

This item is the archived peer-reviewed author-version of:

Exploring three different expression systems for recombinant expression of globins: *Escherichia coli*, *Pichia pastoris* and *Spodoptera frugiperda*

Reference:

Bracke An, Hoogewijs David, Dewilde Sylvia.- Exploring three different expression systems for recombinant expression of globins: *Escherichia coli*, *Pichia pastoris* and *Spodoptera frugiperda*

Analytical biochemistry - ISSN 0003-2697 - 543(2018), p. 62-70

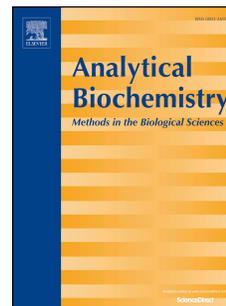
Full text (Publisher's DOI): <https://doi.org/10.1016/J.AB.2017.11.027>

To cite this reference: <https://hdl.handle.net/10067/1499190151162165141>

Accepted Manuscript

Exploring three different expression systems for recombinant expression of globins:
Escherichia coli, *Pichia pastoris* and *Spodoptera frugiperda*

An Bracke, David Hoogewijs, Sylvia Dewilde



PII: S0003-2697(17)30493-1

DOI: [10.1016/j.ab.2017.11.027](https://doi.org/10.1016/j.ab.2017.11.027)

Reference: YABIO 12858

To appear in: *Analytical Biochemistry*

Received Date: 18 September 2017

Revised Date: 29 November 2017

Accepted Date: 30 November 2017

Please cite this article as: A. Bracke, D. Hoogewijs, S. Dewilde, Exploring three different expression systems for recombinant expression of globins: *Escherichia coli*, *Pichia pastoris* and *Spodoptera frugiperda*, *Analytical Biochemistry* (2018), doi: 10.1016/j.ab.2017.11.027.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Exploring three different expression systems for recombinant expression of globins: *Escherichia coli*, *Pichia pastoris* and *Spodoptera frugiperda*

Application to human neuroglobin and androglobin

An Bracke¹, David Hoogewijs², Sylvia Dewilde¹

¹Department of Biomedical Sciences, University of Antwerp, Universiteitsplein 1, Antwerp 2610, Belgium

²Department of Medicine/Physiology, University of Fribourg, Chemin du Musée 5, CH 1700 Fribourg, Switzerland

Corresponding author: Sylvia Dewilde, sylvia.dewilde@uantwerpen.be

Highlights

- Three different systems were explored for the recombinant expression of globins
- Neuroglobin was successfully expressed in these three expression systems
- The full length androglobin was stably expressed in *Spodoptera frugiperda*
- Full protein context is necessary for a stably folded globin domain of androglobin

Abstract

Globins are among the best investigated proteins in biological and medical sciences and represent a prime tool for the study of the evolution of genes and the structure-function relationship of proteins. Here, we explore the recombinant expression of globins in three different expression systems: *Escherichia coli*, *Pichia pastoris* and the baculovirus infected *Spodoptera frugiperda*. We expressed two different human globin types in these three expression systems: I) the well-characterized neuroglobin and II) the uncharacterized, circularly permuted globin domain of the large chimeric globin androglobin. It is clear from the literature that *E. coli* is the most used expression system for expression and purification of recombinant globins. However, the major disadvantage of *E. coli* is the formation of insoluble aggregates. We experienced that, for more complex multi-domain globins, like the chimeric globin androglobin, it is recommended to switch to a higher eukaryotic expression system.

Key words: globin, protein expression, androglobin, neuroglobin

1 Introduction

Globins can be defined as small heme-containing proteins which can mostly reversibly bind oxygen. They occur in the three kingdoms of life: Archaea, Prokaryotes and Eukaryotes and exist as single domain or chimeric proteins [1,2]. The best known single domain globins are the vertebrate hemoglobin (Hb) and myoglobin (Mb), playing respectively a role in oxygen transport and storage. Examples of chimeric globins are the flavohemoglobins (FHbs) in *Escherichia coli* and yeast, consisting of an N-terminal globin domain coupled to a ferredoxin reductase-like domain [3], and the globin coupled sensors (GCS), reported in Bacteria and Archaea, comprising an N-terminal globin domain linked to variable gene regulatory domains [4]. FHbs and GCS do not occur in vertebrates. The recent and rapid accumulation of genomic information has resulted in a substantial increase in newly recognized globins. Bioinformatic genome surveys revealed that more than 50% of the bacterial, approximately 20% of the archaeal and almost 90% of the eukaryote genomes contain genes encoding globins. [1,5–7] Despite their great variety in function and amino acid sequence, most of them share a similar three-dimensional structure, referred to as the globin fold, characterized by a three-over-three alpha helical folding. Few globins, however, display a two-over-two alpha helical folding and are called ‘truncated globins’. [8]

Globins are among the most investigated proteins in biological and medical sciences and represent a major tool for the study of gene evolution and structure-function relationship of proteins. Sperm whale Mb was the first protein of which the three-dimensional structure was revealed by X-ray crystallography [9]. The structure of human Hb followed five years later [10]. In the early times of protein crystallization and structure determination, the investigated proteins were purified from their natural context, e.g. Mb out of sperm whale muscle tissue and Hb out of human erythrocytes. However, under normal physiological conditions, most proteins are present in low cellular concentrations and cannot be purified out of the tissue itself. Since the early eighties it is has been possible to express proteins recombinantly in heterologous host cells like the bacterium *Escherichia coli* (*E.coli*) [11,12] or the yeast *Saccharomyces cerevisiae* (*S.cerevisiae*) [13]. These host cells can produce a large amount of recombinant protein, that enables the characterization of proteins with a low cellular abundance. In addition, this technique facilitates the study of protein-function relationships by site-directed mutagenesis. This revolution in protein biochemistry caused an exponential increase in knowledge on the structure and function of proteins in general.

Neuroglobin (Ngb) was first described in 2000 as a vertebrate globin, preferentially expressed in brain and other nerve tissue [14]. The estimated amount of Ngb in the vertebrate brain under non-pathological conditions is only in the micromolar range. It was possible to perform a biochemical and

structural characterization of NGB subsequent to the recombinant expression of the protein in *E.coli* [15].

Androglobin (Adgb) was described in 2012 as the fifth mammalian globin type (in addition to Hb, Mb, Ngb and cytoglobin (Cygb)) and it displays predominant expression in testis tissue in mice and human [16]. Adgb characteristically comprises four domains, an N-terminal ~350 residues calpain-like cysteine protease domain, a region of ~300 residues without known motifs/domains, followed by an ~150 residue circularly permuted globin domain, and a second, uncharacterized ~750 residues C-terminal domain. Interestingly, the globin domain, which normally consists of eight consecutive α -helices (named A-H), is circularly permuted and split into two parts within Adgb. The part containing helices A and B has been shifted in the C-terminal direction and is separated from the main globin sequence (helices C–H) by a calmodulin-binding IQ motif. Alignments of the split Adgb globin domain with mammalian Mb, Ngb, and Cygb sequences revealed that Adgb—despite its rearrangement—conforms to the criteria of the “globin fold” tertiary structure [16]. This was confirmed by molecular modelling of the human Adgb (ADGB) globin domain 3D structure, showing that the helix C–H segment alone is able to produce a *bona fide* globin fold [16]. *In vitro* structural and functional analysis of ADGB is necessary to better understand its molecular function during spermatogenesis, for this a recombinant form of ADGB is required.

In this paper, we explore three different expression systems for the expression of globins: a bacterial (*E.coli*), a yeast (*Pichia pastoris* (*P.pastoris*)) and an insect cell expression system (baculovirus infected *Sf9*). These three expression systems are widely used and in previous studies they have been compared for their recombinant expression of e.g. G-protein-coupled receptors [17], endostatin [18] and rabbit liver carboxylesterases [19]. We investigate their advantages and disadvantages for the expression of globins in general. Furthermore, we used these three different expression systems to express the two different human globin types that were described above: I) the single domain globin neuroglobin (NGB), which has already been biochemically and structurally characterized and II) the uncharacterized chimeric globin ADGB.

2 Material and methods

2.1 Construction of the recombinant expression vectors

The cDNA sequence of the globin domain of human ADGB (referred to as CH-IQ-AB) and human NGB were cloned into the recombinant expression vectors pET23a (*E.coli*), pPicZ α (*P.pastoris*) and

pDEST10 (*Spodoptera frugiperda* 9 (Sf9)). The cDNA of full length ADGB was only cloned into the pDEST10 vector (Table1). Full length ADGB was RT-PCR amplified, with NcoI and XbaI overhangs (Table2) from human testis cDNA and cloned into the pENTR4 vector. The CH-IQ-AB domain was PCR-amplified for further cloning from the pENTR4 ADGB vector. The NGB domain was PCR-amplified from pET3a NGB construct, which was already available in our laboratory [15]. For the cloning of CH-IQ-AB in the pPicZ α vector, the cDNA of the fragment was synthesized by Invitrogen GeneArt, with optimized codons for expression in *P.pastoris*.

Table 1: Overview of the different expression vectors used in this paper

Vector	Tag	Promotor	Expression host
pET23a	6xHis, C-terminal	T7	<i>Escherichia coli</i> BL21(DE3)pLysS
pPicZ α	6xHis, C-terminal	AOX	<i>Pichia pastoris</i> X33
pDEST10	6xHis, N-terminal	Polyhedrin	<i>Spodoptera frugiperda</i> Sf9

The amplifications of the fragments coding for the CH-IQ-AB and NGB were performed by PCR. PCR amplification was performed as follows: initial denaturation at 94 °C for 5 min, followed by 35 cycles (94 °C, 1 min; 55 °C, 1 min; 72 °C, 1min) and a final elongation at 72 °C for 10 min. The forward and reverse primers' sequences, including restriction sites/attB recombination sites, were in house designed and are listed in Table2. Recombinant pET23a CH-IQ-AB, pET23a NGB, pPicZ α CH-IQ-AB and pPicZ α NGB were created using standard restriction enzyme cloning. Two mutations were made in the pET23a CH-IQ-AB construct resulting in the replacement of two cysteines (on position C3 and E2) using the QuikChange™ site-directed mutagenesis method (Stratagene). The recombinant pDEST10 CH-IQ-AB and pDEST10 NGB vector were created using the Gateway cloning technique: first a pENTR CH-IQ-AB and pENTR NGB vector were created by a BP recombination reaction between attB CH-IQ-AB/NGB PCR product and pDONR 221 vector and second the pDEST10 NGB and pDEST10 CH-IQ-AB were created by a LR recombination reaction between pENTR CH-IQ-AB/pENTR NGB vector and the pDEST10 vector. The pDEST10 ADGB (full length) was created using the Gateway cloning technique: LR recombination reaction between pENTR4 ADGB and the pDEST10 vector. The constructed recombinant expression vectors were verified via direct sequencing.

Table 2: Overview of the in-house designed primers for the cloning of the different recombinant expression vectors

Primer	Sequence	Restriction site / attB
ADGB pENTR4F	CATGCCATGGGCTCAGAGCTCAGCCCTACA	NcoI
ADGB pENTR4R	TGCTCTAGAGCATGGACATCCCCTGGTACTT	XbaI
NGB pET23a F	GGGAATTCATATGGAGCGCCCGAGCCCGAG	NdeI
NGB pET23a R	CCGCTCGAGCTCGCCATCCAGCCTCG	XhoI
CH-IQ-AB pET23a F	GGGAATTCATATGCACATCTGCAGC ATGGTG	NdeI
CH-IQ-AB pET23a R	CCGCTCGAGGCTTTTAAACATTAGTCT	XhoI
NGB pPicZα F	CCGCTCGAGAAAAGAGAGGCTGAAGCTATGGAGCGCCCGAGCCCGAGCTG	XhoI <u>KeX2 cleavage site</u>
NGB pPicZα R	CTAGTCTAGAGGCTCGCCATCCAGCCTCGAC	XbaI
CH-IQ-AB pPicZα R	CCG <u>CTCGAGAAAAGAGAGGCTGAAGCT</u> CACATCTGCAGC ATGGTG	XhoI <u>KeX2 cleavage site</u>
CH-IQ-AB pPicZα R	CTAGTCTAGAGCTTTTAAACATTAGTCT	XbaI
NGB pDONR F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAGCGCCCGAGCCCGAGCTG	attB1
NGB pDONR R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTACTCGCCATCCAGCCTCGACTCATG	attB2
CH-IQ-AB pDONR F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCACATCTGCAGCATGGTGTCAATTTG	attB1
CH-IQ-AB pDONR R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGCTTTTAAACATTAGTCTTAAG	attB2
CH-IQ-AB Cys C3 F	CATATGCACATCTCCAGCATGGTGTCA	Cys → Ser
CH-IQ-AB Cys C3 R	TGACACCATGCTGGAGATGTGCATATG	Cys → Ser
CH-IQ-AB Cys E2 F	GAACCAGAGAGCTCCCGATTACGGAA	Cys → Ser
CH-IQ-AB Cys E2 R	TTCCGTAATCGGGAGCTCTCTGGTTC	Cys → Ser

2.2 Expression in *Escherichia coli*

pPET23a CH-IQ-AB and pPET23a NGB were transformed into competent *E.coli* BL21(DE3)pLysS cells. Subsequently, cells were grown in a shaking incubator (160rpm) at 25°C in TB medium (1.2% bactotryptone, 2.4% yeast extract, 0.4% glycerol, 72mM potassium phosphate buffer, pH7.5) containing 200µg/ml ampicillin, 30µg/ml chloramphenicol, and 1mM δ-amino-levulinic acid. The expression of the C-terminal His-tagged fusion proteins was induced at $A_{550nm} = 0.8$ by the addition of isopropyl-1-thio-D-galactopyranoside (IPTG) to a final concentration of 0.4mM, and expression was continued overnight. The cells were harvested and resuspended in ~1/20 of the initial volume (resuspension buffer = 50mM Tris-HCl pH8.0, 300mM NaCl). The cells were then exposed to three freeze-thaw steps and were sonicated until completely lysed. The extract was clarified by centrifugation (12000xg, 10min, 4°C) and both supernatant and pellet fraction were screened for expression of recombinant protein (CH-IQ-AB, NGB) using SDS-PAGE (section 2.7).

2.3 Expression in *Pichia pastoris*

pPicZα CH-IQ-AB and pPicZα NGB were linearized with restriction endonuclease *SacI* and subsequently 5µg plasmid was transformed into *P.pastoris* X33 electrocompetent cells by electroporation: cells were pulsed using the manufacturer's instructions for *Saccharomyces cerevisiae* (Bio-rad electroporation system). Transformed cells were plated out on YPDS agar plates (1% yeast extract, 2% peptone, 2% dextrose, 1M sorbitol, 2% agar) containing 100µg/ml Zeocin and were incubated 4 days at 30°C. The Mut Phenotype was determined by streaking the colonies out on both MD (1,34% YNB, 4x10⁻⁵% biotin, 2% dextrose) and MM (1,34% YNB, 4x10⁻⁵% biotin, 0,5% methanol) agar plates containing 100µg/ml Zeocin. 8 Mut⁺ colonies, of which the presence of the insert was confirmed by PCR using the recommended sequencing primers, were inoculated in 25ml BMGY medium (1% yeast extract, 2% peptone, 100mM potassium phosphate pH6.0, 1.34% Yeast Nitrogen Base, 4x10⁻⁵% biotin, 1% glycerol) and grown overnight at 30°C in a shaking incubator (200 rpm) until they reached an OD₆₀₀ = 2-5. Subsequently, the cells were harvested by centrifugation (2000xg, 10min, RT) and the cell pellet was resuspended to an OD₆₀₀ of 1 in BMMY medium (1% yeast extract, 2% peptone, 100mM potassium phosphate pH6.0, 1.34% Yeast Nitrogen Base, 4x10⁻⁵% biotin, 0.5% methanol) containing 1mM δ-amino-levulinic acid and further grown at 30°C in a shaking incubator (200rpm). Methanol will induce the expression of the C-terminal His-tagged proteins, so every 24h 100% methanol was added to final concentration of 0.5% to maintain induction. 1ml of expression culture was collected every 24h to analyze expression levels with Western blot (see 2.7).

For this, the cells were centrifuged (4000xg, 20min, RT) ; the supernatant was screened for secreted expression of recombinant protein; the cell pellet was resuspended in SDS-loading buffer and screened for intracellular expression. After 96h all cells were harvested by centrifugation (4000xg, 20min, RT).

2.4 Expression in *Spodoptera frugiperda* 9 insect cells

pDEST10 CH-IQ-AB, pDEST10 NGB and pDEST10 ADGB were transformed into MAX Efficiency[®] DH10Bac[™] competent *E.coli* (ThermoScientific[™]) to generate recombinant baculovirus DNA. The presence of the insert in the baculovirus DNA was verified by PCR using the recommended sequencing primers. 1µg of pure recombinant baculovirus DNA was subsequently transfected into *Sf9* insect cells (ThermoScientific[™]) using Cellfectin[®] II reagent (ThermoScientific[™]) following the manufacturers guidelines. *Sf9* cells were maintained and passaged adherently in cell culture flasks at 28°C with loose caps to allow gas exchange, the cells were cultured in Grace's Insect Cell Culture Medium (ThermoScientific[™]), supplemented with 10% fetal bovine serum (FBS) and 5% Penicillin/Streptomycin. After transfection, cells were incubated at 28°C for 72 hours until signs of viral infections could be seen. Recombinant virus particles were harvested by collecting the medium. The recombinant baculoviral stock was amplified three times (P3). Finally, *Sf9* cell cultures (Grace's Insect Cell medium supplemented with 10% FBS, 5% Penicillin/Streptomycin, 1mM δ-amino-levulinic acid, 14µM E64, 0.15µM Pepstatin A, 4µM CalpainII Inhibitor) were infected with P3 recombinant virus particles and were incubated for 72h at 28°C. Proteins were extracted by resuspending cells in native lysis buffer (25 mM sodium phosphate (pH 7.0), 150 mM NaCl, 1% NP40), followed by centrifugation (12000g, 20min, 4°C). Expression was screened using Western blot (see 2.7).

2.5 Purification

2.5.1 Purification NGB

20mM imidazole was added to the clarified cell lysates of *E.coli* or to the cell medium of *P.pastoris*, prior to loading onto a Ni Sepharose[™] High Performance column (GE Healthcare), pre-equilibrated with equilibration buffer (50mM Tris-HCl pH7.5, 300mM NaCl, 20mM imidazole). Bound proteins were eluted with elution buffer (50mM Tris-HCl pH7.5, 300mM NaCl, 250mM imidazole). The fractions containing the pure recombinant protein were pooled, dialyzed against the proper buffer (50mM Tris-HCl pH8.5, 150mM NaCl, 0.5mM EDTA) and concentrated.

2.5.2 Purification CH-IQ-AB

CH-IQ-AB was purified under denaturing conditions. 8M urea was added to the cell lysate of *E.coli* and the recombinant protein was purified under denaturing conditions (8M urea was added to equilibration and elution buffer). Subsequently, the immobilized nickel affinity chromatography was carried out as described in 2.5.1. Afterwards, the pure unfolded recombinant CH-IQ-AB was refolded *in vitro* by dialysis towards a buffer without urea (50mM Tris-HCl pH8.5, 150mM NaCl, 0.5mM EDTA) and by adding an excess of free heme before starting dialysis (end concentration of 1M hemine). Subsequently, non-incorporated heme was removed by gel filtration (Sephacryl S200 column, running buffer = 50mM Tris-HCl pH8.5, 150mM NaCl, 0.5mM EDTA).

2.6 Optical absorption spectra

To verify a correct heme-incorporation of the purified recombinant globin, UV/visible spectra were taken (200nm – 700nm). Spectral measurements were made with a spectrophotometer (Shimadzu UV-2101PC). The ferrous deoxy sample was obtained by adding an excess of sodium dithionite.

2.7 SDS-Page and Western Blot analysis

Samples were first denatured and reduced by the addition of an equal amount of loading buffer (4% SDS, 10% v/v β -mercapto-ethanol, 125mM TRIS-HCl pH6.8, 20% glycerol) and by incubating them at 95°C for 5min. Subsequently, equal mass samples were separated by SDS-PAGE (15% acrylamide gel for NGB and CH-IQ-AB and 7,5% acrylamide for ADGB). For SDS-PAGE analysis, proteins were stained with Coomassie Brilliant Blue G-250. For Western Blot analysis, proteins were transferred to polyvinylidene difluoride (PVDF) membranes according to the manufacturer's instructions (Immobilon®-P PVDF membrane, pore size 0.45 μ m, Millipore). The membranes were probed with mouse-derived anti-His Tag antibody (6x-His Tag Monoclonal Antibody (4E3D10H2/E3) Invitrogen) or anti-ADGB antibody (polyclonal anti-ADGB antibody HPA036340 SIGMA) (1:5000 in 5% Milk TBS buffer); overnight at 4°C, followed by an incubation at room temperature for 1h with an HRP-conjugated antibody (1:5000 in TBS) (Polyclonal Goat anti-mouse/rabbit Immunoglobulins – HRP, DAKO). Subsequently, Luminata Forte Western HRP substrate (Millipore) was added to the membrane and the targeted proteins were visualized using a SynGene imager.

3 Results and discussion

3.1 Recombinant expression in *Escherichia coli*

The by far most popular host for recombinant protein expression is *E.coli*: 86% of the protein structures entered in the protein databank (PDB) [20] were expressed recombinantly in *E.coli*. This can be explained by the fact that, for a large scale protein production, the *E.coli* system is the easiest, quickest and cheapest method. There are many commercial and non-commercial expression vectors available with different N- and C-terminal tags and many different *E.coli* strains, which are optimized for special applications. The first globins recombinant expressed in *E.coli* were human Mb [21] and Hb [12], followed rapidly by many other globin types of many different organisms. In most cases, globins are small globular proteins, with no specific requirements in folding or post-translational processing and thus *E.coli* is often the perfect host for their large scale recombinant expression.

However, the high degree of accumulation of soluble protein during recombinant expression is not always accepted by the metabolic system of the host. Sometimes, the recombinant expressed protein is toxic to the host cell and causes a cellular stress response, which results in poor and irregular protein expression. An example of such a situation are the expression of microbial FHbs. These chimeric globins possess the ability to generate superoxide and peroxide in the presence of reductant and NAD(P)H and the accumulation of these reactive oxygen species are toxic to the host cell. These problems can be overcome by expressing the FHbs under the control of an inducible promoter. Lewis et al. used the pBAD expression system, which allows tightly controlled, titratable expression of a protein through the regulation of arabinose concentration [22].

Another often encountered problem is the accumulation of target proteins into insoluble aggregates known as inclusion bodies. These proteins are misfolded and biologically inactive. Aggregation prone proteins require the existence of a number of molecular chaperones that interact reversibly with nascent polypeptide chains to prevent aggregation during the folding process. The formation of inclusion bodies could therefore result either from accumulation of high concentrations of folding intermediates or from inefficient processing by molecular chaperones [23]. Refolding from inclusion bodies is in many cases considered undesirable as it results most of the time in a poor recovery of bio-active protein [24]. Most used strategies for shifting the recombinant expression from inclusion bodies to soluble protein are: changing the expression temperature, media and transcription rates, trying various *E.coli* strains, including molecular chaperones, including tRNA complementation

plasmids, expressing fragments, using fusion technology, stabilizing mRNA or screening for soluble variants. For further reading on these strategies we refer to the review paper of Sorensen et al. [23]

The formation of inclusion bodies during recombinant globin expression is a frequently encountered problem. The first recombinant expressed globins (Mb and Hb) were expressed as inclusion bodies and refolded *in vitro* into their functional form [12,21]. Nevertheless, by changing expression conditions most globins can be expressed and purified as soluble proteins [25].

Expression of NGB and CH-IQ-AB in *E.coli*

For the recombinant expression of human NGB and the globin domain of human ADGB (referred to as CH-IQ-AB) in *E.coli* we cloned them into the pET23a expression vector, which uses the strong inducible T7 promoter. The recombinant pET23a constructs were subsequently transformed in BL21(DE3)pLysS *E.coli*, where the recombinant expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG). pET23a expresses the proteins with a C-terminal His-tag, which binds to several types of immobilized metal ions, like nickel. His-tagged proteins can thus easily be purified using nickel affinity chromatography. Both the soluble and insoluble fractions of the cell lysate were screened using SDS-PAGE for the recombinant over-expression of NGB and CH-IQ-AB (Fig1A).

NGB was expressed in a soluble form and was purified with nickel affinity chromatography. After purification of NGB, optical absorption spectra were recorded, which were characteristic for NGB [15]. NGB is a hexacoordinated globin which means that under deoxy conditions the protein forms two coordinative bonds with the iron atom of the heme group. The optical absorption spectrum of a hexacoordinated heme group has large amplitudes for the alpha band (562nm) and the Soret band (428nm). This observation indicates that the heme group was correctly incorporated in the polypeptide chain of NGB (Fig1B), consistent with a previous report [15].

CH-IQ-AB was expressed in *E.coli* as insoluble aggregates, also known as inclusion bodies and it was not possible to purify CH-IQ-AB as a soluble protein even after alternating the conditions (Fig1A). Different strategies were used to shift the expression from inclusion bodies to stable soluble protein. By changing time of induction and lowering the temperature during protein expression we were able to avoid inclusion body formation, however CH-IQ-AB remained unstable and precipitated over time. We tried to increase the solubility of the globin domain by the addition of reducing agents or detergents and by changing the pH or salt concentrations of the resuspension buffer. Because CH-IQ-AB possesses two cysteine residues, which can possibly cause intermolecular disulfide bridge formation and hence precipitation, we mutated the cysteines into serines. Despite these efforts, CH-

IQ-AB remained unstable. As such, CH-IQ-AB was expressed as inclusion bodies, purified under denaturing conditions (8M urea) and the recombinant protein was refolded *in vitro* by dialysis and adding free heme, as previously done for human CYGB [26]. In contrast to the predictions of the computer three-dimensional modeling [16] the refolded globin domain was not stable and precipitated over time. Despite the precipitation, we were able to measure an UV-visible spectrum of the refolded globin domain. The optical absorption spectra of the refolded ferrous deoxy CH-IQ-AB domain displayed the typical spectra of a hexacoordinated globin (high α band (557nm) and Soret band (421nm), low beta band (527nm)), like NGB (Fig1C). However, the purification of the refolded CH-IQ-AB was not efficient and only low yields could be obtained. The refolded globin domain remained unstable, precluding biophysical characterizations, including ligand binding kinetic determination and crystallization.

Folding studies of apomyoglobin (apoMb) [27] and plant hemoglobins [28] showed that the folding of a globin happens based on the contact order of not only local-residue contacts (as is seen with very small proteins), but also on long-range amino acid interactions.. A stepwise-folding process of apoMb has been suggested in the following order; helices G, A, H, B, E, C and D. During this folding pathway, there are two folding intermediates, if one of these intermediates is misfolded, aggregation occurs. We can imagine that the folding of the circular permuted globin of ADGB, where helices A and B are placed after helices C-H, is more complicated than the folding of a typical globin with a normal consecution of the helices, like NGB.

Previously, the effect of circular permutation in Mb was investigated by Ribeiro et al.[29]. A circular permuted Mb was engineered, where helices A and B were shifted towards the C-terminus (similar to the globin domain of ADGB). Subsequently, this protein was expressed in *E.coli* as inclusion bodies and it was afterwards refolded *in vitro*. The refolded circular permuted myoglobin displayed absorption spectra similar to the WT Mb. However, the circular permutation caused a significant decrease in the stability of the protein and a large fraction formed stable aggregates, similar to our experience with the circular permuted globin domain of ADGB.

Most plausible, the globin domain of ADGB needs long-range interactions from the amino-acid sequence of the full-length ADGB to form a stable globin fold. Another possibility is that the human globin domain is simply not able to fold correctly in the prokaryotic *E.coli* system. Therefore, we further investigated if the globin domain can be natively folded in an eukaryotic yeast expression system.

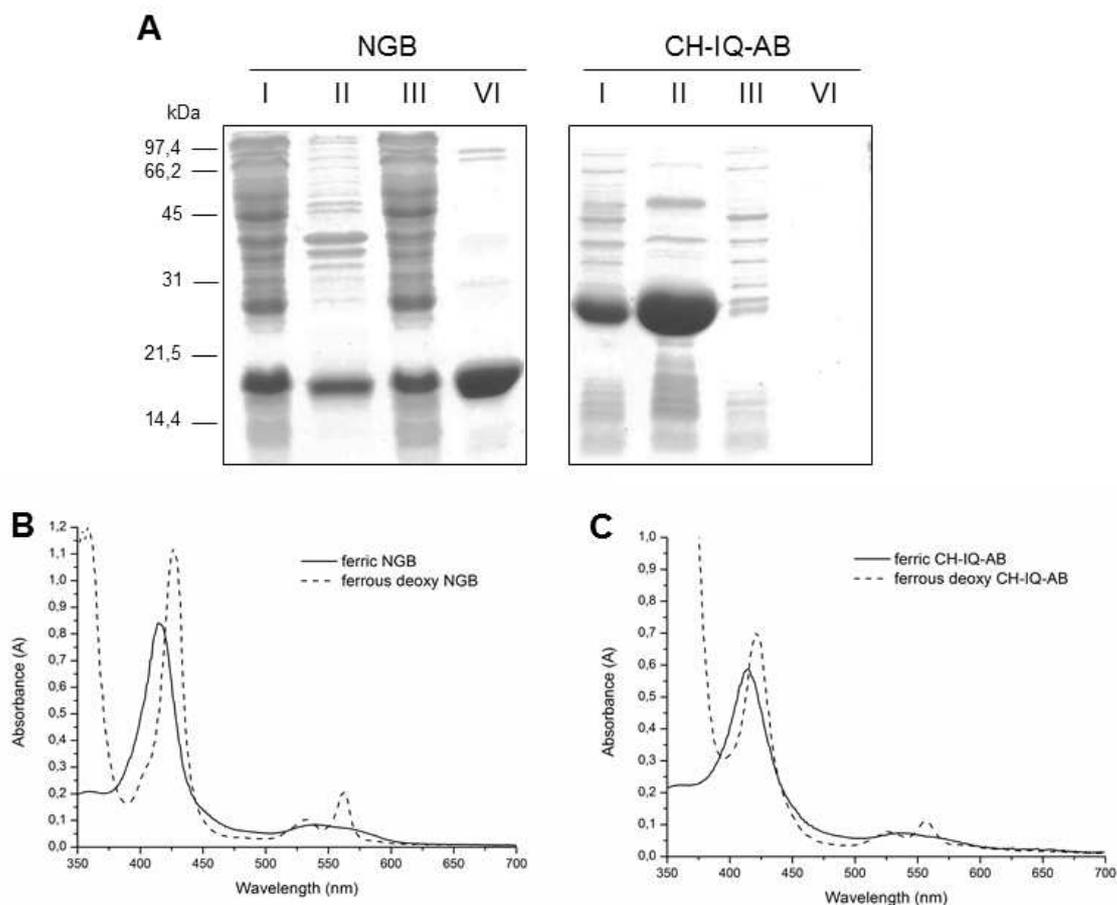


Figure 1: A) SDS-PAGE (15%): Recombinant expression in *E. coli* of CH-IQ-AB (~25kDa) and NGB (~18kDa). I) *E. coli* total cell lysate II) Insoluble fraction of cell lysate III) Soluble fraction of cell lysate IV) Fraction after native purification with Nickel column. B) Optical absorption spectra of NGB expressed in *E. coli* (pET23a) and C) optical absorption spectra of refolded CH-IQ-AB, expressed in *E. coli* (pET23a).

3.2 Recombinant expression in *Pichia pastoris*

E. coli as expression system might be the obvious first choice, but it is not always the best choice, as we have seen for CH-IQ-AB. *E. coli* is a prokaryotic system and does not contain the subcellular machinery of eukaryotes needed for post-translational modifications, including proteolytic processing, folding, disulfide bond formation, glycosylation and phosphorylation. Many eukaryotic proteins need to be modified following translation in order to become active and/or adapt the proper structure. Yeast is an eukaryotic organism and, similar to *E. coli*, can be grown to very high densities, which makes them especially useful for large scale protein production. The two most common yeast strains used for recombinant protein expression are *S.cerevisiae* and the methylotrophic yeast *P.pastoris*.

Only few cases are reported where globins are recombinant expressed in a yeast system [30–32]. An example of a globin that has been successfully expressed in *P.pastoris* is the *Crocodylus siamensis* hemoglobin. This crocodile Hb displays potent antibacterial and anti-inflammation activity, as well as strong antioxidant properties [33–35]. Furthermore, it comprises six cysteine residues in the α -globin chain and three cysteine residues in the β -globin chain [35]. Each of these cysteine residues can form disulfide bonds both of the intermolecular and intramolecular formation. The expression in *E.coli* resulted in low expression yields and precipitation after purification [32]. To overcome these problems the crocodile Hb was successfully expressed in *P.pastoris*.

As the CH-IQ-AB globin domain of ADGB displays similar problems of precipitation after purification and comprises also three cysteine residues (two in the α -helices and one in the IQ motif). We recombinantly expressed CH-IQ-AB in the *P.pastoris* system.

Expression of NGB and CH-IQ-AB in *P.pastoris*

For the expression of NGB and CH-IQ-AB we chose for methylotrophic yeast *P.pastoris* because it is widely used, cost-effective and suitable for large-scale protein production. cDNA of NGB and CH-IQ-AB were cloned into the pPicZ α vector, which places a native *S.cerevisiae* α -secretion signal before the recombinant protein. The advantage of a secreted expression is that *P.pastoris* secretes very low levels of native proteins, facilitating further purification steps. The pPicZ α vector uses the AOX promoter for a high-level, methanol-induced expression. After induction of recombinant expression, every 24h a sample was taken for analysis of expression of CH-IQ-AB and NGB using western blot. NGB is expressed and secreted by *P.pastoris* (Fig2A). 96h post-induction, NGB was purified out of the expression medium and optical absorption spectra were taken (Fig2B). These spectra evidence proper folding and incorporation of heme in NGB in line with previous descriptions [31]. Despite our efforts to optimize the cDNA sequence of CH-IQ-AB to the codon usage of *P.pastoris*, we could not detect recombinant expression of CH-IQ-AB in *P.pastoris* (Fig2A), neither in the secreted fraction nor in the intracellular fraction.

This may be explained by the possibility that CH-IQ-AB cannot be expressed as a stable soluble form because it requires the context of the full length ADGB protein for proper folding as was already suggested in section 3.1. In *E.coli*, misfolded proteins are stored as inclusion bodies. In eukaryotic systems, protein folding occurs in a different way; proteins are folded and processed in the endoplasmic reticulum (ER) and/or the Golgi apparatus and sometimes secreted into the extracellular environment through vesicular transport (as is in the case of recombinant expression of CH-IQ-AB in *P.pastoris*). In eukaryotes aberrant folding properties of the target protein and/or high level production can lead to the accumulation of unfolded or even aggregated protein in the ER,

which can initiate the unfolded protein response (UPR) and ER-associated degradation [36–40]. This can thus possibly explain the fact that we cannot detect expression of CH-IQ-AB in *P.pastoris*.

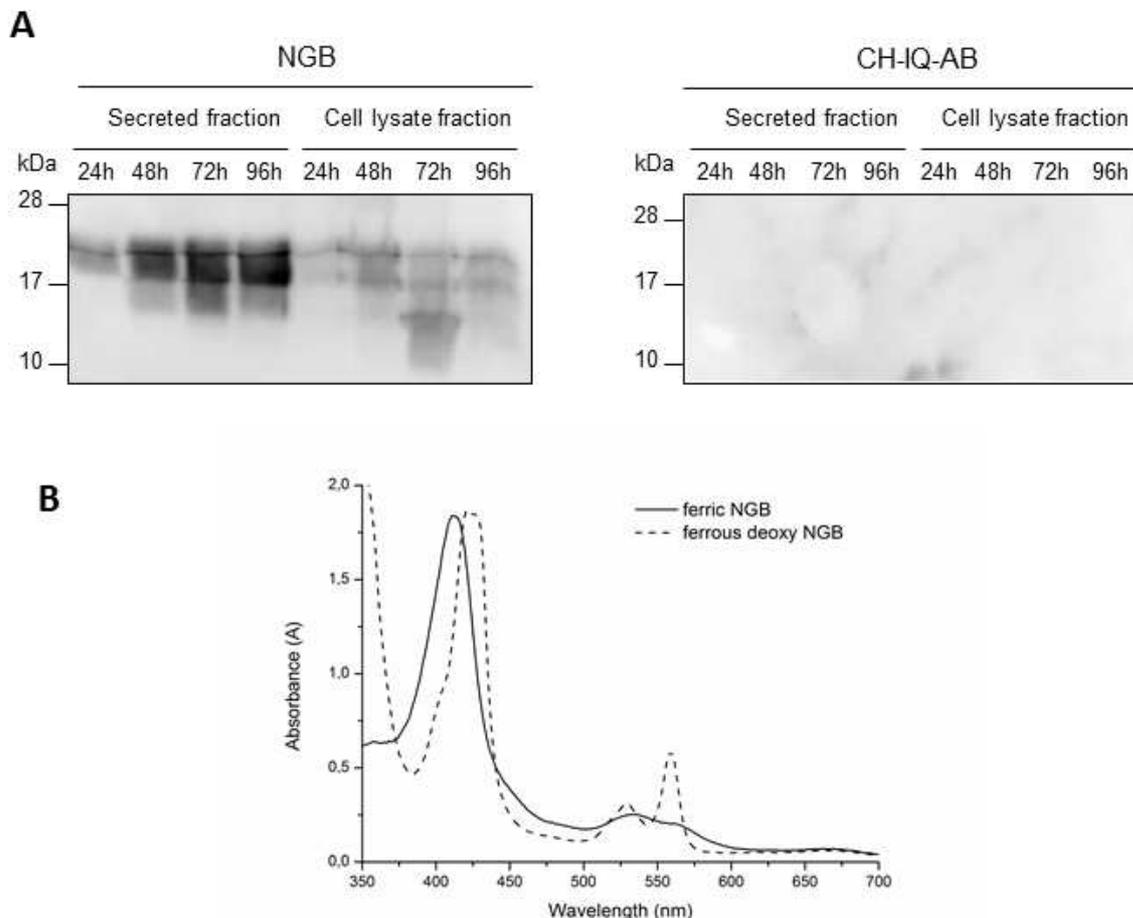


Figure 2: A) Western Blot analysis (anti His-tag) (15% polyacrylamide) of recombinant expression of NGB (~19kDa) and CH-IQ-AB (~26kDa) in X33 *P.pastoris* (pPicZα expression vector). Expression was analyzed every 24h post induction. Both soluble and insoluble fractions were analyzed. B) Optical absorption spectra of NGB, expressed in *P.pastoris* (pPicZα).

3.3 Recombinant expression in *Sf9* insect cells

Finally, we opted for a third expression system, one that is much closer to mammalian cells; the baculovirus-insect cell system. This expression system is a binary system, consisting of a recombinant baculovirus as the vector and lepidopteran insect cells as the host [41]. Insect cells represent a higher eukaryotic system than yeast; they are able to carry out more complex post-translational modifications and have a better folding machinery for mammalian proteins. Furthermore, baculoviral vectors can accommodate large DNA inserts, which allows for the production of multi-subunit protein complexes or even high molecular weight virus-like particles [42–44]. A potential

disadvantage of the baculovirus-insect system is that it is a high-cost and time-consuming method of recombinant protein expression.

So far, there are no reports of globins recombinantly expressed in insect cells. Since the baculovirus-insect cell expression system is widely used for the recombinant expression of high molecular weight proteins, we used it to express the 190kDa full length ADGB protein.

Expression of NGB, CH-IQ-AB and full length ADGB in *Sf9* cells

NGB, CH-IQ-AB and full length ADGB cDNA were cloned into the pDEST10 vector, which possesses the strong transcriptional promoter derived from the very late baculovirus gene *polyhedrin*, able to produce milligrams of proteins. *Sf9* insect cells were infected with recombinant NGB, CH-IQ-AB and ADGB baculovirus particles. 72h after infection, the expression of recombinant protein was analyzed with western blot (Fig3). Expression was detected for NGB and ADGB. The expression level of CH-IQ-AB was however very low, again possibly reflecting that the globin domain would require the context of the full length ADGB for proper folding.

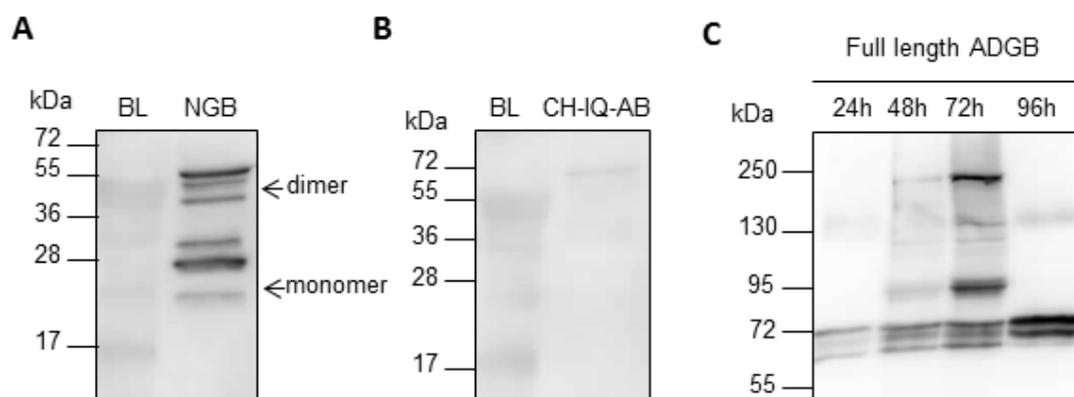


Figure 3: A) Western blot analysis (anti His tag) (15% polyacrylamide) of *Sf9* cell lysate (soluble fraction) infected with recombinant NGB baculovirus particles. Expected molecular weight of NGB is ~21kDa. Multiple bands are detected, most probably caused by dimerization and formation of different stable protein conformations or by post-translational modifications. B) Western blot analysis (anti His tag) (15% polyacrylamide) of *Sf9* cell lysate (soluble fraction) infected with recombinant CH-IQ-AB baculovirus particles. Expected molecular weight of CH-IQ-AB is ~28kDa. C) Western blot analysis (anti ADGB) (7,5% polyacrylamide) of *Sf9* cell lysate (soluble fraction) infected with recombinant ADGB baculovirus particles. Expected molecular weight of ADGB is ~194kDa. ADGB might be (auto)cleaved by internal proteases or by the N-terminal calpain domain of ADGB. BL = Blanc, uninfected *Sf9* cell lysate.

Surprisingly, for NGB multiple bands were detected with western blot analysis (Fig3A). It seems that NGB a) can be detected as dimer and b) appears in three apparent molecular mass forms, in *Sf9*

insect cells. The dimerization could not be prevented by heat denaturation or addition of reducing agents, like β -mercaptoethanol or DTT. Dimerization was also seen for NGB when it was expressed in *E.coli*, but not in *P.pastoris* (Fig4). In *P.pastoris* NGB is secreted into the medium, and is thus present in lower concentrations than when it is intracellularly expressed in *E.coli* or *Sf9*, dimer formation may in this way be concentration dependent. This phenomenon of dimerization, resistant to heat denaturation and reduction, has been observed previously for Ngb in whole protein extracts from mouse brain and retinas [45–48]. Furthermore, it has been shown that three forms of apparent molecular masses (~17, ~19 and ~21 kDa), specifically recognized by anti-Ngb antibodies, were identified in mitochondrial enriched fractions of mouse retina [49]. This pattern of three different molecular mass forms is similar to the one that is seen after expression of NGB in *Sf9* insect cells (Fig3A). These forms might represent atypical migrations of stable protein conformations and it has been hypothesized that that rupture/ formation of the intra molecular (CD7-B5) disulfide bond of NGB [50] can give rise to different apparent molecular mass forms [49]. Post-translational modifications could also occur, such as covalent links of cofactors or phosphorylation. Indeed, it has previously been displayed that Ngb is phosphorylated during hypoxia [51]. As insect cells are known for their capabilities of protein N-glycosylation, we analyzed the possibility that Ngb is N-glycosylated using the NetNGlyc1.0 prediction server [52] and by the addition of PNGaseF, which enzymatically cleaves N-glycans, to the cell lysate. No N-glycosylation sites were predicted in NGB and PNGaseF did not cause a shift in molecular weight on western blot, ruling out the possibility that NGB is N-glycosylated by the insect cell system. It seems that NGB only forms these different molecular mass forms after expression in the higher eukaryotic baculoviral insect cell system and not in *E.coli* or *P.pastoris*, assuming that the latter do not provide the cellular environment for formation of these different conformations or post-translational modifications.

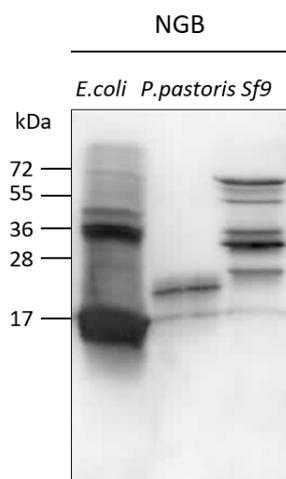


Figure 4: Western blot analysis (anti-His tag) (15% polyacrylamide) of *E.coli*, *P.pastoris* and *Sf9* cell lysates were NGB was recombinantly in expressed. Molecular weight of NGB in *E.coli* should be ~18kDa, in *P.pastoris* ~19kDa and in *Sf9* ~21kDa, due to different C-terminal or N-terminal His-tags. Dimerization of NGB resistant to heat denaturation and reducing agents is seen in both *E.coli* and *Sf9*, not in *P.pastoris*. In *Sf9* NGB appears in three apparent molecular weight forms, possibly due to the formation of different stable protein conformations or by post-translational processing. This is not seen in *E.coli* or *P.pastoris*.

In contrast to the CH-IQ-AB domain, the full length ADGB was significantly expressed in *Sf9* cells, confirming the hypothesis that the full length protein is required for stable expression. However, we observed a time-dependent truncation of the full length protein (Fig3C) and even after scaling-up the expression, the amount of full length ADGB was insufficient for further biophysical characterization. This truncation is most probably a result of proteolysis by intracellular proteases or by auto-cleavage by the N-terminal calpain-like domain, a phenomenon commonly observed among calpains [53–56]. To further evaluate if ADGB autolytic activity would be the source of this potential auto-cleavage we added three different protease inhibitors to the growth medium; the cysteine protease inhibitor E64, the acid protease inhibitor Pepstatin A and the calpain specific Calpain II inhibitor. However, administration of these inhibitors did not prevent the truncation of ADGB. The ADGB calpain-like domain is homologous to calpain 7, for which no specific protease inhibitors exists. Interestingly, substantial autolytic activity has also been demonstrated for calpain-7 [57]. Further research is necessary to understand this putative auto-cleavage of ADGB.

4 Conclusion

In this paper we explored three different expression systems for the recombinant expression of globins; the prokaryote expression system *E.coli* and the two eukaryote expression systems *P.pastoris* and baculovirus infected *Sf9*. In general, globins are small globular proteins, which do not need post-translation modification for a proper folding and in this regard *E.coli* is the obvious first choice for recombinant expression. However, the major disadvantage of recombinant expression in *E.coli* is the formation of insoluble aggregates, inclusion bodies.

We attempted to express the circular permutated globin domain of ADGB in *E.coli*, *P.pastoris* and *Sf9* cells, however, a soluble expression was not possible. This suggest that the globin domain of ADGB needs the full context of the full length protein in order to form a stable globin fold. Furthermore, we displayed that ADGB was truncated, possibly due to auto-cleavage by the internal calpain-like domain. Our study demonstrates that it is not always possible to express globins in *E.coli* and that switching to a (higher) eukaryotic expression system might represent a valid alternative for more complex, multi-domain globins, like the chimeric globin ADGB.

When NGB was recombinantly expressed in insect cells, three apparent molecular mass forms and dimerization were observed, this was not observed in *E.coli* or *P.pastoris*. Interestingly, the same multiple band pattern on western blot has been observed in mitochondrial enriched fractions of mouse retina and in brain tissue. It seems that the baculoviral insect cell expression system provides a cellular environment for folding and processing of proteins that is more similar to the natural context of Ngb (brain and retina) compared to the cellular environments of *E.coli* or *P.pastoris*. In this regard it might be worthwhile to switch to a (higher) eukaryotic expression systems for eukaryotic proteins in general, even if they are easily expressed in a bacterial or yeast expression system.

5 Acknowledgments

M. L. Van Hauwaert and H. Berghmans (University of Antwerp) are gratefully acknowledged for technical assistance.

6 Funding

This work was supported by the Fund for Scientific Research (FWO) and by the German Research Foundation (HO5837/1-1) and the Swiss National Science Foundation (31003A_173000).

7 Conflict of interest

The authors do not have any conflict of interest.

8 References

- [1] S.N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, J. Gough, S. Dewilde, L. Moens, J.R. Vanfleteren, A phylogenomic profile of globins., *BMC Evol. Biol.* 6 (2006) 31. doi:10.1186/1471-2148-6-31.
- [2] S.N. Vinogradov, D. Hoogewijs, X. Bailly, R. Arredondo-Peter, M. Guertin, J. Gough, S. Dewilde, L. Moens, J.R. Vanfleteren, Three globin lineages belonging to two structural classes in genomes from the three kingdoms of life., *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 11385–9. doi:10.1073/pnas.0502103102.
- [3] A. Bonamore, A. Boffi, Flavohemoglobin: Structure and reactivity, *IUBMB Life.* 60 (2008) 19–28. doi:10.1002/iub.9.
- [4] F. Germani, L. Moens, S. Dewilde, Haem-Based Sensors, in: 2013: pp. 1–47. doi:10.1016/B978-0-12-407693-8.00001-7.
- [5] S.N. Vinogradov, D. Hoogewijs, X. Bailly, K. Mizuguchi, S. Dewilde, L. Moens, J.R. Vanfleteren, A model of globin evolution., *Gene.* 398 (2007) 132–42. doi:10.1016/j.gene.2007.02.041.
- [6] S.N. Vinogradov, X. Bailly, D.R. Smith, M. Tinajero-Trejo, R.K. Poole, D. Hoogewijs, Microbial eukaryote globins., *Adv. Microb. Physiol.* 63 (2013) 391–446. doi:10.1016/B978-0-12-407693-8.00009-1.

- [7] S.N. Vinogradov, M. Tinajero-Trejo, R.K. Poole, D. Hoogewijs, Bacterial and archaeal globins - A revised perspective, *Biochim. Biophys. Acta - Proteins Proteomics*. 1834 (2013) 1789–1800. doi:10.1016/j.bbapap.2013.03.021.
- [8] J.B. Wittenberg, M. Bolognesi, B.A. Wittenberg, M. Guertin, Truncated hemoglobins: A new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants, *J. Biol. Chem.* 277 (2002) 871–874. doi:10.1074/jbc.R100058200.
- [9] J.C. KENDREW, G. BODO, H.M. DINTZIS, R.G. PARRISH, H. WYCKOFF, D.C. PHILLIPS, A three-dimensional model of the myoglobin molecule obtained by x-ray analysis., *Nature*. 181 (1958) 662–666.
- [10] H. MUIRHEAD, M.F. PERUTZ, STRUCTURE OF HAEMOGLOBIN. A THREE-DIMENSIONAL FOURIER SYNTHESIS OF REDUCED HUMAN HAEMOGLOBIN AT 5-5 Å RESOLUTION., *Nature*. 199 (1963) 633–638.
- [11] J. Shine, I. Fettes, N.C. Lan, J.L. Roberts, J.D. Baxter, Expression of cloned beta-endorphin gene sequences by *Escherichia coli*., *Nature*. 285 (1980) 456–463.
- [12] K. Nagai, H.C. Thogersen, Generation of beta-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*., *Nature*. 309 (1984) 810–812.
- [13] P. Valenzuela, A. Medina, W.J. Rutter, G. Ammerer, B.D. Hall, Synthesis and assembly of hepatitis B virus surface antigen particles in yeast., *Nature*. 298 (1982) 347–350.
- [14] X. Shi, D.L. Jarvis, Protein N-glycosylation in the baculovirus-insect cell system., *Curr. Drug Targets*. 8 (2007) 1116–25. doi:10.1038/nature13314.
- [15] S. Dewilde, L. Kiger, T. Burmester, T. Hankeln, V. Baudin-Creuzat, T. Aerts, M.C. Marden, R. Caubergs, L. Moens, Biochemical characterization and ligand binding properties of neuroglobin, a novel member of the globin family., *J. Biol. Chem.* 276 (2001) 38949–55. doi:10.1074/jbc.M106438200.
- [16] D. Hoogewijs, B. Ebner, F. Germani, F.G. Hoffmann, A. Fabrizius, L. Moens, T. Burmester, S. Dewilde, J.F. Storz, S.N. Vinogradov, T. Hankeln, Androglobin: a chimeric globin in metazoans that is preferentially expressed in Mammalian testes., *Mol. Biol. Evol.* 29 (2012) 1105–14. doi:10.1093/molbev/msr246.
- [17] V. Sarramegna, F. Talmont, P. Demange, A. Milon, Heterologous expression of G-protein-coupled receptors: Comparison of expression systems from the standpoint of large-scale production and purification, *Cell. Mol. Life Sci.* 60 (2003) 1529–1546. doi:10.1007/s00018-003-3168-7.
- [18] A. Mohajeri, S. Sanaei, F. Kiafar, A. Fattahi, M. Khalili, N. Zarghami, The Challenges of Recombinant Endostatin in Clinical Application: Focus on the Different Expression Systems and Molecular Bioengineering, *Adv. Pharm. Bull.* 7 (2017) 21–34. doi:10.15171/apb.2017.004.
- [19] C.L. Morton, P.M. Potter, Comparison of *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Spodoptera frugiperda*, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase., *Mol. Biotechnol.* 16 (2000) 193–202. doi:10.1385/MB:16:3:193.
- [20] P.E.B. H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, The Protein Data Bank, *Nucleic Acids Res.* (2000) 235–242. <http://www.rcsb.org/pdb/home/home.do> (accessed October 23, 2017).
- [21] R. Varadarajan, A. Szabo, S.G. Boxer, Cloning, expression in *Escherichia coli*, and reconstitution of human myoglobin (gene expression/heme protein), *Biochemistry*. 82 (1985) 5681–5684. doi:10.1073/pnas.82.17.5681.
- [22] M.E.S. Lewis, H.A. Corker, B. Gollan, R.K. Poole, A Survey of Methods for the Purification of Microbial Flavohemoglobins, Elsevier Masson SAS, 2008. doi:10.1016/S0076-6879(08)36009-1.
- [23] H.P. Sørensen, K.K. Mortensen, Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*., *Microb. Cell Fact.* 4 (2005) 1. doi:10.1186/1475-2859-4-1.
- [24] S.M. Singh, A.K. Panda, Solubilization and refolding of bacterial inclusion body proteins, *J. Biosci. Bioeng.* 99 (2005) 303–310. doi:10.1263/jbb.99.303.

- [25] C. Natarajan, X. Jiang, A. Fago, R.E. Weber, H. Moriyama, J.F. Storz, Expression and purification of recombinant hemoglobin in *Escherichia coli*, *PLoS One*. 6 (2011) 1–7. doi:10.1371/journal.pone.0020176.
- [26] D. de Sanctis, S. Dewilde, A. Pesce, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester, M. Bolognesi, Crystal structure of cytoglobin: the fourth globin type discovered in man displays heme hexa-coordination., *J. Mol. Biol.* 336 (2004) 917–927.
- [27] B.C.N. Ishimura, Review Folding of apomyoglobin : Analysis of transient intermediate structure during refolding using quick hydrogen deuterium exchange and NMR, 93 (2017) 10–27.
- [28] S. Nakajima, E. Álvarez-Salgado, T. Kikuchi, R. Arredondo-Peter, Prediction of folding pathway and kinetics among plant hemoglobins using an average distance map method, *Proteins Struct. Funct. Genet.* 61 (2005) 500–506. doi:10.1002/prot.20658.
- [29] E.A. Ribeiro, C.H.I. Ramos, Circular permutation and deletion studies of myoglobin indicate that the correct position of its N-terminus is required for native stability and solubility but not for native-like heme binding and folding, *Biochemistry*. 44 (2005) 4699–4709. doi:10.1021/bi047908c.
- [30] R. Mould, O. Hofmann, T. Brittain, Production of human embryonic haemoglobin (Gower II) in a yeast expression system., *Biochem. J.* 298 Pt 3 (1994) 619–22.
- [31] Q. Ye, Y. Sun, Y. Wu, Y. Gao, Z. Li, W. Li, C. Zhang, *Pichia pastoris* Production of Tat-NGB and Its Neuroprotection on Rat Pheochromocytoma Cells, *Mol. Biotechnol.* 58 (2016) 22–29. doi:10.1007/s12033-015-9898-6.
- [32] P. Anwised, N. Jangpromma, T. Temsiripong, R. Patramanon, S. Daduang, S. Jitrapakdee, T. Araki, S. Klaynongsruang, Cloning, Expression, and Characterization of Siamese Crocodile (*Crocodylus siamensis*) Hemoglobin from *Escherichia coli* and *Pichia pastoris*, *Protein J.* 35 (2016) 256–268. doi:10.1007/s10930-016-9669-7.
- [33] J. Jandaruang, J. Siritapetawee, K. Thumanu, C. Songsiriritthigul, C. Krittanai, S. Daduang, A. Dhiravisit, S. Thammasirirak, The effects of temperature and pH on secondary structure and antioxidant activity of *Crocodylus siamensis* hemoglobin., *Protein J.* 31 (2012) 43–50. doi:10.1007/s10930-011-9372-7.
- [34] S. Phosri, P. Mahakunakorn, J. Lueangsakulthai, N. Jangpromma, P. Swatsitang, S. Daduang, A. Dhiravisit, S. Thammasirirak, An investigation of antioxidant and anti-inflammatory activities from blood components of Crocodile (*Crocodylus siamensis*)., *Protein J.* 33 (2014) 484–492. doi:10.1007/s10930-014-9581-y.
- [35] S. Srihongthong, A. Pakdeesuan, S. Daduang, T. Araki, A. Dhiravisit, S. Thammasirirak, Complete amino acid sequence of globin chains and biological activity of fragmented crocodile hemoglobin (*Crocodylus siamensis*)., *Protein J.* 31 (2012) 466–476. doi:10.1007/s10930-012-9424-7.
- [36] A.L. Vanz, M. Nimtz, U. Rinas, Decrease of UPR- and ERAD-related proteins in *Pichia pastoris* during methanol-induced secretory insulin precursor production in controlled fed-batch cultures, *Microb. Cell Fact.* 13 (2014) 23. doi:10.1186/1475-2859-13-23.
- [37] G. Whiteside, M.J.C. Alcocer, J.R. Kumita, C.M. Dobson, M. Lazarou, R.J. Pleass, D.B. Archer, Native-state stability determines the extent of degradation relative to secretion of protein variants from *Pichia pastoris*., *PLoS One*. 6 (2011) e22692. doi:10.1371/journal.pone.0022692.
- [38] A.R. Hesketh, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures., *Appl. Microbiol. Biotechnol.* 97 (2013) 9747–9762. doi:10.1007/s00253-013-5186-1.
- [39] A.L. Vanz, H. Lunsdorf, A. Adnan, M. Nimtz, C. Gurramkonda, N. Khanna, U. Rinas, Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes., *Microb. Cell Fact.* 11 (2012) 103. doi:10.1186/1475-2859-11-103.
- [40] A.A. Welihinda, W. Tirasophon, R.J. Kaufman, The cellular response to protein misfolding in

- the endoplasmic reticulum., *Gene Expr.* 7 (1999) 293–300.
- [41] M.M. Van Oers, G.P. Pijlman, J.M. Vlak, Thirty years of baculovirus-insect cell protein expression: From dark horse to mainstream technology, *J. Gen. Virol.* 96 (2015) 6–23. doi:10.1099/vir.0.067108-0.
- [42] T.A. Kost, J.P. Condreay, D.L. Jarvis, Baculovirus as versatile vectors for protein expression in insect and mammalian cells., *Nat. Biotechnol.* 23 (2005) 567–575. doi:10.1038/nbt1095.
- [43] K.L. Hefferon, Baculovirus late expression factors., *J. Mol. Microbiol. Biotechnol.* 7 (2004) 89–101. doi:10.1159/000078652.
- [44] I. Berger, D.J. Fitzgerald, T.J. Richmond, Baculovirus expression system for heterologous multiprotein complexes., *Nat. Biotechnol.* 22 (2004) 1583–1587. doi:10.1038/nbt1036.
- [45] A. Fabrizius, D. Andre, T. Laufs, A. Bicker, S. Reuss, E. Porto, T. Burmester, T. Hankeln, Critical re-evaluation of neuroglobin expression reveals conserved patterns among mammals., *Neuroscience.* 337 (2016) 339–354. doi:10.1016/j.neuroscience.2016.07.042.
- [46] M. Schmidt, A. Giessl, T. Laufs, T. Hankeln, U. Wolfrum, T. Burmester, How does the eye breathe? Evidence for neuroglobin-mediated oxygen supply in the mammalian retina., *J. Biol. Chem.* 278 (2003) 1932–1935. doi:10.1074/jbc.M209909200.
- [47] K. Jin, Y. Mao, X. Mao, L. Xie, D.A. Greenberg, Neuroglobin expression in ischemic stroke., *Stroke.* 41 (2010) 557–559. doi:10.1161/STROKEAHA.109.567149.
- [48] Last visited 04/08/2017, Last Visit. 04/08/2017. (n.d.).
https://www.mybiosource.com/prods/Antibody/Polyclonal/NGB/datasheet.php?products_id=129771#QLREF.
- [49] C. Lechauve, S. Augustin, H. Cwerman-thibault, A. Bouaita, V. Forster, C. Célier, P. Rustin, M.C. Marden, J. Sahel, M. Corral-debrinski, *Biochimica et Biophysica Acta Neuroglobin involvement in respiratory chain function and retinal ganglion cell integrity*, 1823 (2012) 2261–2273. doi:10.1016/j.bbamcr.2012.09.009.
- [50] D. Hamdane, L. Kiger, S. Dewilde, B.N. Green, A. Pesce, J. Uzan, T. Burmester, T. Hankeln, M. Bolognesi, L. Moens, M.C. Marden, The Redox State of the Cell Regulates the Ligand Binding Affinity of Human Neuroglobin and Cytoglobin, *J. Biol. Chem.* 278 (2003) 51713–51721. doi:10.1074/jbc.M309396200.
- [51] T. Jayaraman, J. Tejero, B.B. Chen, A.B. Blood, S. Frizzell, C. Shapiro, M. Tiso, B.L. Hood, X. Wang, X. Zhao, T.P. Conrads, R.K. Mallampalli, M.T. Gladwin, 14-3-3 Binding and phosphorylation of neuroglobin during hypoxia modulate six-to-five heme pocket coordination and rate of nitrite reduction to nitric oxide, *J. Biol. Chem.* 286 (2011) 42679–42689. doi:10.1074/jbc.M111.271973.
- [52] E.J. and S.B. R. Gupta, Prediction of N-glycosylation sites in human proteins, (2004).
<http://www.cbs.dtu.dk/services/NetNGlyc/> (accessed October 25, 2017).
- [53] Y. Ono, S.I. Iemura, S.M. Novak, N. Doi, F. Kitamura, T. Natsume, C.C. Gregorio, H. Sorimachi, PLEIAD/SIMC1/C5orf25, a novel autolysis regulator for a skeletal-muscle-specific calpain, CAPN3, scaffolds a CAPN3 substrate, CTBP1, *J. Mol. Biol.* 425 (2013) 2955–2972. doi:10.1016/j.jmb.2013.05.009.
- [54] Y. Ono, M. Shindo, N. Doi, F. Kitamura, C.C. Gregorio, H. Sorimachi, The N- and C-terminal autolytic fragments of CAPN3/p94/calpain-3 restore proteolytic activity by intermolecular complementation, *Proc. Natl. Acad. Sci.* 111 (2014) E5527–E5536. doi:10.1073/pnas.1411959111.
- [55] R.M. Murphy, G.D. Lamb, Endogenous calpain-3 activation is primarily governed by small increases in resting cytoplasmic [Ca²⁺] and is not dependent on stretch, *J. Biol. Chem.* 284 (2009) 7811–7819. doi:10.1074/jbc.M808655200.
- [56] N. Ermolova, I. Kramerova, M.J. Spencer, Autolytic activation of calpain 3 proteinase is facilitated by calmodulin protein, *J. Biol. Chem.* 290 (2015) 996–1004. doi:10.1074/jbc.M114.588780.
- [57] Y. Osako, Y. Maemoto, R. Tanaka, H. Suzuki, H. Shibata, M. Maki, Autolytic activity of human calpain 7 is enhanced by ESCRT-III-related protein IST1 through MIT-MIM interaction., *FEBS J.*

ACCEPTED MANUSCRIPT