J. Microbiol. Biotechnol. (2004), **14**(1), 1–14

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

Design of Bacterial Vector Systems for the Production of Recombinant Proteins in *Escherichia coli*

MERGULHÃO, FILIPE J. M.*, GABRIEL A. MONTEIRO, JOAQUIM M. S. CABRAL, AND M. ANGELA TAIPA

Centro de Engenharia Biológica e Química, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

Received: May 30, 2003
Accepted: October 8, 2003

Abstract More than twenty years have passed since the approval of the first recombinant DNA product for therapeutic use (recombinant human insulin, 1982). However, the biotechnology industry is still facing a shortage of manufacturing capacity due to the increasing demand of therapeutic proteins. This demand has prompted the search for a growing number of biological production systems but, nevertheless, the Gram-negative bacterium *Escherichia coli* remains one of the most attractive production hosts. This review highlights the most important features and developments of plasmid vector design, emphasizing the different reported strategies for improving the expression and secretion of heterologous proteins using the cellular machinery of *E. coli*.

Key words: Recombinant proteins, *Escherichia coli*, plasmid vectors

The Gram-negative bacterium *Escherichia coli* has been the favored host for the purpose of recombinant protein production [2, 84, 196]. This is due to its ability to grow rapidly and at high density on inexpensive substrates, its well-characterized genetics, and the availability of an increasingly large number of cloning vectors and mutant host strains [8, 76]. This organism has the ability to accumulate many recombinant gene products to at least 20% of the total cell protein [133] and, in some cases, to translocate them from the cytoplasm to the periplasm [35].

Proteins like interferons, interleukins, growth hormones, and human serum albumines have been successfully expressed in *E. coli* [97], although this recombinant system cannot be used to produce some large complex proteins, or proteins that require post-translational modification to become biologically active. The expression of complex proteins such as those involving extensive disulfide bond formation requires new approaches in terms of controlled expression systems. Transcriptional control is an effective way to attain graded expression levels, and new promoters are being exploited, with advantages in terms of inducibility and transcriptional activity, towards an optimization of cell factory resources.

Significant progress has been made in the expression and *in vivo* folding of mammalian proteins in *E. coli* [153, 194], and its secretion capacity is being explored with fruitful results [47, 48, 100, 131, 199]. New applications like the synthesis of peptide nucleic acids [125] and the production of recombinant peptides [50, 117] are proof of the renewed interest in this bacterium, suggesting that it will remain one of the most widely used expression hosts in the near future.

The successful production of recombinant proteins in *E. coli* depends on a great number of factors which can be grouped under four major determinants: the host strain, the type of expression vector, the cultivation conditions, and the purification of the target gene product (Fig. 1).

Plasmid Vectors

The term "plasmid" was introduced in 1952 as a generic term for any extrachromosomal genetic element [93]. *E. coli* plasmids have traditionally been used as expression vectors for the overproduction of proteins.

In general, a prokaryotic expression vector contains a set of genetic elements that affect both transcriptional and translational steps of protein production [62], therefore, optimal configuration of these elements should be pursued. The essential architecture of an expression vector includes a promoter, a ribosome binding site (RBS), a start codon, a coding sequence for the target protein, a transcription terminator, an origin of replication, and additionally, it may contain a selective marker. The incorporation of selection schemes into the vector, like the use of antibiotic resistance

*Corresponding author
Phone: 351-218419065; Fax: 351-218419062;
E-mail: filipem@alfa.ist.utl.pt

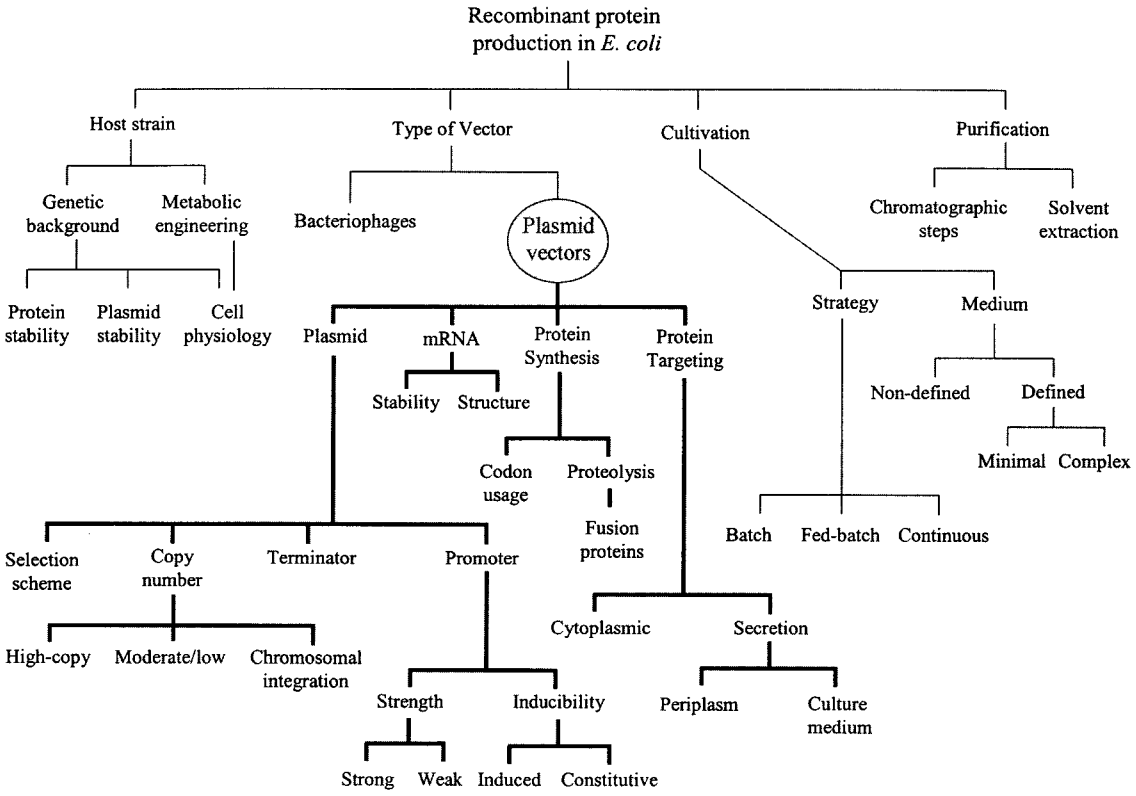

Fig. 1. Factors influencing recombinant protein production in *Escherichia coli.*

genes, has been widespread for many years [154]. The problem associated with antibiotic selection in multicopy plasmids is that the selection agent may be decomposed, deactivated, or tightly bound to the product of the selection gene. Cells with a high content of selection gene reduce the concentration of the antibiotic, thus lowering its effect. Multiple additions of antibiotics during fermentation can solve this problem, however, this solution is not practical in large-scale processes due to its cost and also due to product contamination by antibiotics [6, 37].

Plasmid Copy Number. Multicopy plasmids have been extensively used as vectors for recombinant protein expression. It is known that the amount of a gene product synthesized by a cell can be enhanced by increasing the copy number of the plasmid harboring the gene [65]. Theoretically, the higher the copy number of a plasmid that contains a target gene, the higher will be the gene dosage effect [32]. On the other hand, it is known that the increase in protein synthesis from low to high-copy plasmids is not always proportional to the copy number increase [85, 163]. Although raising the copy number is an effective strategy for increasing gene expression, particularly at low expression levels [109], the plasmid metabolic burden may also contribute to gene expression limitations [24, 37].

There are some situations in which protein expression at very high levels might be deleterious (e.g. when periplasmic secretion is the goal) and, in these cases, either a small number of gene copies cloned in the chromosome [7, 25] or low-copy expression vectors may be sufficient or even desirable [176]. Low-copy plasmids may have a number of advantages over high-copy plasmids such as tight control of gene expression, the ability to replicate large pieces of DNA, and low metabolic burden on the host strains [24].

The origin of replication contained in a vector governs the plasmid copy number. However, it has been reported that some physiological states, like the stringent response [49], or even the cell growth rate [39, 96], may also influence plasmid replication.

Plasmid stability is also a key issue in recombinant protein production and, although naturally occurring *E. coli* plasmids are extremely stable, engineered expression vectors are often lost in the absence of a selective pressure [28, 147, 158, 172], causing significant reductions of production yields [5].

Transcriptional Regulation. Transcriptional regulation is of major importance in recombinant protein production and the choices to be made at this keystone level are crucial to achieve the adequate expression. In the early days of recombinant protein production, strong promoters were used in order to maximize protein expression. However, for production strategies involving protein secretion to the periplasm the expression rate should be

Table 1. Example of promoters that can be used for high-level expression of recombinant proteins in *E. coli*; adapted from [5, 62, 103].

Promoter	Regulator	Induction	Reference
araBAD	*araC*	L-Arabinose	[77, 161]
cst-1	ppGpp	Glucose starvation	[188]
cspA	-	Low temperature	[118, 190, 191]
lac	*lacI*, *lacI^q*	IPTG, allolactose	[48, 104, 131]
lpp	-	Constitutive	[57, 178]
malK	*crp*, *cAMP*	Maltose	[19, 111]
nar	*fnr*	Anaerobiosis	[61, 94]
phoA	*phoB*, *phoR*	Phosphate starvation	[86, 113]
proU	-	Osmolarity	[79, 107]
Protein A	-	Constitutive	[108, 110]
recA	*lexA*	Nalidixic acid	[88, 99]
uspA and *uspB*	*fadR*, IHF	ppGpp	[109, 111, 136]
tac	*lacI*, *lacI^q*	IPTG, allolactose	[95, 145]
trc	*lacI*, *lacI^q*	IPTG, allolactose	[11, 74, 194]
trp	*trpR*	Tryptophan starvation	[26, 197]
T7-*lac* operator	*lacI^q*	IPTG	[12, 100, 153]

Abreviations: IPTG, Isopropyl-β-D-thiogalactopyranoside.

optimized rather than maximized [111, 162]. When post-translational modifications such as the formation of disulfide bonds are necessary for biological activity, the expression rate must be fine-tuned with, for instance, the co-expression of chaperones [77, 129, 153].

Promoters. The promoter key elements are located at -35 and -10 regions [45, 146]. A comparison of promoter sequences has shown a strong homology in these regions among most of *E. coli* promoters, evidencing their role in transcription initiation. Furthermore, it has been demonstrated that the distance between these two regions is also relevant [28].

A large number of promoters can be used for recombinant protein expression [152], and some are listed in Table 1. Several criteria are used to select an appropriate promoter for the expression of a recombinant protein; namely, promoter strength, leakage, inducibility, as well as economical considerations [62, 103]. If inclusion body formation is intended, a strong promoter capable of recombinant protein production in excess of 10-30% of the total cell protein should be used. However, this promoter should display a minimal basal transcriptional activity [55]. A highly repressible promoter is of great importance for minimizing the metabolic burden on the host strain prior to the production phase, particularly if the protein of interest is toxic or detrimental to cell growth. If protein secretion is desired, then the expression rates must be optimized to prevent the saturation of *E. coli* transport machinery [111, 145]. Therefore, in such a case, the promoter strength should be adequate for the export capacity of the target protein in a selected host strain. For low expression rates, constitutive promoters may be a good option, however, if toxic proteins are to be produced, then induced promoters are preferred. In all cases, economical considerations and ease of induction are key factors for promoter choice.

Terminators: Transcriptional terminators determine the points where the complexes formed by mRNA, RNA polymerase, and DNA dissociate, thereby ending transcription. When properly placed downstream of a coding sequence, these elements prevent transcription through another promoter located downstream of the coding sequence [5] and can even inhibit transcription from this second promoter (promoter occlusion).

Efficient termination stabilizes mRNA with positive effects on the expression level. Most expression vectors contain one or several terminators like the T1 and T2, derived from the *rrnB* rRNA operon of *E. coli* [62, 103].

Translational Features. The mRNA 5' untranslated sequence must contain the ribosome binding site (approximately 54 nucleotides between positions -35 (±2) and +19 to +22 of the mRNA coding sequence). Within this region, the Shine-Dalgarno sequence (UAAGGAGG) located 5 to 13 bases upstream of the start codon is essential for the interaction with the 3' end of 16S rRNA during translation initiation. Efficient start codons are AUG, GUG, and UUG, with AUG being the most frequently used start codon in *E. coli* [5]. Bacterial translation is initiated by *N*-formylmethionine, which is deformylated during synthesis but not necessarily removed. Removal is done by an endogenous methionine aminopeptidase but this process is dependent on the side chain of the second amino acid [103]. Since the starting methionine is not always removed, *N*-terminal authenticity is not guaranteed and reduction of biological activity of the expressed protein may occur in some cases [167].

The presence of a stop codon in the mRNA is of great importance in translational termination. Peptide chain liberation is mediated by at least two release factors that recognize the three termination codons. Release factor 1 recognizes UAA and UAG, and release factor 2 recognizes UAA and UGA [122]. Most expression vectors contain all three stop codons in different reading frames to prevent ribosome skipping [154]. The three stop codons differ in their termination efficiencies, and there is a strong bias towards the use of UAA in highly expressed genes. Furthermore, it has been demonstrated that the identity of the nucleotide immediately following the stop codon strongly influences the efficiency of translational termination in *E. coli*. The most efficient terminator sequence for this host organism is UAAU [103].

mRNA Stability. The secondary structure of a single-stranded mRNA is related to its tertiary structure and function. Folding of mRNA molecules is thermodynamically controlled [156, 195] and can influence protein expression in two ways. The formation of stem-loop structures and other localized conformations may influence the half-lives

of certain mRNA molecules, with obvious implications in expression [46]. On the other hand, the secondary structure that a transcript adopts has been shown to play important functional roles in translation of some genes [139, 166], particularly regarding the accessibility of the start codon and Shine-Dalgarno sequence [23].

The degradation of mRNA by host cell RNases is an important factor in post-transcriptional control of recombinant protein expression [163, 187], and it has been reported [24] that mRNA stabilization is one efficient strategy to increase protein expression at all translational levels. Several RNases participate in the degradation process including endonucleases (e.g. RNase E, RNase K, and RNase III) and 3' exonucleases [164]. The stabilization of mRNA can be achieved in three ways:

i) By engineering sequences in the 5' untranslated region, it is possible to obtain changes in the overall secondary structure of the mRNA, thus improving the stability of the transcript [143]. Furthermore, additional stabilizing elements such as *omp*-like leader sequences [41] or RNaseIII cleavage sites can be introduced. It has been demonstrated [132] that processing by RNaseIII can increase the half-life of mRNA 3–4 fold.

ii) By optimization of the 3' untranslated sequences of the mRNA, which may induce the formation of stem-loop structures, thereby blocking the exonucleolytic degradation of the transcript from the 3 terminus [103] or providing accessory protein binding sites that stabilize the mRNA [67].

iii) By using host strains with mutations on the *rnc* or *rnb* genes (encoding for RNase III and RNase II), thus minimizing internal cleavage of mRNA [41] and thereby increasing protein expression.

Codon Usage. Codon usage in *E. coli* is extremely biased as a consequence of a nonrandom usage of synonymous codons [62]. The effect of substitution of rare codons in protein expression has been extensively studied, but definitive conclusions are difficult to draw. Negative effects of the presence of rare codons in the coding sequence have been reported in the expression of several proteins [18, 40, 81, 144, 165, 198]. The explanations found for this negative effect may be either the relatively low abundance of certain transfer RNA species [14, 43, 72, 73] or the different energies of codon-anticodon pairing [54]. The location of rare codons in the transcript [29, 51] and the transcriptional rate [142] have also been reported to influence translation from rare codons. Although codon optimization yielded higher expression levels in a variety of situations [66, 102, 192, 193], it has been proposed that gene expression is usually not limited by rare codons [68], and that the abundance of transfer RNAs is not correlated with codon usage [17]. Some authors have also suggested that the use of rare codons in a gene is a way to naturally slow down the elongation of a peptide chain, thereby allowing the proper folding of specific regions in the nascent peptide [58].

However, for a particular recombinant protein expression system, a codon optimization procedure may be beneficial, not only because the percentage of rare codons is diminished, but also because the optimized sequence may allow the formation of mRNA secondary structures of higher stability.

Protein Targeting

Since *E. coli* is a Gram-negative bacterium, three locations can be chosen for recombinant protein targeting: cytoplasm, periplasm, and culture medium. The protein synthesis occurs in the bacterial cytoplasm where it can accumulate in a soluble form or aggregate in insoluble inclusion bodies [200]. Recombinant proteins can also be secreted to the periplasmic space or to the culture medium, and this ability may be used with several advantages despite the limited capacity of the *E. coli* transport machinery [111, 145]. The targeting to a defined cellular compartment can affect the expression of the gene product in different manners. The most relevant features distinguishing the recombinant protein production in different cellular locations are summarized in Table 2.

Cytoplasmic Production. The formation of high molecular-weight insoluble aggregates, named as inclusion bodies, is often a consequence of high-level protein production in the cytoplasm [4]. Protein properties like the charge average, turn-forming-residue fraction, cysteine and proline fractions, hydrophilicity, and total number of residues, and environmental factors like cultivation temperature, pH, and nutrient supply are known to influence the formation of these aggregates [103, 141].

Inclusion body formation can be desirable in the production of several recombinant proteins like bovine growth hormone or insulin [176]. The main advantages of inclusion body formation include their facile isolation [182], the high protein yield that can be obtained [33], and the simplicity of plasmid constructs. It is often mentioned that expressing proteins in this form is also advantageous,

Table 2. Characterization of recombinant protein production targeted to different cellular locations in *E. coli*.

	Cytoplasm	Periplasm	Medium
Production level	High	Low	Very low
Product stability	High/low[a]	Moderate	High
Biological activity	Inactive	Active	Active
N-terminal authenticity	Unlikely	Possible	Possible
Aggregation	High	Moderate	Low
Host protein contamination	High	Low	Very low
Downstream processing	Complex	Simple	Very simple

[a]Stability is relatively high if inclusion bodies are formed; otherwise rapid degradation occurs due to high cytoplasmic protease level.

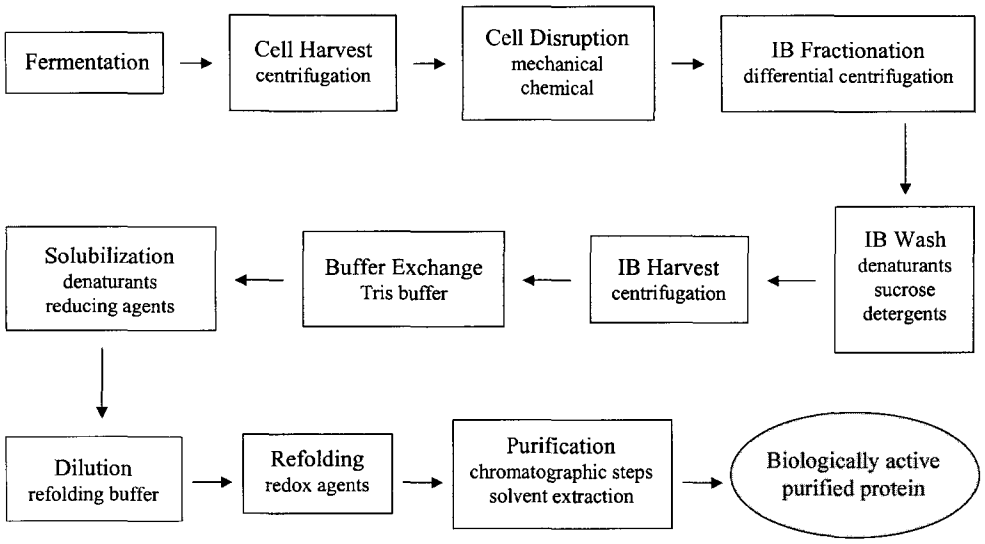

Fig. 2. General strategy for recombinant protein recovery from inclusion bodies.

because the expressed protein is inactive, and therefore, harmless to the host [33]. This is true to some extent, but it is also known that high level production of an heterologous protein is often harmful due to the energy requirements and sequestering of protein synthesis machinery that occurs when a cell is overproducing a protein that it does not need [42, 155]. Another advantage of inclusion body formation is the protective effect against the host proteases [33], although it has been demonstrated that inclusion bodies are not fully protected against protein degradation [197] due to the proteolytic accessibility of solvent-exposed surfaces of inclusion bodies [38].

In most cases, cytoplasmic production of recombinant proteins involves three steps: inclusion body isolation, solubilization of the aggregates, and protein refolding, as described in Fig. 2 [33, 112]. The co-expression of chaperones that are known to be important in the folding pathways [184, 186] increased the production of several proteins [123, 183], however, the success of such approach appears to be protein specific [8]. The co-expression of natural redox agents (e.g. thioredoxin and disulfide isomerases) has also proven to be a good strategy in the expression of correctly folded proteins in the *E. coli* cytoplasm [15].

Protein Secretion. The recovery of a gene product can be greatly simplified, when this product is secreted into the *E. coli* periplasm or to the culture medium (Fig. 3) [75]. Additionally, since secretion often involves the cleavage of a signal sequence [108], the presence of the initial

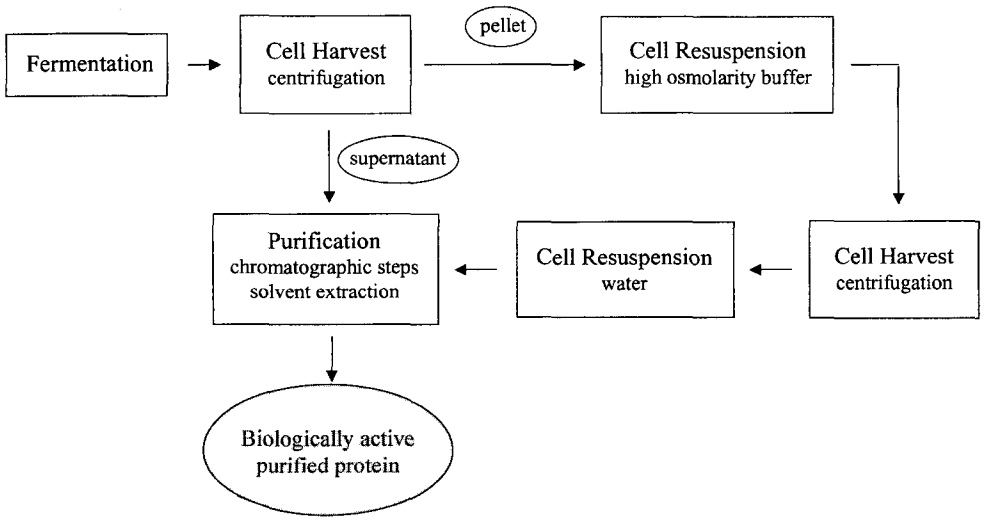

Fig. 3. General strategy for recombinant recovery from periplasm or culture medium.

methionine on a protein that naturally does not contain it can be avoided, thus assuring the *N*-terminal authenticity. Biological activity and stability are dependent on the folding state of the protein, and proper folding is unlikely to occur in the cytoplasm, particularly if disulfide bonds have to be formed, due to its reducing environment [115, 135]. However, protein secretion is a particularly complex process [44, 137], and attempts to secrete recombinant proteins can face several problems; namely, the incomplete translocation across the inner membrane [8], the insufficient capacity of the export machinery [111, 145], and proteolytic degradation [70].

Several factors, including the protein size [89, 131, 151], amino acid composition [3, 21, 80, 173, 174], and the type of leader peptide [121, 124], can affect protein translocation. It has been reported [162] that an optimum translational level exists to achieve high-level secretion of heterologous proteins, otherwise secretion severely drops off. This effect is probably related to the limited secretion capacity of the *E. coli* transport machinery [111, 145]. When this capacity is overwhelmed, the excess of expressed recombinant protein is likely to accumulate in inclusion bodies [71].

Two major mechanisms that are commonly used for recombinant protein secretion in nonpathogenic *E. coli* are known as type I and type II secretion pathways [148]. In Gram-negative bacteria, extracellular secretion involves transportation across two cell membranes. This transport can be done by a single-step mechanism (type I secretion) or by a two-step process (type II secretion) which is mediated by the Sec machinery and can also be used for periplasmic targeting [16, 150]. Extracellular secretion [138] is not always a specific transport mechanism and can occur through periplasmic leakage [63, 108, 159, 160]. Co-secretion of molecular chaperones and medium supplementation with low molecular weight additives have been shown to increase periplasmic secretion and refolding yields in the bacterial periplasm [10, 20, 77, 90, 129, 140].

Protein Stabilization

Even though *E. coli* is an attractive host for recombinant protein production, the accumulation of some proteins has often been proven to be difficult due to degradation by host-specific proteases [13]. Secreted proteins have been found to be extremely susceptible to proteolysis during translocation [70], and product degradation has been considered to be the most serious obstacle to the large-scale production of secreted peptides [106]. The stability of a protein is influenced by its amino acid composition, size [180], folding state [8], cellular location [179], and the presence of proteolytically sensitive sequences [119]. Several strategies have been adopted to circumvent the problem of proteolysis, the most successful being the use of mutant host strains defective in proteases and the expression of the target protein in a fusion form.

Strategies to Improve Protein Stability: Protease-Defective Strains and Fusion Proteins. More than 30 proteases and peptidases have been identified in different cellular compartments of *E. coli*, and some of these are listed in Table 3. One of the strategies to increase the stability of cloned gene products is the use of expression hosts defective in proteases. Protease-deficient hosts in which a single protease (like the *lon*, *clp*, degP) has been inactivated are available [41], and expression strains deficient in multiple proteases have also been reported [9, 106]. The choice of a particular cellular location for recombinant product accumulation dictates the type of mutant strain to be used in each case.

Table 3. Classification of various *E. coli* proteases; adapted from [8, 34, 52, 53, 106, 175].

Name	Description
ATP-dependent*	
Lon	Degrades abnormal proteins
ClpP	Subunit of ClpXP and ClpAP
ClpX	Combines with ClpP to form ClpXP
ClpA	Degrades β-galactosidase fusions
ClpQ	Threonine active site
ClpY	Resembles ClpX, acts with ClpQ
ATP-independent**	
Protease II	Serine active site
OpdA	Degrades signal peptides
PepA	Aminoexopeptidase
PepD	Dipeptidase
PepE	α-Aspartyl dipeptidase
PepM	N-Terminal methionine aminopeptidase
PepN	Aminopeptidase N
PepP	Proline aminopeptidase II
PepQ	Proline dipeptidase
PepT	Nonspecific peptidase
Dcp	Dipeptidyl carboxypeptidase
Ci	Cytoplasmic metalloprotease
Fa	Endoprotease
So	Serine protease
Extracytoplasmic	
LepB	Signal peptidase
LspA	Lipoprotein signal peptidase
SppA	Signal peptidase
DegP	Serine protease
DegQ	Periplasmic serine protease
Protease III	Degrades small peptides (10–30 amino acids)
OmpT	Cuts preferentially at paired basic residues
OmpP	Homologus to OmpT
Iap	N-Terminal Arg-specific aminopeptidase
OrfX	Metalloprotease
Protease VI	Membrane-associated serine protease

*Cytoplasmic location.
**Cytoplasmic or periplasmic location.

Table 4. Applications of genetic fusions in recombinant protein production; adapted from [127].

Application	Reference
Stabilization of cloned gene products	[56, 64, 82, 114, 119, 194]
Facilitated downstream processing	[27, 60, 101, 110, 116, 130]
Solubility of a gene product	[30, 120]
Facilitated *in vivo* folding	[149, 194]
Detection of recombinant proteins	[110, 128]
Bacterial surface display	[36, 87, 170]
Increased therapeutical stability	[168, 169]
Drug targeting	[127]

An alternative and widely used strategy to stabilize an heterologous protein is to express it as a chimera with an homologous protein [13]. Besides protein stabilization, genetic fusions can be highly advantageous in recombinant protein production, facilitating product detection and recovery (Table 4). Recombinant protein expression as a fusion-protein with a peptide tag designed for affinity purification is highly advantageous in terms of downstream processing. The basic principle underlying affinity purification is a specific interaction between the affinity handle and a ligand efficiently immobilized on a gel matrix. This process enables a high purification factor (often greater than 1000) and simultaneous concentration [22]. Several protein-ligand interactions have been used for this purpose [127, 181], and the most common are protein-protein, protein-carbohydrate, protein-metal, and enzyme-substrate interactions. Table 5 lists some commonly used affinity fusion-tag systems and the type of interaction involved in fusion-protein purification. The interaction between the affinity tag and the ligand should be specific and strong enough to enable a one-step recovery procedure, allowing at the same time the use of a mild elution protocol to avoid protein unfolding [1, 22, 126].

For recombinant protein production purposes, the affinity handle should be small, soluble, proteolytically stable and secretion competent, should fold efficiently and independently of the target protein, and should have a structure enabling the specific cleavage of the target protein [127, 128, 130, 168, 169]. This cleavage process can be done by chemical methods that are generally cost-effective and highly scalable. However, the specificity of these agents is usually low [5] and, sometimes, the cleavage procedure requires harsh conditions that may have a deleterious effect on the target product. Engineering the fusion tag to introduce a specific site for enzymatic cleavage is an alternative strategy [78]. Currently used enzymes are enterokinase, subtilisin, factor Xa, thrombin, and tobacco etch virus protease [134, 181]. The enzymatic cleavage procedure allows the specific removal of the fusion partner although it may be more expensive than a chemical method. The purification procedure should then include a first passage through the affinity column to separate the fusion protein from the contaminants, and a second passage after the cleavage reaction, to obtain the mature target protein separated from the handle.

CONCLUSION

The production of heterologous protein molecules in *E. coli* has evolved over the past 30 years, and a strong body of evidence is contributing to a much clearer approach in the optimization of production strategies and expression vector design. However, expression of a particular mRNA in a bacterial host still presents its peculiar set of problems.

Table 5. Commonly used affinity fusion systems and type of interaction on recombinant protein purification; adapted from [22, 127, 181, 189].

Fusion partner	Ligand	Size (kDa)	Type of interaction	Reference
Protein A and derivatives	IgG	7–31		[108, 109, 111]
Protein G	Albumin	5–15	Protein-protein	[169]
S-tag	S-fragment of RNase A	2		[98]
Gluthathione S-transferase	Gluthathione	26	Enzyme-substrate	[157]
c-myc	Anti-c-myc	1		[105]
Flag peptide	Anti-Flag peptide	1–3	Antibody-antigen	[171]
Hemagglutinin	Anti-hemagglutinin	1		[69]
Arginine	Ion exchanger	1	Polyamino acids	[91]
Histidine	IMAC	1		[31,92]
Cellulose-binding domain	Cellulose	3–20	Carbohydrate-protein	[185]
Chitin-binding domain	Chitin	6		[177]
Maltose-binding protein	Amylose	40		[59]
Streptavidin-binding peptide	Streptavidin	4	Other interactions	[83]

Abreviations: IgG, immunoglobulin G; IMAC, immobilized metal affinity chromatography.

Protein expression can be tackled in various ways, but reaching a biologically active product is today, as it was in the past, the major challenge. *E. coli* has traditionally been used as a workhorse for recombinant protein production and, despite successful advances, a comprehensive view of its metabolic capabilities is far from complete. Recombinant protein production in the post-genomic era will bring new challenges in terms of genome manipulation, and it is likely that *E. coli* will continue to play a central role as a recombinant production host.

Acknowledgment

F. J. M. Mergulhão acknowledges the receipt of a post-doctoral fellowship from the Operational Programme for Science Technology and Innovation, Ministério da Ciência e Tecnologia, Portugal.

REFERENCES

1. Abrahmsen, L., T. Moks, B. Nilsson, and M. Uhlen. 1986. Secretion of heterologous gene products to the culture medium of *Escherichia coli*. *Nucleic Acids Res.* **14:** 7487-7500.
2. Andersen, D. C. and L. Krummen. 2002. Recombinant protein expression for therapeutic applications. *Curr. Opin. Biotechnol.* **13:** 117-123.
3. Andersson, H. and G. von Heijne. 1991. A 30-residue-long "export initiation domain" adjacent to the signal sequence is critical for protein translocation across the inner membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **88:** 9751-9754.
4. Baker, K. N., M. H. Rendall, A. Patel, P. Boyd, M. Hoare, R. B. Freedman, and D. C. James. 2002. Rapid monitoring of recombinant protein products: A comparison of current technologies. *Trends Biotechnol.* **20:** 149-156.
5. Balbas, P. 2001. Understanding the art of producing protein and nonprotein molecules in *Escherichia coli*. *Mol. Biotechnol.* **19:** 251-267.
6. Balbas, P. and F. Bolivar. 1990. Design and construction of expression plasmid vectors in *Escherichia coli*. *Methods Enzymol.* **185:** 14-37.
7. Balbas, P. and G. Gosset. 2001. Chromosomal editing in *Escherichia coli*. Vectors for DNA integration and excision. *Mol. Biotechnol.* **19:** 1-12.
8. Baneyx, F. 1999. Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* **10:** 411-421.
9. Baneyx, F. and G. Georgiou. 1991. Construction and characterization of *Escherichia coli* strains deficient in multiple secreted proteases: Protease III degrades high-molecular-weight substrates *in vivo*. *J. Bacteriol.* **173:** 2696-2703.
10. Barth, S., M. Huhn, B. Matthey, A. Klimka, E. A. Galinski, and A. Engert. 2000. Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions. *Appl. Environ. Microbiol.* **66:** 1572-1579.
11. Batisson, I. and M. der Vartanian. 2000. Extracellular DsbA-insensitive folding of *Escherichia coli* heat-stable enterotoxin STa *in vitro*. *J. Biol. Chem.* **275:** 10582-10589.
12. Bayer, M., R. Iberer, K. Bischof, E. Rassi, E. Stabentheiner, G. Zellnig, and G. Koraimann. 2001. Functional and mutational analysis of p19, a DNA transfer protein with muramidase activity. *J. Bacteriol.* **183:** 3176-3183.
13. Belagaje, R. M., S. G. Reams, S. C. Ly, and W. F. Prouty. 1997. Increased production of low molecular weight recombinant proteins in *Escherichia coli*. *Protein Sci.* **6:** 1953-1962.
14. Berg, O. G. and C. G. Kurland. 1997. Growth rate-optimised tRNA abundance and codon usage. *J. Mol. Biol.* **270:** 544-550.
15. Bessette, P. H., F. Aslund, J. Beckwith, and G. Georgiou. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci. USA* **96:** 13703-13708.
16. Binet, R., S. Letoffe, J. M. Ghigo, P. Delepelaire, and C. Wandersman. 1997. Protein secretion by Gram-negative bacterial ABC exporters - a review. *Gene* **192:** 7-11.
17. Bonekamp, F., H. Dalboge, T. Christensen, and K. F. Jensen. 1989. Translation rates of individual codons are not correlated with tRNA abundances or with frequencies of utilization in *Escherichia coli*. *J. Bacteriol.* **171:** 5812-5816.
18. Bonekamp, F. and K. F. Jensen. 1988. The AGG codon is translated slowly in *E. coli* even at very low expression levels. *Nucleic Acids Res.* **16:** 3013-3024.
19. Bostrom, M. and G. Larsson. 2002. Introduction of the carbohydrate-activated promoter P(malK) for recombinant protein production. *Appl. Microbiol. Biotechnol.* **59:** 231-238.
20. Bothmann, H. and A. Pluckthun. 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.
21. Brendel, V., P. Bucher, I. R. Nourbakhsh, B. E. Blaisdell, and S. Karlin. 1992. Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. USA* **89:** 2002-2006.
22. Burgess, R. R. and N. E. Thompson. 2002. Advances in gentle immunoaffinity chromatography. *Curr. Opin. Biotechnol.* **13:** 304-308.
23. Canals, A., M. Ribo, A. Benito, M. Bosch, E. Mombelli, and M. Vilanova. 1999. Production of engineered human pancreatic ribonucleases, solving expression and purification problems, and enhancing thermostability. *Protein Expr. Purif.* **17:** 169-181.
24. Carrier, T., K. L. Jones, and J. D. Keasling. 1998. mRNA stability and plasmid copy number effects on gene expression from an inducible promoter system. *Biotechnol. Bioeng.* **59:** 666-672.
25. Cebolla, A., J. L. Royo, V. De Lorenzo, and E. Santero. 2002. Improvement of recombinant protein yield by a combination of transcriptional amplification and stabilization

of gene expression. *Appl. Environ. Microbiol.* **68:** 5034-5041.
26. Chang, J. R., J. J. Choi, H. K. Kim, and S. T. Kwon. 2001. Purification and properties of *Aquifex aeolicus* DNA polymerase expressed in *Escherichia coli*. *FEMS Microbiol. Lett.* **201:** 73-77.
27. Chang, S. G., D. Y. Kim, K. D. Choi, J. M. Shin, and H. C. Shin. 1998. Human insulin production from a novel mini-proinsulin which has high receptor-binding activity. *Biochem. J.* **329:** 631-635.
28. Chatwin, H. M. and D. K. Summers. 2001. Monomer-dimer control of the ColE1 P(cer) promoter. *Microbiol.* **147:** 3071-3081.
29. Chen, G. F. and M. Inouye. 1990. Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. *Nucleic Acids Res.* **18:** 1465-1473.
30. Chenal, A., P. Nizard, V. Forge, M. Pugniere, M. O. Roy, J. C. Mani, F. Guillain, and D. Gillet. 2002. Does fusion of domains from unrelated proteins affect their folding pathways and the structural changes involved in their function? A case study with the diphtheria toxin T domain. *Protein Eng.* **15:** 383-391.
31. Choe, W. S., R. H. Clemmitt, H. A. Chase, and A. P. Middelberg. 2002. Comparison of histidine-tag capture chemistries for purification following chemical extraction. *J. Chromatogr. A.* **953:** 111-121.
32. Choi, J., K. Ra, and Y. Lee. 1999. Enhancement of bovine growth hormone gene expression by increasing the plasmid copy number. *Biotechnol. Lett.* **21:** 1-5.
33. Clark, E. D. 2001. Protein refolding for industrial processes. *Curr. Opin. Biotechnol.* **12:** 202-207.
34. Cooper, K. W. and F. Baneyx. 2001. *Escherichia coli* FtsH (HflB) degrades a membrane-associated TolAI-II-beta-lactamase fusion protein under highly denaturing conditions. *Protein Expr. Purif.* **21:** 323-332.
35. Cornelis, P. 2000. Expressing genes in different *Escherichia coli* compartments. *Curr. Opin. Biotechnol.* **11:** 450-454.
36. Cornelis, P., J. C. Sierra, A. Lim, Jr., A. Malur, S. Tungpradabkul, H. Tazka, A. Leitao, C. V. Martins, C. di Perna, L. Brys, P. De Baetseller, and R. Hamers. 1996. Development of new cloning vectors for the production of immunogenic outer membrane fusion proteins in *Escherichia coli*. *Biotechnol. (N Y).* **14:** 203-208.
37. Cranenburgh, R. M., J. A. Hanak, S. G. Williams, and D. J. Sherratt. 2001. *Escherichia coli* strains that allow antibiotic-free plasmid selection and maintenance by repressor titration. *Nucleic Acids Res.* **29:** E26.
38. Cubarsi, R., M. M. Carrio, and A. Villaverde. 2001. *In situ* proteolytic digestion of inclusion body polypeptides occurs as a cascade process. *Biochem. Biophys. Res. Commun.* **282:** 436-441.
39. del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62:** 434-464.
40. del Tito, B. J., Jr., J. M. Ward, J. Hodgson, C. J. Gershater, H. Edwards, L. A. Wysocki, F. A. Watson, G. Sathe, and J. F. Kane. 1995. Effects of a minor isoleucyl tRNA on heterologous protein translation in *Escherichia coli*. *J. Bacteriol.* **177:** 7086-7091.
41. Dias, A. 1990. Overproduction of proteins in *Escherichia coli*: Vectors, hosts and strategies. *Methods Enzymol.* **182:** 93-112.
42. Dong, H., L. Nilsson, and C. G. Kurland. 1995. Gratuitous overexpression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J. Bacteriol.* **177:** 1497-1504.
43. Dong, H., L. Nilsson, and C. G. Kurland. 1996. Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J. Mol. Biol.* **260:** 649-663.
44. Economou, A. 1999. Following the leader: Bacterial protein export through the Sec pathway. *Trends Microbiol.* **7:** 315-320.
45. Estrem, S. T., T. Gaal, W. Ross, and R. L. Gourse. 1998. Identification of an UP element consensus sequence for bacterial promoters. *Proc. Natl. Acad. Sci. USA* **95:** 9761-9766.
46. Fargo, D. C., J. E. Boynton, and N. W. Gillham. 1999. Mutations altering the predicted secondary structure of a chloroplast 5' untranslated region affect its physical and biochemical properties as well as its ability to promote translation of reporter mRNAs both in the *Chlamydomonas reinhardtii* chloroplast and in *Escherichia coli*. *Mol. Cell. Biol.* **19:** 6980-6990.
47. Fernandez, L. A. and V. de Lorenzo. 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway. *Mol. Microbiol.* **40:** 332-346.
48. Fernandez, L. A., I. Sola, L. Enjuanes, and V. de Lorenzo. 2000. Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system. *Appl. Environ. Microbiol.* **66:** 5024-5029.
49. Gautam, A. and D. Bastia. 2001. A replication terminus located at or near a replication checkpoint of *Bacillus subtilis* functions independently of stringent control. *J. Biol. Chem.* **276:** 8771-8777.
50. Gavit, P. and M. Better. 2000. Production of antifungal recombinant peptides in *Escherichia coli*. *J. Biotechnol.* **79:** 127-136.
51. Goldman, E., A. H. Rosenberg, G. Zubay, and F. W. Studier. 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.
52. Gottesman, M. E. and W. A. Hendrickson. 2000. Protein folding and unfolding by *Escherichia coli* chaperones and chaperonins. *Curr. Opin. Microbiol.* **3:** 197-202.
53. Gottesman, S. 1996. Proteases and their targets in *Escherichia coli*. *Annu. Rev. Genet.* **30:** 465-506.
54. Grosjean, H. and W. Fiers. 1982. Preferential codon usage in prokaryotic genes: The optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene* **18:** 199-209.

55. Grossman, T. H., E. S. Kawasaki, S. R. Punreddy, and M. S. Osburne. 1998. Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene* **209:** 95-103.
56. Guarente, L., G. Lauer, T. M. Roberts, and M. Ptashne. 1980. Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit beta-globin. *Cell* **20:** 543-553.
57. Guisez, Y., I. Fache, L. A. Campfield, F. J. Smith, A. Farid, G. Plaetinck, J. Van der Heyden, J. Tavernier, W. Fiers, P. Burn, and R. Devos. 1998. Efficient secretion of biologically active recombinant OB protein (leptin) in *Escherichia coli*, purification from the periplasm and characterization. *Protein Expr. Purif.* **12:** 249-258.
58. Guisez, Y., J. Robbens, E. Remaut, and W. Fiers. 1993. Folding of the MS2 coat protein in *Escherichia coli* is modulated by translational pauses resulting from mRNA secondary structure and codon usage: a hypothesis. *J. Theor. Biol.* **162:** 243-252.
59. Hamilton, S. R., J. B. O'Donnell, Jr., A. Hammet, D. Stapleton, S. A. Habinowski, A. R. Means, B. E. Kemp, and L. A. Witters. 2002. AMP-activated protein kinase kinase: Detection with recombinant AMPK alpha1 subunit. *Biochem. Biophys. Res. Commun.* **293:** 892-898.
60. Hammarberg, B., P. A. Nygren, E. Holmgren, A. Elmblad, M. Tally, U. Hellman, T. Moks, and M. Uhlen. 1989. Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II. *Proc. Natl. Acad. Sci. USA* **86:** 4367-4371.
61. Han, S. J., H. N. Chang, and J. Lee. 2001. Characterization of an oxygen-dependent inducible promoter, the nar promoter of *Escherichia coli*, to utilize in metabolic engineering. *Biotechnol. Bioeng.* **72:** 573-576.
62. Hannig, G. and S. C. Makrides. 1998. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol.* **16:** 54-60.
63. Hasenwinkle, D., E. Jervis, O. Kops, C. Liu, G. Lesnicki, C. Haynes, and D. Kilburn. 1997. Very high-level production and export in *Escherichia coli* of a cellulose binding domain for use in a generic secretion-affinity fusion system. *Biotechnol. Bioeng.* **55:** 854-863.
64. Hellebust, H., M. Murby, L. Abrahmsen, M. Uhlen, and S. Enfors. 1989. Different approaches to stabilize a recombinant fusion protein. *Bio/Technology* **7:** 165-168.
65. Herman-Antosiewicz, A., M. Obuchowski, and G. Wegrzyn. 2001. A plasmid cloning vector with precisely regulatable copy number in *Escherichia coli*. *Mol. Biotechnol.* **17:** 193-199.
66. Hernan, R. A., H. L. Hui, M. E. Andracki, R. W. Noble, S. G. Sligar, J. A. Walder, and R. Y. Walder. 1992. Human hemoglobin expression in *Escherichia coli*: Importance of optimal codon usage. *Biochem.* **31:** 8619-8628.
67. Hew, Y., C. Lau, Z. Grzelczak, and F. W. Keeley. 2000. Identification of a GA-rich sequence as a protein-binding site in the 3'-untranslated region of chicken elastin mRNA with a potential role in the developmental regulation of elastin mRNA stability. *J. Biol. Chem.* **275:** 24857-24864.
68. Holm, L. 1986. Codon usage and gene expression. *Nucleic Acids Res.* **14:** 3075-3087.
69. Honey, S., B. L. Schneider, D. M. Schieltz, J. R. Yates, and B. Futcher. 2001. A novel multiple affinity purification tag and its use in identification of proteins associated with a cyclin-CDK complex. *Nucleic Acids Res.* **29:** E24.
70. Huang, H. C., M. Y. Sherman, O. Kandror, and A. L. Goldberg. 2001. The molecular chaperone DnaJ is required for the degradation of a soluble abnormal protein in *Escherichia coli*. *J. Biol. Chem.* **276:** 3920-3928.
71. Ignatova, Z., A. Mahsunah, M. Georgieva, and V. Kasche. 2003. Improvement of posttranslational bottlenecks in the production of penicillin amidase in recombinant *Escherichia coli* strains. *Appl. Environ. Microbiol.* **69:** 1237-1245.
72. 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.
73. Ikemura, T. 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *J. Mol. Biol.* **151:** 389-409.
74. Jeong, K. J. and S. Y. Lee. 2001. Secretory production of human granulocyte colony-stimulating factor in *Escherichia coli*. *Protein Expr. Purif.* **23:** 311-318.
75. Jeong, K. J. and S. Y. Lee. 2002. Excretion of human beta-endorphin into culture medium by using outer membrane protein F as a fusion partner in recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* **68:** 4979-4985.
76. Jeong, K. J. and S. Y. Lee. 2003. Enhanced production of recombinant proteins in *Escherichia coli* by filamentation suppression. *Appl. Environ. Microbiol.* **69:** 1295-1298.
77. Joly, J. C., W. S. Leung, and J. R. Swartz. 1998. Overexpression of *Escherichia coli* oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proc. Natl. Acad. Sci. USA* **95:** 2773-2777.
78. Jonasson, P., J. Nilsson, E. Samuelsson, T. Moks, S. Stahl, and M. Uhlen. 1996. Single-step trypsin cleavage of a fusion protein to obtain human insulin and its C peptide. *Eur. J. Biochem.* **236:** 656-661.
79. Jordi, B. J., T. A. Owen-Hughes, C. S. Hulton, and C. F. Higgins. 1995. DNA twist, flexibility and transcription of the osmoregulated proU promoter of *Salmonella typhimurium*. *EMBO J.* **14:** 5690-5700.
80. Kajava, A. V., S. N. Zolov, A. E. Kalinin, and M. A. Nesmeyanova. 2000. The net charge of the first 18 residues of the mature sequence affects protein translocation across the cytoplasmic membrane of gram-negative bacteria. *J. Bacteriol.* **182:** 2163-2169.
81. Kane, J. F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **6:** 494-500.
82. Kang, Y. and J. W. Yoon. 1994. Effect of modification of connecting peptide of proinsulin on its export. *J. Biotechnol.* **36:** 45-54.
83. Keefe, A. D., D. S. Wilson, B. Seelig, and J. W. Szostak. 2001. One-step purification of recombinant proteins using a

nanomolar-affinity streptavidin-binding peptide, the SBP-Tag. *Protein Expr. Purif.* **23:** 440-446.
84. Kelley, B. D. 2001. Biochemical engineering: Bioprocessing of therapeutic proteins. *Curr. Opin. Biotechnol.* **12:** 173-174.
85. Kim, B. G. 1990. Analysis of pBR322 replication kinetics and its dependency on growth rate. *Biotechnol. Bioeng.* **36:** 233-242.
86. Kim, J., J. Luirink, and D. A. Kendall. 2000. SecB dependence of an exported protein is a continuum influenced by the characteristics of the signal peptide or early mature region. *J. Bacteriol.* **182:** 4108-4112.
87. Kim, Y. S., H. C. Jung, and J. G. Pan. 2000. Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants. *Appl. Environ. Microbiol.* **66:** 788-793.
88. Kitagawa, J., K. Yamamoto, and H. Iba. 2001. Computational analysis of sos response in ultraviolet-irradiated *Escherichia coli*. *Genome Inf.* **12:** 280-281.
89. Koster, M., W. Bitter, and J. Tommassen. 2000. Protein secretion mechanisms in Gram-negative bacteria. *Int. J. Med. Microbiol.* **290:** 325-331.
90. Kurokawa, Y., H. Yanagi, and T. Yura. 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.
91. Kweon, D. H., D. H. Lee, N. S. Han, C. S. Rha, and J. H. Seo. 2002. Characterization of polycationic amino acids fusion systems for ion-exchange purification of cyclodextrin glycosyltransferase from recombinant *Escherichia coli*. *Biotechnol. Prog.* **18:** 303-308.
92. Lai, W. B. and A. P. Middelberg. 2002. The production of human papillomavirus type 16 L1 vaccine product from *Escherichia coli* inclusion bodies. *Bioprocess Biosyst. Eng.* **25:** 121-128.
93. Lederberg, J. 1998. Plasmid (1952-1997). *Plasmid* **39:** 1-9.
94. Lee, J., M. Cho, E. Hong, K. Kim, and J. Lee. 1996. Characterization of the *nar* promoter to use as an inducible promoter. *Biotechnol. Lett.* **18:** 129-134.
95. Lee, S., I. Kim, D. Kim, K. Bae, and S. Byun. 1998. High level secretion of recombinant staphylokinase into periplasm of *Escherichia coli*. *Biotechnol. Lett.* **20:** 113-116.
96. Lee, S. B. and J. E. Bailey. 2000. Analysis of growth rate effects on productivity of recombinant *Escherichia coli* populations using molecular mechanism models. *Biotechnol. Bioeng.* **26:** 66-73.
97. Lee, S. Y. 1996. High cell-density culture of *Escherichia coli*. *Trends Biotechnol.* **14:** 98-105.
98. Lellouch, A. C. and R. A. Geremia. 1999. Expression and study of recombinant ExoM, a beta1-4 glucosyltransferase involved in succinoglycan biosynthesis in *Sinorhizobium meliloti*. *J. Bacteriol.* **181:** 1141-1148.
99. Levine, A. D., S. H. Rangwala, N. A. Horn, M. A. Peel, B. K. Matthews, R. M. Leimgruber, J. A. Manning, B. F. Bishop, and P. O. Olins. 1995. High level expression and refolding of mouse interleukin 4 synthesized in *Escherichia coli*. *J. Biol. Chem.* **270:** 7445-7452.
100. Loo, T., M. L. Patchett, G. E. Norris, and J. S. Lott. 2002. Using secretion to solve a solubility problem: High-yield expression in *Escherichia coli* and purification of the bacterial glycoamidase PNGase F. *Protein Expr. Purif.* **24:** 90-98.
101. Mackin, R. B. 1999. Streamlined procedure for the production of normal and altered versions of recombinant human proinsulin. *Protein Expr. Purif.* **15:** 308-313.
102. Makoff, A. J., M. D. Oxer, M. A. Romanos, N. F. Fairweather, and S. Ballantine. 1989. Expression of tetanus toxin fragment C in *E. coli*: high level expression by removing rare codons. *Nucleic Acids Res.* **17:** 10191-10202.
103. Makrides, S. C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **60:** 512-538.
104. Manosroi, J., C. Tayapiwatana, F. Gotz, R. G. Werner, and A. Manosroi. 2001. Secretion of active recombinant human tissue plasminogen activator derivatives in *Escherichia coli*. *Appl. Environ. Microbiol.* **67:** 2657-2664.
105. McKern, N. M., M. Lou, M. J. Frenkel, A. Verkuylen, J. D. Bentley, G. O. Lovrecz, N. Ivancic, T. C. Elleman, T. P. Garrett, L. J. Cosgrove, and C. W. Ward. 1997. Crystallization of the first three domains of the human insulin-like growth factor-1 receptor. *Protein Sci.* **6:** 2663-2666.
106. Meerman, H. J. and G. Georgiou. 1994. Construction and characterization of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnol. (N Y).* **12:** 1107-1110.
107. Mellies, J., R. Brems, and M. Villarejo. 1994. The *Escherichia coli* proU promoter element and its contribution to osmotically signaled transcription activation. *J. Bacteriol.* **176:** 3638-3645.
108. Mergulhão, F., G. Monteiro, A. Kelly, M. Taipa, and J. Cabral. 2000. Recombinant human proinsulin: A new approach in gene assembly and protein expression. *J. Microbiol. Biotechnol.* **10:** 690-693.
109. Mergulhão, F., G. Monteiro, G. Larsson, A. Sandem, A. Farewell, T. Nystrom, J. Cabral, and M. Taipa. 2003. Medium and copy number effects on the secretion of human proinsulin in *Escherichia coli* using the universal stress promoters *uspA* and *uspB*. *Appl. Microbiol. Biotechnol.* **61:** 495-501.
110. Mergulhão, F. J., G. A. Monteiro, J. M. Cabral, and M. A. Taipa. 2001. A quantitative ELISA for monitoring the secretion of ZZ-fusion proteins using SpA domain as immunodetection reporter system. *Mol. Biotechnol.* **19:** 239-244.
111. Mergulhão, F. J. M., G. A. Monteiro, G. Larsson, M. Bostrom, A. Farewell, T. Nystrom, J. M. S. Cabral, and M. A. Taipa. 2003. Evaluation of inducible promoters on the secretion of a ZZ-Proinsulin fusion protein. *Biotechnol. Appl. Biochem.* **38:** 87-93.
112. Middelberg, A. 2002. Preparative protein refolding. *Trends Biotechnol.* **20:** 437.
113. Mikhaleva, N. I., V. V. Golovastov, S. N. Zolov, M. V. Bogdanov, W. Dowhan, and M. A. Nesmeyanova. 2001.

Depletion of phosphatidylethanolamine affects secretion of *Escherichia coli* alkaline phosphatase and its transcriptional expression. *FEBS Lett.* **493:** 85-90.
114. Misoka, F., T. Miyake, K. Miyoshi, M. Sugiyama, S. Sakamoto, and T. Fuwa. 1989. Overproduction of human insulin-like growth factor-II in *Escherichia coli*. *Biotechnol. Lett.* **11:** 839-844.
115. Missiakas, D. and S. Raina. 1997. Protein misfolding in the cell envelope of *Escherichia coli*: New signaling pathways. *Trends Biochem. Sci.* **22:** 59-63.
116. Moks, T., L. Abrahmsen, E. Holmgren, M. Bilich, A. Olsson, M. Uhlen, G. Pohl, C. Sterky, H. Hultberg, S. Josephson, *et al.* 1987. Expression of human insulin-like growth factor I in bacteria: use of optimized gene fusion vectors to facilitate protein purification. *Biochem.* **26:** 5239-5244.
117. Monzavi-Karbassi, B., G. Cunto-Amesty, P. Luo, and T. Kieber-Emmons. 2002. Peptide mimotopes as surrogate antigens of carbohydrates in vaccine discovery. *Trends Biotechnol.* **20:** 207-214.
118. Mujacic, M., K. W. Cooper, and F. Baneyx. 1999. Cold-inducible cloning vectors for low-temperature protein expression in *Escherichia coli*: Application to the production of a toxic and proteolytically sensitive fusion protein. *Gene* **238:** 325-332.
119. Murby, M., P. A. Nygren, H. Rondahl, U. Hellman, S. O. Enfors, and M. Uhlen. 1991. Differential degradation of a recombinant albumin-binding receptor in *Escherichia coli*. *Eur. J. Biochem.* **199:** 41-46.
120. Murby, M., E. Samuelsson, T. N. Nguyen, L. Mignard, U. Power, H. Binz, M. Uhlen, and S. Stahl. 1995. Hydrophobicity engineering to increase solubility and stability of a recombinant protein from respiratory syncytial virus. *Eur. J. Biochem.* **230:** 38-44.
121. Nakai, K. and M. Kanehisa. 1991. Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins* **11:** 95-110.
122. Nakamura, Y. and K. Ito. 2002. A tripeptide discriminator for stop codon recognition. *FEBS Lett.* **514:** 30-33.
123. Netzer, W. J. and F. U. Hartl. 1998. Protein folding in the cytosol: Chaperonin-dependent and -independent mechanisms. *Trends Biochem. Sci.* **23:** 68-73.
124. Nielsen, H., J. Engelbrecht, S. Brunak, and G. von Heijne. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10:** 1-6.
125. Nielsen, P. E. 1999. Applications of peptide nucleic acids. *Curr. Opin. Biotechnol.* **10:** 71-75.
126. Nilsson, B. and L. Abrahmsen. 1990. Fusions to staphylococcal protein A. *Methods Enzymol.* **185:** 144-161.
127. Nilsson, B., G. Forsberg, T. Moks, M. Hartmanis, and M. Uhlen. 1992. Fusion proteins in biotechnology. *Curr. Opin. Biotechnol.* **3:** 363-369.
128. Nilsson, J., S. Stahl, J. Lundeberg, M. Uhlen, and P. A. Nygren. 1997. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expr. Purif.* **11:** 1-16.
129. Nishihara, K., M. Kanemori, M. Kitagawa, H. Yanagi, and T. Yura. 1998. Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* **64:** 1694-1699.
130. Nygren, P. A., S. Stahl, and M. Uhlen. 1994. Engineering proteins to facilitate bioprocessing. *Trends Biotechnol.* **12:** 184-188.
131. Palacios, J. L., I. Zaror, P. Martinez, F. Uribe, P. Opazo, T. Socias, M. Gidekel, and A. Venegas. 2001. Subset of hybrid eukaryotic proteins is exported by the type I secretion system of *Erwinia chrysanthemi*. *J. Bacteriol.* **183:** 1346-1358.
132. Panayotatos, N. and K. Truong. 1985. Cleavage within an RNase III site can control mRNA stability and protein synthesis *in vivo*. *Nucleic Acids Res.* **13:** 2227-2240.
133. Pines, O. and M. Inouye. 1999. Expression and secretion of proteins in *E. coli*. *Mol. Biotechnol.* **12:** 25-34.
134. Polyak, S. W., G. Forsberg, B. E. Forbes, K. A. McNeil, S. E. Aplin, and J. C. Wallace. 1997. Introduction of spacer peptides N-terminal to a cleavage recognition motif in recombinant fusion proteins can improve site-specific cleavage. *Protein Eng.* **10:** 615-619.
135. Prinz, W. A., F. Aslund, A. Holmgren, and J. Beckwith. 1997. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* **272:** 15661-15667.
136. Prytz, I., A. M. Sandén, T. Nystrom, A. Farewell, A. Wahlstrom, C. Forberg, I. Tubulekas, Z. Pragai, M. Barer, C. Harwood, and G. Larsson. 2002. Fed-batch production of recombinant beta-galactosidase using the universal stress promoters *uspA* and *uspB* in high cell density cultivation. *Biotechnol. Bioeng.* **83:** 595-603.
137. Pugsley, A. P. 1993. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **57:** 50-108.
138. Pugsley, A. P., O. Francetic, O. M. Possot, N. Sauvonnet, and K. Hardie. 1997. Recent progress and future directions in studies of the main terminal branch of the general secretory pathway in Gram-neagtive bacteria - a review. *Gene* **192:** 13-19.
139. Puri, N., K. B. Appa Rao, S. Menon, A. K. Panda, G. Tiwari, L. C. Garg, and S. M. Totey. 1999. Effect of the codon following the ATG start site on the expression of ovine growth hormone in *Escherichia coli*. *Protein Expr. Purif.* **17:** 215-223.
140. Qiu, J., J. R. Swartz, and G. Georgiou. 1998. Expression of active human tissue-type plasminogen activator in *Escherichia coli*. *Appl. Environ. Microbiol.* **64:** 4891-4896.
141. Rinas, U. and J. Bailey. 1992. Protein compositional analysis of inclusion bodies produced in recombinant *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **37:** 609-614.
142. Robinson, M., R. Lilley, S. Little, J. S. Emtage, G. Yarranton, P. Stephens, A. Millican, M. Eaton, and G. Humphreys. 1984. Codon usage can affect efficiency of

translation of genes in *Escherichia coli*. *Nucleic Acids Res.* **12:** 6663-6671.

143. Rosenbaum, V., T. Klahn, U. Lundberg, E. Holmgren, A. von Gabain, and D. Riesner. 1993. Co-existing structures of an mRNA stability determinant. The 5' region of the *Escherichia coli* and *Serratia marcescens* ompA mRNA. *J. Mol. Biol.* **229:** 656-670.
144. Rosenberg, A. H., E. Goldman, J. J. Dunn, F. W. Studier, and G. Zubay. 1993. Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. *J. Bacteriol.* **175:** 716-722.
145. Rosenberg, H. F. 1998. Isolation of recombinant secretory proteins by limited induction and quantitative harvest. *Biotechniques* **24:** 188-191.
146. Ross, W., S. E. Aiyar, J. Salomon, and R. L. Gourse. 1998. *Escherichia coli* promoters with UP elements of different strengths: Modular structure of bacterial promoters. *J. Bacteriol.* **180:** 5375-5383.
147. Rowe, D. C. and D. K. Summers. 1999. The quiescent-cell expression system for protein synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.* **65:** 2710-2715.
148. Salmond, G. P. and P. J. Reeves. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. *Trends Biochem. Sci.* **18:** 7-12.
149. Samuelsson, E., P. Jonasson, F. Viklund, B. Nilsson, and M. Uhlen. 1996. Affinity-assisted *in vivo* folding of a secreted human peptide hormone in *Escherichia coli*. *Nat. Biotechnol.* **14:** 751-755.
150. Sandkvist, M. 2001. Biology of type II secretion. *Mol. Microbiol.* **40:** 271-283.
151. Sauvonnet, N., I. Poquet, and A. P. Pugsley. 1995. Extracellular secretion of pullulanase is unaffected by minor sequence changes but is usually prevented by adding reporter proteins to its N- or C-terminal end. *J. Bacteriol.* **177:** 5238-5246.
152. Sawers, G. and M. Jarsch. 1996. Alternative regulation principles for the production of recombinant proteins in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **46:** 1-9.
153. Schaffner, J., J. Winter, R. Rudolph, and E. Schwarz. 2001. Cosecretion of chaperones and low-molecular-size medium additives increases the yield of recombinant disulfide-bridged proteins. *Appl. Environ. Microbiol.* **67:** 3994-4000.
154. Schenborn, E. and D. Groskreutz. 1999. Reporter gene vectors and assays. *Mol. Biotechnol.* **13:** 29-44.
155. Schmidt, M., E. Viaplana, F. Hoffmann, S. Marten, A. Villaverde, and U. Rinas. 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.
156. Seffens, W. and D. Digby. 1999. mRNAs have greater negative folding free energies than shuffled or codon choice randomized sequences. *Nucleic Acids Res.* **27:** 1578-1584.
157. Serra, E. C., N. Carrillo, A. R. Krapp, and E. A. Ceccarelli. 1993. One-step purification of plant ferredoxin-NADP+ oxidoreductase expressed in *Escherichia coli* as fusion with glutathione S-transferase. *Protein Expr. Purif.* **4:** 539-546.
158. Sharpe, M. E., H. M. Chatwin, C. Macpherson, H. L. Withers, and D. K. Summers. 1999. Analysis of the ColE1 stability determinant Rcd. *Microbiol.* **145:** 2135-2144.
159. Shokri, A., A. M. Sanden, and G. Larsson. 2002. Growth rate-dependent changes in *Escherichia coli* membrane structure and protein leakage. *Appl. Microbiol. Biotechnol.* **58:** 386-392.
160. Shokri, A., A. M. Sandén, and G. Larsson. 2003. Cell and process design for targeting of recombinant protein into the culture medium of *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **60:** 654-664.
161. Siegele, D. A. and J. C. Hu. 1997. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. USA* **94:** 8168-8172.
162. Simmons, L. C. and D. G. Yansura. 1996. Translational level is a critical factor for the secretion of heterologous proteins in *Escherichia coli*. *Nat. Biotechnol.* **14:** 629-634.
163. Smolke, C. D. and J. D. Keasling. 2002. Effect of copy number and mRNA processing and stabilization on transcript and protein levels from an engineered dual-gene operon. *Biotechnol. Bioeng.* **78:** 412-424.
164. Sozhamannan, S. and B. L. Stitt. 1997. Effects on mRNA degradation by *Escherichia coli* transcription termination factor Rho and pBR322 copy number control protein Rop. *J. Mol. Biol.* **268:** 689-703.
165. Spanjaard, R. A. and J. van Duin. 1988. Translation of the sequence AGG-AGG yields 50% ribosomal frameshift. *Proc. Natl. Acad. Sci. USA* **85:** 7967-7971.
166. Stacey, S. N., D. Jordan, A. J. Williamson, M. Brown, J. H. Coote, and J. R. Arrand. 2000. Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA. *J. Virol.* **74:** 7284-7297.
167. Stader, J. A. and T. J. Silhavy. 1990. Engineering *Escherichia coli* to secrete heterologous gene products. *Methods Enzymol.* **185:** 166-187.
168. Stahl, S., J. Nilsson, S. Hober, M. Uhlen, and P. Nygren. 1999. Affinity fusions, gene expression, pp. 49-63. *In* M. Flickinger and Drew S. (eds.), *Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis and Bioseparation*. John Wiley & Sons, New York, U.S.A.
169. Stahl, S. and P. A. Nygren. 1997. The use of gene fusions to protein A and protein G in immunology and biotechnology. *Pathol. Biol. (Paris)* **45:** 66-76.
170. Stahl, S. and M. Uhlen. 1997. Bacterial surface display: Trends and progress. *Trends Biotechnol.* **15:** 185-192.
171. Su, X., A. K. Prestwood, and R. A. McGraw. 1992. Production of recombinant porcine tumor necrosis factor alpha in a novel *E. coli* expression system. *Biotechniques* **13:** 756-762.
172. Summers, D. 1998. Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Mol. Microbiol.* **29:** 1137-1145.
173. Summers, R. G., C. R. Harris, and J. R. Knowles. 1989. A conservative amino acid substitution, arginine for lysine, abolishes export of a hybrid protein in *Escherichia coli*.

Implications for the mechanism of protein secretion. *J. Biol. Chem.* **264:** 20082-20088.
174. Summers, R. G. and J. R. Knowles. 1989. Illicit secretion of a cytoplasmic protein into the periplasm of *Escherichia coli* requires a signal peptide plus a portion of the cognate secreted protein. Demarcation of the critical region of the mature protein. *J. Biol. Chem.* **264:** 20074-20081.
175. Swamy, K. H. and A. L. Goldberg. 1982. Subcellular distribution of various proteases in *Escherichia coli*. *J. Bacteriol.* **149:** 1027-1033.
176. Swartz, J. R. 2001. Advances in *Escherichia coli* production of therapeutic proteins. *Curr. Opin. Biotechnol.* **12:** 195-201.
177. Szweda, P., R. Pladzyk, R. Kotlowski, and J. Kur. 2001. Cloning, expression, and purification of the *Staphylococcus simulans* lysostaphin using the intein-chitin-binding domain (CBD) system. *Protein Expr. Purif.* **22:** 467-471.
178. Taguchi, S., K. I. Nishihama, K. Igi, K. Ito, H. Taira, M. Motoki, and H. Momose. 2000. Substrate specificity analysis of microbial transglutaminase using proteinaceous protease inhibitors as natural model substrates. *J. Biochem. (Tokyo)* **128:** 415-425.
179. Talmadge, K. and W. Gilbert. 1982. Cellular location affects protein stability in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **79:** 1830-1833.
180. Tang, J. and M. Hu. 1993. Production of human proinsulin in *E. coli* in a non-fusion form. *Biotechnol. Lett.* **15:** 661-666.
181. Terpe, K. 2003. Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60:** 523-533.
182. Thomas, J. G. and F. Baneyx. 1996. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins. *J. Biol. Chem.* **271:** 11141-11147.
183. Thomas, J. G. and F. Baneyx. 1997. Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*. *Protein Expr. Purif.* **11:** 289-296.
184. Thomas, J. G. and F. Baneyx. 2000. ClpB and HtpG facilitate de novo protein folding in stressed *Escherichia coli* cells. *Mol. Microbiol.* **36:** 1360-1370.
185. Tomme, P., A. Boraston, B. McLean, J. Kormos, A. L. Creagh, K. Sturch, N. R. Gilkes, C. A. Haynes, R. A. Warren, and D. G. Kilburn. 1998. Characterization and affinity applications of cellulose-binding domains. *J. Chromatogr. B. Biomed. Sci. Appl.* **715:** 283-296.
186. Tomoyasu, T., A. Mogk, H. Langen, P. Goloubinoff, and B. Bukau. 2001. Genetic dissection of the roles of chaperones and proteases in protein folding and degradation in the *Escherichia coli* cytosol. *Mol. Microbiol.* **40:** 397-413.
187. Tsuji, A., H. Koshimoto, Y. Sato, M. Hirano, Y. Sei-Iida, S. Kondo, and K. Ishibashi. 2000. Direct observation of specific messenger RNA in a single living cell under a fluorescence microscope. *Biophys. J.* **78:** 3260-3274.
188. Tunner, J. R. and C. R. Robertson. 1992. Use of glucose starvation to limit growth and induce protein production in *Escherichia coli*. *Biotechnol. Bioeng.* **40:** 271-279.
189. Uhlen, M., T. Moks, and L. Abrahmsen. 1988. Protein engineering to optimize recombinant protein purification. *Biochem. Soc. Trans.* **16:** 111-112.
190. Vasina, J. A. and F. Baneyx. 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.
191. Vasina, J. A., M. S. Peterson, and F. Baneyx. 1998. Scale-up and optimization of the low-temperature inducible cspA promoter system. *Biotechnol. Prog.* **14:** 714-721.
192. Wang, H., D. J. O'Mahony, D. J. McConnell, and S. Z. Qi. 1993. Optimization of the synthesis of porcine somatotropin in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **39:** 324-328.
193. Williams, D. P., D. Regier, D. Akiyoshi, F. Genbauffe, and J. R. Murphy. 1988. Design, synthesis and expression of a human interleukin-2 gene incorporating the codon usage bias found in highly expressed *Escherichia coli* genes. *Nucleic Acids Res.* **16:** 10453-10467.
194. Winter, J., P. Neubauer, R. Glockshuber, and R. Rudolph. 2001. Increased production of human proinsulin in the periplasmic space of *Escherichia coli* by fusion to DsbA. *J. Biotechnol.* **84:** 175-185.
195. Workman, C. and A. Krogh. 1999. No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution. *Nucleic Acids Res.* **27:** 4816-4822.
196. Yokoyama, S. 2003. Protein expression systems for structural genomics and proteomics. *Curr. Opin. Chem. Biol.* **7:** 39-43.
197. Yoon, S. and W. Kang. 1994. Fed-batch operation of recombinant *Escherichia coli* containing the *trp* promoter with controlled specific growth rate. *Biotechnol. Bioeng.* **43.**
198. Zahn, K. 1996. Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth. *J. Bacteriol.* **178:** 2926-2933.
199. Zavialov, A. V., N. V. Batchikova, T. Korpela, L. E. Petrovskaya, V. G. Korobko, J. Kersley, S. MacIntyre, and V. P. Zav'yalov. 2001. Secretion of recombinant proteins via the chaperone/usher pathway in *Escherichia coli*. *Appl. Environ. Microbiol.* **67:** 1805-1814.
200. Zhang, Y., D. R. Olsen, K. B. Nguyen, P. S. Olson, E. T. Rhodes, and D. Mascarenhas. 1998. Expression of eukaryotic proteins in soluble form in *Escherichia coli*. *Protein Expr. Purif.* **12:** 159-165.