DE GRUYTER OPEN

Nova Biotechnol Chim (2017) 16(1): 1-11
DOI: 10.1515/nbec-2017-0001

Nova Biotechnologica *et* Chimica

How to approach heterogeneous protein expression for biotechnological use: An overview

Daniela Jamrichová✉, Lenka Tišáková✉, Veronika Jarábková and Andrej Godány

Department of Biology, Faculty of Natural Sciences, University of SS. Cyril and Methodius in Trnava, Nám. J. Herdu 2, Trnava, SK-917 01, Slovak Republic

Article info

Article history:
Received: 8th June 2017
Accepted: 26th June 2017

Keywords:
Escherichia coli
Gene expression
Protein purification
Heterologous

Abstract

Production of recombinant proteins in *Escherichia coli* expression systems has shown many advantages, as well as disadvantages, especially for biotechnological and other down-stream applications. The choice of an appropriate vector depends on the gene, to be cloned for purification procedures and other analyses. Selection of a suitable production strain plays an important role in the preparation of recombinant proteins. The main criteria for the selection of the host organism are the properties of the recombinant produced protein, its subsequent use and the total amount desired. The most common problems in eukaryotic gene expression and recombinant proteins purification are, for instance, post-translational modifications, formation of disulphide bonds, or inclusion bodies. Obtaining a purified protein is a key step enabling further characterization of its role in the biological system. Moreover, methods of protein purification have been developed in parallel with the discovery of proteins and the need for their studies and applications. After protein purification, and also between the individual purification steps, it is necessary to test protein stability under different conditions over time. Shortly, all the essential points have been briefly discussed, which could be encountered during production and purification of a recombinant protein of interest, especially from eukaryotic source and expressed heterogeneously in prokaryotic production system.

Introduction

Over years, the production of recombinant proteins in the *Escherichia coli* expression system has shown many advantages, as well as disadvantages. There is the evidence of such recombinant protein evolution reflecting in a wide range of commercially available production *Escherichia coli* strains with modified properties for the heterogeneous production of eukaryotic proteins under various conditions. The major advantage of an expression system with *E. coli* as a host organism is time, technical and financial effectivity. *E. coli* cells can be propagated under common laboratory conditions in a simple culture medium with selected ions, salts and carbon sources. In the exponential phase, its generation time is 20 min, in a complex culture medium containing vitamins, salts, microelements, amino acids and glucose, at an intensive passage at 37 °C (Sezonov *et al.* 2007). It is doubtless that the production of recombinant proteins in various microbial systems (especially those of eukaryotic origin), has revolutionized the whole biochemical and biotechnological overview of the basic approach how to obtain high yield of protein

✉ *Corresponding author:* jamrichova.daniela@gmail.com, lenka.tisakova@ucm.sk

in a decent purity. Many scientific studies focused on application output of proteins have encountered obstacles with heterogeneous expression of eukaryotic genes in prokaryotic systems, because of miscellaneous reasons. This paper could serve as brief overview that pays attention to crucial steps and potential hinders in production of eukaryotic recombinant proteins in heterogeneous *E. coli*, methods of isolation, purification and detection of the (overexpressed) recombinant proteins, especially (Fig. 1).

Gene design

When planning production of recombinant protein, it is inevitable to take into consideration several steps in preparation of the gene cloned. Principally, before starting the experimental part, it is important to apply bioinformatics approaches to evaluate the usability/suitability of the gene of interest for the planned downstream procedures. Various *in silico* analyses could be conducted by online programs and tools, such as restriction digestion analysis of the gene, primer design for amplification of the gene, primer melting temperature calculation, verification the open reading frames (ORFs). It should be emphasized that *E. coli* codon usage evaluation is one of the key factors for heterogeneous expression, so it must be performed prior to gene design. Nowadays, several commercial companies provide codon usage optimization as one of their bonus services along with gene synthesis in a selected type of vector. In the case of using a synthetic gene for recombinant protein production, much time could be saved in cloning procedures and the real starting point in the laboratory is the expression part. Last but not least, when thinking about fusion protein preparation, it is good to evaluate, whether to design a (synthetic) gene with or without a stop codon. Finally, hasty decision while designing a gene, cloning the gene or directly ordering a synthetic one, could rapidly decrease the rate of getting relevant results and increase the sources needed to reassess or repeat the procedures.

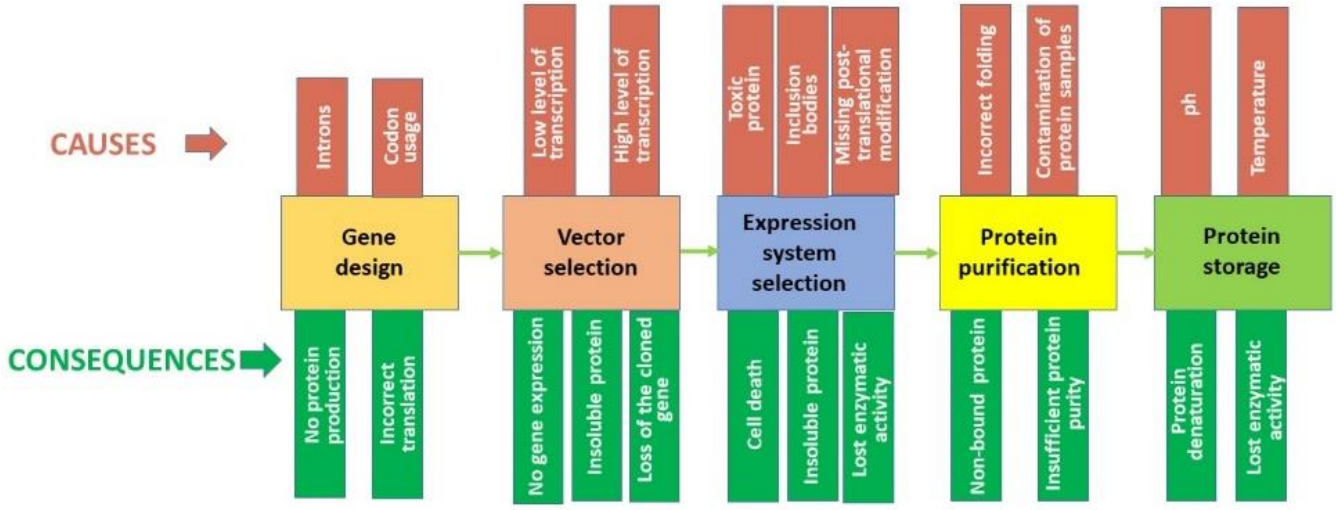

Fig. 1. Overview of the steps (in blocks) for heterogeneous protein production with all the essential steps involved – obstacles during protein production (bars above blocks) and their consequences (bars below the corresponding blocks).

Vector selection

The choice of an appropriate vector for successful down-stream procedures depends on the gene to be cloned. All expression vectors must contain three important signal sequences including promoter, terminator and ribosome binding site (RBS). Currently, the most common expression vectors are designed as a combination of replicon, promoter, RBS, selection marker, multicloning site (MCS), affinity tag to facilitate purification, and sometimes a fusion protein to increase protein solubility. There is a wide range of commercially available expression vectors, so the selection of a suitable

vector for a specific gene of interest to be cloned must be not only complex, but also substantial. Therefore, attention should be paid to the needs of the (individual) gene to be cloned and expressed, when searching for the most suitable expression vector. The promoter is a cardinal part of the expression vector, as a promotor sequence regulates the initial stage of gene expression, RNA polymerase attachment, and controls the rate of mRNA synthesis. There are a number of vectors on the market with different types of promoters, while the most commonly used promoter types are listed in Table 1.

Escherichia coli expression systems

For cloning under laboratory conditions and in industrial scale-up production technologies, non-pathogenic *E. coli* strains have become the starting point for the developing new strains with features optimized for gene cloning or recombinant protein expression and production. The fundamental task in the heterologous protein production in bacterial expression systems is the correct folding of the produced protein. Since the *E. coli* production strains lack the systems necessary to handle more complex eukaryotic proteins or to further modify them (e.g. glycosylation), special strains have been developed, e.g. Origami that supports the formation of disulphide bridges in the cytoplasm (Derman *et al.* 1993) or Rosetta (Novagen, USA) with modified tRNA codes providing easier expression of eukaryotic genes in the prokaryotic systems (Guillaume *et al.* 2006; Nakamura *et al.* 2008). Selection of a suitable production strain plays an important role even for following downstream procedures. The main criteria for the selection of the host organism are the properties of the recombinant protein produced, its subsequent use and the total amount. Inappropriate choice of the production strain may cause the recombinant protein to be inactive, insoluble, degraded, or can only be produced in limited amounts. Therefore, *E. coli* strains have been developed that lack nucleases and proteases, to prevent degradation of the recombinant DNA and the produced protein. A variety of *E. coli* strains with various desired properties have been developed by mutations of specific genes (Baneyx 1999; Yokoyama 2003; Klock *et al.* 2005; Terpe 2006). The most common *E. coli* strains include BL21 (derived from the B strain) and derivatives thereof. BL21 strains are non-pathogenic, capable of growth in simple culture media, allowing higher gene expression under the T7 bacteriophage promoter integrated into the bacterial chromosome of the bacterium (Studier and Moffatt 1986). Nowadays, there are many biotechnologically and therapeutically used proteins produced in *E. coli* cells (Chrastilová *et al.* 2007).

Selection of a host strain

When using *E. coli* as a production organism for expression of recombinant proteins, some drawbacks could be encountered, directly derived from its inherent properties, especially due to its prokaryotic nature. There are some typical and commonly occurring problems in expressing recombinant genes and producing recombinant proteins.

Post-translational modifications

Protein post-translational modifications (phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis) increase the functional diversity of the proteome by the covalent addition of functional groups or proteins, proteolytic cleavage of regulatory subunits or degradation of entire proteins. These modifications generally influence almost all aspects of normal cell biology and pathogenesis, therefore, identifying and understanding PTMs it is critical to identify and understand them, when focusing on gene expression and protein production.

There are no posttranslational modifications occurring during the protein production in *E. coli*, such as N-glycosylation, C-glycosylation, acetylation, phosphorylation, which might be indispensable for proper functions of eukaryotic proteins (Stevens 2000; Yin *et al.* 2007). This problem can be addressed by using evolutionary higher systems, for example, yeasts, or mammalian cells. From another point of view, lack of post-translational modifications can also be

beneficial, when protein production in non-glycosylated form is more suitable for subsequent applications, for instance, protein crystallisation.

Formation of disulphide bonds

To ensure catalysis of protein disulphide bonds (Dsb) formation, there have been evolved elaborate Dsb systems in *Escherichia coli* (Inaba 2009). Disulphide bonds are important for protein structure and functionality but are only formed in the periplasm under catalysis of the Dsb system (Andersen *et al.* 1997; Bardwell 1994). Therefore, if disulphide bond formation is required, the recombinant protein should be directed to the periplasmic space by means of a cleavable signal peptide (e.g. pelB) (Chung *et al.* 1998; Le-Calvez *et al.* 1996). Such a signal is used by W3110 production strains as well as by some vectors, for example the pET26b (Makrides 1996; Matthey *et al.* 1999). The lower abundance of proteins in the periplasm facilitates the recombinant protein purification (Choi and Lee 2004). However, only small yields of the final proteins are usually obtained in the periplasm, which is marked as the major drawback of this type of production. A more convenient approach is the use of specific, commercially available strains, for example Origami strain (Novagen). These strains involve mutated genes encoding for thioredoxine reductase (trxB) and glutathione reductase (Gor) in the Dsb system, allowing cytoplasmic cysteine reduction by thioredoxin and glutathione redoxine (Bessette *et al.* 1999). If further formation of disulphide bonds is needed, the gene for thioredoxin can by fused with the gene of interest and subsequently expressed in the trxB / gor *E. coli* strain (Lavallie *et al.* 1993). If the protein does not get the correct tertiary structure (folding), it is usually insoluble and creates a so-called inclusion bodies within the bacterial host (Chatterjee and Esposito 2006).

Table 1. Digest of the most commonly used promoters in *E. coli* expression vectors.

Promotor Type	Features	References
Lac	Inducer: IPTG. *Lac* promoters and their lacUV5 derivatives are weak promoters therefore not often used for recombinant proteins production.	Müller-Hill *et al.* (1968), Calos (1978), Deuschle *et al.* (1986), Makoff and Oxer (1991)
Trp	Inducer: indolyl-3-acrylic acid. It is used to produce majority of protein.	Hallewell and Emtage (1980), Tacon *et al.* (1983), Yansura (1990)
Tac	Inducer: IPTG. A hybrid type of promoter prepared by joining (-35) *trp* and (-10) *lac* promoter regions. *Tac* is 10 fold stronger than lacUV5 promoter.	De Boer *et al.* (1983)
λP_L	Inducer: heat shock. A strong type of promotor with wide use in industrial biotechnology at higher fermentation temperatures.	Menart *et al.* (2003), Dodd *et al.* (2005), Valdez-Cruz *et al.* (2010)
T7	Inducer: lactose or IPTG. Basal expression is controlled by lacIQ as well as co-expression of T7 lysozyme.	Moffatt and Studier (1987), Baneyx (1999), Graumann and Permstaller (2006), Stano and Patel (2004)

Formation of inclusion bodies

Proteins in inclusion bodies are inactive, aggregated and insoluble. They usually require denaturation and subsequent renaturation,

re-folding, which complicates protein production and reduces yields (Graslund *et al.* 2008). On the other hand, if there is no problem with protein renaturation, the inclusion bodies can be easily separated by centrifugation since they contain only small amounts of contaminants, and ultimately, the higher purity of the recombinant protein (Baneyx and Mujacic 2004; Cabrita and Bottomley 2004). Formation of inclusion bodies can be suppressed by slowing expression of the (trans)gene. This can be achieved in different ways e.g. by reducing the cultivation temperature, using a weaker promoter, application of lower concentrations of inducer. Slowing down the formation of the protein increases the chance for its proper folding.

Presence of lipopolysaccharides

Another obstacle in the production of proteins in *E. coli* may be the presence of lipopolysaccharides which constitute a part of the cell wall of gram-negative bacteria. These lipopolysaccharides have pyrogenic effects and cause complications in the purification of end products. Therefore, de-pyrogenation is needed (Terpe 2006; Yin *et al.* 2007), despite the reduced final yield of the pure protein and possible adverse effect on protein stability and native conformation (Petsch and Anspach 2000).

Toxicity of the recombinant proteins

Toxic (recombinant) proteins cause death of the host cells or complication during their cultivation. Since the toxicity for the given expression strain can be specific, the strain can be replaced by a more compatible one for protein production. Another option is the use of a highly regulated expression system, where the expression of the gene of interest is regulated by an inducible promoter and transcription terminator. Controlled protein production can result from keeping the plasmid copy number low, while protein toxicity can by modified by engineering the recombinant protein sequence (Saida 2007).

Nucleotide sequence of the gene of a foreign origin

The eukaryotic gene cloned may contain introns. Since *E. coli* lacks the necessary splicing mechanism for eukaryotic mRNA, production of active protein poses a serious problem in *E. coli*. This can be overcome by using complementary DNA (cDNA) prepared from mature mRNA or using a synthetic gene (Le Hir *et al.* 2003).

Purification of recombinant proteins

Obtaining a purified protein from expression systems is often a key step enabling its characterization and clarification of its role in biological systems. Purification of recombinant proteins represents a series of procedures that lead to the obtaining of a single protein species from complex bacterial lysates.

Methods of protein purification have been developed in parallel with the discovery of proteins and the need for their further study. Until the beginning of the 20th century, filtration, precipitation and crystallization were mainly used to isolate proteins. The breakthrough occurred in 1903 when Cvet separated plant pigments in a calcium carbonate column and subsequently, in 1906, introduced the concept of chromatography for this method. Nearly two decades later, in 1924, Svedberg proved that proteins can be separated by analytical ultracentrifugation (Philo 1997). In the course of several years, other methods of protein separation were discovered, namely electrophoresis, affinity chromatography (AC) and ion exchange chromatography. Since the mid-20th century, there has been an intensive development of various types of chromatographic matrices. Initially, starch or modified dextran (Sephadex) were only used to allow the separation of proteins by size (Porath and Flodin 1959). To obtain a protein of even a higher purity, other matrices such as polyacrylamide (Raymond and Weintraub 1959), methyl acrylate (Curtain 1963) or agarose (Hjrtén 1964) have been developed that have several advantages over the starch gels like higher stability or easier handling. On the other hand, some disadvantages could emerge, e.g. toxicity of the acrylamide, which comes from a nature of this substance.

Protein and peptide tags

To facilitate detection and purification of the recombinant protein, protein or peptide tags (labels) have been implemented. These tags represent exogenous sequences with high affinities to specific biological or chemical ligands (Bornhorst and Falke 2000; Arnau 2006*a*). Affinity labelling allows for simple detection of proteins with low immunogenicity or no antibodies available (Brizzard 2008). Protein tags have been proven very suitable in protein purification where the labelled protein can be purified directly from the cell lysate or supernatant (Lichty 2005). Affinity tags may be located at the N- or C-terminus of the recombinant protein (Justesen *et al.* 2009), for some specific applications it is possible to place the tag even on both ends. The use of tags can improve the biochemical properties and increase the yield of the produced protein, simplify renaturation conditions, increase the solubility of the produced protein, or prevent its proteolysis. However, the use of affinity tags may also have adverse impacts on labelled proteins since, for example, can alter the conformation of the labelled protein, inhibit enzyme activity, cause toxicity, alter biological function and reduce the protein yield (Arnau 2006*b*). Therefore, when choosing an affinity tag, several inevitable points have to be considered to achieve the desired biochemical features of the protein of interest. These include the bacterial production strain used, the conditions of purification, the desired purity and the final amount of protein.

Table 2. The most commonly used stabilizing agents for protein storage.

Class	Examples	Value / Usage
Buffers	Phosphate buffer, Tris-HCl, HEPES, HEPES – NaOH, etc.	Stability
Salts	KCl, NaCl, $(NH_4)_2SO_4$	Stability
Ionic detergents	Sodium deoxycholate, Sodium cholate, CTAB, SDS	Stability and increase protein solubility
Non-ionic detergents	Triton X-100, Nonidet P-40, Tween 20, Brij 35, etc.	
Cryoprotectants	Glycerol, Saccharose	Storage and stability
Chelating agents	Citric acid, salicylic acid, EDTA, EGTA	Stability
Reducing agents	β-Mercaptoethanol (BME), dithiotreitol (DTT), tris(2-carboxyethyl)phosphine (TCEP)	Stability, prevents formation of disulphide bonds
Ligands, cofactors	Mg^{2+}, ATP phosphate	Stability
Protease inhibitors	PMSF (phenylmethylsulfonyl fluoride), TPCK (tosyl phenylalanyl chloromethyl ketone), TLCK (tosyl-L-lysyl-chloromethane hydrochloride)	Prevents protein degradation

Chromatographic methods for protein purification

Presently, various types of liquid chromatography are used to purify recombinant proteins. These differ in the principle of separation of the purified protein and include affinity-, ion exchange-, gel permeation-, or hydrophobic chromatography. Due to the large number of available methods

and various chromatographic matrices it is necessary to select the purification method prior to the construction of the recombinant molecule. Purification methods based on protein-peptide fusion or on a fusion protein with high affinity for a particular ligand enable using a simple affinity chromatography. Other protein purification methods involve various interactions, such as antibody–antigen, enzyme–substrate, or protein–metal ion. To facilitate purification, modern expression vectors contain specific sequences enabling better detection and/or subsequent purification of expressed proteins, or polypeptide sequences that affect protein solubility.

The FLAG was the first tag and represents a hydrophilic sequence of eight amino acids (DYKDDDDK) recognized by a monoclonal antibody. This antibody allows detection of the fusion protein and its purification by immobilization on a suitable chromatographic matrix (Hopp *et al.* 1988). Currently a histidine tag (His-tag, His6-tag) consisting of 6-12 residues is routinely used for protein purification (Janknecht *et al.* 1991; Lichty 2005). The high-affinity binding of the imidazole group of histidine to a transition metal like Ni^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+} or $Zn^{2+}$is the basis of a method known as Immobilized Metal Affinity Chromatography (IMAC) that has become one of the basic purification procedures for the isolation of recombinant proteins. The metallic ions are immobilized on an agarose or Sepharose matrix by chelation with imidodiacetic acid (IDA), nitrilotriacetic acid (NTA) or tris (carboxymethyl) ethylenediamine (TED) (Porath *et al.* 1975; Hochuli *et al.* 1987). Such a matrix is used for purification of His-tagged proteins under both native and denaturation conditions to yield up to ten mg of protein per 1 mL of a matrix. The protein can be eluted from the matrix with buffer containing imidazole (usually 150–300 mM), EDTA (eluting with metal ions) or by lowering the buffer pH to values in the range 2.5–7.5. The procedure itself and selection of the most suitable metal ion should be optimized separately for each individual protein based on its physicochemical properties (Bolanos-Garcia and Davies 2006). Noteworthy, washing out the metal from the matrix can cause insufficient purity of the protein. This can be avoided using a so called pentadentate (TED) matrix, which binds the metal ions very tightly and thus prevents their leaching, leaving only one free spot for the protein to bind. If affinity chromatography fails to provide the product of the desired purity, other chromatographic methods should be used in subsequent steps, e.g. ion exchange chromatography that separates the molecules by their charge. The stationary phase consists of matrices from dextran derivatives (Sephadex) or agarose (Sepharose) with alkaline functional groups – anionic exchangers, or with acidic functional groups – cationic exchangers. Ion exchange chromatography has been frequently used not only for protein purification but also for its concentrating. Another option for protein purification is a gel permeation chromatography method, also called a molecular sieve, which separates proteins based on their molecular size and also could be used for desalination of protein samples.

Monitoring the concentration and purity of proteins

Determination of the purity of a recombinant protein is of key importance for its subsequent use. However, a direct method for quantification of the purity of the protein sample is missing. The available procedures always start assessing the amount of individual types of impurities such as of nucleic acids, carbohydrates, lipids, isoenzymes, inactive enzymes. For this operation, the proper type of chemical/physical method should be chosen. The protein purity value is dependent on the type of test method, as well as its sensitivity (detection limit). Therefore, several factors are taken into account when deciding for a particular approach, in particular, the amount of protein available (yield and concentration), the nature of the impurities, and the type of solution, in which the protein is re-suspended. The number of analytical methods for monitoring the purity of proteins include colorimetry, electrophoresis, mass spectrometry, dynamic light scattering and enzyme analysis (Rhodes *et al.* 2009).

Storage of purified proteins

In terms of maintaining the protein quality required for further procedures, storage of purified proteins is crucial. Storage conditions depend on the type of protein, therefore there is no general rule and procedure available. Importantly, protein samples should not be exposed to extreme changes in pH or temperature, and other factors that could potentially lead to protein denaturation. Long-term storage of the protein solution at room temperature also requires the use of antibacterial and antimycotic agents to avoid biological contamination and destruction of the protein sample.

During and after protein purification it is necessary to test the stability of the protein under different conditions, simplest on small aliquots of samples, e.g. at room temperature, at 4–6 °C, at lower temperatures –20 °C, and for its enzymatic activity and concentration over time. Several storage conditions may have an adverse effect on protein stability and activity. An illustrative example is storage at –20°C in 50 % glycerol, which ensures protein stability, however, when followed by purification the glycerol has to be removed from the sample (e.g. by dialysis) to avoid dropping the protein amount and activity. Protein storage is enhanced by rapid freezing without the use of a cryoprotectants (such as glycerol, ethylene glycol, sucrose), however in this case protein freezing and defrosting can cause protein denaturation, aggregation and precipitation (Cao *et al.* 2003). Proteins can be unstable and lose their biological activity as a result of various physical and chemical procedures, too, even at low temperatures (Carpenter *et al.* 2002; Simpson 2010). Therefore, when choosing the proper protein storage conditions, it is important to consider what the purified protein will be used for (determination of specific enzymatic activity, protein crystallization, and interaction studies etc.). Protein stabilizers or serum albumins (e.g. BSA) can be added to the protein to extend the shelf life. The presence of albumin in a sample is not an obstacle when the protein activity is tested in the following steps, but can interfere in structural studies. In contrast, the addition of substances that prevent from the growth of microorganisms is recommended, as they reduce the freezing point, or preserve the protein in the solution. When adding such substances, it is necessary to perform protein stability studies over time. It should be mentioned that there has been no universal stabilizing solution developed for storage of all the various protein samples yet. It is also necessary to be cautious about optimal conditions for the specific protein sample that can vary during storage (Tab. 2) since storage may alter intermolecular interactions. Further, optimal protein storage conditions may not be necessarily the same as the conditions for determining the protein activity. Moreover, additives used to ensure protein stability might act as inhibitors of enzymes. That is the reason why interfering substances should be taken into consideration and must be removed or sufficiently diluted before the use of the protein. Currently the market offer comprises many commercial kits with a wide variety of specific solutions to test protein and storage stability under different conditions.

Conclusions

In this paper, all the steps and factors have been summarized and briefly discussed which could be encountered during production and purification of a recombinant protein of interest, especially from eukaryotic source, and expressed heterogeneously in prokaryotic production system. It is inevitable to mention that despite the simplicity and overall effectiveness of *E. coli* production system in combination with various commercially available expression vectors, the final amount and purity of a recombinant protein is insufficient and, many times, contaminated, insoluble or inactive. In many cases, when a soluble protein is inactive or not crystallized after all the separation and purification steps, there is an urgent need to go back to the beginning of the whole protocol, find and solve the origin of a problem and try to optimize the whole process. These findings open a wide range of up-to-date issues to be analysed, as well as possibilities how to improve protein features. However, abovementioned hinders can be overcome by investment of resources into more in-depth studies.

References

Andersen CL, Matthey-Dupraz A, Missiakas D, Raina S (1997) A new *Escherichia coli* gene, dsbG, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. Mol. Microbiol. 26: 121-132.

Arnau J, Lauritzen C, Pedersen J (2006*a*) Cloning strategy, production and purification of proteins with exopeptidase–cleavable His–tags. Nat. Protoc. 1: 2326-2333.

Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006*b*) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr. Purif. 48: 1-13.

Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. Nat. Biotechnol. 22: 1399-1408.

Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. Curr. Opin. Biot. 10: 411-421.

Bardwell JC (1994) Building bridges: Disulphide bond formation in the cell. Mol. Microbiol. 14: 199-205.

Bessette PH, Aslund F, Beckwith J, Georgiou G (1999) Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Proc. Natl. Acad. Sci. USA 24: 13703-13708.

Bolanos-Garcia VM, Davies OR (2006) Structural analysis and classification of native proteins from *E. coli* commonly co–purified by immobilised metal affinity chromatography. Biochim. Biophys. Acta. 1760: 1304-1313.

Bornhorst JA, Falke JJ (2000) Purification of proteins using polyhistidine affinity tags. Methods Enzymol. 245: 245-254.

Brizzard B (2008) Epitope tagging of recombinant proteins. Curr. Protoc. Neurosci. 44: 693-695.

Cabrita LD, Bottomley SP (2004) Protein expression and refolding - A practical guide to getting the most out of inclusion bodies. Biotechnol. Annu. Rev. 10: 31-50.

Calos MP (1978) DNA sequence for a low–level promoter of the lac repressor gene and an 'up' promoter mutation. Nature 274: 762-765.

Cao E, Chen Y, Cui Z, Foster PR (2003) Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions. Biotechnol. Bioeng. 82: 684-690.

Carpenter JF, Chang BS, Garzon-Rodriguez W, Randolph TW (2002) Rational design of stable lyophilized protein formulations: theory and practice. Pharm. Biotechnol. 13: 109-133.

Curtain CC, Nayler WG (1963) The isolation from human blood plasma of a substance having positive inotropic action on the isolated toad heart. Biochem. J. 89: 69-75.

De Boer HA, Comstock LJ, Yansura DG, Heyneker HL (1983) The *tac* promoter: a functional hybrid derived from the trp and lac promoters. Proc. Natl. Acad. Sci. USA 80: 21-25.

Derman AI, Prinz WA, Belin D, Beckwith J (1993) Mutations that allow disulphide bond formation in the cytoplasm of *Escherichia coli*. Science 262: 1744-1747.

Deuschle U, Kammerer W, Gentz R, Bujard H (1986) Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures. EMBO J. 5: 2987-2994.

Dodd IB, Shearwin KE, Egan JB (2005) Revisited gene regulation in bacteriophage lambda. Curr. Opin. Genet. Dev. 15: 145-152.

Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, Dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, Prilusky J, Sussman JL, Stevens RC, Lesley SA, Wilson IA, Joachimiak A, Collart F, Dementieva I, Donnelly MI, Eschenfeldt WH, Kim Y, Stols L, Wu R, Zhou M, Burley SK, Emtage JS, Sauder JM, Thompson D, Bain K, Luz J, Gheyi T, Zhang F, Atwell S, Almo SC, Bonanno JB, Fiser A, Swaminathan S, Studier FW, Chance MR, Sali A, Acton TB, Xiao R, Zhao L, Ma LC, Hunt JF, Tong L, Cunningham K, Inouye M, Anderson S, Janjua H, Shastry R, Ho CK, Wang D, Wang H, Jiang M, Montelione GT, Stuart DI, Owens RJ, Daenke S, Schütz A, Heinemann U, Yokoyama S, Büssow K, Gunsalus KC (2008) Protein production and purification. Nat. Methods 5: 135-146.

Graumann K, Premstaller A (2006) Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnol. J. 1: 164-186.

Guillaume G, Auxilien S, Tuszynska I, Locard M, Gajda MJ, Chaussinand G, Fernandez B, Dedieu A, Grosjean H, Golinelli-Pimpaneau B, Bujnicki JM, Armengaud J (2006) THUMP from archaeal tRNA:m22G10 methyl transferase, a genuine autonomously folding domain. Nucleic. Acids Res. 34: 2483-2494.

Hallewell RA, Emtage S (1980) Plasmid vectors containing the tryptophan operon promoter suitable for efficient regulated expression of foreign genes. Gene 9: 27-47.

Hjertén S (1964) The preparation of agarose spheres for chromatography of molecules and particles. Biochim. Biophys. Acta 30: 393-398.

Hochuli E, Dobeli HA, Schacher A (1987) New metal chelate adsorbent selective for proteins and peptide containing neighbouring histidine residues. J. Chromatogr. 411: 177-184.

Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Cerretti DP, Urdal DL, Conlon PJ (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. J. Biotechnol. 6: 1204-1210.

Chatterjee DK, Esposito D (2006) Enhanced soluble protein expression using two new fusion tags. Protein Expr. Purif. 46: 122-129.

Choi JH, Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl. Microbiol. Biotechnol. 64: 624-635.

Chrastilová Z, Macková M, Šotola J, Král V (2007) Bioléčiva – Jaký je jejich skutečný potenciál? Chem. Listy 101: 25-35.

Chung BH, Sohn MJ, Oh S, Park U, Poo H, Kim BS, Yu MJ, Lee SK (1998) Overproduction of human granulocyte–colony stimulating factor fused to the PelB signal peptide in *Escherichia coli*. J. Ferment. Bioeng. 85: 443-446.

Inaba K (2009) Disulphide bond formation system in *Escherichia coli*. J. Biochem. 146: 591-597.

Janknecht R, De Martynoff G, Lou J, Hipskind RA, Nordheim A, Stunnenberg HG (1991) Rapid and efficient purification of native histidine–tagged protein expressed by recombinant vaccinia virus. Proc. Natl. Acad. Sci. USA 88: 8972-8976.

Justesen SFL, Lamberth K, Nielsen LB, Schafer-Nielsen C, Buus S (2009) Recombinant chymosin used for exact and complete removal of a prochymosin derived fusion tag releasing intact native target protein. Protein Sci. 18: 1023-1032.

Klock HE, White A, Koesema E, Lesley SA (2005) Methods and results for semi–automated cloning using integrated robotics. J. Struct. Funct. Genomic. 6: 89-94.

Lavallie ER, Diblasio EA, Kovacic S, Grant KL, Schendel PF, Mccoy JM (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Bio/Technology 11: 187-193.

Le Hir H, Nott A, Moore MJ (2003) How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci. 28: 215-220.

Le-Calvez H, Green JM, Baty D (1996) Increased efficiency of alkaline phosphatase production levels in *Escherichia coli* using a degenerate PelB signal sequence. Gene 170: 51-55.

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. Protein Expr. Purif. 41: 98-105.

Makoff AJ, Oxer MD (1991) High level heterologous expression in *E. coli* using mutant forms of the lac promoter. Nucleic Acids Res. 19: 2417-2421.

Makrides SC (1996) Strategies for achieving high–level expression of genes in *Escherichia coli*. Microbiol. Rev. 60: 512-538.

Matthey B, Engert A, Klimka A, Diehl V, Barth S (1999) A new series of pET derived vectors for high efficiency expression of *Pseudomonas* exotoxin–based fusion proteins. Gene 229: 145-153.

Menart V, Jevsevar S, Vilar M, Trobis A, Pavko A (2003) Constitutive versus thermoinducible expression of heterologous proteins in *Escherichia coli* based on strong PR, PL promoters from phage lambda. Biotechnol. Bioeng. 83: 181-190.

Moffatt BA, Studier FW (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell 49: 221-227.

Müller-Hill B, Crapo L, Gilbert W (1968) Mutants that make more lac repressor. Proc. Natl. Acad. Sci. USA 59: 1259-1264.

Nakamura E, Kozaki K, Tsuda H, Suzuki E, Pimkhaokham A, Yamamoto G, Irie T, Tachikawa T, Amagasa T, Inazawa J, Imoto I (2008) Frequent silencing of a putative tumor suppressor gene melatonin receptor 1 A (MTNR1A) in oral squamous–cell carcinoma. Cancer Sci. 99: 1390-1400.

Petsch D, Anspach FB (2000) Endotoxin removal from protein solutions. J. Biotechnol. 76: 97-119.

Philo JS (2006) Is any measurement method optimal for all aggregate sizes and types? AAPS J. 8: 564-57.

Porath J, Carlsson J, Olsson I, Belfrage G (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. Nature 258: 598-609.

Porath J, Flodin P (1959) Gel filtration: A method for desalting and group separation. Nature 183: 1657-1659.

Raymond S, Weintraub L (1959) Acrylamide gel as a supporting medium for zone electrophoresis. Science 130: 711-722.

Rhodes DG, Bossio RE, Laue TM (2009) Determination of size, molecular weight, and presence of subunits. Methods Enzymol. 463: 691-723.

Saida F (2007) Overview on the expression of toxic gene products in *Escherichia coli.* Curr. Protoc. Protein Sci. 5.19.11-5.19.13.

Sezonov G, Joseleau-Petit D, D'ari R (2007) *Escherichia coli* physiology in Luria–Bertani broth. J. Bacteriol. 23: 8746-8749.

Simpson RJ (2010) Stabilization of proteins for storage. Cold Spring Harb. Protoc. 5: 110-116.

Stano NM, Patel SS (2004) T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation. J. Biol. Chem. 279: 16136-16143.

Stevens RC (2000) Desing of high–throught methods of protein production for structural biology. Structure 8: 177-185.

Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high–level expression of cloned genes. J. Mol. Biol. 189: 113-130.

Tacon WCA, Bonass WA, Jenkins B, Emtage JS (1983) Expression plasmid vectors containing *Escherichia coli* tryptophan promoter transcriptional units lacking the attenuator. Gene 23: 255-265.

Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. 72: 211-222.

Valdez-Cruz NA, Caspeta L, Pérez NO, Ramírez OT, Trujillo-Roldán MA (2010) Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. Microb. Cell Factor 19: 9-18.

Yansura DG (1990) Expression as trpE fusion. Methods Enzymol. 185: 161-166.

Yin J, Li G, Ren X, Herrler G (2007) Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for forein genes. J. Biotechnol. 127: 335-347.

Yokoyama S (2003) Protein expression systems for structural genomics and proteomics. Curr. Opin. Chem. Biol. 7: 39-43.