Chapter 2

Overcoming the Solubility Problem in *E. coli*: Available Approaches for Recombinant Protein Production

Agustín Correa and Pablo Oppezzo

Abstract

Despite the importance of recombinant protein production in academy and industrial fields, many issues concerning the expression of soluble and homogeneous product are still unsolved. Although several strategies were developed to overcome these obstacles, at present there is no magic bullet that can be applied for all cases. Indeed, several key expression parameters need to be evaluated for each protein. Among the different hosts for protein expression, *Escherichia coli* is by far the most widely used. In this chapter, we review many of the different tools employed to circumvent protein insolubility problems.

Key words Recombinant proteins, Protein expression, *E. coli*, High-throughput screening, Inclusion bodies, Directed evolution

1 Introduction

With the advances in genome sequencing nowadays, over 1,900 genomes are publicly available (http://www.microbesonline.org) generating massive information in this area. A typical microbial genome codes for between 1,500 and 8,000 proteins while in eukaryotic genomes is around 10,000–60,000 proteins. Despite all this information, and in contrast with nucleic acids, obtaining the target protein from the natural host in a soluble, homogeneous state and enough quantities for biochemical and structural studies is very uncommon. This makes the production of the target protein in a recombinant form the method of choice. Different expression hosts are available for recombinant expression, including bacterial, fungal, or eukaryotic host cells. Among these, the use of the enterobacterium *Escherichia coli* is the most commonly employed with approximately 60 % of all recombinant proteins in the literature and nearly 30 % of the currently approved recombinant therapeutic proteins produced on it [1, 2]. This is mainly due to the low cost, fast growth, easy handling, high yield of target protein and the extensive knowledge of the genetics of *E. coli*.

Elena García-Fruitós (ed.), *Insoluble Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1258,
DOI 10.1007/978-1-4939-2205-5_2,

However, when working with eukaryotic proteins, it has being estimated that approximately only 30 % of the cloned genes can be expressed in a soluble form in *E. coli* where the rest of the targets are degraded, expressed as insoluble aggregates known as inclusion bodies (IBs) or undetectable in cell extracts [3]. This is especially the case for membrane proteins or those requiring posttranslational modifications for folding or function. In order to overcome these limitations, several *E. coli* strains were developed as well as vectors carrying promoters with different strengths, fusion of the gene of interest with molecular tags that can aid in the purification and/or soluble production of the target protein, or the co-expression of chaperones or biological partners that can improve protein folding and stability [4, 5]. Furthermore, with the advent of the high-throughput screening (HTS) technology, all these variables can be evaluated in a simultaneous, fast, automated, and reliable manner in order to find the combination of the parameters that enable a soluble protein production [6]. Despite all this, soluble and homogeneous expressions of the target protein are not always the case. In this regard, many efforts were done with some success in the refolding of insoluble proteins from IBs [3, 7].

As an alternative strategy, the introduction of rational or, moreover, random mutations into the gene of interest in order to obtain a variant with stabilized properties or increased soluble expression has shown to be an attractive and effective approach in the soluble expression of target proteins, thus being a suitable last resource when everything else fails [8, 9].

2 Common Problems When Expressing Recombinant Proteins in *E. coli*

One of the main reasons why eukaryotic proteins often fail to be produced as soluble proteins in *E. coli* is the requirement of posttranslational modifications for correct folding. So a first step could be a sequence-based prediction of these modifications. In this regard, the ExPASy server (http://www.expasy.org) contains numerous bioinformatic tools that can estimate with a good accuracy the presence or not of posttranslational modifications like N- or O-glycosylation sites, phosphorylation sites, and protein localization, among others [10]. All this information can give us an idea of the possible success and help us in the strategy to follow for protein expression. Other factors that can have an impact in the soluble expression of the target are the codon usage, the sequence at the translation initiation region (TIR), as well as the correct formation of disulfide bridges. A brief description of the strategies designed to obtain soluble recombinant proteins is given in the next sections, and finally, in-depth information of them will be given in the following chapters of this book.

2.1 Effects of DNA/RNA Sequence in Protein Expression

The presence of uncommon codons for *E. coli* can have a strong influence in the gene expression. Because of the heterologous nature of the target protein, the target gene may have codons that are in low abundance in this host. This can lead into growth arrest, premature translation termination, and low yield of protein production, among others [11].

In order to overcome this problem, two different approaches have been proposed: (1) the substitution of the rare codons present in the gene sequence by de novo gene synthesis or (2) the expression of the gene of interest in an *E. coli* strain that is supplemented by tRNAs that are present in low abundance. In the former case, several algorithms were developed in order to optimize the gene sequence to the host codon usage [12, 13]. More recently, a software was developed in order to not only evaluate the codon frequency but also the codon pair usage or codon context. This approach suggests that codon pair usage and codon context can be as important as the individual codon usage [14]. For the second strategy, several commercially available strains have been developed that co-express tRNAs for rare codons, like BL21 CodonPlus (Novagen) and Rosetta™ (Invitrogen). The use of such strains demonstrated to be effective for the soluble expression of several targets [15, 16].

Finally, changing the rare codons can increase the translation rate, but in some cases this can lead to protein aggregation and misfolding as it was demonstrated for several proteins expressed in *E. coli* [17]. This suggests that translation pauses can be necessary in some cases for proper folding of individual domains [18]; thus the procedure of gene optimization or the use of a codon optimized for an *E. coli* strain cannot be used as a general rule.

Also at the DNA sequence level, it has been shown that the sequence at the 5′ of the gene can have an important impact in the levels of protein expression due to the generation of secondary structures in the messenger RNA that can hamper the translation by the ribosome complex. In this regard, it was shown that sequences immediately after the start codon up to position +25 can have a profound effect in protein expression. For these reasons, there are some programs that enable the optimization of the TIRs in order to improve protein expression by defining silent mutations in the first seven codons [19]. More recently a predictive method for designing synthetic ribosome binding sites was developed where different translation initiation rates can be targeted, thus enabling the rational control and fine tuning of recombinant protein expression [20, 21].

In the same way, in bacteria, the half-life of mRNA is much shorter than in eukaryotic cells. It was shown that mutation in the gene coding for RNaseE confers increased mRNA stability [22]. A BL21 derivative strain containing such mutation is commercialized by Invitrogen under the name of BL21 Star.

2.2 Disulfide Bonds: A Common Posttranslational Modification Implies a Common Problem for Recombinant Protein Expression

Disulfide bonds correspond to a covalent linkage between two sulfur atoms from two cysteine residues. They are frequently essential for proper folding, stability, and/or function of the target protein, thus a very important feature to take into account when expressing a target gene [23]. The presence of disulfide bonds can be predicted by web-based servers in order to estimate if the target protein can require such posttranslational modification [24, 25].

Disulfide bonds are formed in oxidizing environments like the eukaryotic endoplasmic reticulum or the bacterial periplasm. The formation of disulfide bridges in the periplasmic of E. coli, requires the action of DsbC system were DsbA catalyze disulfide formation whiles DsbC catalyze the isomerization of incorrectly formed disulfide bridges. The cycle can be restarted by the actions of the membrane proteins DsbB and DsbD that recycle DsbA and DsbC, respectively [26]. The expression of recombinant proteins in the periplasm of *E. coli* has allowed the correct formation of disulfide bridges of several targets [26, 27]. Purification of proteins from the periplasm is usually easier than purification of proteins from total cell lysates, since the periplasm contains a less complex protein mixture than the cytoplasm [28]. Targeting proteins to the periplasm of *E. coli* can be achieved by the addition of an N-terminal leader peptide that, depending on its nature, can use the Sec (relatively slow, posttranslational translocation) or the SRP (fast, cotranslational translocation) pathways that transport proteins through the inner plasma membrane as unfolded precursors [29, 30]. There is another translocation system: the twin-arginine translocation pathway, named Tat pathway, that, in contrast with the aforementioned pathways, catalyzes the translocation of proteins in their folded state [31].

However, one common drawback of periplasmic expression is that the translocation machinery can be saturated, which can be toxic for the host cell and decrease the final yield of the target protein. By using a strain where the expression intensity can be precisely controlled like Lemo21(DE3) (New England Biolabs), the saturation of the translocation machinery can be avoided, and thus these negative effects are minimized [32, 33].

As an alternative to periplasmic expression, engineered *E. coli* strains that contain a more oxidizing cytoplasm were developed in order to improve disulfide bond formation in this compartment. These strains contain mutations in the genes for glutathione reductase (*gor*) and thioredoxin reductase (*trxB*) involved in the maintenance of the reduced environment in the cytoplasm and a mutation in the peroxiredoxin gene *ahpC* essential for restoring growth in these mutants [23, 34]. One strain containing such mutations and used for the expression of disulfide bridges containing proteins is Origami, commercialized by Novagen [35]. However, a common problem for using such strains is the lack of disulfide bond isomerization. In this regard, a strain containing the trxB/gor/ahpC

mutations that express the DsbC isomerase in the cytoplasm of *E. coli* was developed and commercialized by New England Biolabs known as Shuffle, allowing the soluble expression of some disulfide-containing proteins in its cytoplasm [36]. Recently, by the co-expression of the sulfhydryl oxidase from *S. cerevisiae* Erv1p and the *E. coli* disulfide isomerase DsbC, disulfide bonds were generated in proteins expressed in the *E. coli* cytoplasm with the reducing pathways intact. Moreover, for some cases it was shown that the addition of a catalyst for the formation of disulfide bonds could be more effective than the removal of the reducing pathways [37, 38]. In the same sense, after making N-terminal fusions with DsbC with 28 different small disulfide-rich proteins, it was found that the strain BL21(DE3)pLysS was much more efficient in producing soluble and oxidized folded proteins in comparison to Origami B(DE3)pLysS or Shuffle T7 Express lysY cells [39]. Interestingly, when one of the fusions was used to evaluate if the disulfide bound formation occurred in the cytoplasm of BL21(DE3)pLysS cells or during the extraction and purification steps, it was found that this process occurred ex vivo [39].

3 Boosting Protein Purification and/or Expression

3.1 The Use of Fusion Tags/ Proteins

With the advent of genetic engineering, the target gene can be easily cloned in frame with different affinity and/or solubility-enhancing tags that can be exploited to increase protein solubility and yield and facilitate protein purification or downstream processing. In this regard, we can separate the fusion tags in three main categories. In the first category, referred as affinity tags, we found short tags that can be placed as N-terminus or C-terminus of the partner and can be recognized by special matrices or molecules serving for affinity purification of the fusion protein. In the second category, extremely soluble proteins with chaperone activities or thermostable characteristics in some cases are fused in order to transfer some of these properties to the fusion partner and improve the folding and/or the final yield of the target protein. Usually these tags are expressed as N-terminal fusions and are termed solubility-enhancing tags. Finally, we also have proteins that can offer a double purpose, in one hand, can be recognized by other molecules, thus serving for purification purposes, and, in the other hand, can improve the soluble production of the target protein, thus improving the target protein purity and yield [4, 40, 41].

3.2 Affinity Tags

Among the affinity tags, the His-tag is one of the most commonly used for purification of recombinant proteins in *E. coli*. This small tag (0.84 kDa) consists of 6–10 histidines in tandem and can reversibly interact with metal ions most commonly Ni or Co immobilized in a metal chelate matrix (Ni-NTA, Qiagen; Sepharose 6,

GE; or Talon resins, Clontech) [42], thus allowing mild elution conditions like the use of a competitor such as imidazole. The His-tag has several advantages like its small size and relatively inert nature, making it compatible with most downstream applications. Because the ternary structure of the His-tag is not necessary for metal coordination, it is possible to purify the protein in denaturing conditions or even perform the refolding procedure on column [43, 44]. Also the purification scheme has been automated in small and large-scale formats and has been used widely in HTS protocols, demonstrating the versatility of this tag [45–47].

As a disadvantage, when working with low-yield expressed proteins, it was shown that increasing the culture volume does not correlate with an increase in recovery. Moreover, there is a decrease in recovery because of the presence of small chelators mainly associated with the periplasm of *E. coli* that can decrease the binding capacity of the purification resin [48]. This can be improved by removing the periplasmic material before cell lysis [48]. Also it was shown that several histidine-rich *E. coli* proteins can bind to the column (like ArnA, SlyD, and GlmS), especially when working with low-expressing protein targets [49]. This reduces the purity of the target protein, consequently requiring the addition of more purification steps, thus reducing the final yield.

In this regard, some *E. coli* strains that are mutants in some of these proteins have been developed in order to overcome this issue [50, 51], and one is commercially available as NiCo21 (New England Biolabs).

Another strategy is the use of an alternative affinity tag such as Strep-tag II. This is also a small tag consisting of eight residues (WSHPQFEK) and can be specifically recognized by an engineering version of streptavidin (Strep-Tactin) [52]. Elution can be done as for the case of His-tag using mild conditions by competition with D-desthiobiotin for the Strep-tag II. Despite the binding capacity of the Strep-Tactin containing media can be lower when comparing to Sepharose 6 resins for His-tagged proteins, for example, its greater specificity makes it a good option when working with proteins that are expressed in very low quantities [52, 53]. Purification schemes for the Strep-tag II include prepacked columns as well as 96× well plates (www.iba-lifesciences.com). A variation of the Strep-tag II named Twin-Strep-tag® was recently developed and exhibited higher stability and affinity for the interaction with the Strep-Tactin. This tag consists of two *Strep*-tag®II-binding sequences connected by a short linker and showed to be more suitable for purification of diluted samples [54].

3.3 Solubility-Enhancing Tags

A common strategy to overcome the solubility problem is to fuse the target protein with a very stable and soluble one that can drive the resulting expression. It was shown for many proteins that were not soluble when expressed alone, that when expressed

as a fusion with other protein can be produced in a soluble and homogeneous state. Moreover, after cleavage and removal of the fusion partner, the target remained soluble demonstrating the utility of this approach [6, 39, 55, 56].

Among the commonly used solubility-enhancing fusion proteins, we can find the maltose-binding protein (MBP), glutathione S-transferase (GST), thioredoxin A (TrxA), disulfide isomerase C (DsbC), small ubiquitin-like modifier protein (SUMO), and N-utilization substance A (NusA).

An attractive feature of MBP and GST is that they can be used also as affinity tags. MBP is a 42 kDa protein expressed in the *E. coli* periplasm and can bind strongly to amylose resins, and elution can be done with free maltose [57]. For the case of GST, it is a 26 kDa protein from *Schistosoma japonicum* that can bind to glutathione resins, and elution is achieved by the application of reduced glutathione allowing a single-step purification process [58]. Despite GST protein is widely used, it has been shown to be a poor solubility enhancer, since in many cases after cleavage of the fusion, the target protein precipitates [6, 55, 59]. However, expression can be improved for some proteins or peptides, and the purification by glutathione resins makes it still an attractive option. Vectors for the expression of GST fusions can be found in the pGEX series from GE Healthcare or pET41a-c/pET42a-c from Novagen.

MBP was fused to either N- or C-terminus, where the expression and folding of eukaryotic fusion proteins was increased in many cases [59–61]. Vectors for MBP fusion can be found in the pMAL series from New England Biolabs or pIVEX from Roche. Also if the natural signal peptide of MBP is present ($MalE_{ss}$), expression can be directed to the periplasm of *E. coli*. This was used recently for the successful expression of disulfide-rich venom peptides [27].

TrxA is an 11.6 kDa *E. coli* thermostable (Tm: 85 °C) oxidoreductase that is expressed in very high yields. When used as a fusion tag, some of these properties can be transferred to the target protein improving its folding, solubility, and stability [62–64]. Moreover in a comparative study, all positive hits with Trx-fusions, remained still soluble after tag cleavage [6]. Expression vectors for fusion with Trx are pET32a-c from Novagen.

SUMO is a yeast protein (11.2 kDa) that when used as N-terminal fusion protein during prokaryotic expression can promote folding and soluble expression of the target protein [65–67]. Another advantage of this fusion is that it can be cleaved by a specific and efficient protease (yeast Ulp1) which recognize tertiary structure elements and a Gly-Gly-containing motif in the C-terminus of SUMO and can leave a native N-terminus on the target (except for proline) [66].

The *E. coli* disulfide isomerase DsbC (25 kDa) has isomerase and chaperonin activities [34] and has been successfully used as a fusion partner for the soluble expression of disulfide-containing

targets as mentioned earlier [39, 68]. The pET40 (Novagen) expression vector allows fusion with DsbC.

Finally, the transcription elongation and anti-termination factor of *E. coli* NusA (55 kDa) have also demonstrated to be useful for enhancing soluble protein expression [69]. In a comparative study after using several aggregation-prone target proteins, it was shown that the solubility-enhancing properties of NusA were comparable and similar to the well-studied MBP validating its utility [70]. Fusion with NusA can be achieved with the pET43.1a-c and pET44a-c vector series from Novagen.

Because TrxA, SUMO, DsbC, and NusA do not facilitate purification on their own, they are used in conjunction with small affinity tags like the aforementioned His-tag or Strep-tag II to enable protein purification. It is important to underline that despite some trends in fusion proteins were found in several studies, there is no rule for which is the best suited for the protein of interest, so it is better to test several different fusions in order to find the best option.

3.4 Tag Removal

Once the protein is expressed, in most of the cases, it is necessary to remove the fusion tag. This can be achieved by incorporating an aminoacidic sequence between the fusion tag and the protein of interest that can be recognized by a specific protease. Several proteases appear as possible options for tag removal like enterokinase (DDDDK′X), factor Xa (IE/DGR′X, where X can be any residue except for R or P), thrombin (LVPR′GS), PreScission™ protease (GE Healthcare, LEVLFQ′GP), and tobacco etch virus (TEV) protease (ENLYFQ′G), among others [41]. Between these, TEV protease is a very specific protease with several advantages like that it can be produced in the lab with high yields in *E. coli* [71], and cleavage can be done at 4 °C. Moreover, despite reducing conditions are optimal for cleavage (usually 1 mM DTT), if avoided, cleavage can still occur [27] which is preferable for disulfide-containing proteins. Also, it was demonstrated that the last glycine residue from the cleavage recognition site can be substituted by all residues except for proline, but at the expense of cleavage efficiency, allowing the release of a target protein with a native N-terminus [72].

Finally, fusion proteins were not only used for expression/purification purposes, but they were also used to obtain the crystallographic structures of several targets. This last brings the additional advantage that if the structure of the fusion is known, it can also help in the structure determination process of the target protein. Such is the case for some fusions with MBP, GST, Trx, and GFP, among others [73–76].

3.5 Cloning Methods

In order to succeed in the soluble expression of a “difficult” target protein, a recommended strategy is to test different fusion proteins, which requires the cloning of the gene of interest in several vectors. Doing this by restriction-based methods can be a complicated task,

principally when different restriction sites are present and even further when working with several targets at the same time. Nowadays some methodologies were developed as alternatives to the restriction-based cloning to facilitate the easy transfer of a DNA fragment into several vectors. Commercial kits like Gateway (Invitrogen) [77] and In-FusionTM (Clontech) [78] are efficient recombination-based cloning methods. For the case of Gateway, a suite of vectors for the easy transfer of the same DNA fragment between vectors is available. More recently, a cloning method based only in PCR reactions was developed and initially termed as RF cloning (RF, restriction free) [79]. In this method, the DNA is amplified with primers that contain complementary sequences with the site of insertion in the destination plasmid. So after the first PCR, the generated megaprimer is used in a second PCR to amplify the whole plasmid, inserting in this reaction the gene of interest in the desired position. The advantage is that insertion can be done at any position in the destination vector avoiding extra sequences to be added to the gene of interest, and if several vectors contain the same insertion sequence, the same megaprimer can be used in all of them, facilitating the cloning stage and allowing an automated HTS cloning approach [4, 79]. So by using a vector containing a fusion protein, just by inserting in the same position of the fusion other genes (like MBP, GST, SUMO, etc.), one can easily make its own suite of expression vectors where the site of insertion for the target gene is conserved along all vectors [80]. Recently, an improved protocol for RF cloning termed Transfer-PCR was developed where the generation of the megaprimer and subsequent integration of the PCR product into the destination vector occur in a single PCR reaction [81]. A web-based tool was developed for the correct design of the primers for RF cloning and is freely available (http://www.rf-cloning.org) [82]. The use of this kind of tools for molecular cloning is very useful for the generation of the genetic constructs necessary for finding a condition for soluble expression.

3.6 Expression Conditions

At the culture level, several parameters like induction temperature and medium composition can have an important effect in soluble protein yields. It was shown that lower temperature during induction (16–25 °C) can increase the final yield of soluble protein. It was assumed that a slower translation rate could favor the correct folding of the protein [83]. However, the lower temperature can also decrease the final biomass, so if the protein is well expressed, this can hamper the final yield [6]. In general, it is necessary to evaluate different temperatures to find the optimal condition. At the medium level, several media have been used for protein expression: Luria Broth (LB), 2xYT, Terrific Broth (TB), Super Broth (SB), autoinduction medium, and others. Among these media, the autoinduction medium, developed by Studier [84], has been used with success for protein expression screening in a wide range of

scales because it produces a high level of biomass. Thus, there is no need to monitor the growth; induction of cultures in well plates occurs at a comparable growth phase, which is preferable in HTS experiments; and there is a tighter control of protein induction improving expression of toxic proteins [6, 84]. A disadvantage of this medium is that it is adversely affected by aeration level. This can be reduced by decreasing the level of lacI repressor provided by the expression vector [85]. Recently, it was demonstrated that the oxygen sensitivity of expression in autoinduction medium can be practically obviated. This was achieved by using a glucose fed-batch-based autoinduction medium where the glycerol carbon source was substituted with the EnBase system [86]. This system is based on a soluble polysaccharide component within the medium and slow release of the glucose units from the polymer chain by an added specific biocatalyst [85, 87]. This kind of rich media allows an increase in the biomass production, so expression conditions can be evaluated in a reduced format like a 24× deep wells, enhancing the sensibility of automated HTS screenings for soluble protein production [4, 85, 88]. Also and as it was mentioned along the text, several strains should be used in order to find the proper condition; thus a combination of temperature and strain should be included in the screening. These in conjunction with the use of different constructs (i.e., fusion tags) make a considerable number of conditions to evaluate. In this regard, the HTS methods have had a pivotal role in making this kind of approaches possible [4, 6, 88, 89].

4 Inclusion Bodies' Renaturation

Frequently proteins accumulate, as insoluble aggregates in the cytoplasm or periplasm of *E. coli* known as inclusion bodies (IBs). As dramatically as it seems, this is not always a negative issue. Some advantages of expressing the protein as IBs are the high yield of its expression and the homogeneity in composition where in some cases the recombinant target can account for more than 90 % of the proteins in that fraction, facilitating the purification of the target after renaturation [90]. Renaturation conditions involve the evaluation of several parameters like pH, ionic strength, temperature, and addition of low molecular weight compounds, among others. In this regard, several approaches in a 96× well format have been developed to facilitate the optimization of the refolding conditions, and automated HTS protocols for protein refolding were proposed [7, 91]. Apart from the mentioned parameters, these can be combined with several methods to perform the refolding process like dilution, dialysis, or in-column refolding methods [7, 43, 92].

An attractive and counter-intuitive strategy is to introduce a tag that reduces the solubility of the fusion protein and can direct the expressed protein into insoluble IBs. This is particularly useful if the target protein is toxic to the host when soluble and correctly folded. In this regard, a mutant variant of the N-terminal autoprotease N^{pro}, of classical swine fever virus termed EDDIE, when fused to the N-terminus of the target protein can reduce its solubility in such a way that the fusion accumulates as IBs. When changing from chaotropic to kosmotropic conditions, the protease becomes active and can perform the autocleavage of the fusion, leaving a native N-terminus in the target protein [93, 94]. The comprehension of IBs nature has dramatically changed in the last years. Often, it was assumed that IBs were made of inert aggregates composed of denatured or partially folded polypeptides rather from mature native molecules. Nevertheless, in the last decades, it was shown in several cases that IBs can be made with native and active proteins [95–98]. This opens the possibility of using them in downstream applications without the need of performing protein renaturation in applications where protein aggregation is not an impediment, thus facilitating production/purification and reducing costs [99–101].

5 Protein Characterization

Obtaining the protein in a soluble state does not assure proper folding of the target protein. A common scenario is to find that the protein is soluble but forms aggregates. This is indicative of unfolded regions. A last purification step by size-exclusion chromatography (SEC) is recommendable to not only remove some remaining impurities but also to assess the oligomeric state of the sample. Protein quality assessment, can be implemented at the analytical level, with microgram quantities of protein by coupling for example, Ni Sepharose 6 beads or His MultiTrap FF 96-well plates (GE Healthcare) with the minicolumns for analytical SEC (ASEC), Superdex™ 5/150 GL (GE Healthcare), when still evaluating different expression conditions [88, 102, 103]. By using an autosampler for ASEC, the characterization step can be completely automated requiring only 14 min for each sample [102]. Also, sometimes it is necessary to evaluate different combination of additives, like for the case of membrane proteins, a combination of different detergents and/or lipids and genetic constructs in order to find a condition that gives a soluble and homogeneous sample. This kind of screening requires the purification of microgram to milligram of protein. A very useful alternative is to make GFP covalent fusions with the target protein and performing fluorescence-detection size-exclusion chromatography (FSEC). By using this approach, it is possible to determine the soluble expression

level, oligomeric state, thermostability, and approximate molecular mass using only nanogram quantities of unpurified protein, allowing working directly with the soluble extracts [104, 105]. Recently, a similar approach was developed where instead of fusing the target protein with GFP, a special fluorescent probe that can specifically recognize the His-tag was used, thus overcoming the limitations that can be associated in some cases with GFP fusions like the presence of false positives or protein aggregation issues following fusion cleavage [106].

6 Directed Evolution for Soluble Protein Expression

Despite the evaluation of many expression and growth conditions, it is often impossible to obtain the target protein in a soluble and stable manner. Under these circumstances, instead of exploring more expression parameters, one can change the physical properties of the target by making mutations or deletions in the target sequence in order to improve the solubility/stability of the recombinant protein. When structural and functional information are available, these sequence modifications can be achieved by rationally designed site-directed mutagenesis [107, 108]. Unfortunately, for most of the interesting targets, structural information is not available so rational design is not possible. In these cases, an interesting alternative is the use of directed evolution. This approach is based on an iterative process consisting of a first step of sequence diversification followed by a second step of selection of the improved mutants. The diversification process is usually achieved by random mutagenesis (error-prone PCR, chemical mutagenesis, or a mutator *E. coli* strain) [109] and/or in vitro recombination (DNA shuffling) [110]. In the directed evolution approaches, a library of mutants generated by a random process is screened for the solubility/stability of the target protein. So, after mutation occurs, one must select those few mutants with the improvements in the desired phenotype among the millions of futile mutants generated. In this regard, one can perform the selection by analyzing the activity of a reporter protein (reporter tag) or in special cases the activity of the target protein [111].

One folding reporter tag that was used successfully for the evolution of active and soluble mutant variants is the GFP-folding reporter [112, 113]. In this system, the test protein is expressed as an N-terminal fusion with GFP. So the fluorescence of *E. coli* cells is directly related to the productive folding of the fused protein [112]. In this way, the isolation of the brightest cells in the search for the mutations that improve solubility can be done using simple colony-plating techniques or fluorescence-activated cell sorting (FACS) in a flow cytometer. Later this system was improved even further by the design of a self-complemented split GFP [114]

derived from an exceptionally well-folded variant of GFP, "superfolder GFP" [115]. In this case, the target protein is fused as an N-terminal fusion to a small GFP fragment (residues 215–230, GFP11), while the GFP detector fragment (residues 1–214, GFP1-10) is expressed separately in another vector. So if the target protein is expressed in a soluble form, the GFP11 fragment can interact with GFP1-10, leading to the development of fluorescence [114].

In a different approach, the target protein can be expressed as an N-terminal fusion with a selectable marker such as the chloramphenicol acetyltransferase (CAT; 25 kDa), thus conferring resistance to chloramphenicol. It was observed that if the fusion protein is expressed in a soluble form, the cell is resistant to higher concentrations of chloramphenicol than when it is expressed in an insoluble form [116]. By using this method, it was possible to obtain soluble variants of the membrane-associated human cytochrome P450 (1A2), confirming the usefulness of this method [117].

More recently another antibiotic was used as a selectable marker but in a split manner linking in vivo protein stability to antibiotic resistance. In this case the target protein is inserted into the TEM1-β-lactamase (resistance to β-lactam antibiotics) as part of a tripartite fusion [8]. The antibiotic-resistance gene is separated between residues 196 and 197, for the insertion of the target protein gene. Thus, when protein is expressed in a soluble and stable form, the two fragments of β-lactamase can interact and thereby confer resistance to β-lactam antibiotics [8]. This method showed a low false-positive rate and, as for the CAT, is based on a selection rather than a screening for obtaining improved mutants.

Another elegant approach is the colony filtration (CoFi) blot. This is based in the fact that IBs can be separated from soluble proteins by filtration at the colony level. So after transforming bacteria with the mutant library, colonies are transferred to a filter membrane where protein expression is induced and cells are then lysed. Soluble proteins can diffuse through the filter and bind to the nitrocellulose membrane for detection [118, 119]. An anti-His antibody can be used for the detection of His-tagged soluble variants making it an easy to adopt method. Cornvik and colleagues randomized the N-terminal region of 32 mammalian proteins, and mutants were selected for soluble expression using this methodology. By this approach, the success rate for soluble expression was increased from 34 to 68 %, showing the high potential of this methodology [118].

Just as in the HTS, usually many different expression conditions for the same protein are evaluated; in the directed evolution approach, a library of mutants generated by a random process is screened for the solubilization/stabilization of the target protein. The key issues in this strategy are the diversity of the library and the selection/isolation method employed for finding the mutant with the improved characteristics.

7 Conclusions and Future Perspectives

Although a lot of progress has been made in recombinant protein expression, this field is still far for the generation of a universal protocol, so many different parameters are necessary to be evaluated for each target. The development of robotic technologies has facilitated the evaluation of an important number of different conditions reducing cost and effort through miniaturization of experiments. At present there are novel technological approaches (strain engineering, fusion technologies, and protein purification, among others), which are key factors that should be used in the lab to increase the success for the production of a soluble and homogeneous target protein.

Acknowledgments

This work was financed by a research grant from FCE-7273 and FMV-7323, 2011 from Agencia Nacional de Investigación e Innovación (ANII), Montevideo, Uruguay to P. Oppezzo. A. Correa was financed by a doctoral fellowship from ANII, Uruguay.

References

1. Sorensen HP (2010) Towards universal systems for recombinant gene expression. Microb Cell Fact 9:27
2. Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. J Ind Microbiol Biotechnol 39:383–399
3. Yang Z, Zhang L, Zhang Y et al (2011) Highly efficient production of soluble proteins from insoluble inclusion bodies by a two-step-denaturing and refolding method. PLoS One 6:e22981
4. Correa A, Oppezzo P (2011) Tuning different expression parameters to achieve soluble recombinant proteins in E. coli: advantages of high-throughput screening. Biotechnol J 6:715–730
5. Samuelson JC (2011) Recent developments in difficult protein expression: a guide to E. coli strains, promoters, and relevant host mutations. Methods Mol Biol 705:195–209
6. Vincentelli R, Cimino A, Geerlof A et al (2011) High-throughput protein expression screening and purification in Escherichia coli. Methods 55:65–72
7. Vincentelli R, Canaan S, Campanacci V et al (2004) High-throughput automated refolding screening of inclusion bodies. Protein Sci 13:2782–2792
8. Foit L, Morgan GJ, Kern MJ et al (2009) Optimizing protein stability in vivo. Mol Cell 36:861–871
9. Hart DJ, Waldo GS (2013) Library methods for structural biology of challenging proteins and their complexes. Curr Opin Struct Biol 23:403–408
10. Artimo P, Jonnalagedda M, Arnold K et al (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40:W597–W603
11. Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22: 346–353
12. Puigbo P, Guzman E, Romeu A et al (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35:W126–W131
13. Villalobos A, Ness JE, Gustafsson C et al (2006) Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformatics 7:285
14. Chung BK, Lee DY (2012) Computational codon optimization of synthetic gene for protein expression. BMC Syst Biol 6:134

15. Burgess-Brown NA, Sharma S, Sobott F et al (2008) Codon optimization can improve expression of human genes in Escherichia coli: a multi-gene study. Protein Expr Purif 59:94–102
16. Tegel H, Tourle S, Ottosson J et al (2010) Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). Protein Expr Purif 69: 159–167
17. Rosano GL, Ceccarelli EA (2009) Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. Microb Cell Fact 8:41
18. Marin M (2008) Folding at the rhythm of the rare codon beat. Biotechnol J 3:1047–1057
19. Voges D, Watzele M, Nemetz C et al (2004) Analyzing and enhancing mRNA translational efficiency in an Escherichia coli in vitro expression system. Biochem Biophys Res Commun 318:601–614
20. Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol 27:946–950
21. Salis HM (2011) The ribosome binding site calculator. Methods Enzymol 498:19–42
22. Makino T, Skretas G, Georgiou G (2011) Strain engineering for improved expression of recombinant proteins in bacteria. Microb Cell Fact 10:32
23. Salinas G, Pellizza L, Margenat M et al (2011) Tuned Escherichia coli as a host for the expression of disulfide-rich proteins. Biotechnol J 6:686–699
24. Ferre F, Clote P (2005) DiANNA: a web server for disulfide connectivity prediction. Nucleic Acids Res 33:W230–W232
25. Lin HH, Tseng LY (2010) DBCP: a web server for disulfide bonding connectivity pattern prediction without the prior knowledge of the bonding state of cysteines. Nucleic Acids Res 38:W503–W507
26. Berkmen M (2012) Production of disulfide-bonded proteins in Escherichia coli. Protein Expr Purif 82:240–251
27. Klint JK, Senff S, Saez NJ et al (2013) Production of recombinant disulfide-rich venom peptides for structural and functional analysis via expression in the periplasm of E. coli. PLoS One 8:e63865
28. Mergulhao FJ, Summers DK, Monteiro GA (2005) Recombinant protein secretion in Escherichia coli. Biotechnol Adv 23:177–202
29. den Blaauwen T, Driessen AJ (1996) Sec-dependent preprotein translocation in bacteria. Arch Microbiol 165:1–8
30. Luirink J, Sinning I (2004) SRP-mediated protein targeting: structure and function revisited. Biochim Biophys Acta 1694:17–35
31. Natale P, Bruser T, Driessen AJ (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. Biochim Biophys Acta 1778:1735–1756
32. Wagner S, Klepsch MM, Schlegel S et al (2008) Tuning Escherichia coli for membrane protein overexpression. Proc Natl Acad Sci U S A 105:14371–14376
33. Schlegel S, Rujas E, Ytterberg AJ et al (2013) Optimizing heterologous protein production in the periplasm of E coli by regulating gene expression levels. Microb Cell Fact 12:24
34. de Marco A (2009) Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. Microb Cell Fact 8:26
35. Bessette PH, Aslund F, Beckwith J et al (1999) Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. Proc Natl Acad Sci U S A 96: 13703–13708
36. Lobstein J, Emrich CA, Jeans C et al (2012) SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact 11:56
37. Hatahet F, Nguyen VD, Salo KE et al (2010) Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of E. coli. Microb Cell Fact 9:67
38. Nguyen VD, Hatahet F, Salo KE et al (2010) Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of E. coli. Microb Cell Fact 10:1
39. Nozach H, Fruchart-Gaillard C, Fenaille F et al (2013) High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in E. coli. Microb Cell Fact 12:37
40. Walls D, Loughran ST (2011) Tagging recombinant proteins to enhance solubility and aid purification. Methods Mol Biol 681:151–175
41. Young CL, Britton ZT, Robinson AS (2012) Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol J 7:620–634
42. Murphy MB, Doyle SA (2005) High-throughput purification of hexahistidine-tagged proteins expressed in E. coli. Methods Mol Biol 310:123–130

43. Zhu XQ, Li SX, He HJ et al (2005) On-column refolding of an insoluble His6-tagged recombinant EC-SOD overexpressed in Escherichia coli. Acta Biochim Biophys Sin (Shanghai) 37:265–269
44. Li M, Su ZG, Janson JC (2004) In vitro protein refolding by chromatographic procedures. Protein Expr Purif 33:1–10
45. Schafer F, Romer U, Emmerlich M et al (2002) Automated high-throughput purification of 6xHis-tagged proteins. J Biomol Tech 13:131–142
46. Vincentelli R, Canaan S, Offant J et al (2005) Automated expression and solubility screening of His-tagged proteins in 96-well format. Anal Biochem 346:77–84
47. Steen J, Uhlen M, Hober S et al (2006) High-throughput protein purification using an automated set-up for high-yield affinity chromatography. Protein Expr Purif 46:173–178
48. Magnusdottir A, Johansson I, Dahlgren LG et al (2009) Enabling IMAC purification of low abundance recombinant proteins from E. coli lysates. Nat Methods 6:477–478
49. 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
50. Robichon C, Luo J, Causey TB et al (2011) Engineering Escherichia coli BL21(DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. Appl Environ Microbiol 77:4634–4646
51. Andersen KR, Leksa NC, Schwartz TU (2013) Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins 81:1857–1861
52. Schmidt TG, Skerra A (2007) The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. Nat Protoc 2:1528–1535
53. Lichty JJ, Malecki JL, Agnew HD et al (2005) Comparison of affinity tags for protein purification. Protein Expr Purif 41:98–105
54. Schmidt TG, Batz L, Bonet L et al (2013) Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. Protein Expr Purif 92:54–61
55. Hammarstrom M, Hellgren N, van Den Berg S et al (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. Protein Sci 11:313–321
56. Esposito D, Chatterjee DK (2006) Enhancement of soluble protein expression through the use of fusion tags. Curr Opin Biotechnol 17:353–358
57. Pattenden LK, Thomas WG (2008) Amylose affinity chromatography of maltose-binding protein: purification by both native and novel matrix-assisted dialysis refolding methods. Methods Mol Biol 421:169–189
58. Smith DB, Johnson KS (1988) Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 67:31–40
59. Dyson MR, Shadbolt SP, Vincent KJ et al (2004) Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression. BMC Biotechnol 4:32
60. Kapust RB, Waugh DS (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 8:1668–1674
61. Cho HJ, Lee Y, Chang RS et al (2008) Maltose binding protein facilitates high-level expression and functional purification of the chemokines RANTES and SDF-1alpha from Escherichia coli. Protein Expr Purif 60:37–45
62. LaVallie ER, Lu Z, Diblasio-Smith EA et al (2000) Thioredoxin as a fusion partner for production of soluble recombinant proteins in Escherichia coli. Methods Enzymol 326:322–340
63. Kim S, Lee SB (2008) Soluble expression of archaeal proteins in Escherichia coli by using fusion-partners. Protein Expr Purif 62: 116–119
64. LaVallie ER, DiBlasio EA, Kovacic S et al (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. Biotechnology (N Y) 11:187–193
65. Marblestone JG, Edavettal SC, Lim Y et al (2006) Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO. Protein Sci 15:182–189
66. Malakhov MP, Mattern MR, Malakhova OA et al (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J Struct Funct Genomics 5:75–86
67. Butt TR, Edavettal SC, Hall JP et al (2005) SUMO fusion technology for difficult-to-express proteins. Protein Expr Purif 43:1–9
68. Zhang Z, Li ZH, Wang F et al (2002) Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in Escherichia coli. Protein Expr Purif 26:218–228

69. De Marco V, Stier G, Blandin S et al (2004) The solubility and stability of recombinant proteins are increased by their fusion to NusA. Biochem Biophys Res Commun 322:766–771
70. Nallamsetty S, Waugh DS (2006) Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. Protein Expr Purif 45:175–182
71. van den Berg S, Lofdahl PA, Hard T et al (2006) Improved solubility of TEV protease by directed evolution. J Biotechnol 121:291–298
72. Kapust RB, Tozser J, Copeland TD et al (2002) The P1′ specificity of tobacco etch virus protease. Biochem Biophys Res Commun 294:949–955
73. Moon AF, Mueller GA, Zhong X et al (2010) A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. Protein Sci 19:901–913
74. Suzuki N, Hiraki M, Yamada Y et al (2010) Crystallization of small proteins assisted by green fluorescent protein. Acta Crystallogr D Biol Crystallogr 66:1059–1066
75. Smyth DR, Mrozkiewicz MK, McGrath WJ et al (2003) Crystal structures of fusion proteins with large-affinity tags. Protein Sci 12:1313–1322
76. Corsini L, Hothorn M, Scheffzek K et al (2008) Thioredoxin as a fusion tag for carrier-driven crystallization. Protein Sci 17: 2070–2079
77. Esposito D, Garvey LA, Chakiath CS (2009) Gateway cloning for protein expression. Methods Mol Biol 498:31–54
78. Berrow NS, Alderton D, Sainsbury S et al (2007) A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. Nucleic Acids Res 35:e45
79. Unger T, Jacobovitch Y, Dantes A et al (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. J Struct Biol 172:34–44
80. Correa A, Ortega C, Obal G, Alzari P, Vincentelli R, Oppezzo P (2014) Generation of a vector suite for protein solubility screening. Front Microbiol. 5: 67
81. Erijman A, Dantes A, Bernheim R et al (2011) Transfer-PCR (TPCR): a highway for DNA cloning and protein engineering. J Struct Biol 175:171–177
82. Bond SR, Naus CC (2012) RF-Cloning.org: an online tool for the design of restriction-free cloning projects. Nucleic Acids Res 40:W209–W213
83. Vera A, Gonzalez-Montalban N, Aris A et al (2007) The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol Bioeng 96:1101–1106
84. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. Protein Expr Purif 41:207–234
85. Blommel PG, Becker KJ, Duvnjak P et al (2007) Enhanced bacterial protein expression during auto-induction obtained by alteration of lac repressor dosage and medium composition. Biotechnol Prog 23:585–598
86. Ukkonen K, Mayer S, Vasala A et al (2013) Use of slow glucose feeding as supporting carbon source in lactose autoinduction medium improves the robustness of protein expression at different aeration conditions. Protein Expr Purif 91:147–154
87. Krause M, Ukkonen K, Haataja T et al (2010) A novel fed-batch based cultivation method provides high cell-density and improves yield of soluble recombinant proteins in shaken cultures. Microb Cell Fact 9:11
88. Vincentelli R, Romier C (2013) Expression in Escherichia coli: becoming faster and more complex. Curr Opin Struct Biol 23:326–334
89. Koehn J, Hunt I (2009) High-throughput protein production (HTPP): a review of enabling technologies to expedite protein production. Methods Mol Biol 498:1–18
90. Ventura S, Villaverde A (2006) Protein quality in bacterial inclusion bodies. Trends Biotechnol 24:179–185
91. Dechavanne V, Barrillat N, Borlat F et al (2010) A high-throughput protein refolding screen in 96-well format combined with design of experiments to optimize the refolding conditions. Protein Expr Purif 75:192–203
92. Clark EDB (1998) Refolding of recombinant proteins. Curr Opin Biotechnol 9:157–163
93. Achmuller C, Kaar W, Ahrer K et al (2007) N(pro) fusion technology to produce proteins with authentic N termini in E. coli. Nat Methods 4:1037–1043
94. Ke T, Liang S, Huang J et al (2012) A novel PCR-based method for high throughput prokaryotic expression of antimicrobial peptide genes. BMC Biotechnol 12:10
95. Tokatlidis K, Dhurjati P, Millet J et al (1991) High activity of inclusion bodies formed in Escherichia coli overproducing Clostridium thermocellum endoglucanase D. FEBS Lett 282:205–208
96. Garcia-Fruitos E, Gonzalez-Montalban N, Morell M et al (2005) Aggregation as bacterial inclusion bodies does not imply inactiva-

tion of enzymes and fluorescent proteins. Microb Cell Fact 4:27

97. de Groot NS, Ventura S (2006) Protein activity in bacterial inclusion bodies correlates with predicted aggregation rates. J Biotechnol 125:110–113
98. Peternel S, Grdadolnik J, Gaberc-Porekar V et al (2008) Engineering inclusion bodies for non denaturing extraction of functional proteins. Microb Cell Fact 7:34
99. Garcia-Fruitos E (2010) Inclusion bodies: a new concept. Microb Cell Fact 9:80
100. Garcia-Fruitos E, Vazquez E, Diez-Gil C et al (2012) Bacterial inclusion bodies: making gold from waste. Trends Biotechnol 30:65–70
101. Villaverde A, Garcia-Fruitos E, Rinas U et al (2012) Packaging protein drugs as bacterial inclusion bodies for therapeutic applications. Microb Cell Fact 11:76
102. Low C, Moberg P, Quistgaard EM et al (2013) High-throughput analytical gel filtration screening of integral membrane proteins for structural studies. Biochim Biophys Acta 1830:3497–3508
103. Sala E, de Marco A (2010) Screening optimized protein purification protocols by coupling small-scale expression and mini-size exclusion chromatography. Protein Expr Purif 74:231–235
104. Hattori M, Hibbs RE, Gouaux E (2012) A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. Structure 20:1293–1299
105. Kawate T, Gouaux E (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. Structure 14:673–681
106. Backmark AE, Olivier N, Snijder A et al (2013) Fluorescent probe for high-throughput screening of membrane protein expression. Protein Sci 22:1124–1132
107. Dale GE, Broger C, Langen H et al (1994) Improving protein solubility through rationally designed amino acid replacements: solubilization of the trimethoprim-resistant type S1 dihydrofolate reductase. Protein Eng 7:933–939
108. Eijsink VG, Bjork A, Gaseidnes S et al (2004) Rational engineering of enzyme stability. J Biotechnol 113:105–120
109. Rasila TS, Pajunen MI, Savilahti H (2009) Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, Escherichia coli mutator strain, and hydroxylamine treatment. Anal Biochem 388:71–80
110. Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389–391
111. Roodveldt C, Aharoni A, Tawfik DS (2005) Directed evolution of proteins for heterologous expression and stability. Curr Opin Struct Biol 15:50–56
112. Waldo GS, Standish BM, Berendzen J et al (1999) Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol 17:691–695
113. Pedelacq JD, Piltch E, Liong EC et al (2002) Engineering soluble proteins for structural genomics. Nat Biotechnol 20:927–932
114. Cabantous S, Terwilliger TC, Waldo GS (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. Nat Biotechnol 23:102–107
115. Pedelacq JD, Cabantous S, Tran T et al (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24:79–88
116. Maxwell KL, Mittermaier AK, Forman-Kay JD et al (1999) A simple in vivo assay for increased protein solubility. Protein Sci 8:1908–1911
117. Sieber V, Martinez CA, Arnold FH (2001) Libraries of hybrid proteins from distantly related sequences. Nat Biotechnol 19: 456–460
118. Dahlroth SL, Nordlund P, Cornvik T (2006) Colony filtration blotting for screening soluble expression in Escherichia coli. Nat Protoc 1:253–258
119. Cornvik T, Dahlroth SL, Magnusdottir A et al (2005) Colony filtration blot: a new screening method for soluble protein expression in Escherichia coli. Nat Methods 2: 507–509

http://www.springer.com/978-1-4939-2204-8

Insoluble Proteins
Methods and Protocols
García-Fruitós, E. (Ed.)
2015, XVI, 425 p. 72 illus., 41 illus. in color., Hardcover
ISBN: 978-1-4939-2204-8
A product of Humana Press