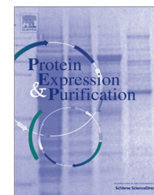
Protein Expression and Purification 94 (2014) 22–32

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep



Review

Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: A general review

Christos P. Papaneophytou, George Kontopidis *

Veterinary School, University of Thessaly, Trikalon 224, Karditsa 43100, Greece
Institute for Research and Technology – Thessaly (I.RE.TE.TH.), The Centre for Research & Technology Hellas (CE.R.TH.), Technology Park of Thessaly, 1st Industrial Area, Volos 38500, Greece

ARTICLE INFO

Article history:
Received 3 October 2013
and in revised form 23 October 2013
Available online 5 November 2013

Keywords:
Solubility enhancement
Statistically designed experiments
Recombinant protein
Escherichia coli
Response surface methodology (RSM)
Fractional factorial

ABSTRACT

The supply of many valuable proteins that have potential clinical or industrial use is often limited by their low natural availability. With the modern advances in genomics, proteomics and bioinformatics, the number of proteins being produced using recombinant techniques is exponentially increasing and seems to guarantee an unlimited supply of recombinant proteins. The demand of recombinant proteins has increased as more applications in several fields become a commercial reality. *Escherichia coli* (*E. coli*) is the most widely used expression system for the production of recombinant proteins for structural and functional studies. However, producing soluble proteins in *E. coli* is still a major bottleneck for structural biology projects. One of the most challenging steps in any structural biology project is predicting which protein or protein fragment will express solubly and purify for crystallographic studies. The production of soluble and active proteins is influenced by several factors including expression host, fusion tag, induction temperature and time. Statistical designed experiments are gaining success in the production of recombinant protein because they provide information on variable interactions that escape the "one-factor-at-a-time" method. Here, we review the most important factors affecting the production of recombinant proteins in a soluble form. Moreover, we provide information about how the statistical design experiments can increase protein yield and purity as well as find conditions for crystal growth.

© 2013 Elsevier Inc. All rights reserved.

Contents

Introduction 23
Factors affecting soluble expression of recombinant proteins 23
  Effect of medium composition 23
  Choice of expression hosts 24
  Fusion tags 24
  Rate of protein synthesis 24
    Timing of induction 24
    Temperature and duration of induction 25
    Inducer concentration 25
    Other factors: lysis buffer and additives 25
Statistical approaches for maximize soluble protein production 25
Full and fractional factorial designs 26
  Full factorial designs 26
  Fractional factorial designs 26
  Response surface methodology 26
    Screening of variables 27
    Choice of the experimental design and mathematical–statistical treatment of data 27
    Evaluation of the fitted model 28
    Determination of the optimal conditions 28
    Application of RSM in production of recombinant protein 28

* Corresponding author. Address: Laboratory of Biochemistry, Veterinary School, University of Thessaly, Trikalon 224, Karditsa 43100, Greece. Tel.: +30 24410 66081; fax: +30 24410 66041.
E-mail address: gkontopidis@vet.uth.gr (G. Kontopidis).

1046-5928/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.pep.2013.10.016

Incomplete fractional designs . . . 30
Response surface methodology vs incomplete factorial approaches . . . 31
Conclusions . . . 31
References . . . 31

Introduction

Production of soluble recombinant proteins is vital for structure–function analysis and therapeutic applications. Pharmaceutical protein development requires the ability to express and purify recombinant proteins having desired pharmacokinetics and physicochemical properties [1]. Recombinant proteins are required in biological research to investigate enzyme activity, ligand binding, protein interactions, or other functions *in vitro*. Many proteins are also potential pharmaceutical agents [2,3]. Major advances in genetic engineering have resulted in the development of bacterial expression systems, particularly those in *Escherichia coli*, capable of producing large amounts of proteins from cloned genes [4]. However, two challenges in the production of heterologous proteins in *E. coli*, the workhorse of protein expression systems, are poor or low expression, and the mis-folding of the expressed protein into insoluble aggregates called inclusion bodies [5]. Protein expression is no longer considered a major limiting step and protein purification techniques have improved dramatically in the past decade. Although, producing soluble proteins for purification has continued to be a major bottleneck in the field [6]. Insoluble recombinant proteins are a major issue for both structural genomics and enzymology research. More than 30% of recombinant proteins expressed in *E. coli* appear to be insoluble [7].

E. coli expression system continuous to dominate the bacterial expression systems and remain the first choice for laboratory investigations and initial development in commercial activities. The main purpose of recombinant protein expression is often to obtain a high degree of accumulation of soluble product in the bacterial cell [8]. Many of the most biochemically interesting families of proteins, including kinases, phosphatases, membrane-associated proteins and many other enzymes, are extremely difficult to produce as soluble proteins in *E. coli* [6].

Successful expression and solubility of target protein dependent on the amino acid composition of the protein, and primary sequence analysis can be used to guide the design and choice of expression system [9]. Often small differences in the amino acid sequence itself, or in length of the construct, can transform a protein that fails to express into one that expresses, purifies and crystallizes readily [10,11].

In an ideal situation, the recombinant protein is expressed from a strong promoter, highly soluble, and recovered in high yield and activity. Unfortunately, it is quite common that the overproduced recombinant protein is either detrimental to the cell or simply compartmentalized into insoluble inclusion bodies [12]. In some cases, the recombinant protein can be recovered in an active form after denaturation and subsequent renaturation [13]. However, this is less than desirable because it is often uncertain whether the refolded protein has regained full function. In general, expression and solubility can be optimized by varying expression conditions such as post-induction temperature, type of cultivation media and the type of *E. coli* strain. When a protein is insoluble multiple rescue procedures may be undertaken including: refolding of denatured proteins [14] creating fusion protein constructs such as maltose binding protein [15]. Moreover, in an attempt to increase the solubility of recombinant proteins, they have often been co-expressed in the presence of chaperones [12] or at low temperature [16]. Even though several theoretical and empirical methods to improve soluble production have been suggested, there is to date no universally accepted protocol.

The production of recombinant proteins is generally performed using a trial-and-error approach, with the different expression variables being tested independently from each other. Therefore, variable interactions are lost which makes the trial-and-error approach time-consuming. As significant amount of protein is required for every structural biology projects the traditional trial-and-error method has been progressively replaced by factorial approaches (full factorial, incomplete factorial and sparse matrix) at every step of process ranging from gene expression to crystallization.

In this study, we attempt to illustrate the effect of main factors influencing both recombinant protein expression and solubilisation, and report on those materials and technologies we have found most useful for our own projects. In addition, we provide information and references for a more detailed introduction to statistical analysis in experimentation. To our knowledge this will be the first review, which extensively examined the use of statistical approaches on recombinant protein production.

Factors affecting soluble expression of recombinant proteins

To facilitate cloning and expression of target genes for improved solubility in *E. coli*, a variety of vectors and methods are available. However, a number of criteria must be considered when optimizing conditions for the high-level expression of a recombinant protein in *E. coli*. In the following paragraphs, we examine the main factors affecting soluble protein expression and their influences are studied with statistical designed experiments.

Effect of medium composition

To optimize the level of soluble expression, the first parameters to tune are the culture conditions and/or culture medium because this is easy, cheap, and has been proven to have an impact on protein solubility levels [17,18]. In several cases, medium composition, specifically the concentration of some salts, peptone and yeast, can increase the concentration of recombinant protein [19,20]. However, Vincentelli et al. [21] reported that culture medium composition (SB; 2YT; TB) is not a major determinant of protein solubility for both prokaryotic and eukaryotic targets. Overall, the solubility was the same per cell, and the higher the biomass the more protein produced. Culture medium composition (LB; TB; 2YT) also had a minimal impact on recombinant RANKL solubility [22].

On the other hand, the addition of prosthetic groups or co-factors which are essential for proper folding or for protein stability in the culture medium can prevent the formation of inclusion bodies [23]. The addition of such co-factors or binding partners to the cultivation media may increase the yield of soluble protein dramatically. The aggregation of proteins secreted into the periplasmic space can be suppressed by growing the cells in the presence of relatively high concentrations of polyols (e.g., sorbitol) or sucrose, a non-metabolizable sugar for *E. coli*. The increase in osmotic pressure caused by these factors leads to the accumulation of osmoprotectants (e.g., glycine betaine, threholose) in the cell, which stabilize the native protein structure [23]. Other growth additives, that can have a beneficial effect on soluble protein

Download English Version:

https://daneshyari.com/en/article/2020504

Download Persian Version:

https://daneshyari.com/article/2020504

Daneshyari.com