Fig. 1C. The use of biotin-free culture medium also enabled the efficient purification of secreted re-avidin by 2-iminobiotin column, but we have not quantified the recovery.

The production of re-avidin by Sf9 cells was also studied by immunoelectron microscopy. Immunolabeling showed high amounts of re-avidin inside vesicular compartments, which most likely participate in the secretory pathway of infected Sf9 cells. Large numbers of the vesicles were observed close to the surrounding plasma membrane, suggesting that re-avidin is in the process of being secreted from the cells (Fig. 2). These results, together with the biochemical data, also suggest that the natural signal sequence of chicken avidin was capable of directing re-avidin to the secretory pathway of the insect cells.

### *Characterization of the Insect Re-Avidin*

The properties of purified re-avidin were compared to those of egg-white avidin. Immunoblot analysis using avidin antibody revealed that the purified insect cell material separated into at least three main components, a major form at about $M_r$ 18,000 and two minor forms at about $M_r$ 17,000 and 15,500 (Fig. 1B). The majority of the secreted re-avidin (Sec) fit with the $M_r$ 17,000 form, but forms which fit with $M_r$ 15,500 and 18,000 also were present in the sample from the supernatant fluids. Whether the different species of secreted

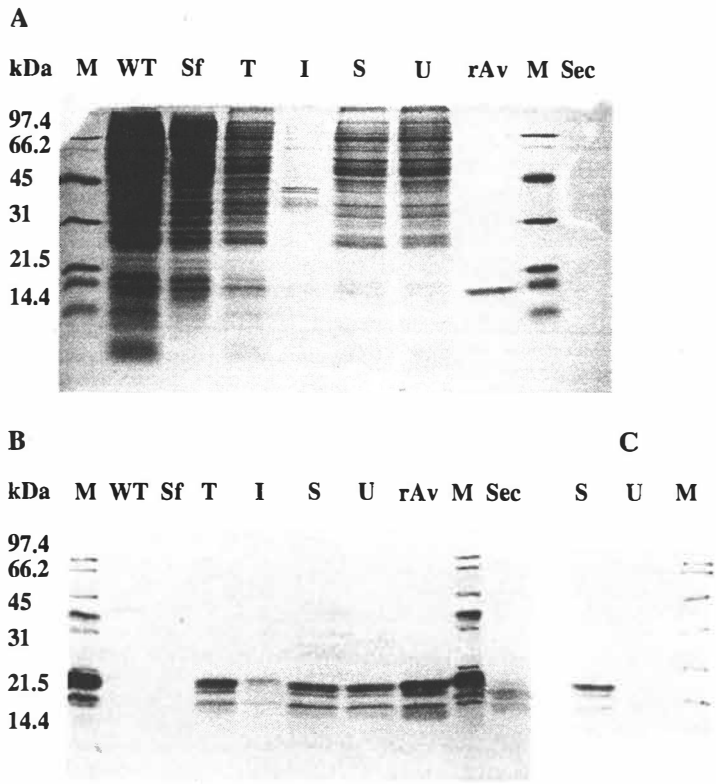



**FIG. 1.** Expression of re-avidin in baculovirus-infected Sf9 cells. Samples were diluted 1:2 in sample buffer, heated at 97°C for 5 min, and analyzed by SDS/15% PAGE. Samples from A and B were obtained from cells grown in the usual (biotin-containing) serum-free culture medium; samples from C were derived from cells grown in biotin-free medium. In A and B, each sample corresponds to protein obtained from $5 \times 10^4$ Sf9 cells (except sample Sec, see below); in C, samples represent material from $3 \times 10^4$ Sf9 cells. Proteins were detected either by staining with Coomassie brilliant blue R-250 (A) or by immunoblotting (B and C) using avidin antibodies. Lanes: **M**, low-molecular-weight standard to which 1 $\mu$g of egg-white avidin was added; **WT**, total cell protein from wild-type virus-infected Sf9 cells; **Sf**, total cell protein from uninfected Sf9 cells; **T**, total cell protein from recombinant virus-infected Sf9 cells; **I**, insoluble fraction from lysed recombinant virus-infected Sf9 cells; **S**, soluble fraction from lysed recombinant virus-infected Sf9 cells, before affinity chromatography on iminobiotin agarose; **U**, "unbound" fraction of the latter (material which failed to bind to the affinity column); **rAv**, affinity-purified re-avidin (1.5 $\mu$g); **Sec**, sample from 50 ml of culture medium of recombinant virus-infected Sf9 cells ($5 \times 10^7$) corresponding to 20 $\mu$l of medium.

re-avidin were actually identical to the intracellular forms remains to be studied more carefully. As expected for a truncated protein, the $M_r$ 18,000 form migrated slightly faster than the commercial egg-white avidin; the $M_r$ 15,500 form migrated to a position near the theoretical molecular weight of re-avidin without carbohydrate moiety (13,829). This suggests that the $M_r$ 15,500 band represents the carbohydrate-free form of re-avidin and the two other, perhaps, represent different forms of glycosylated re-avidin. Indeed, some commercial preparations of egg-white avidin are known to comprise various species, both glycosylated and deglycosylated (28). The glycosylated state of re-avidin was confirmed by treating the protein with Endo $H_f$ and PNGase F glycosidases (Fig. 3). Endo $H_f$ treatment removed the $M_r$ 18,000 band, but seemed to have little effect on the other two forms of re-avidin. PNGase F treatment removed both of the high-molecular-weight forms, leaving only the $M_r$ 15,500 band (Fig 3). These results collectively imply that the Sf9 insect cells were capable of glycosylating re-avidin. This is in agreement with previous studies that N-glycosylated

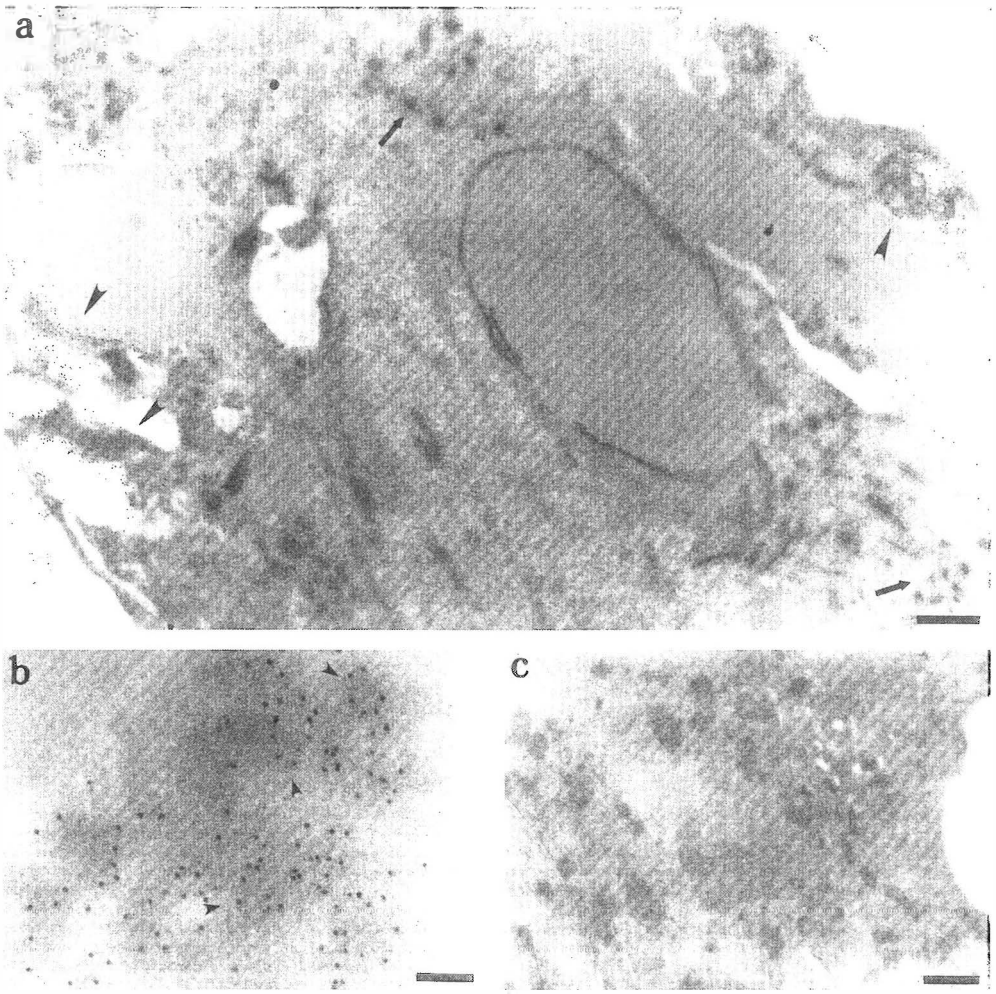



**FIG. 2.** Thin frozen sections of the recombinant virus-infected Sf9 cells. Sections were labeled with avidin antibodies and stained with 10-nm protein A–gold particles. (a) Low-magnification micrograph of an infected cell after 48 hr of infection. At this stage, the cell exhibits typical characteristics of baculovirus infection, such as accumulation of virions (arrows) in the cytoplasm. High amounts of re-avidin are located primarily close to the periphery of the cell (arrowheads). (b) A detailed view showing the peripheral cytoplasm of an infected cell. Note that avidin appears to be confined to vesicular structures, the majority of which are likely to belong to the secretory apparatus of the cell. The limiting membrane of the vesicles is typically seen as white against darker background (arrowheads). (c) Control Sf9 cells are essentially unstained by avidin antibodies. Bars: 500 nm (a), 100 nm (b), 500 nm (c).

proteins are frequently glycosylated in insect cells (16,29–32). Like natural avidin, most of the purified glycosylated re-avidin contained Endo H-sensitive oligosaccharide chains ($M_r$ 18,000 form), but an Endo H-resistant, glycan-containing band ($M_r$ 17,000 ) was also evident in immunoblots. The majority of the purified re-avidin was, in fact, glycosylated, as demonstrated by Coomassie staining (Fig. 1A). The distinct $M_r$ 18,000

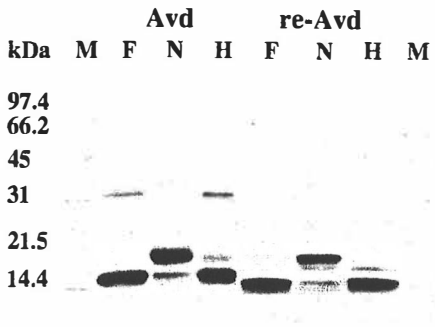



**FIG. 3.** Glycosylation of re-avidin by baculovirus-infected insect cells. Samples of avidin (Avd) and re-avidin (re-Avd) were treated with Endo $H_f$ (H) or PNGase F (F). Following SDS/15% PAGE, the contents of the gel were transferred to nitrocellulose paper, and the blotted proteins were analyzed immunochemically using avidin antibodies. Each lane contains 0.5 $\mu$g of protein, diluted 1:2 in sample buffer; samples were heated at 97°C for 5 min prior to electrophoresis. M, low-range molecular weight standard; N, untreated sample.

band suggests a relatively discrete oligosaccharide moiety, the size of which appears to be similar to that of egg-white avidin. It will therefore be interesting to study the composition of the carbohydrate residues of re-avidin.

Avidin aggregates extensively when mixed at ambient temperatures with buffers which contain sodium dodecyl sulfate, e.g., a sample buffer for SDS–PAGE (22). The resultant aggregates fail to penetrate polyacrylamide gels during electrophoresis, unless heated at high temperatures (>70°C) before loading onto gels. Heating disrupts the tetrameric structure of avidin, and the monomeric form is mainly observed in polyacrylamide gels after electrophoresis. The extent of the dissociation of tetramers to monomers is temperature dependent, and biotin stabilizes the tetramer so that the transition temperature is higher for biotin-saturated avidin (22). Thus, biotin-free avidin migrates into the gel when treated at 80°C for 20 min in sample buffer, while it penetrates poorly when saturated with biotin. The thermostability test was used to analyze the purified re-avidin (Fig. 4). The results showed that, like the egg-white protein, biotin-saturated re-avidin is stabilized at 80°C, suggesting that re-avidin binds biotin in a manner similar to the native protein.

The authenticity of the re-avidin was also studied by comparing its biotin-binding activity with that of egg-white avidin (data not shown). For this purpose, an ELISA-like assay was performed based on the interaction of avidin with biotinyl alkaline phosphatase. The biotin-binding activity of re-avidin was estimated to be about 90% that of the native egg-white protein. These data further support the biological viability of baculovirus-produced re-avidin.

We have recently developed a method to study the quaternary state of avidin in SDS–PAGE (22). The method is based on the acetylation of samples, which inhibits aggregation of the positively charged protein in SDS-containing sample buffer. Upon boiling, avidin migrates in SDS–PAGE as the dissociated monomer; at ambient temperatures, the acetylated avidin is capable of penetrating to the stacking gel, and avidin migrates at a position consistent with the intact tetramer. This approach was used to study the quaternary state of re-avidin. As shown in Fig. 5, under undenaturing conditions, most of the re-avidin was in the tetramer state, the remainder being the free monomer. Interestingly, the re-avidin sample did not contain dimer or oligo-tetramer forms, which were observed in the commercial egg-white avidin standard. Oligomerization of the avidin tetramer is commonly observed in some, but

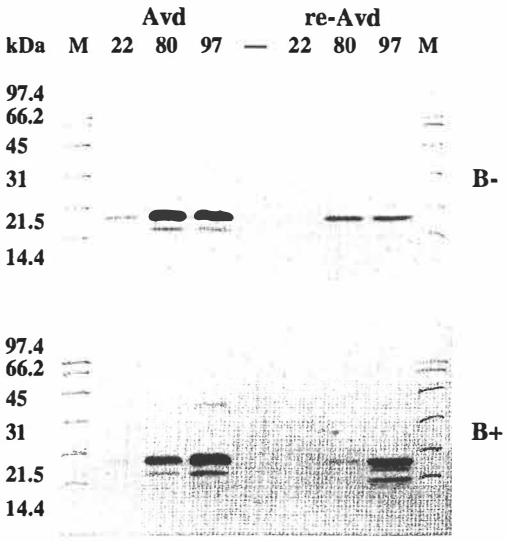



**FIG. 4.** Thermostability properties of re-avidin versus egg-white avidin. Samples of egg-white avidin (Avd) and re-avidin (re-Avd) were heated for 20 min in sample buffer at the indicated temperature. Samples were then subjected to SDS/15% PAGE and immunoblotting procedures. At 22°C, both avidin and re-avidin formed characteristic aggregates in the presence of SDS, which remained in the stacking gel (not shown). Note that biotin-free samples of both avidin and re-avidin are dissociated into monomers at 80°C, whereas biotin-complexed samples are less prone to dissociate into the monomers at this temperature. This demonstrates that biotin stabilizes re-avidin, similar to its stabilization of the egg-white protein. M, low-range molecular weight standard; B–, without biotin; B+, with biotin; empty lane.

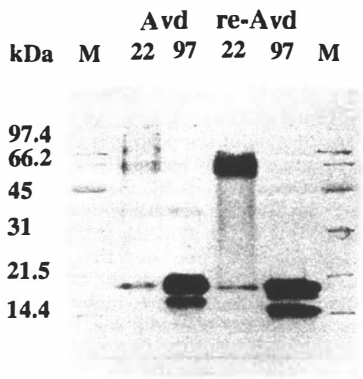



**FIG. 5.** Quaternary state of re-avidin. Samples of egg-white avidin (Avd) and re-avidin (re-Avd) were acetylated using acetyl *N*-hydroxysuccinimide ester. The samples were then treated for 20 min at the indicated temperature prior to SDS/15% PAGE and immunoblotting. Note that at 22°C the commercial preparation of avidin contains oligo-tetrameric forms, whereas re-avidin consists mainly of the tetramer. At 97°C, both samples show both glycosylated and unglycosylated forms of avidin. M, low-range molecular weight standard.

not all, commercial preparations of avidin, and is considered to reflect an artifact of the purification procedure. In any event, these data indicate that, similar to native avidin, re-avidin assembles into a stable tetramer in solution.

In conclusion, the baculovirus expression system was shown to offer an efficient way to produce biologically active and structurally viable re-avidin. The availability of glycosylated and fully active re-avidin in this system will enable further studies on the importance of carbohydrates in the correct folding of subunits and their possible role in the assembly of the tetramer. This system might be particularly appropriate for studying whether basic amino acids interfere with folding in the absence of carbohydrates. Finally, unlike the *E. coli* expression system in which the recombinant proteins usually appear in inclusion bodies, the proteins in the insect cell system are either commonly secreted or in a readily solubilized form. Our studies may thus enable more efficient production of active forms of mutagenized avidins and avidin fusion proteins for biotechnological purposes.

## ACKNOWLEDGMENTS

We thank Irene Helkala, Anne Ettala, Arja Mansikkaviita, Raija Vassinen, and Paavo Niutanen and Tikva Kulik for first-class technical assistance. This work was supported by grants from the Developmental Foundation of the Province of Central Finland, Jyväskylä, Finland, and from the Ministry of Education and the Academy of Finland, Helsinki, Finland, as well as from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.

## REFERENCES

1. Green, N. M. (1975) Avidin. *Adv. Prot. Chem.* **29,** 85–133.
2. Chaiet, L., and Wolf, F. J. (1964) The properties of streptavidin, a biotin-binding protein produced by *Streptomyces. Arch. Biochem. Biophys.* **106,** 1–5.
3. 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.
4. Bruch, R. C., and White, H. B. (1982) Compositional and structural heterogeneity of avidin glycopeptides. *Biochemistry* **21,** 5334–5341.
5. Hiller, Y., Gershoni, J. M., Bayer, E. A., and Wilchek, M. (1987) Biotin binding to avidin: Oligosaccharide side chain not required for ligand association. *Biochem. J.* **248,** 167–171.
6. Gope, M. L., Keinänen, R. A., Kristo P. A., Conneely, O. M., Beattie, 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.
7. Wallén, M. J., Laukkanen, M. O., and Kulomaa, M. S. (1995) Cloning and sequencing the chicken egg-white avidin-encoding gene and its relationship with the avidin-related genes *Avr1-Avr5. Gene* **161,** 205–209.
8. Livnah, O., Bayer, E. A., Wilchek, M., and Sussman, J. L. (1993) Three-dimensional structures of avidin and the avidin-biotin complex. *Proc. Natl. Acad. Sci. USA* **90,** 5076–5080.
9. Pugliese, L., Coda, A., Malcovati, M., and Bolognesi, M. J. (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.
10. White, H. B., III, Orth, W. H., III, Schreiber, R. W., Jr., and Whitehead, C. C. (1992) Availability of avidin-bound biotin to chicken embryo. *Arch. Biochem. Biophys.* **298,** 80–83.
11. Wilchek, M., and Bayer, E. A. (1990) Avidin-biotin technology. *Methods Enzymol.* **184,** 14–45.
12. Sano, T., and Cantor, C. R. (1990) Expression of a cloned streptavidin gene in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **87,** 142–146.
13. Nagarajan, V., Ramaley, R., Albertson, H., and Chen, M. (1993) Secretion of streptavidin from *Bacillus subtilis. Appl. Environ. Microbiol.* **59,** 3894–3898.
14. Airenne, K. J., Sarkkinen, P., Punnonen, E-L. and Kulomaa, M. S. (1994) Production of recombinant avidin in *Escherichia coli. Gene* **144,** 75–80.
15. Luckow, V. A., and Summers, M. D. (1988) Trends in the development of baculovirus expression vectors. *Biotechnology* **6,** 47–55.
16. O'Reilly, D. R., Miller, L. K., and Luckov, V. A. (1994) "Baculovirus Expression Vectors. A Laboratory Manual," Oxford University Press, New York.
17. Airenne, K. J., and Kulomaa, M. S. (1995) Rapid purification of recombinant proteins fused to chicken avidin. *Gene,* **167,** 63–68.
18. Barnes, W. M. (1987) Sequencing DNA with dideoxyribonucleotides as chain terminators: Hints and strategies for big projects. *Methods Enzymol.* **152,** 538–556.
19. Heery, D. M., Gannon, F., and Powell, R. (1990) A simple method for subcloning DNA fragments from gel slices. *TIG* **6,** 173.
20. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.
21. Joensuu, T., Tuohimaa, P., and Vilja, P. (1991) Avidin and ovalbumin induction by progesterone in chicken oviduct detected by

sensitive immunoenzymometric assays. *J. Endocrinol.* **130,** 191–197.

22. 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.
23. Durance, T. D., and Wong, N. S. (1992) Kinetics of thermal inactivation of avidin. *Food Res. Int.* **25,** 89–92.
24. Donovan, J. W., and 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* **12,** 512–517.
25. Marjomäki, V. S., Huovila, A. J., Surkka, M. A., Jokinen, I., and Salminen, A. (1990) Lysosomal trafficking in rat cardiac myocytes. *J. Histochem. Cytochem.* **38,** 1155–1164.
26. Luckow, V. A., Lee, S. C., Barry, G. F., and Olins, P. O. (1993) Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Esherichia coli. J. Virol.* **67,** 4566–4579.
27. Bayer, E. A., Ben-Hur, H., Gitlin, G., and Wilchek, M. (1986) An improved method for the single-step purification of streptavidin. *J. Biochem. Biophys. Methods* **13,** 103–112.
28. Bayer, E. A., De Meester, F., Kulik, T., and Wilchek, M. (1995) Deglycosylation of egg-white avidin. *Appl. Biochem. Biotechnol.* **53,** 1–9.
29. James, D. C., Freedman, R. B., Hoare, M., Ogonah, O. W., Rooney, B. C., Larionov, O. A., Dobrovolsky, V. N., Lagutin, O. V., and Jenkins, N. (1995) N-Glycosylation of recombinant human interferon-γ produced in different animal expression systems. *Biotechnology* **13,** 592–596.
30. Aeed, A. P., and Elhammer, Å. P. (1994) Glycosylation of recombinant prorenin in insect cells: The insect line Sf9 does not express the mannnose 6-phosphate recognition signal. *Biochemistry* **33,** 8793–8797.
31. Davidson, D. J., and Castellino, F. J. (1991) Asparagine-linked oligosaccharide processing in lepidopteran insect cells. Temporal dependence of the nature of the oligosaccharides assembled on asparagine-289 of recombinant human plasminogen produced in baculovirus vector infected *Spodoptera frugiperda* (IPLB-SF-21AE) cells. *Biochemistry* **30,** 6167–6174.
32. Jarvis, D. L., and Finn, E. E. (1995) Biochemical analysis of the N-glycosylation pathway in baculovirus-infected lepidopteran insect cells. *Virology* **212,** 500–511.

# IV

**Avidin provides an efficient tag for fusion protein production in baculovirus-infected insect cells**

by

Airenne, K. J., Laitinen, O. H., Alenius, H., Mikkola, J., Palosuo, T, Kalkkinen, N., Arif, S. A. M., Yeang, H. Y. & Kulomaa, M. S.

Biotechniques (submitted)

https://doi.org/10.1006/prep.1999.1123