Protein Expression and Purification 133 (2017) 160–169

Contents lists available at ScienceDirect

Protein Expression and Purification

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

Streamlining workflow and automation to accelerate laboratory scale protein production

Jennifer Konczal, Christopher H. Gray*

Drug Discovery Program, CRUK Beatson Institute, Garscube Estate, Switchback Road, Glasgow G61 1BD, United Kingdom

ARTICLE INFO

Article history:
Received 20 February 2017
Accepted 17 March 2017
Available online 19 March 2017

Keywords:
Novel vectors and cloning
Strain engineering
Higher throughput expression screening
Automated protein purification
Parallel purification
Multidimensional chromatography
Protein conditioning

ABSTRACT

Protein production facilities are often required to produce diverse arrays of proteins for demanding methodologies including crystallography, NMR, ITC and other reagent intensive techniques. It is common for these teams to find themselves a bottleneck in the pipeline of ambitious projects. This pressure to deliver has resulted in the evolution of many novel methods to increase capacity and throughput at all stages in the pipeline for generation of recombinant proteins. This review aims to describe current and emerging options to accelerate the success of protein production in *Escherichia coli*. We emphasize technologies that have been evaluated and implemented in our laboratory, including innovative molecular biology and expression vectors, small-scale expression screening strategies and the automation of parallel and multidimensional chromatography.

Contents

1. Introduction

Despite the appreciated limitations, recombinant protein expression in *Escherichia coli* remains the most cost effective and convenient method for most laboratories [1]. This is especially true in programs where substantial amounts of recombinant material are routinely and repeatedly required for reagent intensive structural and biophysical methods. Protein production facilities are often tasked to produce diverse arrays of proteins for these demanding techniques [2]. Thus protein production teams often

* Corresponding author.
E-mail address: c.gray@beatson.gla.ac.uk (C.H. Gray).

http://dx.doi.org/10.1016/j.pep.2017.03.016

find themselves a bottleneck in the pipeline of ambitious projects. This pressure to deliver has resulted in the evolution of many novel methods to increase capacity and throughput.

Original solutions have appeared in all stages of the protein production pipeline from cloning and expression, to purification (areas addressed in this review are presented in Fig. 1). Improvements in conventional plasmid expression vectors have added performance and functionality to recombinant proteins. The introduction of novel tags and solubilizing fusions give access to higher yields of protein that are more amenable to rapid and efficient purification [3]. Streamlined expression screening workflows, using miniaturized expression culturing and automated or semi-automated analysis for recombinant expression have allowed laboratories with modest resources to investigate significant numbers of candidate expression constructs and strains leading to improved rates of success. This uplift in output from expression screening puts pressure on purification following scale up and this is generating innovative approaches to laboratory prep scale protein purification, pushing the capacity and functionality of modern FPLC systems [4].

This review aims to present current techniques in lab scale *E. coli* protein expression and purification with an emphasis on technologies that afford significant improvements in workflow performance and throughput. Particular emphasis is placed on automation of prep scale protein purification. These approaches go some way to assisting busy protein facilities in the alleviation of the protein production bottleneck.

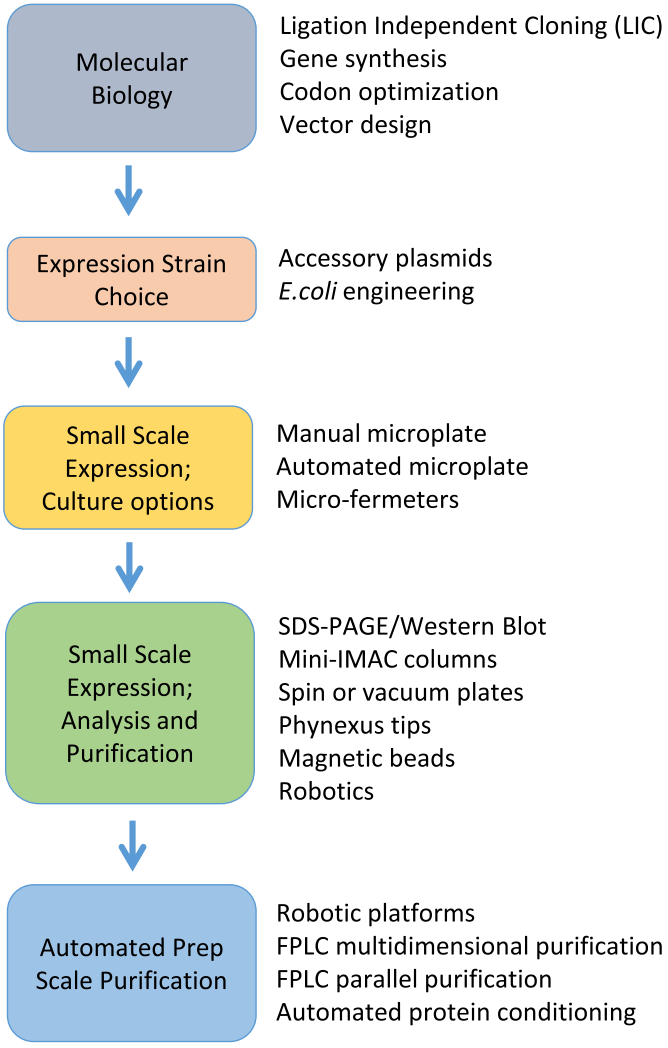

Fig. 1. Core protein screening, optimization and production variables considered in the BICR Drug Discovery Program noting potential areas offering enhanced efficiency and success.

2. Improving success rates using innovative molecular biology and strains

2.1. Cloning, gene synthesis and novel vector options

Many laboratories now forego the traditional in-house PCR/T4 ligase cloning process in favor of more reliable and faster methods. The introduction of ligation independent cloning (LIC) and Gateway™ cloning (Invitrogen) [5] was a substantial step forward in this area, increasing the ease of generating larger numbers of expression clones, and Gateway™ continues to be used widely. Indeed the Gateway™ system has evolved to include access to a diverse array of destination vectors that include all typical tags and protease sites [6,7] as well as suites with distinct replication origins and resistance markers for co-transformation and expression of multi-protein complexes [8]. Methods that are more recent have expanded the use of homologous recombination. Scientists at the Max-Plank Institute Core Facility have developed Sequence and Ligation Independent Cloning (SLIC) method that makes use of homologous recombination to facilitate insertion of expression cDNA, coupled with a toxic ccdB gene that affords a background free result by conventional insertional inactivation [9]. Similarly, the homologous recombination capability of yeast has been employed to efficiently generate bacterial expression constructs using a two–transformation approach where the recombination is achieved in yeast cells then the plasmid rescued and transformed into *E. coli*, again resulting in a system with negligible background [10].

Now, relatively cheap commercial gene synthesis is often a preferred option, providing an outsourced service that affords the benefit of codon-optimized cDNAs that can be custom cloned into the required expression vectors by the supplier [11]. Consequently, a substantial amount of molecular biology is removed from the laboratory, freeing up personnel for expression and purification tasks. The flexibility of gene synthesis has allowed much easier design of innovative arrangements of tags and fusion constructs that have greatly improved the efficiency of production and the functionality of products. For example, we have recently described the double 8-histidine tagged system that we use extensively in our laboratory, principally because this tag allows a robust immobilization on nickel NTA chips in our SPR fragment screening platform [4]. Using gene synthesis, we have expanded on this initial pBDDP-SPR3 plasmid to generate a suite of vectors that provide this His_8-His_8 functionality in concert with additional affinity tags or fusion proteins (Fig. 2). This provides us with a substantial panel of appropriate clones that can enter into our small-scale expression screening workflow.

Most *E. coli* expression vectors continue to make use of the inducible T7 promotor system typified in the long established pET vector series [13]. Whilst highly specific and powerful, this system can lead to excess transcription and translation resulting in poor folding and inclusion body formation. The inducible T7 system is relatively insensitive to modulation or fine tuning. Varying the IPTG concentration for induction often has limited success. Alternative promotors that offer improved response to tuning can provide solutions when the T7 system proves unfruitful. The Qiagen pQE vector system uses the T5 promotor system. This has the advantage of improved repression as it is silenced more effectively by excess lac repressor, giving more effective control of toxic protein expression. However, fine-tuning of the IPTG response is still limited [14]. Sugar