Research Article

FINE TUNING SOLUBLE EXPRESSION OF A HEME PROTEIN IN *ESCHERICHIA COLI* USING LACTOSE AS AN INDUCER

Tajith B. Shaik and Biswajit Pal*

CSIR-Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India.

Abstract: Availability of genome sequences from many different species poses a huge challenge to the biochemists and structural biologists to decipher the structure-function relationship in proteins. Milligram quantities of proteins are required to achieve this goal. Recombinant DNA technology and heterologous expression of proteins aid these challenges and *Escherichia coli* based expression hosts are extensively used. However, several issues like solubility and incorporation of co-factors in the apo-proteins need to be addressed. Therefore, establishment of simple, efficient and cost effective methods to express and purify these proteins are required. In this report, we present a novel and cost effective method of expression and simple and efficient purification method for a heme protein from *Vibrio cholerae*, which is a sensor of NO. The key feature of this method is that lactose was used to express the heme-b containing holoprotein. A simple, two step purification process, combining ion-exchange and size exclusion chromatography, was used to purify the protein. Overall yield of the properly folded heme protein was about 12 mg per litre culture.

Key words: Expression and purification; Lactose induction; Heme; H-NOX; SONO; Nitric oxide sensor.

Introduction

Nearly 4000 genome sequences from different organisms are publicly available today (*http://www.microbesonline.org/*). A typical eukaryotic genome is comprised of 10,000-60,000 proteins where as the number is between 1,500 and 8,000 for microbes (Kim *et al.*, 2011). This poses an immense challenge for biochemists to express, purify and characterize these large numbers of proteins. Recombinant DNA technology along with prokaryotic expression system is most widely used to express these proteins for purification and characterization. However, when proteins are expressed at high levels in bacteria, it is very common that many of the proteins become insoluble hindering direct

Corresponding Author: **Biswajit Pal**
E-mail: pal@ccmb.res.in

Received: June 12, 2015

Accepted: August 11, 2015

Published: August 16, 2015

purification. Further, when a protein containing a cofactor is expressed, many a times only the apo-protein is expressed. One such example is expression of heme proteins. Heme-proteins constitute a large class of proteins and play a significant role in diverse biological functions. Overproduction of heme proteins in bacteria is hindered by the lack of incorporation of heme prosthetic group in the recombinant proteins. Several strategies have been used to overcome this issue (Woodard and Daily, 1995; Jung *et al.* 2001; Karow *et al.* 2004; Varnado and Goodwin, 2004; Liu *et al.* 2007; Sudhamsu *et al.* 2010). These include co-expression of other enzymes related to heme biosynthesis (Sudhamsu *et al.*, 2010), addition of heme or precursor of heme (δ-aminolevulinic acid or δ-ALA) in the bacterial culture (Karow *et al.*, 2004; Liu *et al.*, 2007), construction of heme protein expression (HPEX) system (Varnado and Goodwin, 2004), etc. Further, specialized tuner cells have been used along with δ-ALA to control the over expression

of heme proteins using controlled intake of Isopropyl β-D-thiogalactopyranoside (IPTG) (Karow *et al.*, 2004). However, these methods either need expensive chemicals like IPTG and δ-ALA or specialized expression techniques (Woodard and Daily, 1995; Jung *et al.* 2001; Karow *et al.* 2004; Varnado and Goodwin, 2004; Liu *et al.* 2007; Sudhamsu *et al.* 2010). On the other hand, lactose, a disaccharide present in milk, has been used instead of IPTG as an inducer in many over expressions studies previously (Hwang *et al.*, 2000; Kim *et al.*, 2007). However, lactose was rarely used to express heme proteins, although it has several advantages over IPTG. Only one example was found in the literature, where lactose was used to express cytochrome P450 in a fermenter (Woyski and Cupp-Vickery, 2001). Since then, there is no example of expression of heme proteins using lactose as an inducer at the laboratory scale except a very recent report (Zhang *et al.*, 2014).

To find a simple, low cost, efficient, protein expression system we tested lactose as an inducer for heme proteins and we selected sensor of nitric oxide (SONO) from *V. cholerae,* which belongs to the H-NOX family of heme proteins, as a target protein to test our hypothesis. SONO has already been cloned, purified and partially characterized (Karow *et al.*, 2004). This protein is a bacterial homologue of the heme binding domain of soluble Guanylate Cyclase (sGC) and senses nitric oxide (NO). Although the biological function of this SONO is not well understood, it has been proposed that this protein modulates symbiotic colonization in *Vibrio* species (Wang *et al.*, 2010). In this report, we present the expression of this highly soluble holo-protein using lactose and simple two step purification of the protein, to achieve higher yields.

Materials and methods

Genomic DNA of *V. cholerae* (strain MO10) was kind gift from Dr. Ranjan Nandi, NICED, Kolkata, India. This gDNA was used to amplify the SONO from *V. cholerae*. Chemicals were either from Sigma (USA) or local sources. Primers were synthesized at Bioserve, Hyderabad, India. Restriction and other enzymes were from NEB (USA). Preparation of media and other basic molecular biology experiments were performed following the protocols described by Sambrook and Russell, 2001. Chromatography materials for protein purification were from GE Healthcare (USA).

Construction of expression vector containing SONO from V. cholerae gene: Construction of truncated version (residues 11-191) of sensor of nitric oxide from *V. cholerae* was done as per literature (Karow *et al.*, 2004). In brief, putative sequence of this gene (VCA0720) was amplified by PCR using 5'-GGAATTCCATATGCAAGGCA TTATCTATACCGTTCTC-3' as forward primer and 5'-ATAGTTTAGCGGCCGCTCATGATGGC AAAAATTCCACTAC-3'as reverse primer. Amplified products were inserted into pET 20b expression vector using *Nde*I and *Not* I site. A stop codon was used at the end and the construct was without 6x-His tag.

Expression and purification of SONO from V. cholerae: The expression vector containing the gene was transformed into *E. coli* BL21 (DE3) pLysS. Single colony was chosen for over expression and was grown overnight in 25 ml LB media containing 100 µg/ml ampicillin and 34µg/ml chloramphenicol. 500 ml 2XYT media containing same antibiotics were inoculated with 1% of the overnight pre-culture and grown at 37°C with 200 rpm till the OD_{600} reached about 1.5. Then lactose (1.0 mM to 5 mM, final concentration) was added in the culture and further shaken for 5-7 h at 37°C at 100 rpm. Cells were then harvested and kept frozen at -80°C till further use.

Frozen cells were thawed on ice and lysed in 50 mM phosphate buffer, pH 6.0 containing 5 mM 2-mercaptoethanol (2-ME) (lysis buffer). After complete lysis, 10µg/ml DNase and 5 mM $MgCl_2$ was added and incubated till all DNA was digested. Then the lysate was centrifuged at 17,000 rpm for 30 min and the brown colored, clear supernatant was decanted and loaded on a DEAE sepharose column pre-equilibrated with the lysis buffer. Unbound *E. coli* proteins were washed out first with lysis buffer and then with lysis buffer containing 50 mM NaCl. Finally proteins were eluted in a linear gradient of 50-1000 mM NaCl. Fractions containing protein of interest were concentrated using amicon

concentrator, reduced using pinch of solid dithionite and centrifuged and the supernatant was applied on a superdex 75 (nearly 120 ml) column, pre-equilibrated with 50 mM Glycine-NaOH buffer, pH 10.5 containing 10 mM 2-ME and 10% glycerol. Fractions containing brown color were checked on SDS-PAGE and stained with Coomassie brilliant blue for visualization using standard protocol. Protein estimation was done using Bradford's reagent supplied by BioRad using BSA solution at varied concentrations as standards. Heme protein ratio and the state of the heme iron of the purified protein were judged by UV-Vis spectra in quartz cuvette of 1 cm path length. To prepare NO-complex, purified protein was taken in a glass-vial of nearly 1 cm path length with rubber septum. Protein solution inside the vial was degassed and purged with nitrogen gas several times. Small volume of degassed and nitrogen purged sodium dithionite solution was added. Finally, chemically prepared NO gas was added. Spectra were recorded in the visible region. Molecular mass was calculated from the elution profile of gel filtration column using standard markers.

CD Spectroscopy: The secondary structure of protein was determined using far UV CD spectroscopy. CD spectra of 0.1 mg/ml protein between 195 nm to 250 nm wavelengths range were recorded in 0.1 cm path length cuvette using a JASCO J-815 spectropolarimeter. Wavelength scans were carried out in the ellipticity mode at a scan speed of 100 nm/min, bandpass of 2 nm, at response time of 2 s and wavelength step of 0.1 nm. All spectra recorded are an average of 4 accumulations. All spectra were corrected for buffer baseline by subtracting the respective blank spectra recorded identically without the protein.

Homology Modelling: Full sequence of SONO from *V. cholerae* containing 191 residues was submitted for automated modelling in Swiss-Pro modelling server (Schwede *et al.* 2003; Arnold *et al.* 2006). Automated search resulted in a model containing 179 residues and 10 residues from the N-terminus and 2 residues from the C-terminus were removed by the software. Model was visualized using PyMol.

Results and Discussions

Cloning, expression and purification

Multiple sequence alignment of SONO from *V. cholerae* with other H-NOX proteins showed that ten extra N-terminal residues did not align with other proteins used (Karow *et al.*, 2004). The eleventh residue, which happened to be a methionine and the remaining sequence aligned well with other H-NOX domains from bacterial origin. Therefore, N-terminally truncated protein containing 181 residues were cloned in pET 20b without any purification tag. Expressed protein was highly soluble and purified using simplified two step purification protocol. Different lactose concentrations (1mM – 5mM) resulted in similar levels of expressions. The theoretical pI of the protein is around 5 and therefore, the pH of binding buffer of DEAE anion exchange chromatography was reduced to 6.0 to get rid of the *E. coli* proteins. This resulted in highly enriched desired protein after ion exchange column purification. Further size exclusion chromatography on Superdex 75 at pH of 10.5 resulted in several well resolved peaks separating the SONO protein from other impurities (Figure 1A). The protein eluted at 66 ml as a sharp and symmetric peak and had a dark brown colour (Figure 1A). SONO from *V. cholerae* thus purified was highly pure and homogenous as judged on SDS-PAGE and appeared as a single band of about 21 kDa and matched well with the calculated molecular mass of 20.6 kDa (Figure 1B). The yield was ~12 mg/l of initial culture which was nearly double compared to the yield reported earlier (Karow *et al.*, 2004). Molecular mass calculation using gel filtration showed that the protein was about 40 kDa, indicating dimer in solution. Electronic absorption spectra showed the Soret band at 423 nm (Figure 2), with A_{423}/A_{280} of more than 2.2. The Q-band appeared between 500 and 600 nm. In NO bound state, the Soret band appeared at 399 nm indicating the cleavage of proximal Fe-His bond (Pal *et al.* 2013). This resembles sGC as this proximal bond cleavage at the N-terminal structurally rearranges the catalytic site at the C-terminal (Pal and Kitagawa, 2004). Previously, this protein was purified as a mixture of ferrous and ferric heme with a Soret band at 423 nm which shifted to 429
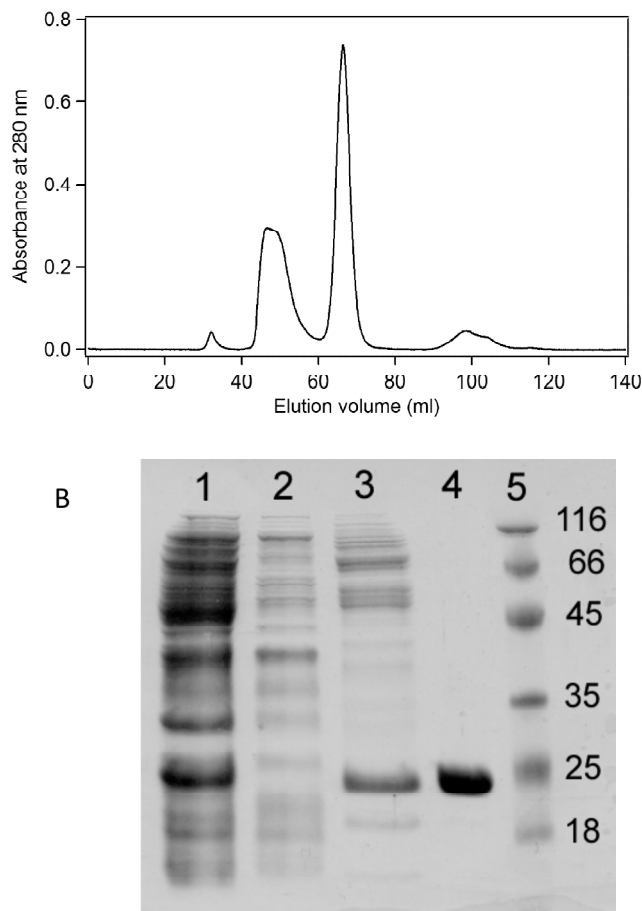
Figure 1: **Chromatographic and electrophoretic analysis of SONO. A. Size exclusion chromatography profile of SONO from *V. cholerae*. B. SDS-PAGE analysis of purification of SONO from *V. cholerae* Lane1 : cell lysate, lane 2 : flow through after loading on DEAE column, 3: elute from DEAE column, 4: elution from Superdex 75 column, 5: Molecular mass marker (kDa).**
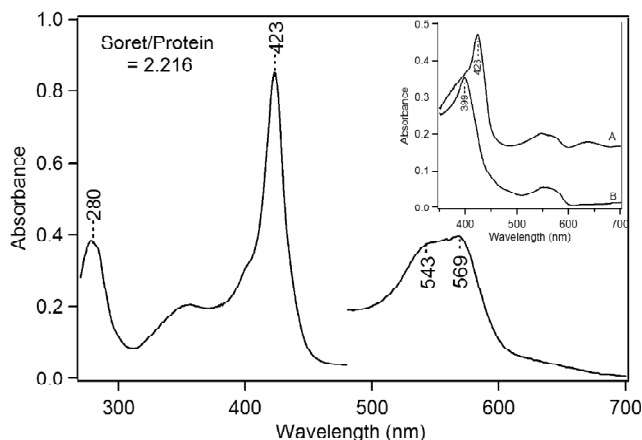

Figure 2: **Absorption spectra of SONO from *V. cholerae* in 50 mM glycine-NaOH buffer, pH 10.5. Spectra were recorded using quartz cuvette of 1 cm path length in a Perkin Elmer Lambda 25 UV-VIS spectrophotometer. 481nm-700 nm region has been enhanced 5 times. Inset shows the visible region of the protein in the absence (A) and presence (B) of NO.**

nm upon reduction with dithionite. Comparison of our results with the previously reported results by Karow *et al.*, 2004, shows that the heme is incorporated stoichiometrically into the polypeptide chain.

Incorporation of heme into the protein

Incorporation of heme during expression is carried out using one of the widely used expression system, pET20b vector containing T7 promoter. Normally, expression is induced by the addition of IPTG, which will start the production of T7 RNA polymerase, which in turn, starts transcription of gene under the control of T7 promoter producing the desired protein of interest. IPTG binds irreversibly to the *lac* repressor and thus induces the conformational change of the repressor structure and preventing it from binding with DNA. Thus, the effect of IPTG is irreversible and not controllable. Therefore, in another study, tuner cells were used to control the level of induction by controlling the intake of IPTG inside the cell (Karow *et al.*, 2004). As an alternative to this method, lactose could be used as an inducer and it is consumed by the cells. Therefore, the induction using lactose is a reversible phenomenon. This could give us a tool to control the level of expression of protein as well as the biosynthesis of heme and eventually the proper incorporation of heme. Our results show that this protein could be expressed and purified as a holo-protein with stoichiometrically incorporated heme using lactose as an inducer. To our best of our knowledge, this is the first report of expression of a heme protein using lactose as an inducer at the laboratory scale (during review process it was found that one more report has been published by Zhang *et al.*, 2014, showing similar results).

CD spectrum

There is no assay to characterize the function of this protein. This protein is a homolog of heme binding domain of sGC, a mammalian protein. sGC has two non-identical chains, each chain having a sensory domain, a PAS-like domain, a coiled coil/dimerization domain and a catalytic domain. The sensory domain of beta-subunit binds heme and the SONO from *V. cholerae*

represents only the sensory domain of beta subunit. Since there is no catalytic domain, this protein does not have any activity. Other catalytic activity of this protein, if any, is not yet known. Therefore, to check the proper folding of the protein, circular dichroism spectrum in the far UV wavelength range between 195 and 250 nm was measured. CD spectrum in this range is mostly contributed by the backbone conformation of the polypeptide chain indicating secondary structural contents of the protein. CD spectrum shows a negative maximal peak around 210 nm (Figure 3) indicating the protein is predominantly helical (Sreerama and Woody, 2004).
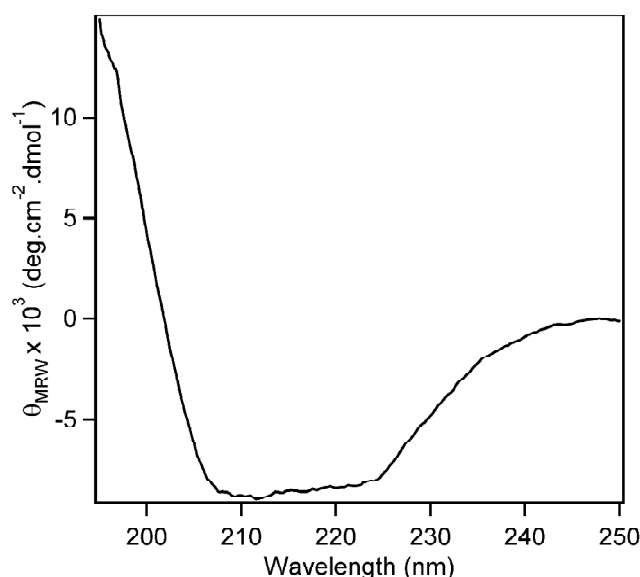

Figure 3: **CD spectrum of SONO from *V. cholerae*. Spectrum presented here is collected using a JASCO J-815 spectropolarimeter and is an average of 4 accumulations. The spectrum was collected at a protein concentration of 0.1 mg/ml, path length of 0.1 cm, scan speed of 100 nm/min, response time of 2 s and wavelength step of 0.1 nm .**

Homology Model

To compare the CD results with structure, a homology model of SONO from *V. cholerae* was prepared. Full length SONO from *V. cholerae* had 191 amino acid residues and all the residues were submitted to Swiss-Pro modelling server in automated mode. This resulted in a model containing 179 residues from 11 to 189 of the original sequence. The template selected (PDB ID 2o09B, H-B-chain of HNOX domain from *Nostoc sp.* PCC 7120) had a sequence identity of 25% for 179 residues. Heme was added in the model based on the template (Figure 4) and the His114 residue
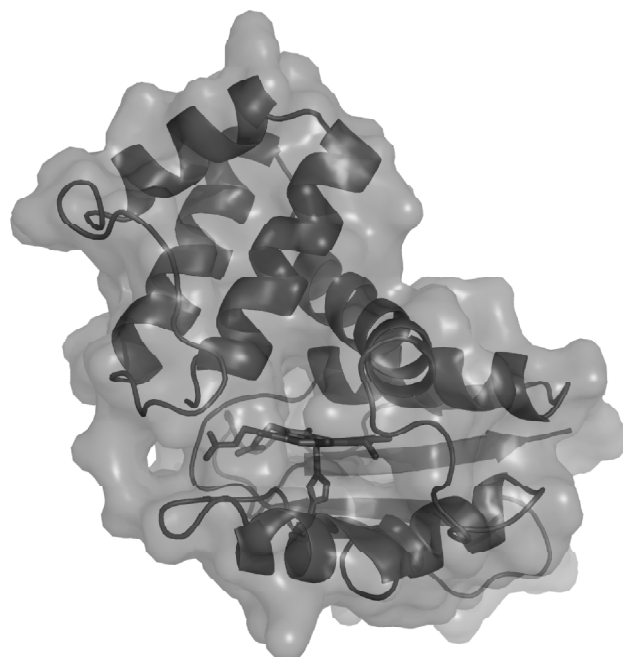

Figure 4: **Homology model of SONO from *V. cholerae* (residues 11-189). Light gray shows the envelope of the protein. Dark gray shows the secondary structural elements of the protein.**

was within a bonding distance of 2.32 Å from the heme iron. Modelled protein was predominantly helical with 44% α-helix, 18% β-sheet and the rest 38% was not well-defined. This matched well with the CD spectrum of SONO from *V. cholerae*.

In conclusion, we have optimized the expression and purification of a heme protein in a simple, cost effective way with higher yield. Heme is incorporated into the protein in stoichiometric ratio. This could serve as a platform for other heme protein expression and purification and could be extremely useful for industrial application.

Abbreviations

δ-ALA, δ-aminolevulinic acid; IPTG, Isopropyl β-D-thiogalactopyranoside; NO, nitric oxide; SONO, sensor of nitric oxide; 2-ME, 2-mercaptoethanol; sGC, soluble Guanylate Cyclase.

References

Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics* 22, 195-201.

Hwang, T., Fu, H., Lin, L. and Hsu, W. (2000). High-level expression of *Trigonopsis variabilis* D-amino acid oxidase in *Escherichia coli* using lactose as inducer. *Biotechnol Lett.* 22, 655-658.

Jung, Y., Kwak, J. and Lee, Y. (2001). High-level production of heme-containing holoproteins in *Escherichia coli*. *Appl Microbiol Biotechnol.* 55, 187–191.

Karow, D. S., Pan, D., Tran, R., Pellicena, P., Presley, A., Mathies, R. A. and Marletta, M. A. (2004). Spectroscopic characterization of the soluble Guanylate Cyclase-like heme domains from *Vibrio cholerae* and *Thermoanaerobacter tengcongensis. Biochemistry* 43, 10203-10211.

Kim, M., Elvin, C., Brownlee, A. and Lyons, R. (2007). High yield expression of recombinant pro-resilin: Lactose-induced fermentation in *E. coli* and facile purification. *Protein Expr Purif.* 52, 230-236.

Kim, Y., Babnigg, G., Jedrzejczak, R., Eschenfeldt, W. H., Li, H., Maltseva, N., Hatzos-Skintges, C., Gu, M., Makowska-Grzyska, M., Wu, R., An, H., Chhor, G. and Joachimiak, A. (2011). High-throughput protein purification and quality assessment for crystallization. *Methods* 55, 12–28.

Liu, W., Rogge, C. E., Kamensky, Y., Tsai, A. L. and Kulmacz, R. J. (2007). Development of a bacterial system for high yield expression of fully functional adrenal cytochrome b_{561}. *Protein Expr Purif.* 56, 145–152.

Pal, B. and Kitagawa, T. (2005). Interactions of soluble guanylate cyclase with diatomics as probed by resonance Raman spectroscopy, *J. Inorg. Biochem.* 99, 267–279.

Pal B., Tanaka, K., Takenaka, S., Shaik, T.B. and Kitagawa, T. (2013). Structural characterization of nitric oxide-bound soluble Guanylate Cyclase using resonance Raman spectroscopy, *J. Porphyrins Phthalocyanines* 17, 240–246.

Sambrook. J. and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server, *Nucleic Acids Res.* 31, 3381-3385.

Sreerama, N. and Woody, R. W. (2004). Computation and analysis of protein Circular Dichroism spectra. *Methods Enzymol.* 383, 318-351.

Sudhamsu, J., Kabir, M., Airola, M. V., Patel, B. A., Yeh, S. R., Rousseau, D. L. and Crane, B. R. (2010). Co-expression of ferrochelatase allows for complete heme incorporation into recombinant proteins produced in *E. coli. Protein Expr Purif.* 73, 78-82.

Varnado, C. L. and Goodwin, D. C. (2004). System for the expression of recombinant hemoproteins in *E. coli. Protein Expr Purif.* 35, 76-83.

Wang, Y., Dufour, Y. S, Carlson, H.K., Donohue, T.J., Marletta, M. A. and , Ruby, E. G. (2010). H-NOX-mediated nitric oxide sensing modulates symbiotic colonization by *Vibrio fischeri*. Proc Natl Acad Sci USA. 107, 8375-8380.

Woodard, S. I. and Dailey, H. A. (1995). Regulation of heme biosynthesis in *Escherichia coli. Arch Biochem Biophys.* 316, 110-115.

Woyski, D. and Cupp-Vickery, J. R. (2001). Enhanced expression of cytochrome P450s from *lac*-based plasmids using lactose as the inducer. *Arch Biochem Biophys.* 388, 276-280.

Zhang, R. Li, Z. Shi, W. and Xu, R. (2014) Expression, purification, and characterization of fusion protein TAT-cytoglobin, *Sheng Wu Gong Cheng Xue Bao.* 30, 1247-1255.