Protein Expression and Purification 144 (2018) 12–18

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*

Irshad Ahmad[a], Nighat Nawaz[b], Nizam M. Darwesh[c], Sadeeq ur Rahman[d], Mohammad Z. Mustafa[e], Sher B. Khan[f], Simon G. Patching[a,*]

[a] *Leeds, UK*
[b] *Department of Chemistry, Islamia College Peshawar, Peshawar, Pakistan*
[c] *Institute for Health Research, University of Bedfordshire, Luton, UK*
[d] *College of Animal Husbandry and Veterinary Sciences, Abdul Wali Khan University, Mardan, Pakistan*
[e] *Centre for Advanced Studies in Vaccinology and Biotechnology, University of Balochistan, Quetta, Pakistan*
[f] *Department of Animal Health, The University of Agriculture, Peshawar, Pakistan*

ARTICLE INFO

Keywords:
Bacteria
E. coli
Membrane proteins
Protein expression
Recombinant proteins

ABSTRACT

A thorough characterisation of the genetics, physiology and metabolism of *Escherichia coli* has led to the availability of a large number of strains and vectors suitable for recombinant protein expression. Despite the relative ease in using *E. coli* for achieving amplified expression of many recombinant proteins, for some proteins this can be a frustrating and time-consuming process leading to very low expression or no expression at all. This is especially true for membrane proteins, which introduce additional challenges. A number of factors can be considered and optimised for achieving required levels of amplified expression of recombinant proteins in *E. coli* that are broadly classified as host strain, expression vector and growth conditions. In this paper we summarise these factors and consolidate the common challenges encountered and approaches to overcome them, focusing in particular on cases where there is low amplified expression or no expression at all of the desired recombinant protein, due to various reasons.

1. Introduction

Structural and functional studies on proteins requires production of sufficient quantities of protein, typically in milligram quantities, as the essential first step. The natural expression levels of many proteins, especially membrane proteins [1], are usually too low, so amplified expression of recombinant proteins must be achieved. Owing to its easy accessibility, genetic manipulation and handling, *Escherichia coli* is still the most widely used organism for amplifying the expression of recombinant proteins, including membrane proteins [2–14]. Theoretically, recombinant protein production in *E. coli* is relatively straightforward, beginning with target identification, cloning of the gene into an appropriate vector, transformation of the construct into a suitable host strain, induction for amplified expression and then protein purification and characterisation (sequencing, purity, structural integrity, stability, activity). But in practice, a number of obstacles can be encountered along the pipeline that must be overcome. These can include poor growth of the host strain, protein instability or toxicity, aggregation and inclusion body formation, unsuitability of environmental conditions (temperature, pH, salt concentration, etc) and even no amplified expression at all. Whilst continuous progress has been made in the field of recombinant protein expression using *E. coli*, including execution of structural genomics programmes with prokaryotic membrane proteins [15–19], the amplified expression of recombinant proteins, especially membrane proteins [20–27], is still often a matter of "trial and error". We have therefore summarised these factors and consolidated the common challenges encountered and approaches to overcome them, focusing in particular on cases where there is low amplified expression or no expression at all of the desired recombinant protein.

2. Expression host

The different types of expression host available for producing recombinant proteins include bacteria, filamentous fungi, yeasts, unicellular algae and insect and mammalian cells. Each type of host has different advantages and disadvantages and their choice often depends on the protein of interest [28–30]. For producing the large majority of recombinant proteins, especially those of bacterial origin, *E. coli* is by far the most widely preferred host organism. The advantages of using *E.*

* Corresponding author. Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK.
E-mail address: sgp_uk2000@yahoo.co.uk (S.G. Patching).

https://doi.org/10.1016/j.pep.2017.11.005
Received 27 October 2017; Received in revised form 15 November 2017; Accepted 21 November 2017
Available online 24 November 2017
1046-5928/ © 2017 Published by Elsevier Inc.

coli as the host organism are well recognised: (i) fast growing [31]; (ii) high cell density is easily attained [32]; (iii) inexpensive complex media can be used for growth [33]; (iv) well-characterised genetics, physiology and metabolism has led to the availability of a large number of cloning and expression strains and vectors [34]; (v) transformation with foreign DNA can be straightforward and quick [35]; (vi) easy to scale for fermentation. *E. coli* is also tractable to incorporating various labelling types, including stable isotopes for NMR studies (e.g. ^{2}H, ^{13}C, ^{15}N) [36–41], radiolabels [42–44] (e.g. ^{3}H, ^{14}C, ^{35}S) and non-natural amino acids (e.g. Se-Met for crystallography) [45–48]. *E. coli* strain BL21(DE3) and its derivatives are by far the most widely used strains for recombinant protein expression [49–51]. BL21(DE3) is protease-deficient, lacking Lon protease (cytoplasm) and OmpT protease (outer membrane), has low acetate production at high glucose levels and has enhanced permeability. It also contains the λDE3 lysogen that carries the gene for T7 RNA polymerase under control of the *lac*UV5 promoter, which can direct high-level expression of cloned genes under control of the T7 promoter.

3. Expression vector

The choice of expression vector is a key factor in the successful amplified expression of a target protein in *E. coli*. A huge number of expression vectors are available, but the most commonly used ones are plasmid vectors that contain a number of sequence elements including an origin of replication, a promoter, a multi-cloning site, affinity tags, a terminator and selection markers (Fig. 1). Choosing a suitable vector requires a good understanding of these features and their usefulness can be carefully evaluated according to the characteristics of the target protein. One of the important parameters to consider is the copy number. Whilst a high copy number can, and often does, lead to more recombinant protein in the cell, occasionally a high copy number can drop the bacterial growth rate by imposing a metabolic burden and can cause plasmid instability and decreased protein synthesis [52,53]. A variety of expression vectors carrying different promoters are commercially available and, for an ideal expression system, the vector must have a tightly regulated strong promoter with a low basal expression level [54,55]. In some cases, the use of a very strong promoter has been reported lethal for the expression host and tends to result in inclusion body formation [56,57].

The *lac* promoter of the *lac* (lactose) operon has historically provided key knowledge in development of prokaryotic expression vectors [58], but for recombinant protein production, the *lac* promoter and its derivatives are weak and consequently not very useful [59,60]. Instead, a synthetic combination of different promoters has proven to be advantageous. For example, the *tac* promoter comprised of the −35 region of the *trp* (tryptophan) promoter and the −10 region of the *lac* recognition site for the repressor, is around 10-fold stronger than the standard *lac*UV5 promoter [61]. The *tac* promoter can be used in combination with *lacI*^{Q}, which is the mutated repressor of the *lac*I gene that controls basal expression and achieves higher expression levels (around 10-fold) than *lac*I [62]. The T7 promoter system, which is found in pET vectors, is commonly used [63,64]. Here, the gene of interest is cloned behind a promoter recognised by the phage T7 RNA polymerase and is often placed in the bacterial genome under transcriptional control of a *lac*UV5 promoter, as in BL21(DE3) [49], thus allowing induction by lactose or its non-hydrolysable analog isopropyl β-*D*-1-thiogalactopyranoside (IPTG).

Vectors based on the arabinose promoter system (pBAD vectors) are also widely used. Such vectors contain the P_{BAD} promoter of the *araBAD* (arabinose) operon and the gene encoding the positive and negative regulator of this promoter, *araC* [65]. AraC inhibits expression from the *araBAD* promoter in the absence of *L*-arabinose or in the presence of glucose and activates transcription in the presence of *L*-arabinose and in the absence of glucose [66]. Whilst pBAD vectors generally do not achieve the very high levels of expression possible with pET vectors, they can be strongly induced to achieve moderately high-level expression. Because its kinetics of induction and repression are very rapid, the P_{BAD} system can be rapidly and efficiently turned on and off without changes in temperature, therefore allowing extremely tight regulation of gene expression. The P_{BAD} system also allows modulation of expression over a wide range of *L*-arabinose concentrations, which can be reduced to extremely low levels by glucose. This makes it possible to achieve synthesis levels similar to those of the wild-type gene [65]. A further advantage is that *L*-arabinose is relatively inexpensive (compared with IPTG), making large-scale protein production more feasible. Because wild-type *E. coli* strains can catabolise *L*-arabinose, it is sensible to use host strains that are mutant for *L*-arabinose catabolism (e.g. TOP10, LMG194).

For selection of anticipated cells carrying the desirable plasmid and to prevent the growth of plasmid-free cells, a resistance marker is generally included in the plasmid. Antibiotic resistance genes are usually used in the *E. coli* system for this purpose; for example, resistance to ampicillin is conferred by the *bla* gene. A multiple cloning site containing many (up to around 20) restriction sites is a key feature of the vector that allows foreign DNA to be inserted. For detecting recombinant protein expression and for purification from crude extracts, a variety of affinity tags have been developed [67–72]. Small peptide tags such as hexahistidine (His_6) are less likely to affect properties of the recombinant protein; but some tags can have adverse effects on the biological activity or structure of the protein [29]. For achieving amplified expression of various types of bacterial and archaeal membrane proteins, the high copy number plasmid pTTQ18 (*tac* promoter, *lacI*^{Q} repressor, *rrn*B terminator, ampicillin resistance) [73] with BL21(DE2)-type strains has proved to be very successful [3,67,74–88], including scale-up to fermentor cultures [3,89,90].

4. Growth conditions

Recombinant protein expression can be enhanced by optimisation of *E. coli* growth conditions [32,91–94] and there are a number of main parameters that can be altered. (i) *Composition of the growth medium* – a complex or defined medium can be used with different carbon sources and other supplements. (ii) *Temperature profile* – this affects the rate of cell growth and has implications for recombinant protein toxicity, stability and aggregation, especially post-induction. (iii) *Cell density at the*

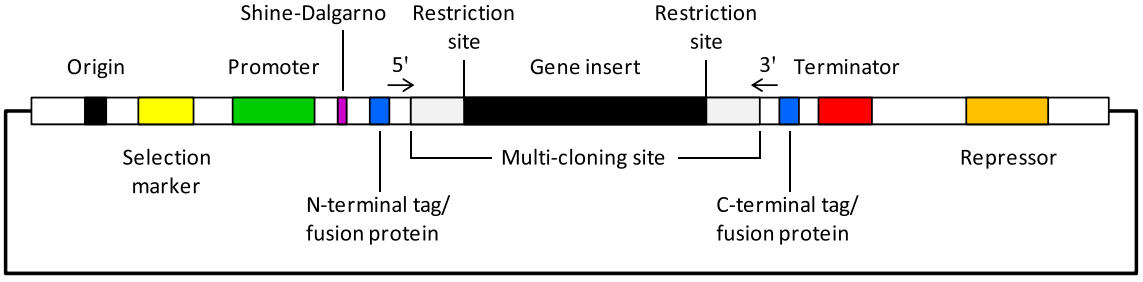

Fig. 1. General features of a plasmid vector for amplified expression of recombinant proteins using *E. coli*.

Download English Version:

https://daneshyari.com/en/article/8359529

Download Persian Version:

https://daneshyari.com/article/8359529

Daneshyari.com