

# Efficient Expression and Mutation of Avidin and Streptavidin as Host Proteins for Enantioselective Catalysis

Andrea Zocchi<sup>a\*</sup>, Nicolas Humbert, Temugin Berta, and Thomas R. Ward<sup>b\*</sup>

**Abstract:** Avidin and structurally related streptavidin are attractive host proteins for the creation of artificial metalloenzymes displaying features reminiscent both of homogeneous catalysts and enzymes. The main advantages are that both proteins have been cloned and expressed in several organisms and possess a deep hydrophobic binding pocket capable of hosting biotinylated catalyst precursors. An optimized artificial avidin gene in *Pichia pastoris* is expressed. The high level of active protein produced in the extracellular medium is suitable for the performance of high-throughput screening in 96-well plate format. Biologically active recombinant streptavidin is expressed in *E. coli*. Mutations have been introduced both in avidin and streptavidin genes and both wild type and mutated proteins have been utilized to explore the role of the second coordination sphere in enantioselective catalysis.

**Keywords:** Avidin · *Escherichia coli* · *Pichia pastoris* · Site-directed and random mutagenesis · Streptavidin

As outlined in the preceding article [1], the research in the Ward group is centered on the incorporation of active catalyst precursors in a well-defined protein environment to afford selective catalysts. Rather than covalent anchoring, the biotin-avidin technology is a straightforward method that ensures precise localization of a biotinylated moiety without the need for a chemical coupling step between the host protein and the organometallic moiety. For synthetic organometallic chemists, avidin and structurally related streptavidin are attractive host proteins for several reasons:

- i) Derivatization of the valeric acid side chain of biotin does not significantly affect the stability constant of the biotin-(strept) avidin couple ( $K_a = \sim 10^{14} \text{ M}^{-1}$ ) [2].
- ii) The proteins have been cloned and expressed in several organisms.

- iii) The proteins are easy to purify by affinity chromatography on an iminobiotin resin [3].
- iv) The proteins are very robust (temperature, pH, organic solvents, etc.) [2].
- v) The proteins possess a deep hydrophobic binding pocket capable of hosting biotinylated catalyst precursors [4][5].
- vi) The proteins are easily quantifiable by titration with fluorescent biotinylated probes [6].
- vii) The protein can be immobilized and/or recycled.

Because of the prohibitive cost of purchasing either avidin or streptavidin, we set out to express these host proteins. Two major issues were addressed in this context: the development of efficient expression systems and the generation of protein diversity (*i.e.* mutagenesis).

## Expression Systems

The modern techniques of molecular biology, the availability of genes and the increasing knowledge of cellular mechanisms allow the routine utilization of several living systems for the expression of exogenous proteins.

**Avidin** is a glycoprotein found in avian, reptilian and amphibian egg white, which

most likely plays a role as a host defense factor [7][8]. The active form of hen-avidin is a homotetramer composed of four singly glycosylated subunits, containing one single intramonomeric disulfide bridge. Each monomer can bind a (+)-biotin (vitamin H) molecule with an extraordinary affinity ( $K_a = \sim 10^{14} \text{ M}^{-1}$ ) [7]. Each subunit consists of 128 amino acids and its primary sequence was determined in 1971 [9]. The cDNA of the chicken oviduct avidin gene was documented by Gope *et al.* [10] and a genomic clone was isolated by Keinänen *et al.* [11]. The primary source for the commercial production of avidin has been extraction and purification from chicken egg white, but biologically active recombinant isoforms have been produced in several expression systems, including *E. coli* [12] [13], baculovirus-infected cells [14] and maize [15]. The methylotrophic yeast *Pichia pastoris* offers an attractive alternative to the above expression systems. This expression system combines the advantages of microbial and eukaryotic organisms, namely the formation of disulfide bridges, the glycosylation and the high-density cell fermentation [16].

We expressed an optimized artificial avidin gene in *Pichia pastoris*. The expression vector is integrated into the yeast genome. The plasmid we chose drives the

\*Correspondence: Dr. A. Zocchi<sup>a</sup>, Prof. T.R. Ward<sup>b</sup>  
Institute of Chemistry  
University of Neuchâtel  
Av. Bellevaux 51  
Case Postale 2  
CH-2007 Neuchâtel  
<sup>a</sup>Tel.: +41 32 718 2734  
Fax: +41 32 718 2511  
E-Mail: andrea.zocchi@unine.ch  
<sup>b</sup>E-Mail: thomas.ward@unine.ch

extracellular secretion of the recombinant protein. After transformation, high-level expressing clones are roughly selected growing transformants on nitrocellulose filters. The filters are utilized for the dot blot procedure [17] and colonies that stain most are utilized for further expression.

The recombinant avidin is efficiently produced both in flasks at a concentration of ~40 mg/l and in a high-cell density fermentation at a concentration of ~350 mg/l [18]. In this system, the purification of recombinant avidin only requires the concentration of the culture medium and its purification *via* an affinity chromatography on an iminobiotin column.

**Streptavidin** is the prokaryotic counterpart of avidin, with the same tridimensional eight-stranded  $\beta$ -barrel structure, but only 32% protein sequence homology [8]. It is produced by *Streptomyces avidinii* in its the extracellular medium and has antimicrobial activity [19]. As for avidin, the active form of streptavidin is a homotetramer. Each of subunit can bind a (+)-biotin molecule with almost the same affinity as avidin [8][19]. Each monomer consists of 159 amino acids and the gene of streptavidin was isolated by Argaraña *et al.* [20]. The primary source for the commercial production of streptavidin has been the isolation from cultures of *Streptomyces avidinii* [21] or the expression in recombinant *E. coli* [22]. Biologically active recombinant streptavidin has been also secreted in the extracellular medium in the *Bacillus subtilis* expression system [22].

The streptavidin gene is expressed in *E. coli* both in shaking flasks and in a 15 l

fermentor. The episomal vectors target the protein into the cytoplasm. This expression system was optimized in our group. The multi-step purification involves a lysis of the cells by heat shock, an enzymatic digestion of nucleic acids and a denaturing-renaturing step prior to affinity purification of the active protein. Under optimized conditions, we routinely obtain a concentration of active streptavidin of up to 120 mg/l of fermenting culture.

The expression of avidin in this prokaryotic system (*i.e.* *E. coli*) is more problematic, essentially because it is produced in inclusion bodies [13]. The purification from inclusion bodies requires an obligatory denaturing-renaturing step and avidin is not prone to renaturing [13].

### Mutagenesis

In order to introduce variability in either avidin or streptavidin, four different approaches were investigated: i) site-directed mutagenesis, ii) cassette mutagenesis, iii) DNA shuffling and iv) artificial gene synthesis (Table).

(i) The **site-directed mutagenesis** is utilized for direct deletion, insertion or modification of sequences. The procedure, based on PCR, involves two normal primers located at the beginning and the end of the target sequence, and one other longer primer, complementary to the sequence portions flanking the zone to modify and bearing the mutation (mutagenic primer) (Scheme). All the PCRs required in this technique are performed with a proof-reading DNA polymerase. Two PCRs are required. For the

first PCR, an external normal primer and a mutagenic internal primer are used. After purification of the product, a second PCR is performed. This PCR contains as template the unmodified target sequence, the product of the previous PCR acting as megaprimer, and one of the external normal primers. The result is a complete functional sequence containing the desired mutation [24]. A variant of this technique is performed on the whole expression vector containing the protein gene. Starting from the site in which the mutation is introduced, two opposite primers direct a PCR that resynthesizes the whole plasmid. This variant usually is carried out with the 'Quickchange Mutagenesis Kit' (Stratagene) and is considerably less time-consuming.

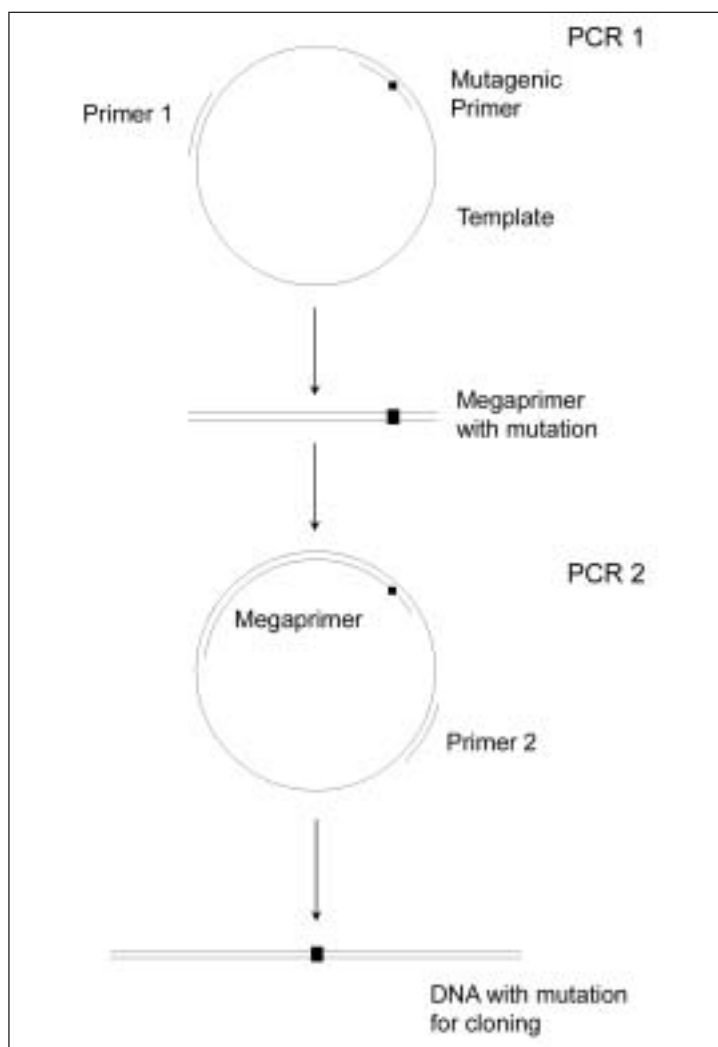
Site-directed mutagenesis is essentially utilized in the *E. coli*-streptavidin system. The mutations (single or multiple) are located mainly in loop connecting  $\beta$ -sheets around the biotin-binding pocket. In most cases, the original amino acid was exchanged with a glycine, to confer more flexibility to the loop.

ii) The **cassette mutagenesis** is utilized to completely change one specific region of a gene. The gene is roughly subdivided according the position of functional or structural domains.

Random or site-directed mutations can be introduced in one or more of these partial sequences (cassettes). Their extremities maintain a high homology with the original gene. One or more cassettes are introduced into the gene by homologous recombination or by cloning utilizing restriction sites present in the original sequence. This method requires a stringent selection of

Table. Summary of the mutagenesis methods utilized for the generation of (strept)avidin diversity.

Mutagenesis method	Template	DNA treatment	Remarks
Site Directed Mutagenesis	Plasmid	Quick change procedure (Stratagene protocol)	Sequencing required
Cassette Mutagenesis	Linear DNA or plasmid	Excision of sequences corresponding to functional protein regions + Site-directed mutagenesis or erratic PCR + Cloning of PCR product	High throughput screen or sequencing required
DNA shuffling	Linear DNA or plasmid	Random cut of sequences (DNase I) + Mutagenic and recombinogenic PCR + Cloning of PCR products	High throughput screen required
Gene synthesis	Synthesis of coding fragments, based on back-translation of amino acid sequence	Ligation of coding fragments + Amplification (PCR) + Cloning of PCR product	Sequencing required



Scheme. General scheme of the site-directed mutagenesis procedure for streptavidin

expressing organism (<http://www.kazusa.or.jp/codon>). The resulting coding strand is subdivided into several fragments. The length of these fragments along the gene is chosen to minimize the formation of secondary structures. Shorter oligomers, on the antisense strand, are synthesized across couples of adjacent coding fragments. These oligomers have the function to align the coding fragments in the right order. The coding fragments are subsequently covalently ligated to give a full-length coding strand. A high-fidelity PCR is performed to synthesize the antisense strand and obtain enough material for sequencing and cloning. This method was utilized to test a hypothesis concerning the factors responsible for enantioselection. Indeed, streptavidin proved to be far a better chiral inducer than avidin in the vast majority of experiments. By way of example, the best enantiomeric excess (ee) for the reduction of acetamidoacrylic acid with mutant streptavidin as host protein affords acetamidoalanine with 96% ee (*R*) [28]. The best ee with mutant avidin for the same substrate was only 69% (*S*) [29]. Analysis of the quaternary structure of both proteins reveals that the main structural difference arises from loop 5-6 which is in close proximity to the catalytic event (Fig.). We thus set out to synthesize a chimeric gene, by incorporating the 5-6 loop of streptavidin in the avidin gene. The resulting chimeric avidin was expressed in baffled flasks, yielding after purification, 10 mg/l of functional chimeric avidin.

## Outlook

By optimizing both the expression levels and the mutagenesis methods, we have been able to generate diversity in host proteins for enantioselective catalysis. By combining the chemical diversity of the biotinylated catalysts with the genetic diversity of the protein, it is possible to produce artificial metalloenzymes which are amenable to optimization by high throughput screening methods.

functional mutated protein [25]. For both avidin and streptavidin, the purification step on an iminobiotin column fulfills this selection step as only those proteins that bind to iminobiotin are purified, isolated and subsequently screened for catalytic activity in the presence of a biotinylated catalyst precursor.

iii) The *DNA shuffling* technique [26] [27] is utilized to introduce random mutations throughout the avidin or streptavidin gene. The method implies a random fragmentation of the gene or of the genes to be modified through digestion with an unspecific endonuclease. Re-assembly PCR is performed combining the fragment isolated above. In the case of mixing streptavidin with avidin, the genes of which display only 30% homology, the first 10–15 cycles of the PCR are performed manually, alternating a fast freezing of the denatured DNA with annealing, elongation and denaturation. After this *in vitro* recombination, an aliquot of this PCR is utilized to perform another normal PCR with primer designed to insert appropriate restriction sites allowing the creation of a mutant library.

An important step of this procedure is the screening of the mutant library. In this phase, functional proteins with new or improved characteristics are selected for the following cycle of shuffling, starting again with the random cut. It is clear that in the case of this technique the screening step must be a high throughput step. For this purpose, a 96 well-plate fluorescence quenching titration with biotin-4-fluorescein allows the binding properties and the activity towards biotin of the mutant proteins to be assessed. It is also possible to perform the shuffling protocol *in vivo*, exploiting the promptness of *Pichia* to recombination [28]. In this case random fragments are introduced into yeast cell already containing a gene homologous with the fragments.

iv) The *synthesis of an artificial gene* is utilized to optimize the expression of streptavidin or avidin in a given expression system. For this purpose, the amino acid sequence of the gene of interest is back-translated and the degenerate codons are selected according to the frequency of the triplet per amino acid, found in the chosen



Fig. a) Superposition of avidin (light) and streptavidin (dark) monomers emphasizing their structural similarity; b) Zoom on the 5-6 loop (Biotin: ball and stick).

### Acknowledgements

This work was supported by the Swiss National Science Foundation (Grants 620-57866.99 and 4047-057532) as well as the Canton of Neuchâtel. We thank Prof. J.-M. Neuhaus and P. Schürmann for their support for this work. We thank Belovo Egg Science and Technology for a generous gift of egg white avidin.

We thank Prof. C.R. Cantor for the gift of the streptavidin cDNA and Prof. P. Arosio and Dr. P. Santambrogio (San Raffaele Scientific Institute, Milan, Italy), for their valuable advice in streptavidin expression. We acknowledge Dr. Z. Lei and Prof. J.-L. Reymond (Chemistry Department, University of Bern) for their advice in site-directed mutagenesis.

Received: July 14, 2003

- [1] T.R. Ward, J. Collot, J. Gradinaru, A. Loosli, M. Skander, C. Letondor, E. Joseph, G. Klein, *Chimia* **2003**, *57*, 586.
- [2] N.M. Green, 'Avidin and Streptavidin', in 'Avidin-Biotin Technology', Eds. M. Wilchek, E.A. Bayer, Academic Press, San Diego, **1990**, p. 51.
- [3] B. Fudem-Goldin, G.A. Orr, *Methods in Enzymology* **1990**, *184*, 167.
- [4] M.E. Wilson, G.M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306.
- [5] C.C. Lin, C.W. Lin, A.S.C. Chan, *Tetrahedron: Asymmetry* **1999**, *10*, 1887.
- [6] H.J. Gruber, G. Kada, M. Marek, K. Kaiser, *Biochim. Biophys. Acta* **1998**, *1381*, 203.
- [7] J. Korpela, *Med. Biol.* **1984**, *62*, 5.
- [8] N.M. Green, *Adv. Prot. Chem.* **1975**, *29*, 85.
- [9] R.J. De Lange, T.S. Huang, *J. Biol. Chem.* **1971**, *246*, 698.
- [10] M.L. Gope, R.A. Keinanen, P.A. Kristo, O.M. Conneely, W.G. Beattie, T. Zarucki-Schulz, B.W. O'Malley, M.S. Kulomaa, *Nucl. Ac. Res.* **1987**, *15*, 3595.
- [11] R.A. Keinanen, M.-L. Laukkanen, M.S. Kulomaa, *J. Ster. Biochem. Mol. Biol.* **1988**, *30*, 17.
- [12] K.J. Airene, P. Sarkkinen, E.-L. Punnonen, M.S. Kulomaa, *Gene* **1994**, *144*, 75.
- [13] E. Nardone, C. Rosano, P. Santambrogio, F. Cumis, A. Corti, F. Magni, A.G. Siccardi, G. Paganelli, R. Losso, B. Aprea, M. Bolognesi, A. Sidoli, P. Arosio, *Eur. J. Biochem.* **1998**, *256*, 453.
- [14] K.J. Airene, C. Oker-Blom, V.S. Marjomäki, E.A. Bayer, M. Wilchek, M.S. Kulomaa, *Protein Express. Purif.* **1997**, *9*, 100.
- [15] E.E. Hood, D.R. Witcher, S. Maddock, T. Meyer, C. Baszczyński, M. Bailei, P. Flynn, J. Register, L. Marshall, D. Bond, E. Kulisek, A. Kusnadi, R. Evangelista, Z. Nikolov, C. Wooge, R.J. Mehig, R. Hernan, W.K. Kappel, D. Ritland, C.P. Li, J.A. Howard, *Mol. Breeding* **1997**, *3*, 291.
- [16] C.A. Scorer, J.J. Clare, W.R. McCombie, M.A. Romanos, K. Sreerishna, *Biotechnology* **1994**, *12*, 181.
- [17] J. Sambrook, D.W. Russel, 'Molecular Cloning. A laboratory manual', Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA **2001**.
- [18] A. Zocchi, A.M. Jobé, J.-M. Neuhaus, T. Ward, *Protein Express. Purif.* **2003**, in press.
- [19] L. Chaiet, F.J. Wolf, *Archives of Biochemistry and Biophysics* **1964**, *106*, 1.
- [20] C.E. Argaraña, I.D. Kuntz, S. Birken, R. Axel, C.R. Cantor, *Nucl. Ac. Res.* **1986**, *14*, 1871.
- [21] M. Suter, J. Cazin, Jr, J.E. Butler, D.M. Mock, *J. of Immunol. Meth.* **1988**, *113*, 83.
- [22] T. Sano, C.R. Cantor, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 142.
- [23] V. Nagarajan, R. Ramaley, H. Albertson, M. Chen, *Applied and Environmental Microbiology* **1993**, *59*, 3894.
- [24] J. Brøns-Poulsen, N.E. Petersen, M. Hør-

der, K. Kristiansen, *Molecular and Cellular Probes* **1998**, *12*, 345.

- [25] M. Kammann, J. Laufs, J. Schell, B. Gronenborn, *Nucl. Ac. Res.* **1989**, *17*, 5404.
- [26] W.P.C. Stemmer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747.
- [27] A. Crameri, S.-A. Raillard, E. Bermudez, W.P.C. Stemmer, *Nature* **1998**, *391*, 288.
- [28] V. Abécassis, D. Pompon, G. Truan, *Nucl. Ac. Res.* **2000**, *28*, e88.
- [29] J. Collot, J. Gradinaru, N. Humbert, M. Skander, A. Zocchi, T. Ward, *J. Am. Chem. Soc.* **2003**, *125*, 9030.