Appl Biochem Biotechnol (2011) 165:523–537
DOI 10.1007/s12010-011-9272-5

Optimization of Human Granulocyte Macrophage-Colony Stimulating Factor (hGM-CSF) Expression Using Asparaginase and Xylanase Gene's Signal Sequences in *Escherichia coli*

Yogender Pal Khasa · Amardeep Khushoo · Suman Tapryal · K. J. Mukherjee

Received: 30 June 2010 / Accepted: 25 April 2011 /
Published online: 12 May 2011

Abstract The toxicity of the recombinant protein towards the expression host remains a significant deterrent for bioprocess development. In this study, the expression of human granulocyte macrophage-colony stimulating factor (hGM-CSF), which is known to be toxic to its host, was enhanced many folds using a combination of genetic and bioprocess strategies in *Escherichia coli*. The N terminus attachment of endoxylanase and asparaginase signal sequences from *Bacillus subtilis* and *E. coli*, respectively, in combination with and without His-tag, considerably improved expression levels. Induction and media optimization studies in shake flask cultures resulted in a maximal hGM-CSF concentration of 365 mg/L in the form of inclusion bodies (IBs) with a specific product yield ($Y_{P/X}$) of 120 mg/g dry cell weight in case of the asparaginase signal. Culturing the cells in nutrient rich Terrific broth maintained the specific product yields ($Y_{P/X}$) while a 6.6-fold higher volumetric concentration of both product and biomass was obtained. The purification and refolding steps were optimized resulting in a 95% pure protein with a fairly high refolding yield of 45%. The biological activity of the refolded protein was confirmed by a cell proliferation assay on hGM-CSF dependent human erythroleukemia TF-1 cells. This study demonstrated that this indeed is a viable route for the efficient production of hGM-CSF.

Y. P. Khasa
Department of Microbiology, University of Delhi South Campus, New Delhi 110021, India

Y. P. Khasa · A. Khushoo · S. Tapryal · K. J. Mukherjee (✉)
School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India
e-mail: kjmukherjee@mail.jnu.ac.in

A. Khushoo
Oregon Health and Science University, Portland 97239, USA

S. Tapryal
National Institute of Immunology, New Delhi 110067, India

Keywords hGM-CSF · Asparaginase · Endoxylanase · Signal sequence · Recombinant protein · *E. coli*

Introduction

Recent advances in our understanding of the function, regulation, and interaction of cellular gene products together with the availability of new genetic tools have allowed *Escherichia coli* as the preferred host for recombinant protein expression [1–3]. The secretory production of recombinant proteins using different signal sequences provide several advantages like reduced level of proteolytic degradation, ease of purification, and disulfide bond formation [4, 5], whereas a protein fusion at N terminus can also reduce the rates of degradation brought about by the presence of specific amino acid residues at the amino terminus [6–8]. The *E. coli* native signal sequences such as PhoA, OmpA, OmpT, and OmpF or *Bacillus subtilis* endoxylanases or pelB of *Erwinia carotovora* are known for efficient export of recombinant proteins to the periplasmic space or culture medium [4, 9, 10]. Interestingly, the extracellular expression of *E. coli* asparaginase protein using its native signal sequence has been reported in the culture supernatant [11]. The asparaginase gene's signal sequence has also been used for efficient release of cholesteryl ester transfer protein C terminus fragment and hirudin III protein extracellularly [12]. Similarly, the endoxylanase gene signal sequence from *B. subtilis* is also effective in the secretion of recombinant proteins like human granulocyte-colony stimulating factor (hG-CSF), alkaline phosphatase, and human leptin to the *E. coli* periplasm [13–15].

Cytokines are low molecular weight proteins secreted by a variety of cells of the immune system. Of this human granulocyte macrophage-colony stimulating factor (hGM-CSF) is a multi-lineage cytokine involved in the proliferation, differentiation, and survival of macrophages, neutrophils, and bone marrow associated cells [16]. It has numerous clinical applications ranging from treatment of leukemia, neutropenia, and aplastic anemia to cure bacterial and fungal infections. Furthermore, as a supplement during bone marrow transplantation, hGM-CSF is known to accelerate neutrophil formation which in turn prevents infection [17].

The toxicity of hGM-CSF towards its hosts is the major bottleneck hampering its overexpression in most recombinant hosts. Therefore, determining the critical factors responsible for poor expression is a challenging task for researchers. The soluble cytoplasmic expression of hGM-CSF adversely affects the cell physiology along with cell lysis and poor yield in *E. coli.* On the other hand, inclusion bodies formation requires a subsequent refolding step often resulting in compromised final specific activities and poor refolding yields typically in the range of 10–20% [18, 19]. The use of signal sequences such as SP1 and SP3 has been shown to increase expression levels to 0.45 μg/mL/OD in the periplasmic space [20]. The fusion of a commercially available signal peptide pelB and MBP-tag at the N terminus resulted in a product concentration of 40 μg/mL and 69.5 μg/mL, respectively [6]. Recently, hGM-CSF expression has been reported using salt inducible GJ1158 *E. coli* cells, where a soluble expression of 20 μg/mL was obtained [21]. Use of alternate expression system like plant cells (0.783 μg/mL) [22], transgenic animals (0.180 μg/mL) [23], and baculovirus system [24] has shown marginal increase in expression but appeared to be unacceptable for large-scale production.

In this study, we describe the use of xylanase (*B. subtilis*) and asparaginase (*E. coli*) gene signal sequences with and without His-tag for efficient and high-yield production of hGM-CSF in *E. coli* shake flask culture.

Materials and Methods

Bacterial Strains and Plasmids

Different vectors such as pRSET-A (Amp^R) (Invitrogen, USA), pET22-b (Amp^R) (Novagen, USA), and pET28-a (Kan^R) (Novagen) were used for development of different expression plasmids. The hGM-CSF gene was cloned under the T7 promoter. *E. coli* DH5α (Amersham Biosciences, USA) was used for cloning and plasmid propagation. Expression studies were conducted in *E. coli* BLR (DE3) strain (Novagen). The previously reported pRSET-GMCSF plasmid having hGM-CSF gene under T7 promoter was used as control where no fusion tag was incorporated [6]. The media constituents for growth were purchased from Hi-Media, India. The dNTPs, antibiotics, and IPTG were purchased from Sigma, USA.

PCR Amplification and Cloning

The primers used in this study were obtained from Microsynth (Switzerland). All PCR reactions were subjected to a program consisting of initial DNA denaturation step at 95 °C for 5 min, followed by 95 °C for 1 min, 48 °C for 1 min, and 72 °C for 45 s for a total of 30 cycles. The primers used in this study have been given in Table 1 with their respective restriction sites. The PFU DNA polymerase, restriction, and modifying enzymes were obtained from MBI Fermentas (USA). The sequence of recombinant expression plasmids was confirmed by PCR, restriction analysis, and DNA sequencing.

Table 1 Primers used in this study to construct various expression plasmids for hGM-CSF in *E. coli*

Primer number	Primer sequences (5′–3′)
1	GGA GGT AA**C ATA TGT TTA AGT TTA AAA AG** forward *Nde*I (Xyl sequence)
2	GGC CAG CTG CAG AGG CGG TTG CCG AAA AC reverse *Pst*I (Xyl sequence)
3	GGC CAG **CTG CAG GCA CCC GCC CGC TCG** forward *Pst*I (hGM-CSF gene)
4	C **GAA TTC TCA CTC CTG GAC TGG** reverse *Eco*RI (hGM-CSF gene)
5	GCC ATC AC**C CAT GGA GGC GGT TGC CGA AAA** C reverse *Nco*I (Xyl sequence)
6	GTG CAG CA**C ATA TGA GTT TTT CAA AAA GAC GGC AC** forward *Nde*I (Asp sequence)
7	CGG CAC TG**C CAT GGT GCC AAT GCT GCT GCA CCA CT** reverse *Nco*I (Asp sequence)
8	CGG CAC TG**C CAT GGA GTT TTT CAA AAA GAC GGC AC** forward *Nco*I (Asp sequence)
9	GTG CAG CA**C ATA TGT GCC AAT GCT GCT GCA CCA CT** reverse *Nde*I (Asp sequence)
10	CCT GCA GC**C ATA TGG CAC CCG CCC GCT TCG** forward *Nde*I (hGM-CSF gene)
11	GG**G GAT CCT CAC TCC TGG ACT GG** reverse *Bam*HI (hGM-CSF gene)
12	CGG CAC TG**C CAT GGA GTT TTT CAA AAA GAC GGC AC** forward *Nco*I (Asp sequence)
13	CGC GAT AC**C ATA TGA ATC CTT CCC TCG ATG TAC TGA TTG AAG ATC TG** reverse *Nde*I (complete asparaginase gene)

Shake Flask Expression Studies

To determine the expression of hGM-CSF in shake flasks, *E. coli* BLR (DE3) cells were transformed with different plasmids and grown in 15 mL LB (Luria Bertani) medium containing yeast extract 5 g/L, tryptone 10 g/L, NaCl 10 g/L, pH 7.2, and 15 mL Terrific broth (TB) containing yeast extract 24 g/L, tryptone 12 g/L, KH_2PO_4 2.31 g/L, K_2HPO_4 12.54 g/L, and 0.4% glycerol, pH 7.2. In LB medium, cells were induced at OD_{600} of ~0.8, whereas in TB medium cells were induced at different growth phase, i.e., early, mid-, and late log phase using 1 mM IPTG. The antibiotics were used at a concentration of 100 μg ampicillin/mL and 50 μg kanamycin/mL. The cell biomass was determined by taking 1 mL of culture sample in a pre-weighed centrifuge tube and centrifuging at 12,000×*g* at 4 °C for 5 min. The supernatant was removed, the cells were washed with saline, and dry cell weight (DCW) was measured by drying the wet pellet to a constant weight in an 80 °C oven.

Quantitative ELISA for hGM-CSF

The hGM-CSF concentration was measured using ELISA Duo set as per manufacturer's instructions (R&D Systems, USA). The total product concentration was quantified using a standard curve for hGM-CSF (R&D Systems) with the appropriately diluted samples for linear range of ELISA reading. Sample preparation for ELISA has been comprehensively elaborated elsewhere [25].

SDS–PAGE and Western Blot Analysis

For SDS–PAGE analysis, the induced culture was normalized by dilution to an OD_{600} of 1 and 50 μL of this was centrifuged at 10,000×*g* for 10 min at 4 °C to obtain uniform-size pellets. The cells were then re-suspended in 80 μL of MilliQ water and 20 μL of standard 5× SDS loading buffer [25 mM Tris–HCl (pH 6.8), 0.1 M 2-mercaptoethanol, 10% SDS, and 50% glycerol]. The samples were boiled at 95 °C for 5 min, centrifuged for 2 min at 10,000×*g* to remove cell debris, and 5 μL of this sample was subsequently loaded on a 15% SDS–PAGE gel. Western blot analysis was done by separating protein samples on 15% SDS–PAGE followed by electrophoretic transfer to a nitrocellulose membrane (Sigma) using buffer consisting of 25 mM Tris–HCl, 192 mM glycine, and 20% methanol. The membrane was blocked using 20 mL PBST solution (PBS+0.1% Tween 20) and 1% BSA. To detect the hGM-CSF bands, mouse anti-hGM-CSF monoclonal antibody (R&D Systems) was used at a dilution of 1:10,000 in PBST buffer, followed by three 10-min washes in the same buffer. The membrane was further incubated with rabbit anti-mouse IgG secondary antibody, conjugated to horseradish peroxidase (HRPO) (Sigma) at a dilution of 1:10,000. Each antibody solution was kept in contact with the membrane for 1 h at room temperature. The Western blot was developed with 10 mM 3,3′-diaminobenzidine (Sigma) in 100 mM Tris–HCl, pH 7.6, until the bands developed to the desired intensity.

Development of Various Constructs Using Xylanase and Asparaginase Signal Sequence

Construction of pRSETXyl-GM and pETXH-GM Expression Plasmid Using the Xylanase Signal Sequence

To clone the endoxylanase signal sequence upstream of the hGM-CSF gene, it was PCR amplified from pUC19-xylanase plasmid using P1 (forward/*Nde*I) and P2 (reverse/*Pst*I)

primers [26]. The gel purified PCR fragment was double digested with *Nde*I/*Pst*I restriction enzymes and ligated to the pRSET-B vector to obtain pRSET-Xyl plasmid. The hGM-CSF gene was then PCR amplified from pETHis-GM [16] using P3 (forward/*Pst*I) and P4 (reverse *Eco*RI) primers. The gel eluted hGM-CSF fragment was double digested with *Pst*I and *Eco*RI restriction enzymes and ligated into pRSETXyl plasmid digested with same enzymes to generate pRSETXyl-GM plasmid (Fig. 1a).

To generate the recombinant hGM-CSF under xylanase signal sequence with His-tag, the signal sequence was PCR amplified using P1 (forward/*Nde*I) and P5 (reverse/*Nco*I) primers and ligated to pET22-b vector digested with *Nde*I and *Nco*I restriction enzymes to generate pETXyl plasmid. The hGM-CSF gene fragment with N terminus His tag was obtained via *Nco*I and *Bam*HI double digestion of pETHis-GM plasmid [6] and subsequently ligated into pETXyl plasmid to generate the pETXH-GM construct (Fig. 1b).

Construction of pETAH-GM, pETA-GM, and pETAF-GM Expression Plasmid Using Asparaginase Signal Sequences

The asparaginase signal sequence was PCR amplified using P6 (forward/*Nde*I) and P7 (reverse/*Nco*I) primers from pT7Asp plasmid containing the complete asparaginase gene [12]. The PCR-amplified asparaginase signal sequence was then double digested with *Nde*I and *Nco*I restriction enzymes and ligated to pET22-b vector to generate pET22Asp plasmid. The hGM-CSF gene fragment with N terminus His-tag was obtained from pETHis-GM plasmid by *Nco*I and *Bam*HI double digestion and then ligated into pET22Asp plasmid to generate pETAH-GM construct with asparaginase signal sequence upstream of the hGM-

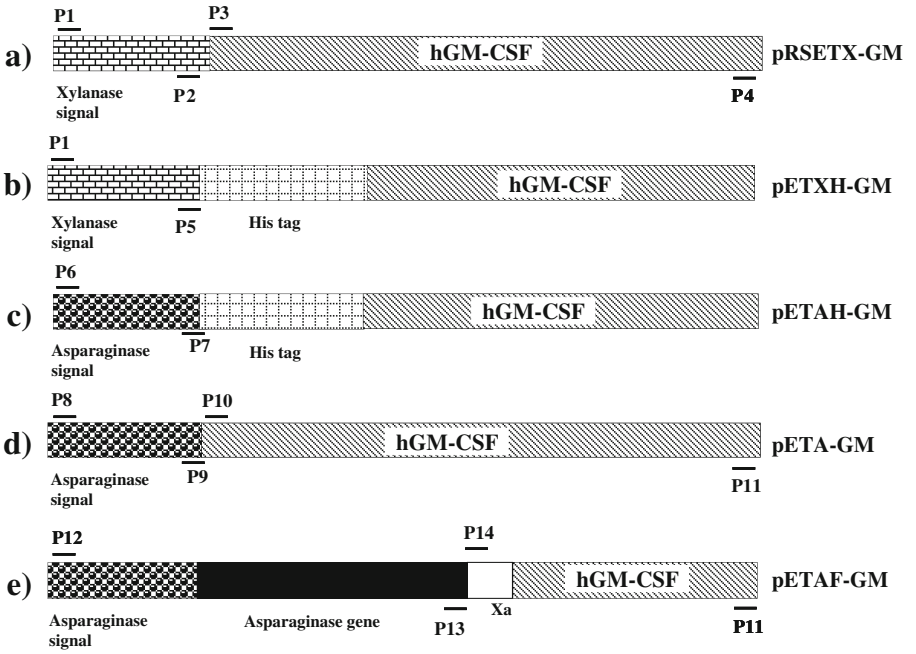

Fig. 1 Schematic representation of different expression plasmids having N terminus endoxylanase and asparaginase gene's signal sequence with and without His-tag. **a** pRSETXyl-GM; **b** pETXH-GM; **c** pETAH-GM; **d** pETA-GM; **e** pETAF-GM

CSF gene (Fig. 1c). The hGM-CSF gene was also cloned without the N terminus His-tag in a pET28-a vector. The asparaginase signal sequence was PCR amplified using P8 (forward/*Nco*I) and P9 (reverse/*Nde*I) primers from pET29Asp plasmid [12]. The PCR-amplified asparaginase signal sequence was double digested with *Nco*I and *Nde*I restriction enzymes and ligated to pET28-a vector to generate pETAsp plasmid. The hGM-CSF gene was PCR amplified from pETHis-GM recombinant plasmid [6] using P10 (forward/*Nde*I) and P11 (reverse/*Bam*HI) and ligated to pETAsp plasmid digested with *Nde*I/*Bam*HI restriction enzymes to generate pETA-GM plasmid without His-tag (Fig. 1d).

To further investigate the effect of fusion of the complete asparaginase gene at N terminus of hGM-CSF, the full-length asparaginase gene was PCR amplified using P12 (forward/*Nco*I) and P13 (reverse/*Nde*I) primers from pET29Asp recombinant plasmid. The factor Xa cleavage site was introduced at the 3′ end of the asparaginase gene through reverse primer P13. This PCR-amplified fragment was cloned into the *Nco*I and *Nde*I digested pET28-a expression vector to generate the pETAspFus plasmid. To generate the chimeric plasmid, the hGM-CSF gene was PCR amplified from pETHis-GM using P10 (forward/*Nde*I) and P11 (reverse/*Bam*HI) primers. The *Nde*I/*Bam*HI double-digested hGM-CSF fragment was ligated into pETAspFus plasmid to generate the pETAF-GM construct which had full-length asparaginase gene (Fig. 1e).

Plasmid Stability Determination

In shake flask studies, the pre- and post-induction culture samples were taken at regular time intervals, and appropriately diluted cells were plated on LB agar plates without antibiotics. On average, 100 isolated colonies were picked up from the LB agar plate by toothpicks and transferred on to LB agar plates supplemented with an appropriate antibiotic (ampicillin or kanamycin) under sterile conditions. The fraction of the colonies that did not grow on the antibiotic plate gave the fraction of non-recombinants in the cell population. In this study, we measured only segregational instability. All measurements were done in triplicates and average values have been reported.

Purification and Refolding of hGM-CSF

Inclusion body (IB) isolation was done by taking 30 mL culture from mid-log phase induced cells from the pETA-GM clone. Cells were lysed by sonication (30-s burst/30-s cooling/20 mHz) for 15 cycles using a sonicator from Misonix, Inc., USA. The inclusion bodies were isolated by centrifuging the cell lysate at 12,000×*g* for 30 min at 4 °C. The pellet was resuspended in 1% deoxycholic acid (DCA) and incubated at 25 °C with slow shaking for 1 h after which the sample was centrifuged at 12,000×*g* for 10 min. This was followed by washing with Tris buffer (50 mM Tris–HCl, 5 mM EDTA, 1 mM PMSF, pH 8.0). The IB washing step was repeated twice and final washing was done with distilled water. The hGM-CSF gene sequence with asparaginase signal sequence at its 5′ end was translated at http://www.ExPASy.com and the isoelectric point (p*I*) of the translated protein sequence was calculated to be 5.8. Accordingly, the protein was purified using anion exchange chromatography at pH 8.5. Approximately 9 mg inclusion bodies were resuspended in column buffer (50 mM Tris, 8 M urea, pH 8.0) at a concentration of 1 mg/mL and passed through a 5-mL anion exchange column (Mono Q HR 5/5; Bio-Rad, USA) equilibrated with column buffer. The NaCl concentration in the elution buffer was optimized over a range of 0.1 M to 1 M NaCl, in step gradient elution experiments (data not shown). The protein was eluted maximally at 0.4 M NaCl concentration in a single step. The column was washed with column buffer and

bound hGM-CSF was eluted with elution buffer (50 mM Tris–HCl, 8 M urea, 0.4 M NaCl) in 10 fractions (1 mL each). Different fractions were run on 15% SDS–PAGE. All 10 fractions were pooled together and subjected to dilution refolding by adding them dropwise into 100 mL of renaturation buffer [50 mM Tris–HCl, 2 M urea, 10% sucrose, 1 mM EDTA, 1 mM PMSF, 3 mM glutathiones (reduced), 0.5 mM (oxidized), 0.4 M NaCl, 0.5 M arginine, pH 8.0]. The pulse dilution refolding was done at 4 °C with slow stirring overnight. The refolded hGM-CSF was dialyzed against the renaturation buffer without glutathiones. The urea concentration was decreased stepwise to zero during dialysis (with gradual decrease of 0.1 M at a time). The refolded protein was further concentrated using centricon membrane of 3.5 kDa MWCO. Protein concentration was quantified following standard method using BSA as standard.

Biological Activity of hGM-CSF

The biological activity assay of recombinant hGM-CSF was performed on hGM-CSF dependent human erythroleukemia TF-1 cells (ATCC, CRL 2003). The cells were cultured in RPMI-1640 medium containing 10% v/v fetal bovine serum. The standard hGM-CSF, as a positive control, was commercially purchased from R&D Systems. Cells were washed with incomplete RPMI medium, counted, and plated in a 96-well plate in triplicates at 1×10^5 cells per well and cultured (with 5% CO_2 in a 37 °C incubator) for 48 h. The samples were divided into three groups: negative (no hGM-CSF), positive [hGM-CSF expressed and refolded in this study at a varying concentration (0 to 64 ng/mL)], and standard control (2 ng/mL standard hGM-CSF from R&D Systems). Cells maintained in media without hGM-CSF was treated as negative control. The hGM-CSF growth promoting activity of TF-1 cells was evaluated by MTT colorimetric assay [19, 27]. The experiments were carried out in triplicate and average values had been reported.

Results

Expression Studies Using Xylanase Signal Sequence Clones

The expression of hGM-CSF in shake flask culture was determined by growing transformed *E. coli* BLR cells with the pRSETXyl-GM and pETXH-GM constructs (Fig. 1) in 15 mL LB medium. Cells were induced at an optical density (OD_{600}) of 0.8 with 1 mM IPTG. Following induction, the cells stopped growing and a decline in biomass was observed with both clones (Fig. 2a). After 7 h of induction, a final OD_{600} of 0.51 and 0.72 was achieved for pRSETXyl-GM and pETXH-GM, respectively. The expression using xylanase signal sequence clone, i.e., pRSETXyl-GM, resulted in poor soluble hGM-CSF concentration of 7 mg/L; however, fusion of xylanase signal and N terminus His-tag, i.e., pETXH-GM clone, increased its expression to 27 mg/L (primarily as inclusion bodies) (Fig. 2b). The PRSET-GMCSF control plasmid without any fusion tag yielded an hGM-CSF concentration of 0.8 mg/L of soluble protein; however, fusion of only His-tag improved its expression to 5 mg/L ($Y_{P/X}$ of 7.5 mg/g DCW) (data not shown). The low protein expression level led to protein solubility with deleterious effects on cell physiology, whereas improved expression resulted in protein aggregates (Fig. 2b). Thus, optimization studies with pETXH-GM clone were conducted to further improve hGM-CSF expression

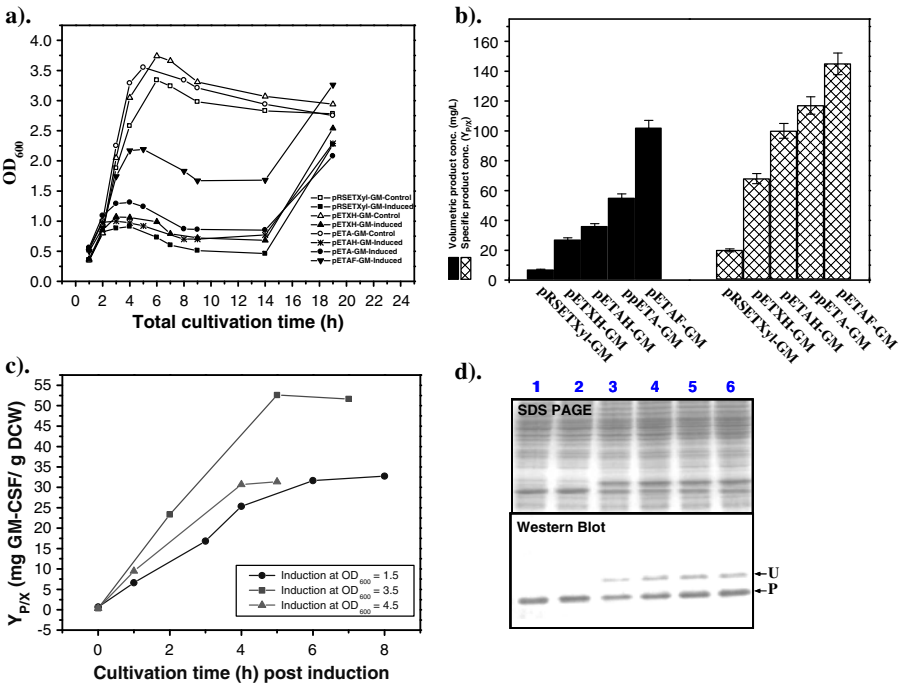

Fig. 2 Expression of hGM-CSF using xylanase and asparaginase gene's signal sequence in BLR (DE3) cells having pRSETXyl-GM, pETXH-GM, pETAH-GM, pETA-GM, and pETAF-GM recombinant plasmids. **a** Growth profile (OD_{600}) in LB medium. **b** Volumetric and specific product concentration in LB medium. **c** Specific product concentration in TB medium (cells were induced in different phase of growth, i.e., OD_{600} of 1.5, 3.5, and 4.5). **d** SDS–PAGE and Western blot analysis of recombinant hGM-CSF from pETXH-GM plasmid in LB and TB medium. *Lanes 1* and *2* represent the second and fourth hour post-induction samples in LB medium. *Lanes 3*, *4*, *5*, and *6* represent post-induction sample of first, second, third, and fourth hours, respectively, in TB medium. *U* unprocessed signal sequence, *P* processed signal sequence

The nutrient-rich TB medium is known to alleviate the stress associated with recombinant protein expression. A plot of ln OD_{600} vs. time for growth in TB medium showed that the specific growth rate (μ) declined with time, it was 0.6 h^{-1} when the OD_{600} was 1.5 and it fell to 0.2 h^{-1} at the late log phase (OD_{600}=4.5). Therefore, *E. coli* cells containing pETXH-GM plasmid were further tested in 15 mL TB medium and induced at different cell densities of OD_{600} ~1.5 (early log phase), OD_{600} ~3.5 (mid-log phase) and OD_{600} ~4.5 (late log phase) using 1 mM IPTG. The cell growth continued unabated in the culture induced at OD_{600} ~1.5 till 3 h of induction and then it gradually declined. Cells induced at higher OD_{600} of 3.5 and 4.5 showed a marginal decline in cell growth as compared to un-induced control. The final cell biomass (OD_{600}) of 3.16, 3.5, and 3.85 was achieved in early, mid-, and late log phase of induction, respectively, compared to un-induced control where an OD_{600} of 4.4 was observed. The early and late log phase induction resulted in a maximum hGM-CSF concentration of 42 mg/L ($Y_{P/X}$=30 mg/g DCW) and 50 mg/L ($Y_{P/X}$=37 mg/g DCW), respectively (Fig. 2c). The mid-log phase induced culture showed maximum expression where a product concentration of 68 mg/L of hGM-CSF with a specific product yield ($Y_{P/X}$) of 52 mg/g DCW was obtained in cytoplasm after 5 h of induction. The $Y_{P/X}$ yields achieved was 1.75-fold and 1.4-fold higher compared to early and late log phase induction. Interestingly, TB medium gave a 2.5-fold higher volumetric product

concentration over LB medium, demonstrating the contribution of richer medium in obtaining higher cell biomass albeit with similar final product yields.

The expressed hGM-CSF showed an expected molecular weight of ~18 kDa on SDS–PAGE (Fig. 2d). However, a faint band at ~21 kDa was also observed, demonstrating inefficient processing of the xylanase signal sequence. The fact that the 21-kDa band was indeed the hGM-CSF protein, it was further confirmed by Western blot analysis (Fig. 2d) and N terminus protein sequencing.

Expression Studies Using the Asparaginase Signal Sequence

The three *E. coli* clones containing the asparaginase signal sequence (pETAH-GM, pETA-GM, and pETAF-GM) were grown in LB medium and induced at an OD_{600} of 0.8 with 1 mM IPTG. From the growth profile (Fig. 2a), a sharp decline in biomass was observed in case of pETAH-GM, whereas cell growth was maintained till 3 h post-induction in pETA-GM. Interestingly, hGM-CSF expressed as a fusion with the complete asparaginase protein supported the maximum cell growth. A remarkable recovery in cell growth was observed after 12 h of induction which could be because of plasmid-free cells. The plasmid's segregation stability was monitored and the culture was found to be stable before induction (95%), which rapidly declined post-induction to a value of 65–70%. No such decline was observed with the un-induced control. The rapid drop in plasmid stability is typically associated with the metabolic burden of expressing a toxic recombinant protein. The pETAH-GM clone gave an hGM-CSF concentration of 36 mg/L with a specific product yield of 100 mg/g DCW. The fusion of the complete asparaginase gene at the N terminus resulted in a maximum volumetric product concentration of 102 mg/L ($Y_{P/X}$=145 mg/g DCW) as compared to 55 mg/L ($Y_{P/X}$= 117 mg/g DCW) for the pETA-GM clone containing only asparaginase signal peptide (Fig. 2b). However, the molecular ratio of asparaginase (~36 kDa) to hGM-CSF (~15 kDa) in fusion protein is ~2.4 implying that the effective volumetric product concentration of hGM-CSF was only 42 mg/L in the fusion protein. The distribution of soluble and insoluble fractions of hGM-CSF in the cytoplasm is given in Table 2. From SDS–PAGE analysis, two bands of ~21 kDa and ~18 kDa were observed due to the additional asparaginase signal sequence and also the 6× His-tag (which increased the molecular weight of hGM-CSF by ~3 kDa each; from the original molecular weight of ~15 kDa). Though expression was obtained as inclusion bodies, it was felt that this was the appropriate strategy for hGM-CSF expression to circumvent the metabolic stress, associated with product toxicity. Therefore, further studies on hGM-CSF expression were carried out using the pETA-GM clone.

Optimization of Induction Strategy Using pETA-GM Construct in TB Medium

To further optimize the expression level using pETA-GM plasmid, the cells were induced at different growth phase, i.e., early log (OD_{600}=2), mid-log (OD_{600}=4), and late log (OD_{600}=5.5)

Table 2 Distribution of hGM-CSF in soluble and inclusion bodies with different clones under the asparaginase gene's signal sequence after 4 h of induction

Clone	Insoluble fraction in cytoplasm (inclusion bodies) (mg/L)	Soluble fraction on cytoplasm (mg/L)
pETAH-GM	32 (~94%)	2 (~6%)
pETA-GM	52 (~97%)	1.7 (~3%)
pETAF-GM	96 (~95%)	5 (~5%)

phase using 1 mM IPTG in 15 mL TB medium. Cells induced at early log phase showed the maximum decline in specific growth rate; however, this decline was moderate in mid- and late log phase induction. The un-induced control culture grew normally without exhibiting any growth inhibition (Fig. 3a). Induction at early log phase resulted in a cell density of 3.98 after 10 h of growth, whereas the final cell density obtained in mid- and late log phase induction was comparable to the un-induced culture which flattened out at an OD_{600} of 6.6. The maximum

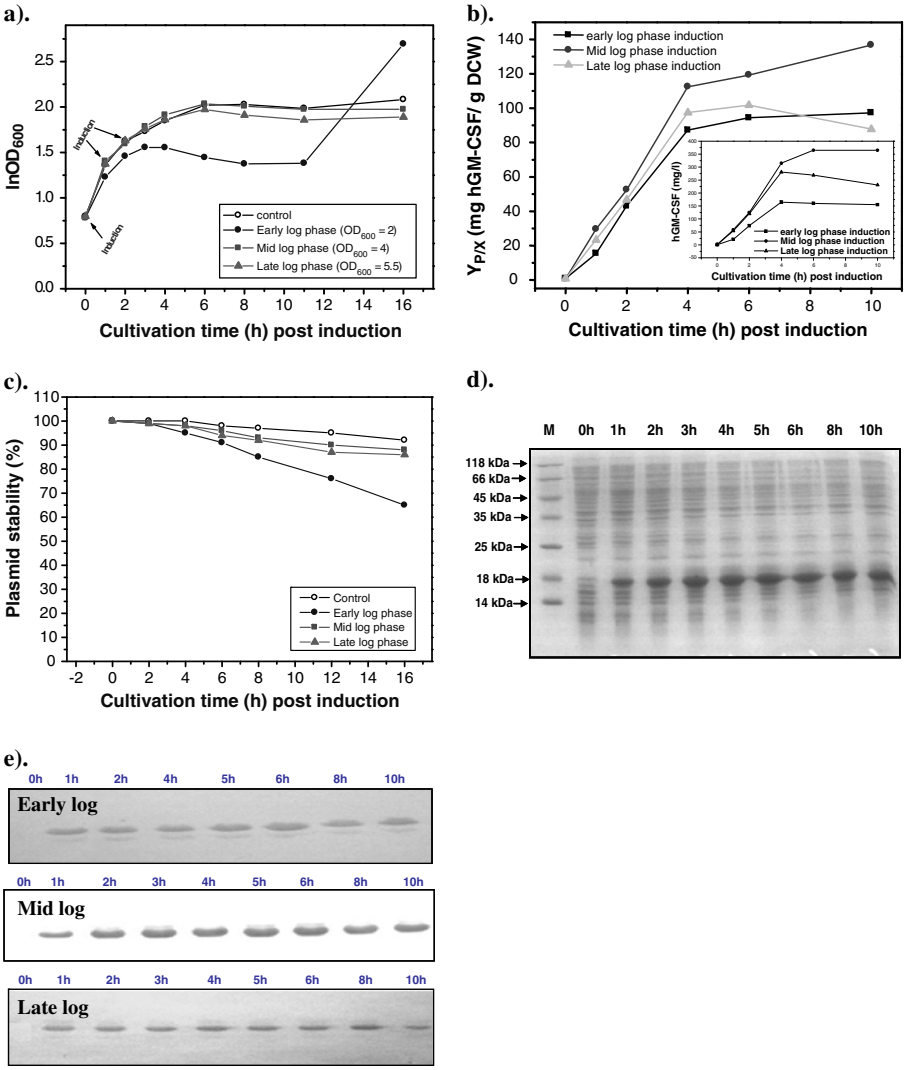

Fig. 3 Expression studies using pETA-GM plasmid in TB medium. **a** Growth profile (ln OD_{600}). **b** Volumetric and specific product concentration of hGM-CSF using pETA-GM plasmid in TB medium (inside hGM-CSF concentration where cells were induced in different phase of growth, i.e., OD_{600} of 2, 4, and 5.5). **c** Plasmid stability profile of early, mid- and late log phase induced culture. **d** SDS–PAGE analysis of hGM-CSF expression in mid-log phase induced culture having pETA-GM plasmid. **e** Western blot analysis of different phase induced cells, i.e., early, mid-, and late log phase

volumetric product concentration was obtained in mid-log phase induced culture at 365 mg/L, which was 2.4 and 1.6 times higher than early and late log phase induction, respectively (Fig. 3b). Interestingly, the specific product concentration ($Y_{P/X}$) in mid-log phase was slightly higher at 120 mg/g DCW, which clearly reflected the enhanced production capacity of individual cells in nutrient-rich medium. To determine the optimum IPTG concentration, induction was done at mid-log phase with varying IPTG concentrations ranging from 0.1 mM to 1.6 mM (doubling the IPTG concentration in a geometric series). There was a slight increase in $Y_{P/X}$ from 0.1 mM to 0.4 mM IPTG after which no further increase was observed which was also similar to previous studies [28]. We therefore decided to continue our experiment with 1 mM IPTG induction.

To investigate the remarkable recovery in cell growth after 10 h of induction in early log phase (Fig. 3a), the plasmid segregational instability was determined at regular time intervals. The un-induced culture showed a maximum stability of 92±5% which decreased to 64±5% when cells were induced at early log phase. However, the plasmid remained reasonably stable in mid- (88±5%) and late (84±5%) log phase induction even after 16 h of growth (Fig. 3c).

From SDS–PAGE analysis, the maximum recombinant protein band intensity was around 25% of the total cellular protein after 4 h of induction (Fig. 3d). The most interesting observation from Western blot was the slight signal processivity of hGM-CSF which was observed only in early log phase, whereas no processed band of hGM-CSF was observed in mid- and late log phase induced culture (Fig. 3e). However, in spite of the signal sequence processivity, the protein was obtained in the soluble cytoplasmic fraction which might be the reason behind the high depression of cell growth rate as well as higher plasmid instability in early log phase induced culture.

Inclusion Body Purification and Refolding of hGM-CSF

Inclusion body purification was performed with IPTG induced culture of pETA-GM clone from TB medium (Fig. 4a). Briefly, the inclusion bodies washed with deoxycholic acid resulted in ~90% protein purity as an estimate from SDS–PAGE (Fig. 4b). The purification was performed under denaturing condition in the presence of 8 M urea. A total of 9 mg of protein was loaded onto the column at a concentration of 1 mg/mL. After washing, the bound protein was eluted using optimized single-step salt gradient of 0.4 M NaCl. A total of 10 fractions of 1 mL each were collected and analyzed on 15% SDS–PAGE (Fig. 4c). The anion-exchange chromatography in the presence of 8 M urea resulted in ~95% purity. The whole purification procedure, started with 9 mg protein (inclusion bodies), resulted in ~95% pure protein with 45% recovery after completion of folding process. Refolding was done by pulse dilution in the presence of arginine and glutathiones followed by dialysis where urea was removed stepwise. The refolded protein was concentrated using centricon membrane, and no significant protein loss was observed.

Biological Activity of hGM-CSF

The biological activity of refolded hGM-CSF was estimated by measuring its ability to support the growth of hGM-CSF dependent human erythroleukemia TF-1 cell line in MTT assay. The samples were divided into three groups: negative (no hGM-CSF), positive (hGM-CSF expressed and refolded in this study), and standard control (standard hGM-CSF from R&D Systems). Cells maintained in media without hGM-CSF were treated as

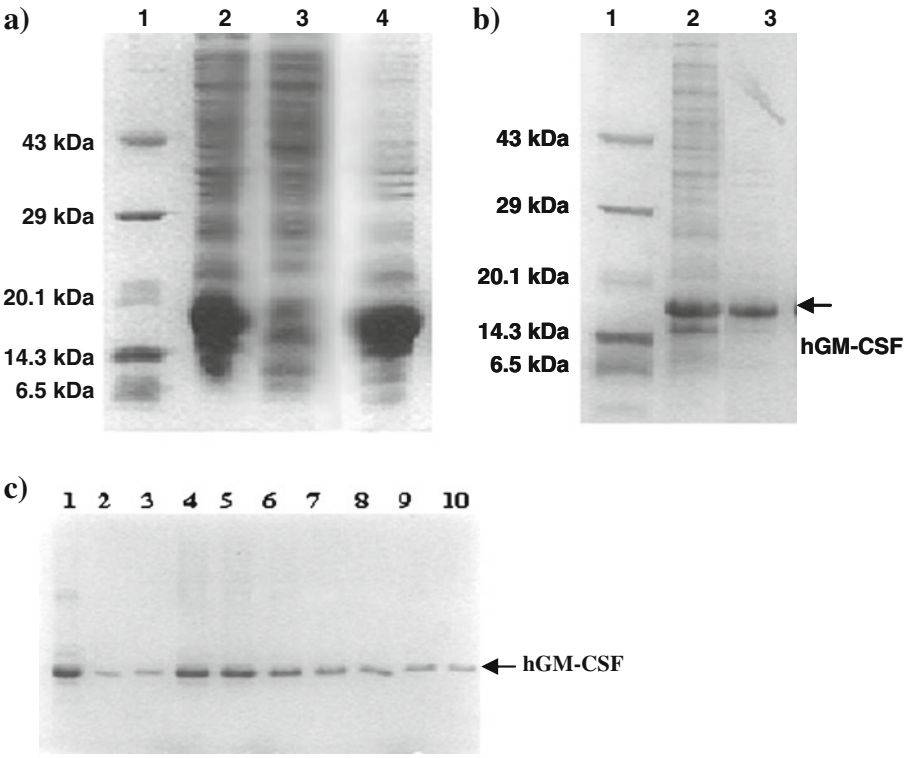

Fig. 4 **a** SDS–PAGE analysis of purified and refolded hGM-CSF inclusion bodies produced in BLR (DE3) *E. coli* cells. **a** hGM-CSF inclusion body isolation. *Lane 1*: mol. wt. marker; *lane 2*: total cell lysate; *lane 3*: soluble fraction; *lane 4*: insoluble fraction. **b** Purified hGM-CSF IBs after 1% DCA wash. *Lane 1*: mol. wt. marker; *lane 2*: total cell lysate; *lane 3*: 1% DCA washed hGM-CSF IBs. **c** Anion exchange purified hGM-CSF under denaturating conditions. *Lane 1*: 8 M urea denatured hGM-CSF IBs; *lanes 2–10*: various fractions of column eluted protein

negative control. From the results, it was observed that the anion exchange purified and refolded recombinant hGM-CSF showed cell proliferation over negative control. There was no significant difference between proliferation activities of standard hGM-CSF purchased from R&D Systems (positive control) over the recombinant hGM-CSF developed in this study. The refolded hGM-CSF concentrations tested were from 0 to 64 ng/mL; however, 2 ng/mL appeared to be the saturation limit which is also the recommended maximum concentration for standard obtained from R&D Systems (data not shown).

Discussion

The critical factor impeding successful commercialization of hGM-CSF has been extremely low expression levels and poor purifications yields in a range of hosts including *E. coli* [20, 29]. To address this issue, we optimized various genetic and expression strategies, essentially using xylanase (*B. subtilis*) and asparaginase (*E. coli*) signal sequences either with or without His-tag. The essential idea behind this strategy was to facilitate the export of hGM-CSF to the periplasm or even the extracellular medium [9, 10], thereby preventing its interaction with cellular components. We had previously observed that hGM-CSF toxicity was the primary reason

behind poor yields [6], and export seemed an elegant solution to this problem. However, the most surprising part of the results was that a significantly higher hGM-CSF yield (as cytoplasmic IBs) was obtained in the case of pETA-GM carrying the asparaginase signal compared to the case when the protein was expressed without a signal peptide [6]. The high level expression is possibly the reason behind protein misfolding and formation of inclusion bodies in the cytoplasm [30]. It is well known that the inclusion body formation sequesters the protein inside the cell and blocks its interaction with the cellular machinery, effectively negating the toxic effect of its expression [30–32]. The introduction of the asparaginase signal thus had the fortuitous outcome of overcoming the toxicity associated with hGM-CSF expression by the completely unexpected route of IB formation. It is clear from the post-induction growth profiles that the cells producing hGM-CSF in inclusion body form were not significantly retarded compared to the un-induced culture. It is important to point out that our previous experiments with other leader peptides like pelB and fusion tags like Maltose Binding Protein (MBP-tag) did not lead to IB formation [6] possibly because misfolding occurs only above a certain rate of expression. Clearly, the maximum increase in expression rates was obtained when the asparaginase signal was used since it is native to *E. coli*. This strategy of directing expression towards IB formation has the additional advantages of ease of purification and reasonably high yields if refolding is not difficult [33, 34]. Thus, hGM-CSF, a comparatively small, single domain protein with only two disulfide bonds, was refolded by pulse dilution at a very high efficiency. The final purification and refolding yields obtained were 95% and 45%, respectively, which is comparable if not better than earlier reports [19, 35]. Recently, Schwanke and co-workers also reported purification and refolding of hGM-CSF from *E. coli*, where a protein concentration of 88.7 mg/L was reported; however, the homogeneous hGM-CSF yield was 0.4 mg/L of bioactive protein [19]. The purified and refolded protein in this study supported efficient TF1 cell proliferation, similar to standard hGM-CSF commercially available from R&D Systems.

It has been postulated by many authors [36–38] that the decline in growth associated with recombinant protein expression is due to the diversion of metabolites for product formation, instead of growth which leads to a metabolic burden on the host cells. Clearly by this argument, higher expression rates should lead to higher metabolic burden and hence faster decline in growth rates. Instead, we observed that low levels of soluble expression led to a steep fall in growth, whereas a much higher production rate as IBs did not significantly affect growth rates. Thus, rather than diversion of metabolites, it is the toxicity associated with the presence of a foreign protein inside the cell which plays a significant role in growth decline. This study of hGM-CSF expression is a clear demonstration that the nature and form of expression (soluble or IBs) are really the critical factors controlling growth and expression rates. Similar findings had also been reported earlier where soluble expression of toxic proteins led to poor cell growth and compromised expression levels [31].

The specific growth rate (μ) at the time of induction is a critical parameter for recombinant protein production [36, 37]. We observed in the shake flask experiments with TB that there was an optimum μ for maximum expression; both high and low μ values (early and late log phase induction) resulted in poor product yields. The TB medium is known to enhance recombinant protein yields possibly by enhancing the exogenous supply of amino acids and increasing specific growth rates in recombinant cultures [28]. The plasmid instability observed during soluble hGM-CSF expression was difficult to explain since such high rates of emergence of plasmid-free cells seem unlikely. One possibility could be the emergence of plasmid-free cells during growth on the antibiotic free plates, which we ruled out on methodological grounds [39]. Secondly, it has also been observed

that the cellular viability in terms of colony-forming ability is often lost by cells expressing and accumulating toxic proteins and thus these cells get under-represented upon plating, concomitantly leading to an over-representation of plasmid-free cell colonies on the plate [6, 40].

The bioprocess optimization using high cell density cultures of the pETA-GM clone is under progress, where the specific issues related to plasmid instability and growth in complex media need to be properly handled [41]. A high volumetric productivity is targeted by maintaining a high specific growth rate and hence a high specific product yield.

Acknowledgment The authors are thankful to Prof. R. C. Kuhad, Department of Microbiology, University of Delhi South Campus, for his advice during the preparation of manuscript.

References

1. Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., et al. (2002). *Science, 298*, 1790–1793.
2. Zhang, Z., Gildersleeve, J., Yang, Y. Y., Xu, R., Loo, J. A., Uryu, S., et al. (2004). *Science, 303*, 371–373.
3. Zoonens, M., & Miroux, B. (2010). *Method Molecular Biology, 601*, 49–66.
4. Makrides, S. C. (1996). *Microbiological Reviews, 60*, 512–538.
5. Hannig, G., & Makrides, S. C. (1998). *Trends in Biotechnology, 16*, 54–60.
6. Bhattacharya, P., Pandey, G., Srivastava, P., & Mukherjee, K. J. (2005). *Molecular Biotechnology, 30*, 103–116.
7. Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino terminal residue. *Science, 234*, 179–186.
8. Nygren, P. A., Stahl, S., & Uhlen, M. (1994). Engineering proteins to facilitate bioprocessing. *Trends in Biotechnology, 12*, 184–188.
9. Ni, Y., & Chen, R. (2009). *Biotechnological Letters, 31*, 1661–1670.
10. Yoon, S. H., Kim, S. K., & Kim, J. F. (2010). *Recent Patents on Biotechnology, 4*, 23–29.
11. Gaofu, Q., Jie, L., Rongyue, C., Xin, Y., Dan, M., Jie, W., et al. (2005). *Journal of Immunological Methods, 299*, 9–19.
12. Khushoo, A., Pal, Y., & Mukherjee, K. J. (2005). *Applied Microbiology and Biotechnology, 68*, 189–197.
13. Choi, J. H., Jeong, K. J., Kim, S. C., & Lee, S. Y. (2000). *Applied Microbiology and Biotechnology, 53*, 640–645.
14. Jeong, K. J., & Lee, S. Y. (2000). *Biotechnology and Bioengineering, 67*, 398–407.
15. Jeong, K. J., & Lee, S. Y. (2001). *Protein Expression and Purification, 23*, 311–318.
16. Metcalf, D. (1985). *Science, 229*, 16–22.
17. Armitage, J. O. (1998). *Blood, 92*, 4491–4508.
18. Ling, M., Zou, M., Xu, M., Wang, J., & Ma, X. (1995). *Chinese Journal of Biotechnology, 11*, 157–62.
19. Schwanke, R. C., Renard, G., Chies, J. M., Campos, M. M., Junior, E. L., Santos, D. S., et al. (2009). *International Journal of Biological Macromolecules, 45*, 97–102.
20. Berges, H., Joseph-Liauzun, E., & Fayet, O. (1996). *Applied and Environmental Microbiology, 62*, 55–60.
21. Srinivasa, B. K., Muthukumaran, T., Antony, A., Singh, P., Samuel, S. D., Balamurali, M., et al. (2009). *Biotechnological Letters, 31*, 659–664.
22. Lee, J. H., Kim, N. S., Kwon, T. H., Jang, Y. S., & Yang, M. S. (2002). *Journal of Biotechnology, 96*, 205–211.
23. Ryoo, Z. Y., Kim, M. O., Kim, K. E., Bahk, Y. Y., Lee, J. W., Park, S. H., et al. (2001). *Transgenic Research, 10*, 193–200.
24. Au, L. C., Liu, T. J., Shen, H. D., Choo, K. B., & Wang, S. Y. (1996). *Journal of Biotechnology, 51*, 107–113.
25. Tapryal, S., Khasa, Y. P., & Mukherjee, K. J. (2010). *Biotechnology Journal, 5*, 1078–1089.
26. Srivastava, P., & Mukherjee, K. J. (2001). *Preparative Biochemistry & Biotechnology, 31*, 389–400.
27. Kitamura, T., Tange, T., Terasawa, T., Chiba, S., Kuwaki, T., Miyagawa, K., et al. (1989). *Journal of Cellular Physiology, 140*, 323–334.

28. Khushoo, A., Pal, Y., Singh, B. N., & Mukherjee, K. J. (2004). *Protein Expression and Purification, 38*, 29–36.
29. Hua, Z., Jie, L., & Zhu, D. (1994). *Biochemistry and Molecular Biology International, 34*, 621–626.
30. Gnoth, S., Jenzsch, M., Simutis, R., & Lübbert, A. (2008). *Bioprocess and Biosystems Engineering, 31*, 41–46.
31. Saida, F., Uzan, M., Odaert, B., & Bontems, F. (2006). *Current Protein & Peptide Science, 7*, 47–56.
32. Panda, A. K. (2005). *Method Molecular Biology, 308*, 155–161.
33. Baneyx, F., & Mujacic, M. (2004). *Nature Biotechnology, 22*, 1399–1408.
34. Jungbauer, A., & Kaar, W. (2007). *Journal of Biotechnology, 128*, 587–596.
35. Belew, M., Zhou, Y., Wang, S., Nystrom, L. E., & Janson, J. C. (1994). *Journal of Chromatography A, 679*, 67–83.
36. Yazdani, S. S., & Mukherjee, K. J. (2002). *Bioprocess and Biosystems Engineering, 24*, 341–346.
37. Kayser, A., Weber, J., Hecht, V., & Rinas, U. (2005). *Microbiology, 151*, 693–706.
38. Kopetzki, E., Schumacher, G., & Buckel, P. (1989). *Molecular & General Genetics, 216*, 149–155.
39. Sorensen, S. J., Bailey, M., Hansen, L. H., Kroer, N., & Wuertz, S. (2005). *Nature Reviews. Microbiology, 3*, 700–710.
40. Vethanayagam, J. G., & Flower, A. M. (2005). *Microbial Cell Factories, 4*, 3.
41. Shiloach, J., & Fass, R. (2005). *Biotechnology Advances, 23*, 345–357.