Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies

Sudhir Sahdev · Sunil K. Khattar · Kulvinder Singh Saini

Abstract Among the various expression systems employed for the over-production of proteins, bacteria still remains the favorite choice of a Protein Biochemist. However, even today, due to the lack of post-translational modification machinery in bacteria, recombinant eukaryotic protein production poses an immense challenge, which invariably leads to the production of biologically in-active protein in this host. A number of techniques are cited in the literature, which describe the conversion of inactive protein, expressed as an insoluble fraction, into a soluble and active form. Overall, we have divided these methods into three major groups: *Group-I*, where the factors influencing the formation of insoluble fraction are modified through a stringent control of the cellular milieu, thereby leading to the expression of recombinant protein as soluble moiety; *Group-II*, where protein is refolded from the inclusion bodies and thereby target protein modification is avoided; *Group-III*, where the target protein is engineered to achieve soluble expression through fusion protein technology. Even within the same family of proteins (e.g., tyrosine kinases), optimization of standard operating protocol (SOP) may still be required for each protein's over-production at a pilot-scale in *Escherichia coli*. However, once standardized, this procedure can be made amenable to the industrial production for that particular protein with minimum alterations.

S. Sahdev (✉) · S. K. Khattar · K. S. Saini
Department of Biotechnology & Bioinformatics, New Drug
Discovery Research, Ranbaxy Research Laboratories-R&D-3,
20-Sector 18 Udyog Vihar, Gurgaon 122001, India
e-mail: sudhir.sahdev@ranbaxy.com

Introduction

In the last two decades or so, prokaryotic expression systems, particularly *E. coli*, have been exploited for the production of a variety of therapeutic proteins, on an industrial-scale. The challenges faced by the biotechnologists to achieve a finer balance between the quality and the yield of the desired protein, along with well-documented strategies to troubleshoot some of these hurdles, are discussed at length in this review.

Prokaryotic cells (*E. coli*) are normally the preferred host for the expression of foreign proteins because they offer (i) inexpensive carbon source requirements for growth, (ii) rapid biomass accumulation, (iii) amenability to high-cell density fermentation, and (iv) simple process scale up. However, lack of post-translational machinery and production of inactive protein due to the formation of inclusion bodies, offer a significant challenge in these expression systems.

A number of current protocols are available which describe various strategies for the conversion of inactive protein, expressed as insoluble inclusion bodies, into a soluble and active fraction. Overall, these strategies can be sub-divided under three major headings:

Group-I: The factors influencing the formation of insoluble fraction are modified through a tight control of the cellular milieu, thereby leading to soluble protein expression.
Group-II: Where expressed protein is refolded from the inclusion body fraction.
Group-III: The desirable protein expression is obtained in a soluble fraction through fusion protein production.

Before we discuss these three approaches at length, we would like to outline various genetic as well as

Table 1 The least used codons in *E. coli*, yeast, *drosophila*, *dictyostelium* and primates: [1, 2]

E. coli	Yeast	Drosophila	Dictyostelium	Primates	Amino acid
AGG, AGA, CGA, CGG	AGG, CGA, CGG, CGC	AGA, CGA, CGG	AGG, CGA, CGG, CGC	CGA, CGG, CGC, CGU	Arginine
AUA		AUA			Isoleucine
CUA	CUC	UUA	CUG		Leucine
CCC	CCG		CCG	CCG	Proline
UCG		AGU	UCG	UCG	Serine
	GCG		GCG	GCG	Alanine
	ACG		ACG	ACG	Threonine
		GGG	GGG		Glycine
		UGU			Cysteine
			CAG		Glutamine
			GUG		Valine

Table 2 Biotechnological innovations for producing biologically active proteins through *E. coli*

The challenges in *E. coli* expression system		Counter strategies	References
1. The Codon bias	Rare tRNA genes combination plasmids	pRIG, pRARE, pACYC, pSC101	[2–6]
	Translation of heterologous GC-rich genes	*BL*21-CodonPlus (DE3)-RIPL cells	[7–9]
	Cells with extra copies of rare tRNAs	Rosetta-gami strain	[10, 11]
2. Disulfide bond formation	Formation of disulfide bonds in *E. coli* periplasm	Dsb system	[12–15]
	Formation of disulfide bonds in *E. coli* cytoplasm	Host *trxB* and *gor* gene disruptions	[16–19]
	In vitro disulfide bond formation	Protein disulfide isomerase (PDI) as catalyst	[20–22]
	Non-specific disulfide-bond formation	1. S-sulfonation, 2. Glutathione, 3. Cysteamine-cystamine redox buffer	[23, 24]
3. Protein phosphorylation	Lack of most of the eukaryotic post-translational machinery	Kinase and its substrate co-expression via separate or single plasmid vectors	[25, 26]
	Endogenous phosphorylation	Recombinant protein phosphorylation by endogenous protein kinases	[27]
4. Protein refolding	Lack of cellular organelles required for glycosylation	Non-enzymatic glycosylation resulting in advanced glycation end product	[28]
	Large-scale production with position specific glycosylation	Co-translational synthesis by genetically encoded modified amino acids	[29]

environmental factors associated with bacterial expression systems, which in one way or the other influence successful production of biologically active recombinant proteins (Tables 1, 2).

Recombinant protein expression: the challenge to overcome the codon bias

More than one codon encodes most of the amino acids and all the available amino acid codons are utilized as per the codon bias of each organism. Transfer RNA of every cell closely reflects the codon bias of its mRNA [30, 31]. The requirement for one or more rare tRNAs during heterologous target gene over-expression in *E. coli* results in lower translation rate due to differences in codon usage [32–34]. If high-level *E. coli* expression vector systems are utilized, the presence of small number of rare codons in the

heterologous gene does not lead to significant reduction in the target protein synthesis. However, the same does not hold true when a gene encoding clusters of rare *E. coli* codon is over-expressed, especially when the N-terminus of a coding sequence consists of successive rare codons [1].

Rare codons in *E. coli*

Assessment of the codon usage in all *E. coli* genes reveals a number of codons that are under represented (Table 1). The codon usage of abundantly expressed genes exhibit a trend where the low usage codons [viz. AGA, AGG, and CGA (Arg); AUA (Ile) and CUA (Leu)] signifying less than 8% of their corresponding codon partners, are avoided and the codons GGA (Gly), CGG (Arg), and CCC (Pro) fall to less than 2% of their respective populations. Depending on the status of these codons, *E. coli* expresses the target genes at

a similar or higher level to those of well-expressed endogenous genes during its normal growth [1, 2].

Modification of the culture milieu (e.g., media composition, lower temperature) does not have any significant role in shifting the codon usage levels of most of the tRNA isoacceptors corresponding to the rare codon. These codon-usage bias-related translation problems will increase if over-expressed protein is abundant in a particular amino acid. Supplying the limiting amino acid in the culture medium leads to improved expression in some cases [1, 31, 32], and some of these issues are outlined below.

"On demand" supply of the rare codons

Expression of genes containing the rare codons can be improved with a relative increase of the associated tRNA within the E. coli host by manipulating the copy number of respective tRNA gene through a plasmid. Several studies have reported a significant increase in protein yield when E. coli hosts are utilized with additional argU (AGG/AGA), glyT (GGA), ileX (AUA), leuW (CUA), or proL (CCC) expression by increasing gene dosage of the respective tRNAs [2–6].

To optimize the expression of AT or GC rich genes with the corresponding codon usage bias, several combination plasmids, e.g., the pRIG, pRARE, pACYC, pSC101, etc. encoding various rare tRNA genes (argU, ileX, glyT and leuW) under the control of their native promoters, are being utilized. The BL21-CodonPlus (DE3)-RIPL cells (Stratagene, USA) have the extra copies of tRNAs (argU, ileY, leuW, and proL tRNA) to circumvent the most frequently restricted translation of heterologous GC-rich genes. It appears that this strain rescues the expression of heterologous proteins from organisms that have either AT-or GC-rich genomes thereby reducing the chances of truncated protein formation and resulting in full-length active proteins [7–9]. Despite these advances, the optimal utilization of rare codons for the over-production of a catalytically active eukaryotic protein at an industrial level, remains a challenge for the Protein Biochemist.

Protein folding and disulfide bond formation

Expression of recombinant proteins in E. coli is largely directed to three different locations namely; the cytoplasm, the periplasm, or the growth medium through secretion. Various advantages and disadvantages are associated with respect to directing the recombinant protein to a specific location. Expression in the cytoplasm is preferred as a norm since in this case, the production yields are usually high.

Disulfide bond formation in E. coli is an isolated and actively catalyzed reaction, which is performed in the periplasm through the Dsb system [12–15] where thioredoxin and glutaredoxin facilitate reduction of cysteines. However, thioredoxin and glutaredoxin are kept reduced by thioredoxin reductase (trxB) and by glutathione, respectively. Glutathione in turn is reduced by glutathione reductase (gor). Formation of disulfide bonds in the E. coli cytoplasm has been achieved by disrupting the trxB and gor genes. Furthermore, co-expression of either DsbC in strains lacking trxB as well as gor, or thioredoxin as a fusion partner can lead to improved folding and disulfide bond formation [16–19].

Apart from the Dsb system, protein disulfide isomerase (PDI), small peptides containing the active site of PDI or chemically synthesized dithiol molecules mimicking PDI function in combination with a redox system, have been utilized as catalysts in assisting disulfide bond formation during in vitro protein refolding, thereby improving the refolding rate as well as yield [20–22].

Reducing agents, such as dithiothreitol (DTT) or β-mercaptoethanol, are utilized for disrupting non-native disulfide bonds during solubilization of the inclusion body proteins by chaotropic agents (e.g., guanidinium hydrochloride, urea). Formation of native disulfide bonds following inclusion body solubilization can be achieved by adding low-molecular weight thiols (e.g., glutathione) in their reduced as well as oxidized state to the refolding buffer [35–38].

The protein aggregation problems associated with non-specific disulfide-bond formation, especially during the early stages of refolding, can be solved either by S-sulfonation or by altering the free cysteines through the addition of an oxidized form of a thiol reagent (e.g., glutathione) or a cysteamine-cystamine redox buffer. These chemical modifications introduce charge to the protein residues, which reduce aggregate formation as well as induce better inter-molecular interactions leading to the correct refolding [23, 24].

Eukaryotic protein phosphorylation in E. coli

Escherichia coli lacks most of the eukaryotic post-translational machinery, including serine/threonine/tyrosine protein kinases, which is considered a key disadvantage for producing the eukaryotic phosphoproteins. To circumvent this, introduction and expression of a protein kinase and its substrate from two separate or single plasmid vectors in the same E. coli host cell, has been well documented. This approach may be used for the co-expression of other mammalian protein modifying enzymes, such as protein methylases and acetylases and

their substrates. Co-expression of such enzymes and their substrates results in the production of recombinant proteins that closely resemble native eukaryotic proteins [25, 26].

Interestingly, few research findings have observed that *E. coli* could perform protein phosphorylation by endogenous protein kinases by utilizing adenosine triphosphate as a phosphoryl donor. Endogenous phosphorylation of isocitrate dehydrogenase; p68 RNA helicase; Single-stranded DNA-binding proteins (SSBs); are a few examples which clearly show that recombinant proteins expressed in *E. coli* are phosphorylated at multiple residues by endogenous protein kinases [27, 39, 40].

Protein glycosylation in *E. coli*

Glycosylation is a common but complex form of post-translational modification, which leads to formation of cellular glycans often attached to the proteins and lipids. Glycosyltransferases and glycosidases are responsible for glycosylation of many proteins. These enzymes significantly differ in carrying out dimerization, proteolytic digestion, secretion, and auto-glycosylation processes. Glycoproteins, which are commonly distributed in eukaryotic cells, appear to be rare in prokaryotic organisms, as cellular organelles required for glycosylation are missing [41].

Until recently, non-enzymatic glycosylation (glycation) was thought to affect only the eukaryotic proteins. Glycoprotein synthesis has been observed in prokaryotes with certain structural variations from their eukaryote counterparts and additional data on glycation in *E. coli* has now become available. Occurrence of non-enzymatic glycosylation associated with some post-translational modifications results in advanced glycation end (AGE) product formation during the normal growth cycle of *E. coli*. In addition, recombinant human interferon-γ (rhIFN-γ) expressed in *E. coli* has also been shown to be associated with early glycation products [28, 42].

Apart from natural glycosylation processing in *E. coli*, synthesis of selectively glycosylated proteins can also be achieved by co-translating the genetically encoded modified amino acids. This approach has been utilized for large-scale production of myoglobin containing the position specific β-N-acetylglucosamine (β-GlcNAc)-serine in *E. coli* [29].

Recombinant proteins: trans-membrane transport

Trans-membrane transport is normally mediated by the N-terminal signal peptides, which target the expressed protein to a specific transporter complex in the membrane. A considerable reduction in the amount of protein impurities can be achieved by targeting protein production in the host's periplasm. Other benefits of trans-membrane transport include a much higher probability of obtaining target protein with a correct N-terminus, decreased proteolysis due to lesser contaminants and simplified protein release by routine osmotic shock procedures [43, 44].

Periplasmic leader sequences frequently used for potential export of the recombinant proteins are derived from ompT, ompA, pelB, phoA, malE, lamB, and β-lactamase [15, 45–52]. Few proteins are exported across the inner membrane to the periplasm by utilizing a well-known Sec translocase apparatus [53, 54].

Recombinant protein folding modulators in the periplasm

One of the features that differentiate the periplasm from the cytoplasm is its oxidizing environment. In *E. coli*, proteins with stable disulfide-bond formation occur only in the cell envelope, catalyzed by thiol-disulfide oxidoreductases known as the Dsb proteins, as discussed earlier in this review. Potential periplasmic export and subsequent enhanced disulfide bond formation of the target proteins can be achieved via fusion to DsbA or DsbC [15, 55, 56]. Furthermore, the *E. coli* periplasmic chaperone Skp, a holdase, supports un-folded proteins as they surface from the Sec translocase apparatus for assisting the folding and membrane insertion of outer membrane proteins. Other periplasmic folding modulators include the PPIases, Sur A, FkpA, PpiA, PpiD, etc.; among these, FkpA has the most desirable folding activity with a combination of PPIase and chaperone functions [57–62].

Another soluble periplasmic protein, DsbA, with an active site embedded in a thioredoxin-like fold, promotes disulfide transfer to substrate proteins by the formation of mixed disulfide species. Here, DsbC may also come to the rescue of incorrect disulfide bond formation by seizing folding intermediates and catalyzing isomerization in a process involving mixed disulfide intermediates [63–65].

Recombinant proteins: secretory mode of expression

Most *E. coli* strains are characterized by the absence of well-organized pathways for protein translocation through the outer membrane. However, some proteins having periplasmic leader sequences for export may seep out into the extra-cellular medium. The signal recognition particle (SRP)-dependent pathway has been utilized for

exporting a subset of proteins [66, 67]. Bacterial SRP, a GTPase, can bind either to the signal sequence of certain secretary proteins (provided that it is highly hydrophobic) or to trans-membrane segments of inner membrane proteins as they emerge from the ribosome. The SRP-bound ribosome nascent chain complex (RNC) is then targeted to the membrane [68–70]. In other studies, the ompA signal sequence has been used for translocating recombinant peptide to the periplasm for probable secretion into the growth medium [71–73]. In addition, the co-expression of two secretion factors (secE and secY) has been shown to further increase the secretion of certain proteins [74–76].

Strategies for the production of active proteins through a tight control of the *E. coli* cellular milieu—Group-I

Induction of recombinant protein expression in *E. coli* involves synergy from host's transcriptional and translational machinery. As the expressed protein some times accounts up to $\approx 30\%$ of the total cell protein, metabolic load on the *E. coli* expression machinery is tremendous. Constant improvements in the protein solubilization and refolding procedures have become feasible through innovative R&D practices on the structure, function, and regulation of recombinant protein aggregation in *E. coli*. Some proteins adversely affect the host through their catalytic properties and sometimes over-production of protein may also cause toxicity to its host. Therefore, the host cell's capability to express proteins in soluble fraction also depends on the total metabolic load and can be regulated by a number of factors leading to the formation of soluble protein [77, 78].

These factors include:

Protein expression at lower temperatures

Protein expression in *E. coli* growing at low temperature has demonstrated its success in improving the solubility of a number of difficult proteins. Generally, expression at low temperature conditions leads to the increased stability and correct folding patterns, which is due to the fact that hydrophobic interactions determining inclusion body formation are temperature dependent. Apart from this, any expression associated toxic phenotype observed at 37°C incubation conditions, gets suppressed at lower temperatures [79–81]. The increased expression and activity at lower growth temperatures has also been associated with increased expression of a number of chaperones in *E. coli* [82]. It has been observed that heat shock proteases induced during over-expression are poorly active at lower

temperature conditions. Therefore, growth at a temperature range of 15–23°C, leads to a significant reduction in degradation of the expressed protein [83, 84].

Lower growth temperature conditions strongly affect the efficiency of traditional promoters, which are used in routine vectors for recombinant protein expression in *E. coli* [85]. To improve upon this shortcoming, a cspA promoter-based *E. coli* expression vector has been generated for stronger expression of recombinant proteins at lower temperatures. Production of proteins having membrane spanning domains, or otherwise unstable gene product, has been successfully carried out utilizing this versatile expression tool [85–87].

Protein expression data from our lab indicates that induction of human Phosphodiesterases (PDE-3A, PDE-5A) and p38-α map kinase expression at 22 and 18°C respectively, increased the production of functionally active enzymes as compared to growth temperatures of 30–37°C (unpublished data). Clearly, there will be disadvantages associated with growth at low temperatures, which might include reduced transcription and translation, ultimately leading to a poor rate of turnover of the recombinant protein.

Genetically modified *E. coli* strains for achieving improved protein solubility

*BL*21 *E. coli* strain (Novagen, USA) is an ideal organism for routine protein expression. *BL*21 and its derivatives are deficient in lon and OmpT proteases, a genetic modification, which is responsible for increased protein stability [88]. *BL*R, the recA$^-$ derivative of *BL*21, has been shown to stabilize target plasmids, which are having repetitive sequences or where such products may lead to the loss of DE3 prophage [89–91]. In the lac ZY deletion mutant of *BL*21 (Tuner^{TM}) one can achieve uniform and adjustable protein expression in all the cells. Furthermore, a genetically modified lac permease (lacY) mutation strain has been created, which offers uniform entry of IPTG into all the cells in the population [92], thereby, homogenous and concentration dependent levels of IPTG induction can be tightly regulated.

Relatively new entrants, Origami^{TM}2 (Novagen) host strains harbor both the *trxB* and *gor* gene mutations, thus, rendering greatly enhanced disulfide bond formation in their cytoplasm. Rosetta-gami strains, in addition to the above-mentioned features, also have an over expressing rare tRNA expression vector set for overcoming codon bias associated problems, as discussed earlier in this review. Origami-B host strains are derived from a lac ZY mutant of *BL*21 to enable precise control of expression levels by fine tuning the concentration of IPTG, as is in the case of lacY

mutation. In addition to trxB and gor mutations, these strains include the Ion and OmpT deficiencies of *BL*21, thereby enhancing the solubility of the recombinant proteins [10, 93].

Modification in media composition to obtain soluble protein

The level of intracellular accumulation of a recombinant protein in *E. coli* is dependent on its—specific activity—growth factor requirements as well as -the final cell density [11, 94–96]. Production of recombinant protein during batch culture requires nutrients for growth from the very beginning since there is a limited control on the growth parameters. This process often leads to changes in pH, concentration of dissolved oxygen, substrate depletion as well as accumulation of inhibitory intermediates from various metabolic pathways. These changes are detrimental for the production of soluble as well as correctly folded active protein. Depending on the nature of the expressed protein, proper and efficient protein folding might also require specific cofactors in the growth media, for example, metal ions such as iron-sulfur and polypeptide-cofactors e.g., flavin-mononucleotide (FMN). Addition of these factors to the batch culture considerably increases the yield as well as the folding rate of the soluble proteins [97, 98].

Development of the best expression media formulation involves both experimental and various permutations & combinations of some of the processes outlined above. Currently, for experimental plans, reliable statistical techniques are also in use, which provide best possible media compositions. A precise and comprehensive research scheme is needed for the successful prediction of optimal medium, which can then lead to the production of an active recombinant protein [99, 100].

The composition of growth medium affects the relative levels of soluble protein production. We have carried out a comparative study to standardize the expression of soluble and active protein by utilizing LB, SOC, Terrific broth, Super broth and M9 media. Compositions of various media were modified for achieving better biological activities of human PDE-3A, PDE-5A (unpublished data), and p38-α Map kinase [101] enzymes expressed in *BL*21 *E. coli* strain containing glycerol (1–2%) and glucose (0.8–1.0%), respectively. These modifications reduced the expression time, increased the soluble fraction yield and significantly enhanced the biological activity of these enzymes (Table 3).

Co-expression of molecular chaperones

Molecular chaperones are proteins adapted to assist *de novo* protein folding, facilitate expressed polypeptide's proper conformation attainment, and/or cellular localization without becoming a part of the final structure. Molecular chaperone co-expression strategy has been followed in many instances for preventing inclusion body formation, thereby leading to an improved solubility of the recombinant protein. Chaperones, like trigger factor, help in recombinant protein refolding where these polypeptides continue to attain folding into the native state even after their release from the protein-chaperone complex [103, 104]. Many other chaperones have also been shown to be involved in preventing protein aggregation [105–109]. Molecular chaperones have been divided into 3-functional sub-classes based on their mechanism of action:

1. *Folding chaperones* (e.g., DnaK and GroEL) mediate the net refolding/un-folding of their substrates through an ATP-dependent conformational change, thereby preventing the inclusion body formation via reduction in aggregation and promoting the proteolysis of misfolded proteins [103, 104, 108, 109].

2. *Holding chaperones* (e.g., IbpA, IbpB) can work in association with the folding chaperones to hold

Table 3 Modifications of various bacterial growth media compositions for enhancing solubility of the expressed proteins

Media	Composition	Modifications
LB Media	LB media Bacto tryptone, Yeast extract and NaCl	2 X concentrations (with Mg, Mn, Zn, etc.) (Unpublished data)
SOC Media	Tryptone, Yeast extract, NaCl, KCl, MgSO$_4$, and MgCl$_2$	Addition of glycylglycine dipeptide for increased solubility ([102], unpublished data)
Terrific Broth	Tryptone, Yeast extract, K$_2$HPO$_4$, KH$_2$PO$_4$, and Glycerol	Higher concentration (1–2%) of glycerol, for increased solubility & yield ([99], unpublished data)
SuperBroth	Tryptone, Yeast extract, NaCl, and Glycerol	Higher concentration (1–2%) of glycerol
M9 Media	Na$_2$HPO$_4$.7H$_2$O, KH$_2$PO$_4$, NaCl, NH$_4$Cl, MgSO$_4$ CaCl$_2$, and Casamino acids	Addition of Glucose at 0.8–0.10%, for increased solubility and yield ([100, 101], unpublished data)

partially folded proteins on their surface and under certain circumstances, protect heat-denatured proteins from irreversible aggregation [110, 111].

3. *Disaggregate chaperones* (ClpB, HtpG) promote the solubilization of stress-induced aggregated proteins [112–114].

Co-expression of various chaperone-encoding genes and recombinant target proteins has proven to be effective in getting the protein expressed in a soluble fraction [103]. As an example, GroEL–GroES and DnaK–DnaJ–GrpE chaperone system co-expression along with the trigger factor has been shown to further stimulate solubility of the expressed proteins [115, 116].

The major bottlenecks which are frequently encountered during the expression of eukaryotic proteins in bacterial system include: (i) inherent property of these recombinant proteins, which may interfere with the host's metabolic pathways, thereby causing toxicity; (ii) Expression at temperatures in the range of 37°C, leading to over-production of protein which invariably goes to the insoluble form; (iii) Lack of availability of optimal nutrient media during induction of the recombinant protein generally leads to majority of the protein getting in as insoluble fraction. In summary, the cell mass, growth factor requirements and specific activity of the bacterially expressed protein determines its physical outcome; whether the protein is going to be expressed in a soluble form or as an insoluble one.

Strategies for the production of active proteins by inclusion body refolding—Group-II

Inclusion bodies are refractive, intracellular protein aggregates usually observed when the target gene is over-expressed in the cytoplasm of *E. coli*. This aggregation of proteins leads to a highly unfavorable protein-folding environment. Formation of inclusion bodies in recombinant expression systems occurs as a result of erroneous equilibrium between in vitro protein aggregation and solubilization. Inclusion bodies accumulate in the cytoplasm or the periplasm depending on whether or not a recombinant protein has been engineered for periplasmic localization or secretion [113, 116, 117].

Re-solubilization and refolding of *E. coli* inclusion body proteins

The recombinant protein expression as inclusion bodies could be a blessing-in-disguise since it can be easily purified from the *E. coli* homogenate. Higher expression levels as well as differences in their compactness, compared to other cellular impurities, leads to their easier

physical separation e.g., through centrifugation. Apart from this, resistance to cellular proteases confers lower degradation and minimal impurities, therefore fewer purification steps are required, which drastically reduce the time and effort in getting the final product. Inclusion bodies facilitate straightforward purification of the protein of interest, although the loss of bioactivity of the expressed protein could be a major bottelneck. Despite this drawback, recombinant proteins expressed as inclusion bodies in *E. coli* have been most widely used for the commercial production of therapeutic proteins as the loss in protein recovery is greatly compensated by a very high level of expression [118–120].

So far, a major bioprocess challenge has been to convert these inactive and insoluble inclusion bodies into more efficient, soluble, and correctly folded product. The major drawbacks during the refolding of inclusion body proteins into bioactive components are; (a) reduced recovery, (b) the requirement for rigorous optimization of refolding conditions for each target protein, and (c) the possibility that the re-solubilization procedure could affect the activity of the refolded protein. Therefore, at times the purification of highly expressed soluble protein is less expensive and time saving option as compared to the refolding and purification from inclusion bodies. Exploiting the production of recombinant proteins in a soluble form still remains a preferable alternative to the in vitro refolding procedures.

Isolation and solubilization of inclusion bodies

A high degree of purification of the recombinant protein is achieved by isolation of the inclusion bodies. Under these circumstances, treatment with lysozyme along with EDTA before cell homogenization is carried out to facilitate cell disruption. Inclusion bodies are recovered by low speed centrifugation of bacterial cells mechanically disrupted either by using ultrasonication for -small, French press for -medium, or high pressure homogenization for -large scale cultures. Bacterial cell envelope as well as the outer membrane proteins, which co-precipitate with the insoluble cellular fractions during culture processing, form major fraction of the crude inclusion body impurities [102, 121, 122]. These contaminants can be easily removed by adding detergents such as Triton X-100 and/or low concentrations of chaotropic compounds, either prior to the mechanical disruption of the bacterial cells, or during crude inclusion body washing steps [123–127].

After the isolation and removal of the above-mentioned impurities, inclusion bodies are commonly solubilized by various concentrations of chaotropic agents such as

guanidinium hydrochloride, urea, or several other agents, like for example, arginine, which is an aggregation suppressor. Although expensive, guanidinium hydrochloride is generally favored due to its better chaotropic properties [128–133]. In comparison, inclusion body solubilization by urea is pH dependent and therefore, determination of the optimum pH conditions for each protein becomes a prerequisite. In addition, utilization of urea solutions has been shown to produce cyanate, which can carbamylate the amino groups of the expressed proteins, thereby negatively affecting the activity of the protein [134–136]. Inclusion body proteins solubilized under mild denaturant conditions have been shown to possess a native secondary structure and retain their biological activity up to certain extent [137–140]. It has been postulated that milder solubilization conditions lead to superior refolding yields compared to the solubilization by high concentrations of guanidinium hydrochloride or urea [141].

There are also reports suggesting that extreme pH conditions, in the presence or absence of low concentrations of denaturants [119, 142], frequently lead to the solubilization of inclusion bodies. However, extreme pH treatment often results in irreversible modifications such as deamidation and alkaline desulfuration of cysteine residues, which often alters bioactivity of the resultant protein [143–145] and therefore is not the choicest of method among the protein scientists.

In addition to the above-mentioned solubilizing agents, low molecular weight thiol reagents viz. DTT or 2-mercaptoethanol are often used during the inclusion body solubilization to reduce non-native inter- and intra-molecular disulfide bonds [136, 146–148]. Appropriate redox conditions are required for proteins having disulfide bonds in their native state. Mild alkaline pH offers optimum conditions for disruption of existing disulfide bonds. Before starting the refolding procedure, residual reducing substances, which negatively affect the refolding process, are removed (e.g., by dialysis). As an alternative, immobilized reducing agents (e.g., VectraPrime-Fluorochem Limited, UK; Immobilized DTT-DSH-Biovectra, USA) are also in the vogue, as they simplify removal of these reducing agents. This can easily be achieved by a centrifugation step after the solubilization is complete. To avoid the non-specific disulfide bond formation, the pH of the solution containing solubilized protein should be lowered before removing the reducing agents [136].

Aggregation and precise refolding of solubilized and un-folded proteins

Usually, the methods used for inclusion body solubilization could lead to a non-native conformation of the expressed protein. To achieve native form, the target protein needs to undergo proper refolding procedures during low denaturant concentrations. Protein refolding follows first order kinetics and has been shown to involve intra-molecular interactions. Protein aggregation, on the other hand, involves non-native inter-molecular interactions between protein folding intermediates, and is preferred at high protein concentrations. Higher concentration of the un-folded protein often leads to decreased refolding yields, irrespective of which refolding method has been applied [149–151]. Since refolding process does not implicate these aggregation-prone intermediates, folding is favored. Therefore, it is desirable to keep the initial un-folded protein concentration to a minimum level for achieving higher and correct refolding outputs. Apart from this, by avoiding the hydrophobic inter-molecular interactions during the first few steps of refolding, renaturation at high protein concentrations can be achieved successfully.

Techniques for protein refolding and activation

Dilution methods

Formation of the native structure of the protein can be achieved by diluting the protein-denaturant solution into a refolding buffer to support refolding instead of the aggregation pathway. However, this technique has considerable limitations if scale-up through huge industrial grade refolding vessels needs to be carried out, which generally makes this process expensive and cumbersome.

Protein refolding technology got a tremendous boost when the importance of adding the solubilized, denatured protein into the refolding buffer at very low concentration and slow rate was recognized [152–155]. Under these circumstances, an appropriate knowledge of the folding kinetics of the target protein is a prerequisite. It has been shown, for example in case of bone morphogenetic protein-2 and lysozyme, that the addition of concentrated protein-denaturant solution at a slower rate avoids the accumulation of aggregation-prone folding intermediates [152, 154–156]. Furthermore, during the refolding of proteins through pulse addition protocols, it has been suggested that to achieve maximum efficiency, 80% of the refolding yield should be reached, before adding the next pulse. Other factors which should be given due importance while following this strategy include increasing the concentration of the denaturant with each cycle, and the amount of protein added per pulse. These protocols should be standardized in batch cultures to minimize aggregation [136, 157].

Dialysis methods

Utilization of dialysis and diafiltration techniques, which involves gradual change of denaturing to native buffer conditions, convert solubilized and un-folded protein to its native structure [157, 158]. Mostly, these methods lead to more aggregation during refolding compared to the direct dilution method, as during dialysis, the protein has to pass through different concentrations of denaturant. In addition, non-specific adsorption of refolding protein to the dialysis membrane may also negatively affect the refolding yields. However, according to the requirements of the target protein, use of appropriate denaturant removal process generally leads to increased refolding yields even at high protein concentrations, as seen in the case of recombinant carbonic anhydrase as well as single-chain fragment variable (scFv) fusion protein refolding [159–161].

Chromatographic methods

Size exclusion chromatography (SEC) Buffer exchange for denaturant removal can also be carried out by performing SEC. Here, the denaturant-protein solution is injected into a refolding buffer pre-equilibrated column and the resultant refolded protein in the eluate fraction comes out at considerably higher concentration compared to the simple dilution technique, as seen during the refolding process of platelet derived growth factor, avian erythroblastosis virus E2 oncogene homolog 1 protein (ETS-1), bovine ribonuclease A and *E. coli* integration host factor [162–165]. Depending upon the kinetics, refolding may get completed in the column or occur in the eluate fraction particularly for the proteins with slow folding kinetics.

Aggregate formation can be reduced by physical partitioning of the aggregation prone folding intermediates in the absorbent gel [164]. Apart form this, re-solubilization of already existing aggregates can be achieved by running the denaturant at slower rates [136]. For proteins e.g., lysozyme, which exhibits better re-folding yield during gradual denaturant removal process, elution is performed, preferably by using a reducing denaturant gradient [161, 166–168]. An extra advantage of the SEC method is that during the refolding process, simultaneous purification of the target protein can be achieved [162]. Furthermore, some of the recent applications have shown the feasibility of using SEC for continuous processes of protein refolding (for example, using bovine alpha-lactalbumin as a model protein). The SEC has further been improved by coupling this process to ultra filtration and recycling units, which takes care of the carry over of re-solubilized aggregates formed during the refolding process [169].

Solid-support assisted protein re-folding To avoid the unwanted inter-molecular interaction between aggregation-prone folding intermediates, the solubilized and un-folded protein is attached to a solid support prior to changing conditions from denaturing to native buffer composition. Under these circumstances, attachment requires a stable protein-matrix complex formation, which can withstand the presence of chaotropic agents, and on the contrary, protein should be able to easily detach after changing to native buffer conditions. Several binding matrices have been used in various combinations for supporting binding of the un-folded protein and combining the process of target protein renaturation and its purification due to its selective binding properties [170–172].

Proteins with a naturally occurring charged patch in the un-folded chain can be refolded by first binding to an ion exchange resin followed by the employment of refolding parameters described by Li et al. [168, 173]. Protein refolding has been achieved by incorporating N-or C-terminal peptide tags, such as an anion exchange matrix binding denaturant-resistant glutathione-S-transferase [172]; metal ion binding hexa-histidine [168, 171, 174]; poly-anionic support binding hexa-arginine [175]; or the cellulose matrix binding domain [176]. Once the protein is associated with the matrix; techniques such as dilution, dialysis, or chromatographical buffer exchange can be used for refolding. The resultant protein can be separated from the matrix by using buffers with high ionic strength; with EDTA or through imidazole, depending on the type of protein-matrix interaction involved [168, 170–172, 175–177].

Hydrophobic interaction mediated refolding Hydrophobic Interaction Chromatography (HIC) involves attachment of the protein's hydrophobic region to the HIC matrix leading to the formation of micro domains and thereby allowing the native structure formation in its vicinity. At high salt concentrations, un-folded proteins attach to the hydrophobic matrix and facilitate refolding. Attainment of the native conformation of the un-folded protein is regulated by salt concentration and hydrophobicity of the intermediate(s), through several steps of adsorption and desorption during its migration in the column. Protein impurities are removed simultaneously, while the renaturation process is going on during HIC [178–180].

Strategies for production of active proteins through fusion—Group-III

Target recombinant proteins cannot always be obtained in a soluble form by the strategies outlined above. For proteins

with a tendency to go into insoluble fractions, it has been shown that, metabolic engineering through fusion protein technology usually leads to its soluble expression.

Fusion protein production

In order to simplify the expression, solubilization and purification of recombinant proteins, a wide range of protein fusion partners have been developed. These fusion proteins or chimeric proteins usually include a partner or "tag" which may or may not be linked to the target protein by a recognition site-specific protease. Most fusion partners are utilized for the purpose of specific affinity purification. These fusion tags offer additional advantages by protecting the partner protein from intracellular proteolysis [181, 182], enhance solubility [183–185] or they can be used as specific expression reporters viz. green-fluorescent protein (GFP) independently or in combination with blue-fluorescent protein (BFP) for performing fluorescence resonance energy transfer (FRET) assays [186, 187]. High expression levels often transfer from an N-terminal fusion tag, to a poorly expressing partner, most likely due to mRNA stabilization [188].

Common affinity tags are the poly-histidine tag (His-tag), which are compatible with immobilized metal affinity chromatography (IMAC) and the glutathione S-transferase (GST) tag for purification through glutathione based resins, as discussed earlier. Several other affinity tags exist and have been extensively studied [189–191]. Fusion partners of particular interest with regard to optimization of recombinant expression, include the E. coli maltose binding protein (MBP) and E. coli N-utilizing substance-A (NusA). MBP (40 kDa) and NusA (54.8 kDa) are specifically utilized to increase the solubility of proteins, which tend to form inclusion bodies. MBP is considered to be much better than GST and thioredoxin proteins with respect to its solubility enhancing properties. Comparatively, NusA provides high expressivity as well as good solubility characteristics, the properties which are considered as major advantages for the target protein expression [192–195].

Recently, a highly soluble N-terminal fragment of translation initiation factor, (IF2N, 17.4 kDa), has been utilized as a good solubility enhancing partner [185]. Utilizing such smaller fusion-tags often reduces the amount of energy required to obtain a certain level of expression by reducing steric hindrance. However, the result of fusion to a solubility tag may not necessarily lead to prevention of the inclusion-body formation, since it is more or less a target protein specific phenomenon.

Role of site-specific proteases

The recombinant purified protein is separated from its fusion partners such as affinity tags, solubility enhancers or expression reporters by the use of site-specific proteases. These enzymes cleave at recognition sequences present in various expression vectors. Selection of a specific protease and its optimal cleavage conditions mostly depend upon the amino acid sequences of target protein. Therefore, fusion tags with protease cleavage sequences similar to the one present in recombinant protein should be avoided.

Two serine proteases, namely factor Xa and thrombin, are widely utilized for site-specific fusion protein cleavage. Thrombin's recognition sequence is LVPR/G and Factor Xa recognizes the amino acid sequence IEGR/X for cleavage, (X denotes any amino acid except arginine or proline). Enterokinase, which recognizes DDDDK/X (X can be any amino acid except proline); Precision Protease (Amersham Biosciences), which cleaves at LEVLFQ/GP and the Tobacco Etch Virus protease (TEV), which cleaves at the sequence ENLYFQ/G, are a few examples of more specific proteases used for the intracellular processing of fusion proteins [196–199].

Perspectives and prospective

Recombinant proteins expressed as inclusion bodies in E. coli have been widely used for the production of therapeutic proteins in an industrial setting. New technologies have become available, which allow the production of biologically active proteins through a diverse array of in vitro refolding procedures. Despite these advances, production of recombinant proteins in a soluble form remains the method of choice for a Protein Biochemist.

How the biotechnology strategies, outlined in this review can be fully optimized during the scale-up of therapeutically important proteins remains to be seen. Industrial production of enzymatic proteins, antibodies, hormones, and other pharmaceutical biotech products in various expression systems still remains "an art." At least modern technologies and know-how now seems to have unraveled some of these mysteries with respect to the production of biologically active proteins. The vast accumulation of knowledge of upstream expression and downstream purification protocols has certainly helped us improve the quality as well as yield of the target protein in a desired manner. Hopefully, in the near future, we will be able to "tailor made" the overexpressing recombinant protein in a prokaryotic expression system, and also scale up its production from lab to the industrial level with minimum trouble shooting efforts.

References

1. Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res 29:292

2. Sharp PM, Devine KM (1989) Codon usage and gene expression level in *Dictyostelium discoideum*: highly expressed genes do 'prefer' optimal codons. Nucleic Acids Res 17:5029–5039

3. Rosenberg AH, Goldman E, Dunn JJ et al (1993) Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. J Bacteriol 175:716–722

4. Brinkmann U, Mattes RE, Buckel P (1989) High-level expression of recombinant genes in *Escherichia coli* is dependent on the availability of the DnaY gene product. Gene 85:109–114

5. Cinquin O, Christopherson RI, Menz RI (2001) A hybrid plasmid for expression of toxic malarial proteins in *Escherichia coli*. Mol Biochem Parasitol 117:245–247

6. Yu ZB, Jin JP (2007) Removing the regulatory N-terminal domain of cardiac troponin I diminishes incompatibility during bacterial expression. Arch Biochem Biophys 461:138–145

7. Trundova M, Celer V (2007) Expression of porcine circovirus 2 ORF2 gene requires codon optimized *E. coli* cells. Virus Genes 34:199–204

8. Kirienko NV, Lepikhov KA, Zheleznaya LA et al (2004) Significance of codon usage and irregularities of rare codon distribution in genes for expression of BspLU11III methyltransferases. Biochemistry (Mosc) 69:527–535

9. Sorensen HP, Sperling-Petersen HU, Mortensen KK (2003) Production of recombinant thermostable proteins expressed in *Escherichia coli*: completion of protein synthesis is the bottleneck. J Chromatogr B Analyt Technol Biomed Life Sci 786:207–214

10. Neubauer A, Soini J, Bollok M et al (2007) Fermentation process for tetrameric human collagen prolyl 4-hydroxylase in *Escherichia coli*: improvement by gene optimisation of the PDI/beta subunit and repeated addition of the inducer anhydrotetracycline. J Biotechnol 128:308–321

11. Peti W, Page R (2007) Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. Protein Expr Purif 51(1):1–10. Review

12. Andersen CL, Matthey-Dupraz A, Missiakas D et al (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

13. Liu X, Wang CC (2001) Disulfide-dependent folding and export of *Escherichia coli* DsbC. J Biol Chem 276:1146–1151

14. Kurokawa Y, Yanagi H, Yura T (2000) Overexpression of protein disulfide isomerase DsbC stabilizes multiple-disulfide-bonded recombinant protein produced and transported to the periplasm in *Escherichia coli*. Appl Environ Microbiol 66:3960–3965

15. Kurokawa Y, Yanagi H, Yura T (2001) Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human

16. nerve growth factor in *Escherichia coli*. J Biol Chem 276:14393–14399

17. Levy R, Weiss R, Chen G et al (2001) Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* trxB gor mutants via the coexpression of molecular chaperones. Protein Expr Purif 23:338–347

18. Jurado P, de Lorenzo V, Fernandez LA (2006) Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of *Escherichia coli*: evidence that chaperone activity is the prime effect of thioredoxin. J Mol Biol 357:49–61

19. 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 USA 96:13703–13708

20. Lilie H, McLaughlin S, Freedman R et al (1994) Influence of protein disulfide isomerase (PDI) on antibody folding *in vitro*. J Biol Chem 269:14290–14296

21. Klappa P, Hawkins HC, Freedman RB (1997) Interactions between protein disulphide isomerase and peptides. Eur J Biochem 248:37–42

22. Winter J, Klappa P, Freedman RB et al (2002) Catalytic activity and chaperone function of human protein-disulfide isomerase are required for the efficient refolding of proinsulin. J Biol Chem 277:310–317

23. Chew CC, Magallon T, Martinat N et al (1995) The relative protein disulphide isomerase (PDI) activities of gonadotrophins, thioredoxin and PDI. Biochem Soc Trans 23:394S

24. Masuda K, Kamimura T, Kanesaki M et al (1996) Efficient production of the C-terminal domain of secretory leukoprotease inhibitor as a thrombin-cleavable fusion protein in *Escherichia coli*. Protein Eng 9:101–106

25. Mukhopadhyay A (2000) Reversible protection of disulfide bonds followed by oxidative folding render recombinant hCGbeta highly immunogenic. Vaccine 18:1802–1810

26. Yue BG, Ajuh P, Akusjarvi G et al (2000) Functional coexpression of serine protein kinase SRPK1 and its substrate ASF/SF2 in *Escherichia coli*. Nucleic Acids Res 28:E14

27. Gosse ME, Padmanabhan A, Fleischmann RD et al (1993) Expression of Chinese hamster cAMP-dependent protein kinase in *Escherichia coli* results in growth inhibition of bacterial cells: a model system for the rapid screening of mutant type I regulatory subunits. Proc Natl Acad Sci USA 90:8159–8163

28. Mijakovic I, Petranovic D, Macek B et al (2006) Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. Nucleic Acids Res 34:1588–1596

29. Mironova R, Niwa T, Handzhiyski Y et al (2005) Evidence for non-enzymatic glycosylation of *Escherichia coli* chromosomal DNA. Mol Microbiol 55:1801–1811

30. Zhang Z, Gildersleeve J, Yang YY et al (2004) A new strategy for the synthesis of glycoproteins. Science 303:371–373

31. Ikemura T (1981) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes. J Mol Biol 146:1–21

32. Dong H, Nilsson L, Kurland CG (1996) Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. J Mol Biol 260:649–663

33. Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6:494–500

34. Kurland C, Gallant J (1996) Errors of heterologous protein expression. Curr Opin Biotechnol 7:489–493

35. Goldman E, Rosenberg AH, Zubay G et al (1995) Consecutive low-usage leucine codons block translation only when near the 5′ end of a message in *Escherichia coli*. J Mol Biol 245:467–473

36. Ahmed AK, Schaffer SW, Wetlaufer DB (1975) Non-enzymic reactivation of reduced bovine pancreatic ribonuclease by air

oxidation and by glutathione oxido-reduction buffers. J Biol Chem 250:8477–8482

36. Wetlaufer DB, Branca PA, Chen GX (1987) The oxidative folding of proteins by disulfide plus thiol does not correlate with redox potential. Protein Eng 1:141–146

37. Gough JD, Gargano JM, Donofrio AE et al (2003) Aromatic thiol pKa effects on the folding rate of a disulfide containing protein. Biochemistry 42:11787–11797

38. DeCollo TV, Lees WJ (2001) Effects of aromatic thiols on thiol-disulfide interchange reactions that occur during protein folding. J Org Chem 66:4244–4249

39. Garnak M, Reeves HC (1979) Phosphorylation of Isocitrate dehydrogenase of Escherichia coli. Science 203:1111–1112

40. Yang L, Liu ZR (2004) Bacterially expressed recombinant p68 RNA helicase is phosphorylated on serine, threonine, and tyrosine residues. Protein Expr Purif 35:327–333

41. El-Battari A, Prorok M, Angata K et al (2003) Different glycosyltransferases are differentially processed for secretion, dimerization, and autoglycosylation. Glycobiology 13:941–953

42. Mironova R, Niwa T, Hayashi H et al (2001) Evidence for non-enzymatic glycosylation in Escherichia coli. Mol Microbiol 39:1061–1068

43. Jonasson P, Liljeqvist S, Nygren PA et al (2002) Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli. Biotechnol Appl Biochem 35:91–105

44. Ewis HE, Lu CD (2005) Osmotic shock: a mechanosensitive channel blocker can prevent release of cytoplasmic but not periplasmic proteins. FEMS Microbiol Lett 253:295–301

45. Ignatova Z, Mahsunah A, Georgieva M et al (2003) Improvement of posttranslational bottlenecks in the production of penicillin amidase in recombinant Escherichia coli strains. Appl Environ Microbiol 69:1237–1245

46. Malik A, Rudolph R, Sohling B (2006) A novel fusion protein system for the production of native human pepsinogen in the bacterial periplasm. Protein Expr Purif 47:662–671

47. Oelschlaeger P, Lange S, Schmitt J et al (2003) Identification of factors impeding the production of a single-chain antibody fragment in Escherichia coli by comparing in vivo and in vitro expression. Appl Microbiol Biotechnol 61:123–132

48. Griswold KE, Mahmood NA, Iverson BL et al (2003) Effects of codon usage versus putative 5′-mRNA structure on the expression of Fusarium solani cutinase in the Escherichia coli cytoplasm. Protein Expr Purif 27:134–142

49. Adler J, Bibi E (2002) Membrane topology of the multidrug transporter MdfA: complementary gene fusion studies reveal a nonessential C-terminal domain. J Bacteriol 184:3313–3320

50. Obukowicz MG, Turner MA, Wong EY et al (1988) Secretion and export of IGF-1 in Escherichia coli strain JM101. Mol Gen Genet 215:19–25

51. Arie JP, Miot M, Sassoon N et al (2006) Formation of active inclusion bodies in the periplasm of Escherichia coli. Mol Microbiol 62:427–437

52. Jappelli R, Brenner S (1998) Changes in the periplasmic linker and in the expression level affect the activity of ToxR and lambda-ToxR fusion proteins in Escherichia coli. FEBS Lett 423:371–375

53. Tinker JK, Erbe JL, Holmes RK (2005) Characterization of fluorescent chimeras of cholera toxin and Escherichia coli heat-labile enterotoxins produced by use of the twin arginine translocation system. Infect Immun 73:3627–3635

54. Froderberg L, Houben EN, Baars L et al (2004) Targeting and translocation of two lipoproteins in Escherichia coli via the SRP/Sec/YidC pathway. J Biol Chem 279:31026–31032

55. Rietsch A, Belin D, Martin N et al (1996) An in vivo pathway for disulfide bond isomerization in Escherichia coli. Proc Natl Acad Sci USA 93:13048–13053

56. Ostermeier M, De Sutter K, Georgiou G (1996) Eukaryotic protein disulfide isomerase complements Escherichia coli dsbA mutants and increases the yield of a heterologous secreted protein with disulfide bonds. J Biol Chem 271:10616–10622

57. Chen R, Henning UA (1996) Periplasmic protein (SKP) of Escherichia coli selectively binds a class of outer membrane proteins. Mol Microbiol 19:1287–1294

58. Schafer U, Beck K, Muller M (1999) SKP, a Molecular chaperone of gram negative bacteria is required for formation of soluble periplasmic intermediates of outer membrane proteins. J Biol Chem 274:24567–24574

59. Missiakas D, Betton JM, Raina S (1996) New components of protein folding in extra-cytoplasmic components of Escherichia coli SurA, FkpA and Skp/OmpH. Mol Microbiol 21:871–884

60. Arie JP, Sasson N, Betton JM (2001) Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of Escherichia coli. Mol Microbiol 39:199–210

61. Bothmann H, Pluckthun A (2000) The periplasmic Escherichia coli peptidylprolyl cis,trans-isomerase FkpA. I. Increased functional expression of antibody fragments with and without cis-prolines. J Biol Chem 275:17100–17105

62. Dartigalongue C, Raina S (1998) A new heat-shock gene, ppiD, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in Escherichia coli. EMBO J 17:3968–3980

63. Segatori L, Murphy L, Arredondo S et al (2006) Conserved role of the linker alpha-helix of the bacterial disulfide isomerase DsbC in the avoidance of misoxidation by DsbB. J Biol Chem 281:4911–4919

64. Segatori L, Paukstelis PJ, Gilbert HF et al (2004) Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in Escherichia coli: Reconciling two competing pathways. Proc Natl Acad Sci USA 101:10018–10023

65. Debarbieux L, Beckwith J (1998) The reductive enzyme thioredoxin 1 acts as an oxidant when it is exported to the Escherichia coli periplasm. Proc Natl Acad Sci USA 95:10751–10756

66. Powers T, Walter P (1997) Co-translational protein targeting catalyzed by the Escherichia coli signal recognition particle and its receptor. EMBO J 16:4880–4886

67. Powers T, Walter P (1996) The nascent polypeptide-associated complex modulates interactions between the signal recognition particle and the ribosome. Curr Biol 6:331–338

68. Bacher G, Lutcke H, Jungnickel B et al (1996) Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. Nature 381:248–251

69. Spanggord RJ, Siu F, Ke A et al (2005) RNA-mediated interaction between the peptide-binding and GTPase domains of the signal recognition particle. Nat Struct Mol Biol 12:1116–1122

70. Miller JD, Wilhelm H, Gierasch L A et al (1993) GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. Nature 366:351–354

71. Humphreys DP, Sehdev M, Chapman AP et al (2000) High-level periplasmic expression in Escherichia coli using a eukaryotic signal peptide: importance of codon usage at the 5′ end of the coding sequence. Protein Expr Purif 20:252–264

72. Rockenbach SK, Dupuis MJ, Pitts TW et al (1991) Secretion of active truncated CD4 into Escherichia coli periplasm. Appl Microbiol Biotechnol 35:32–37

73. Bolla JM, Lazdunski C, Inouye M et al (1987) Export and secretion of overproduced OmpA-beta-lactamase in Escherichia coli. FEBS Lett 224:213–218

74. Neumann-Haefelin C, Schafer U, Muller M et al (2000) SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. EMBO J 19:6419–6426

75. Eichler J, Wickner W (1998) The SecA subunit of *Escherichia coli* preprotein translocase is exposed to the periplasm. J Bacteriol 180:5776–5779

76. Kim J, Lee Y, Kim C et al (1992) Involvement of SecB, a chaperone, in the export of ribose binding protein. J Bacteriol 174:5219–5227

77. Hunt I (2005) From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. Protein Expr Purif 40(1):1–22. Review

78. Donovan RS, Robinson CW, Glick BR (2000) Optimizing the expression of a monoclonal antibody fragment under the transcriptional control of the *Escherichia coli* lac promoter. Can J Microbiol 46:532–541

79. 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

80. Chesshyre JA, Hipkiss AR (1989) Low temperatures stabilize interferon a-2 against proteolysis in *Methylophilus methylotrophus* and *Escherichia coli*. Appl Microbiol Biotechnol 31:158–162

81. Niiranen L, Espelid S, Karlsen CR et al (2007) Comparative expression study to increase the solubility of cold adapted Vibrio proteins in *Escherichia coli*. Protein Expr Purif 52:210–218

82. Ferrer M, Chernikova TN, Yakimov MM (2003) Chaperons govern growth of *Escherichia coli* at low temperature. Nat Biotechnol 21:1266–1267

83. Hunke S, Betton JM (2003) Temperature effect on inclusion body formation and stress response in the periplasm of *Escherichia coli*. Mol Microbiol 50:1579–1589

84. Spiess C, Beil A, Ehrmann M (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell 97:339–347

85. Vasina JA, Baneyx F (1997) Expression of aggregation prone recombinant proteins at low temperatures: a comparative study of the *Escherichia coli* cspA and tac promoter systems. Protein Expr Purif 9:211–218

86. Phadtare S, Severinov K (2005) Extended -10 motif is critical for activity of the cspA promoter but does not contribute to low-temperature transcription. J Bacteriol 187:6584–6589

87. Fang L, Xia B, Inouye M (1999) Transcription of cspA, the gene for the major cold-shock protein of *Escherichia coli*, is negatively regulated at 37 degrees C by the 5′-untranslated region of its mRNA. FEMS Microbiol Lett 176:39–43

88. Shaw MK, Ingraham JL (1967) Synthesis of macromolecules by *Escherichia coli* near the minimal temperature for growth. J Bacteriol 94:157–164

89. Schmidt M, Viaplana E, Hoffmann F (1999) Secretion-dependent proteolysis of heterologous protein by recombinant *Escherichia coli* is connected to an increased activity of the energy-generating dissimilatory pathway. Biotechnol Bioeng 66:61–67

90. Studier FW (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J Mol Biol 219:37–44

91. Dubendorff JW, Studier FW (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. J Mol Biol 219:45–59

92. Picaud S, Olsson ME, Brodelius PE (2007) Improved conditions for production of recombinant plant sesquiterpene synthases in *Escherichia coli*. Protein Expr Purif 51:71–79

93. Saejung W, Puttikhunt C, Prommool T et al (2006) Enhancement of recombinant soluble dengue virus 2 envelope domain III protein production in *Escherichia coli* trxB and gor double mutant. J Biosci Bioeng 102:333–339

94. Boccazzi P, Zanzotto A et al (2005) Gene expression analysis of *Escherichia coli* grown in miniaturized bioreactor platforms for high-throughput analysis of growth and genomic data. Appl Microbiol Biotechnol 68:518–532

95. Losen M, Frolich B, Pohl M et al (2004) Effect of oxygen limitation and medium composition on *Escherichia coli* fermentation in shake-flask cultures. Biotechnol Prog 20:1062–1068

96. Broedel SE Jr, Papciak SM et al (2001) The selection of optimum media formulations for improved expression of recombinant proteins in *E. coli*, vol 2. Athena Environmental Sciences, Inc

97. Bruser T, Yano T, Brune DC et al (2003) Membrane targeting of a folded and cofactor-containing protein. Eur J Biochem 270:1211–1221

98. Apiyo D, Wittung-Stafshede P (2002) Presence of the cofactor speeds up folding of *Desulfovibrio desulfuricans* flavodoxin. Protein Sci 11:1129–1135

99. Van Dien SJ, Iwatani S, Usuda Y et al (2006) Theoretical analysis of amino acid-producing *Escherichia coli* using a stoichiometric model and multivariate linear regression. J Biosci Bioeng 102:34–40

100. Lee KM, Gilmore DF (2006) Statistical experimental design for bioprocess modeling and optimization analysis: repeated-measures method for dynamic biotechnology process. Appl Biochem Biotechnol 135:101–116

101. Khattar SK, Gulati P, Kundu PK et al (2007) Enhanced soluble production of biologically active recombinant human p38 mitogen-activated-protein kinase (MAPK) in *Escherichia coli*. Protein Pept Lett 14(8):756–760

102. Rathore AS, Sobacke SE, Kocot TJ et al (2003) Analysis for residual host cell proteins and DNA in process streams of a recombinant protein product expressed in *Escherichia coli* cells. J Pharm Biomed Anal 32:1199–1211

103. Deuerling E, Patzelt H, Vorderwulbecke S et al (2003) Trigger factor and DnaK possess overlapping substrate pools and binding specificities. Mol Microbiol 47:1317–1328

104. Nishihara K, Kanemori M, Yanagi H et al (2000) Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. Appl Environ Microbiol 66:884–889

105. Ehrnsperger M, Graber S, Gaestel M et al (1997) Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. EMBO J 16:221–229

106. Veinger L, Diamant S, Buchner J et al (1998) The small heat-shock protein IbpB from *Escherichia coli* stabilize stress-denatured proteins for subsequent refolding by a multi-chaperone network. J Biol Chem 273:11032–11037

107. Hu X, O'Hara L, White S et al (2007) Optimisation of production of a domoic acid-binding scFv antibody fragment in *Escherichia coli* using molecular chaperones and functional immobilisation on a mesoporous silicate support. Protein Expr Purif 52:194–201

108. Kondo A, Kohda J, Endo Y et al (2000) Improvement of productivity of active horseradish peroxidase in *Escherichia coli* by coexpression of Dsb proteins. J Biosci Bioeng 90:600–606

109. Xu Y, Hsieh MY, Narayanan N et al (2005) Cytoplasmic overexpression, folding, and processing of penicillin acylase precursor in *Escherichia coli*. Biotechnol Prog 21:1357–1365

110. Kitagawa M, Miyakawa M, Matsumura Y et al (2002) *Escherichia coli* small heat shock proteins, IbpA and IbpB, protect enzymes from inactivation by heat and oxidants. Eur J Biochem 269:2907–2917

111. Kuczynska-Wisnik D, Kedzierska S, Matuszewska E et al (2002) The *Escherichia coli* small heat-shock proteins IbpA and

IbpB prevent the aggregation of endogenous proteins denatured *in vivo* during extreme heat shock. Microbiology 148:1757–1765

112. Zolkiewski M (1999) ClpB cooperates with DnaK, DnaJ, and GrpE in suppressing protein aggregation. A novel multi-chaperone system from *Escherichia coli*. J Biol Chem 274:28083–28086

113. Thomas JG, Baneyx F (2000) ClpB and HtpG facilitate *de novo* protein folding in stressed *Escherichia coli* cells. Mol Microbiol 36:1360–1370

114. Zavilgelsky GB, Kotova VY, Mazhul MM et al (2002) Role of Hsp70 (DnaK-DnaJ-GrpE) and Hsp100 (ClpA and ClpB) chaperones in refolding and increased thermal stability of bacterial luciferases in *Escherichia coli* cells. Biochemistry (Mosc) 67:986–992

115. Xu HM, Zhang GY, Ji XD et al (2005) Expression of soluble, biologically active recombinant human endostatin in *Escherichia coli*. Protein Expr Purif 41:252–258

116. Yoshimune K, Ninomiya Y, Wakayama M et al (2004) Molecular chaperones facilitate the soluble expression of N-acyl-D-amino acid amidohydrolases in *Escherichia coli*. J Ind Microbiol Biotechnol 31:421–426

117. Ghosh S, Rasheedi S, Rahim SS et al (2004) Method for enhancing solubility of the expressed recombinant proteins in *Escherichia coli*. Biotechniques 37:418–423

118. Patra AK, Gahlay GK, Reddy BV et al (2000) Refolding, structural transition and spermatozoa-binding of recombinant bonnet monkey (*Macaca radiata*) zona pellucida glycoprotein-C expressed in *Escherichia coli*. Eur J Biochem 267:7075–7081

119. Patra AK, Mukhopadhyay R, Mukhija R et al (2000) Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from *Escherichia coli*. Protein Expr Purif 18:182–192

120. Kwon KS, Lee S, Yu MH (1995) Refolding of alpha 1-antitrypsin expressed as inclusion bodies in *Escherichia coli*: characterization of aggregation. Biochim Biophys Acta 1247:179–184

121. Venkiteshwaran A, Heider P, Matosevic S et al (2007) Optimized removal of soluble host cell proteins for the recovery of met-Human growth hormone inclusion bodies from *Escherichia coli* cell lysate using crossflow microfiltration. Biotechnol Prog May 5 [Epub ahead of print]

122. Batas B, Schiraldi C, Chaudhuri JB (1999) Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. J Biotechnol 68:149–158

123. Oneda H, Inouye K (1999) Refolding and recovery of recombinant human matrix metalloproteinase 7 (matrilysin) from inclusion bodies expressed by *Escherichia coli*. J Biochem (Tokyo) 126:905–911

124. Sunitha K, Chung BH, Jang KH et al (2000) Refolding and purification of Zymomonas mobilis levansucrase produced as inclusion bodies in fed-batch culture of recombinant *Escherichia coli*. Protein Expr Purif 18:388–393

125. Lopez-Vara MC, Gasset M, Pajares MA (2000) Refolding and characterization of rat liver methionine adenosyltransferase from *Escherichia coli* inclusion bodies. Protein Expr Purif 19:219–226

126. Guisez Y, Demolder J, Mertens N et al (1993) High-level expression, purification, and renaturation of recombinant murine interleukin-2 from *Escherichia coli*. Protein Expr Purif 4:240–246

127. Lim WK, Smith-Somerville HE, Hardman JK (1989) Solubilization and renaturation of overexpressed aggregates of mutant tryptophan synthase alpha-subunits. Appl Environ Microbiol 55:1106–1111

128. Rinas U, Hoffmann F, Betiku E et al (2007) Inclusion body anatomy and functioning of chaperone-mediated *in vivo* inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. J Biotechnol 127:244–257

129. Lu H, Yu M, Sun Y et al (2007) Expression and purification of bioactive high-purity mouse monokine induced by IFN-gamma in *Escherichia coli*. Protein Expr Purif Apr 14 [Epub ahead of print]

130. Liu X, He Z, Zhou M et al (2007) Purification and characterization of recombinant extracellular domain of human HER2 from *Escherichia coli*. Protein Expr Purif 53:247–254

131. Li M, Huang D (2007) On-column refolding purification and characterization of recombinant human interferon-lambda1 produced in *Escherichia coli*. Protein Expr Purif 53:119–123

132. Bruser T (2007) The twin-arginine translocation system and its capability for protein secretion in biotechnological protein production. Appl Microbiol Biotechnol [Epub ahead of print]

133. Takahashi S, Ogasawara H, Watanabe T et al (2006) Refolding and activation of human prorenin expressed in *Escherichia coli*: application of recombinant human renin for inhibitor screening. Biosci Biotechnol Biochem 70:2913–2918

134. Hagel P, Gerding JJT, Fieggen W et al (1971) Cyanate formation in solutions of urea I. Calculation of cyanate concentrations at different temperature and pH. Biochim Biophys Acta 243:366–373

135. Boyle DM, McKinnie RE, Joy WD et al (1999) Evaluation of refolding conditions for a recombinant human interleukin-3 variant (daniplestim). Biotechnol Appl Biochem 30:163–170

136. De Bernardez Clark E, Schwarz E, Rudolph R (1999) Inhibition of aggregation side reactions during *in vitro* protein folding. Methods Enzymol 309:217–236

137. Kurucz I, Titus JA, Jost CR et al (1995) Correct disulfide pairing and efficient refolding of detergent-solubilized single-chain Fv proteins from bacterial inclusion bodies. Mol Immunol 2:1443–1452

138. Tsumoto K, Umetsu M, Kumagai I et al (2003) Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine. Biochem Biophys Res Commun 312:1383–1386

139. Goulding CW, Perry LJ (2003) Protein production in *Escherichia coli* for structural studies by X-ray crystallography. J Struct Biol 142:133–143. Review

140. 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

141. Puri NK, Crivelli E, Cardamone M et al (1992) Solubilization of growth hormone and other recombinant proteins from *Escherichia coli* by using a cationic surfactant. Biochem J 285:871–879

142. Khan RH, Appa Rao KBC, Eshwari ANS et al (1998) Solubilization of recombinant ovine growth hormone with retention of native-like secondary structure and its refolding from the inclusion bodies of *Escherichia coli*. Biotechnol Prog 14:722–728

143. Hale JE, Butler JP, Gelfanova V et al (2004) A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis. Anal Biochem 333:174–181

144. Cindric M, Cepo T, Skrlin A et al (2006) Accelerated on-column lysine derivatization and cysteine methylation by imidazole reaction in a deuterated environment for enhanced product ion analysis. Rapid Commun Mass Spectrom 20:694–702

145. Righetti PG, Chiari M, Casale E et al (1989) Oxidation of alkaline immobiline buffers for isoelectric focusing in immobilized pH gradients. Appl Theor Electrophor 1:115–121

146. Danek BL, Robinson AS (2004) P22 tailspike trimer assembly is governed by interchain redox associations. Biochim Biophys Acta 1700:105–116

147. Lee SH, Carpenter JF, Chang BS et al (2006) Effects of solutes on solubilization and refolding of proteins from inclusion bodies with high hydrostatic pressure. Protein Sci 15:304–313

148. Tran-Moseman A, Schauer N, De Bernardez Clark E (1999) Renaturation of *Escherichia coli*-derived recombinant human macrophage colony-stimulating factor. Protein Expr Purif 16:181–189

149. Rudolph R, Zettlmeissl G, Jaenicke R (1979) Reconstitution of lactic dehydrogenase. Non-covalent aggregation vs. reactivation. 2. Reactivation of irreversibly denatured aggregates. Biochemistry 18:5572–5575

150. Goldberg ME, Rudolph R, Jaenicke R (1991) A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme. Biochemistry 30:2790–2797

151. De Bernardez Clark E, Hevehan D, Szela S et al (1998) Oxidative renaturation of hen egg-white lysozyme. Folding vs. aggregation. Biotechnol Prog 14:47–54

152. Vallejo LF, Rinas U (2004) Optimized procedure for renaturation of recombinant human bone morphogenetic protein-2 at high protein concentration. Biotechnol Bioeng 85:601–609

153. Buchner J, Pastan I, Brinkmann U (1992) A method for increasing the yield of properly folded recombinant fusion proteins: Single chain immunotoxins from renaturation of bacterial inclusion bodies. Anal Biochem 205:263–270

154. Fischer B, Perry B, Sumner I et al (1992) A novel sequential procedure to enhance the renaturation of recombinant protein from *Escherichia coli* inclusion bodies. Protein Eng 5:593–596

155. Terashima M, Suzuki K, Katoh S (1996) Effective refolding of fully reduced lysozyme with a flow-type reactor. Process Biochem 31:341–345

156. Ho JGS, Middelberg APJ, Ramage P et al (2003) The likelihood of aggregation during protein renaturation can be assessed using the second virial coefficient. Protein Sci 12:708–716

157. Maeda Y, Koga H, Yamada H (1995) Effective renaturation of reduced lysozyme by gentle removal of urea. Protein Eng 8:201–205

158. Varnerin JP, Smith T, Rosenblum CI et al (1998) Production of leptin in *Escherichia coli*: a comparison of methods. Protein Expr Purif 14:335–342

159. West SM, Chaudhuri JB, Howell JA (1998) Improved protein refolding using hollow-fibre membrane dialysis. Biotechnol Bioeng 57:590–599

160. Yoshii H, Furuta T, Yonehara T et al (2000) Refolding of denatured/reduced lysozyme at high concentration with diafiltration. Biosci Biotechnol Biochem 64:1159–1165

161. Gu Z, Weidenhaupt M, Ivanova N et al (2002) Chromatographic methods for the isolation of, and refolding of proteins from, *Escherichia coli* inclusion bodies. Protein Express Purif 25:174–179

162. Müller C, Rinas U (1999) Renaturation of heterodimeric platelet derived growth factor from inclusion bodies of recombinant *Escherichia coli* using size-exclusion chromatography. J Chromatogr A 855:203–213

163. Werner MH, Clore GM, Gronenborn AM et al (1994) Refolding proteins by gel filtration chromatography. FEBS Lett 345:125–130

164. Batas B, Chaudhuri JB (1996) Protein refolding at high concentration using size-exclusion chromatography. Biotechnol Bioeng 50:16–23

165. Batas B, Chaudhuri JB (1999) Considerations of sample application and elution during size-exclusion chromatography-based protein refolding. J Chromatogr A 864:229–236

166. Gu Z, Su Z, Janson J-C (2001) Urea gradient size-exclusion chromatography enhanced the yield of lysozyme refolding. J Chromatogr A 918:311–318

167. Fahey EM, Chaudhuri JB (2000) Refolding of low molecular weight urokinase plasminogen activator by dilution and size exclusion chromatography-a comparative study. Sep Sci Technol 35:1743–1760

168. Li M, Poliakov A, Danielson UH et al (2003) Refolding of a recombinant full-length non-structural (NS3) protein from hepatitis C virus by chromatographic procedures. Biotechnol Lett 25:1729–1734

169. Schlegl R, Iberer G, Machold C et al (2003) Continuous matrix-assisted refolding of proteins. J Chromatogr A 1009:119–132

170. Zouhar J, Nanak E, Brzobohatý B (1999) Expression, single-step purification, and matrix-assisted refolding of a maize cytokinin glucoside-specific β-glucosidase. Protein Expr Purif 17:153–162

171. Rehm BHA, Qi Q, Beermann BB et al (2001) Matrix assisted *in vitro* refolding of Pseudomonas aeruginosa class II polyhydroxylalkanoate synthase from inclusion bodies produced in recombinant *Escherichia coli*. Biochem J 358:263–268

172. Cho TH, Ahn SJ, Lee EK (2001) Refolding of protein inclusion bodies directly from *E. coli* homogenate using expanded bed adsorption chromatography. Bioseparation 10:189–196

173. Li M, Zhang G, Su Z (2002) Dual gradient ion-exchange chromatography improved refolding yield of lysozyme. J Chromatogr A 959:113–120

174. Negro A, Onisto M, Grassato L et al (1997) Recombinant human TIMP-3 from *Escherichia coli*: Synthesis, refolding physicochemical and functional insights. Protein Eng 10:593–599

175. Stempfer G, Höll-Neugebauer B, Rudolph R (1996) Improved refolding of an immobilized fusion protein. Nat Biotechnol 14:329–334

176. Berdichevsky Y, Lamed R, Frenkel D et al (1999) Matrix-assisted refolding of single-chain Fv-cellulose binding domain fusion proteins. Protein Expr Purif 17:249–259

177. Suttnar J, Dyr JE, Hamšíková E et al (1994) Procedure for refolding and purification of recombinant proteins from *Escherichia coli* inclusion bodies using a strong anion exchanger. J Chromatogr B 656:123–126

178. Geng X, Chang X (1992) High-performance hydrophobic interaction chromatography as a tool for protein refolding. J Chromatogr A 599:185–194

179. Bai Q, Kong Y, Geng X (2003) Studies on renaturation with simultaneous purification of recombinant human pro-insulin from *E. coli* with high performance hydrophobic interaction chromatography. J Liq Chromatogr Relat Technol 26:683–695

180. Gong B, Wang L, Wang C et al (2004) Preparation of hydrophobic interaction chromatographic packings based on monodisperse poly(glycidylmethacrylate-*co*-ethylenedimethacrylate) beads and their application. J Chromatogr A 1022:33–39

181. Jacquet A, Daminet V, Haumont M et al (1999) Expression of a recombinant *Toxoplasma gondii*ROP2 fragment as a fusion protein in bacteria circumvents insolubility and proteolytic degradation. Protein Expr Purif 17:392–400

182. Martinez A, Knappskog PM, Olafsdottir S et al (1995) Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. Biochem J 306:589–597

183. Davis GD, Elisee C, Newham DM et al (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. Biotechnol Bioeng 65:382–388

184. 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

185. Sørensen HP, Sperling-Petersen HU, Mortensen KK (2003) A favorable solubility partner for the recombinant expression of streptavidin. Protein Expr Purif 32:252–259

186. Cabantous S, Waldo GS (2006) In vivo and in vitro protein solubility assays using split GFP. Nat Methods 3:845–854

187. Philipps B, Hennecke J, Glockshuber R (2003) FRET-based in vivo screening for protein folding and increased protein stability. J Mol Biol 327:239–249

188. Arechaga I, Miroux B, Runswick MJ et al (2003) Overexpression of Escherichia coli F1F(o)-ATPase subunit a is inhibited by instability of the uncB gene transcript. FEBS Lett 547:97–100

189. Lee NP, Tsang S, Cheng RH et al (2006) Increased solubility of integrin betaA domain using maltose-binding protein as a fusion tag. Protein Pept Lett 13:431–435

190. Kim JS, Valencia CA, Liu R et al (2007) Highly-efficient purification of native polyhistidine-tagged proteins by multivalent NTA-modified magnetic nanoparticles. Bioconjug Chem 18:333–341

191. Hammarstrom M, Woestenenk EA, Hellgren N et al (2006) Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein. J Struct Funct Genomics 7:1–14

192. Feeney B, Soderblom EJ, Goshe MB et al (2006) Novel protein purification system utilizing an N-terminal fusion protein and a caspase-3 cleavable linker. Protein Expr Purif 47:311–318

193. Klein S, Geiger T, Linchevski I et al (2005) Expression and purification of active PKB kinase from Escherichia coli. Protein Expr Purif 41:162–169

194. de Marco A (2006) Two-step metal affinity purification of double-tagged (NusA-His6) fusion proteins. Nat Protoc 1:1538–1543

195. Cabrita LD, Dai W, Bottomley SP (2006) A family of E. coli expression vectors for laboratory scale and high throughput soluble protein production. BMC Biotechnol 6:12

196. Panagbko C, Morley S, Neely S et al (2002) Expression and refolding of recombinant human alpha-tocopherol transfer protein capable of specific alpha-tocopherol binding. Protein Expr Purif 24:395–403

197. Donath MJ, de Laaf RT, Biessels PT et al (1995) Characterization of des-(741–1668)-factor VIII, a single-chain factor VIII variant with a fusion site susceptible to proteolysis by thrombin and factor Xa. Biochem J 312:49–55

198. Prachayasittikul V, Isarankura Na Ayudhya C, Piacham T et al (2003) One-step purification of chimeric green fluorescent protein providing metal-binding avidity and protease recognition sequence. Asian Pac J Allergy Immunol 21:259–267

199. Kapust RB, Waugh DS (2000) Controlled intracellular processing of fusion proteins by TEV protease. Protein Expr Purif 19:312–318