

## *Basic Science in Medicine*

# IMPROVEMENT OF hGM-CSF EXPRESSION USING A FUSION SYSTEM BY DIRECTING PROTEIN TO THE PERIPLASMIC SPACE

MANA OLOOMI\* AND SAEID BOUZARI

*From the Molecular Biology Unit, Pasteur Institute of Iran, Tehran, I.R. Iran.*

### ABSTRACT

Bacterial expression systems provide economic and logistic advantages in manufacturing proteins for human therapeutic purposes. However, most such proteins accumulate in insoluble biologically inactive form when overexpressed in bacterial cells. This is the case while attempting to produce recombinant human granulocyte-macrophage colony stimulating factor (hGM-CSF) in *E. coli*. In this report, different strong promoters are used to compare the yield of expression when the protein is expressed as insoluble and soluble forms. The results obtained indicated that the level of expression was independent of the nature of promoter used. Moreover experimental data presented here suggest that the fusion system which directs hGM-CSF to the periplasmic space not only can prevent inclusion body formation but also improve the level of expression significantly.

*MJIRI, Vol. 18, No. 2, 141-146, 2004.*

**Keywords:** *Escherichia coli*, hGM-CSF, Expression, Inclusion body, Fusion protein.

### INTRODUCTION

Upon overexpression of recombinant proteins, inclusion bodies can be observed in several host systems, for example, prokaryotes, yeast or higher eukaryotes. Even endogenous proteins if overexpressed can accumulate as inclusion bodies,<sup>1</sup> suggesting that in most cases inclusion body formation is the consequence of high expression rate, regardless of the system or protein used. There is no direct correlation between the propensity of inclusion body formation of a certain protein and its intrinsic properties, such as molecular weight, hy-

drophobicity, folding pathways, and so on. Only in the case of disulfide bonded proteins inclusion body formation is anticipated if the protein is produced in the bacterial cytosol, as formation of disulfide bonds usually do not occur in this reducing cellular compartment.<sup>2</sup> It is not clear why high intracellular concentrations of nascent or newly synthesized recombinant protein tend to misfold and aggregate. The yield of correctly folded, active protein can often be increased by reduced expression levels, slower growth, alternative host strains, redesigning the recombinant construct and fusion to a more soluble protein to facilitate correct folding.<sup>3</sup>

Colony-stimulating factors are a group of molecules that can stimulate hemopoiesis *in vitro*. Human recombinant GM-CSF is one of these factors, which has been expressed in a number of host/vector systems. It was approved by the FDA (Food and Drug Administration) in 1991 under the generic name sargramostim and is indicated to accelerate myeloid recovery in selected patients

**Abbreviations:** hGM-CSF (Human Granulocyte-Macrophage colony stimulating factor), FDA (Food and Drug Administration), IPTG (Isopropyl- $\beta$ -D-thiogalactopyranoside), IAA (Indolyl-3-acrylic acid).

\*To whom correspondence should be addressed at Pasteur Institute of Iran, Pasteur Ave, Tehran, Iran. 13164.

Fax: (9821) 6465132, E-mail: Oloomi@institute.pasteur.ac.ir.

## Improving hGM-CSF Expression

whom are undergoing autologous bone marrow transplant.<sup>4</sup> Glycosylated GM-CSF was expressed in yeast<sup>5</sup> and mammalian cells.<sup>6</sup> Unglycosylated GM-CSF is expressed cytoplasmically in *E. coli* at a high level although in insoluble form.<sup>7</sup> GM-CSF was expressed from a high-level secretion vector (pIN-III-ompA3) in which GM-CSF exports to the periplasmic space. However this treatment did not result in solubilization. This result provided evidence for the high degree of insolubility of human GM-CSF, which accumulated in the periplasmic space.<sup>8-9</sup>

The periplasm affords ease and cost-effectiveness of purification of the target protein from a significantly smaller pool of bacterial proteins compared with the cytoplasm. In addition, the oxidizing environment of the periplasm facilitates proper protein folding, and the *in vivo* cleaving of the signal peptide during translocation to the periplasm is more likely to yield the authentic N-terminus of the target protein. It was shown that certain fusion partners could greatly improve the solubility of recombinant proteins that would otherwise accumulate within inclusion bodies in the cell cytoplasm. It should be notified that all fusion partners are not equally proficient in alleviating inclusion body formation.<sup>10</sup> Much of protein refolding research is focused on trying to minimize aggregation reactions while keeping the yield and concentration of the protein folding reaction as high as possible.

Since in our previous attempt GM-CSF was obtained as inclusion body,<sup>11-12</sup> therefore in the present study, the efficiency of different promoters to express human GM-CSF, as native or fusion form in the cytoplasm and periplasm was employed to improve the yield of protein expression and the important findings are discussed.

## MATERIAL AND METHODS

### Bacterial strains and plasmids

pET-23a (+) (Novagen, Madison, WI, USA), pACT7 (kindly provided by Dr. Richinsky), pMALp2, and pMALc2 (New England BioLabs), and pATH11 (Kindly provided by Dr. Pellegrini) were used. The constructs

were transferred to competent cells. Bacterial strains JM105 (thi-1 rpsL end A sbcBC hsdR4  $\Delta$ (lac-proAB) [F' traD36 proAB lacIqZ $\Delta$ M15]), BL21 (DE3)pLysS (F' ompT, hsdSB9rB<sup>+</sup>mB<sup>+</sup>) dcm gal(DE3)pLys Cm<sup>r</sup> and DH5 $\alpha$  (sup E44  $\Delta$ lacU169 ( $\Phi$ 80lacZ $\Delta$ M15) hsdR17recA1 gyrA96 thi-1 relA1 ) were used as competent cells for transformation.

### Construction of hGM-CSF gene in different vectors

HGM-CSF gene in pET23a was cloned in our previous study, hence for any further construction the source of GM-CSF gene was from this construct.<sup>11-12</sup> Then hGM-CSF gene was cloned in these commercial vectors. (Table 1)

### Culture conditions and induction

The expression was induced by isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) except for pACT7 that was induced by heat shock and for pATH11 that was induced by Indoly1-3-acrylic acid (IAA) (Sigma Chemical Co.). LB medium was used to inoculate cultures for analysis. Cultures were grown at 37°C with vigorous shaking except for heat induction that was performed at 30°C. Cultures were induced at A 600=0.2-0.5 by addition of IPTG (0.5 mM final concentration), IAA (1:1000 dilution 20mg/mL) or growth at 42°C (for heat inducible vector) then incubation continued for an additional 4h. Cells were harvested by centrifugation, the pellet was resuspended in 1 mL lysis buffer (Tris-HCl 50mM, NaCl 100mM, EDTA 2mM), and disrupted by sonication. The soluble and insoluble fractions were separated by centrifugation and insoluble material was solubilized by addition of 8M urea to lysis buffer.

### Detection of recombinant protein

Expressed protein was analysed on 15% SDS-PAGE as described by Laemmli.<sup>13</sup> Total protein concentration was determined by the method of Lowry et al.<sup>14</sup> Immunoblotting was used for detection of expressed protein employing polyclonal antibody against GM-CSF (R&D systems) and the expressed protein was detected by Hyperfilm ECL (Amersham Pharmacia Biotech).

**Table I.** Plasmids and bacterial strains used as expression vector.

Vector	Promoter	Inducer	Protein expression	Bacterial strains used for transformation	GM-CSF cloning sites
pET	T7	IPTG	Intact	BL21(DE3)pLysS	NdeI—HindIII
pMAL-p2	Ptac	IPTG	Periplasmic space	JM105	BamHI—HindIII
pMAL-c2	Ptac	IPTG	Cytoplasmic	JM105	BamHI—HindIII
pATH11	TrpE	IAA	Secretion	DH5 $\alpha$	EcoRI—BamHI
PACT7	PR	Heat	Cytoplasmic	JM105	BamHI—SphI

Enzyme linked immunosorbent assay was used to measure the expression level (BIOTRAK kit Amersham Pharmacia Biotech). In this assay quantitative sandwich enzyme immunoassay technique was employed in order to measure the amount of recombinant product using standard curve prepared for diluted hGM-CSF by measuring absorbency at 450 nm.

### Protein purification

After induction, in pET-GM construct, the bacterial cell pellet was resuspended in a lysis buffer and disrupted by sonication. Osmotic lysis of bacteria was used to isolate the periplasmic expressed protein.<sup>15</sup> High urea (6M) was used for solubilization of the produced protein. The denatured protein was refolded by dialysis in serial decreasing dilution of urea until complete changing of the buffer. At first 4M urea, then 2M and finally tris buffer without urea was used for gradual refolding of the protein. Expressed fusion protein by pMAL-p2 system was excised by Factor Xa as described by the manufacturer.

The expressed protein was purified by immunoaffinity chromatography, using polyclonal neutralizing anti-human GM-CSF antibody (R&D systems) coupled to CNBr-activated Sepharose 4B (Pharmacia), according to the manufacturer's protocol (The activated form of sepharose is suspended in HCl to swell, the resulting gel was immediately transferred to a solution of antibody. After blocking and washing steps the antibody-agarose conjugate was ready for use). The fractions were eluted with 0.2M glycine (pH=3) and analysed again by ELISA.

### Biological assay

Biological activity was assessed by cell proliferation ELISA kit as described in the manufacturer's protocol (Roche Molecular Biochemicals). Colorimetric immunoassay for the quantification of cell proliferation, based on the measurement of BrdU incorporation during DNA synthesis was used which is a non-radioactive alternative to the [<sup>3</sup>H]-thymidine incorporation assay. The KG-1 cell line was used for this assay.

## RESULTS AND DISCUSSION

In our previous attempts hGM-CSF was expressed<sup>11</sup> and purified to homogeneity,<sup>12</sup> but one of the shortcomings of previous experiments was inclusion body formation, which has its own advantages and disadvantages. The advantage of inclusion body formation is the accumulation of the recombinant protein at more than 50% of the total cell protein. The inclusion bodies can be separated from the cell lysate by centrifugation as they are very dense particles, which pellet more rapidly than the rest of the cell debris and soluble components. However, inclusion bodies are insoluble aggregates of the

recombinant protein, which are held together by noncovalent forces. Solubilization involves disruption of the noncovalent interactions and the use of reducing conditions to break the disulfide bonds.<sup>16</sup>

In the present investigation the hGM-CSF gene was cloned and expressed in different systems to improve the expression of this cytokine. The schematic presentations of the constructs are shown in Fig. 1.

A promoter can have fundamental impact on the quality and yield of a recombinant protein.<sup>17</sup> Those employed in both research laboratories and industry to drive heterologous gene expression include the promoters from the lac and the tryptophan (trp) biosynthetic operon, the tac promoter, as well as phage promoters such as the  $\lambda$ pL, and T7 promoter from bacteriophage. All of these have been used to produce large numbers of recombinant proteins that attain levels of at least 5%-10% of the total cell protein. The induction ratio for these established expression systems is reported to be 50-80 fold for P<sub>trp</sub>, 1000 fold for P<sub>tac</sub>, >300 fold for  $\lambda$ pL and >1000-fold for T7.<sup>17-18</sup> The hGM-CSF gene was cloned and expressed in these different vectors and expression level was measured by ELISA method (Fig. 2).

The results obtained indicated that in the crude extracts of bacteria containing these different constructs, the highest expression was obtained in pMAL-p2, followed by pET then pATH11, pMAL-c2, and pACT7 system. However after sonication, the supernatant was measured for expressed protein and pMAL-p2 showed highest expression, followed by pMAL-c2 and in pATH11, pACT7 and pET vector, no expression was detected. Moreover in the pellet of these constructs highest level of expression was noticed in pET, then pATH11, pMAL-p2 followed by pMAL-c2 and pACT7. Western blotting was performed to analyse the expression of the protein (Fig. 3).

The overall results revealed that for pET the pellet was the fraction that contained the expressed protein and in case of pMAL-p2 the supernatant showed the highest amount of it. The tac promoter, in pMAL system is quite strong and routinely allows the accumulation of polypeptides to about 15-30% of the total cell protein.<sup>10</sup> The results of ELISA and immunoblotting showed the nature of protein expressed in these two systems. Since it is well known that the popular pET vector with T7 promoter causes accumulation of the desired protein at very high concentrations (40-50% of the total cell protein), but it is not without drawbacks.

There is the possibility of leaky expression, plasmid or expression instability. Furthermore, even pET plasmids are toxic to *E. coli* in the presence of IPTG.<sup>19</sup> As previously mentioned an additional limitation of T7 and other strong promoters like bacteriophage  $\lambda$ pL promoter, is that the target protein is often unable to obtain a native conformation and either partially or completely

# Improving hGM-CSF Expression

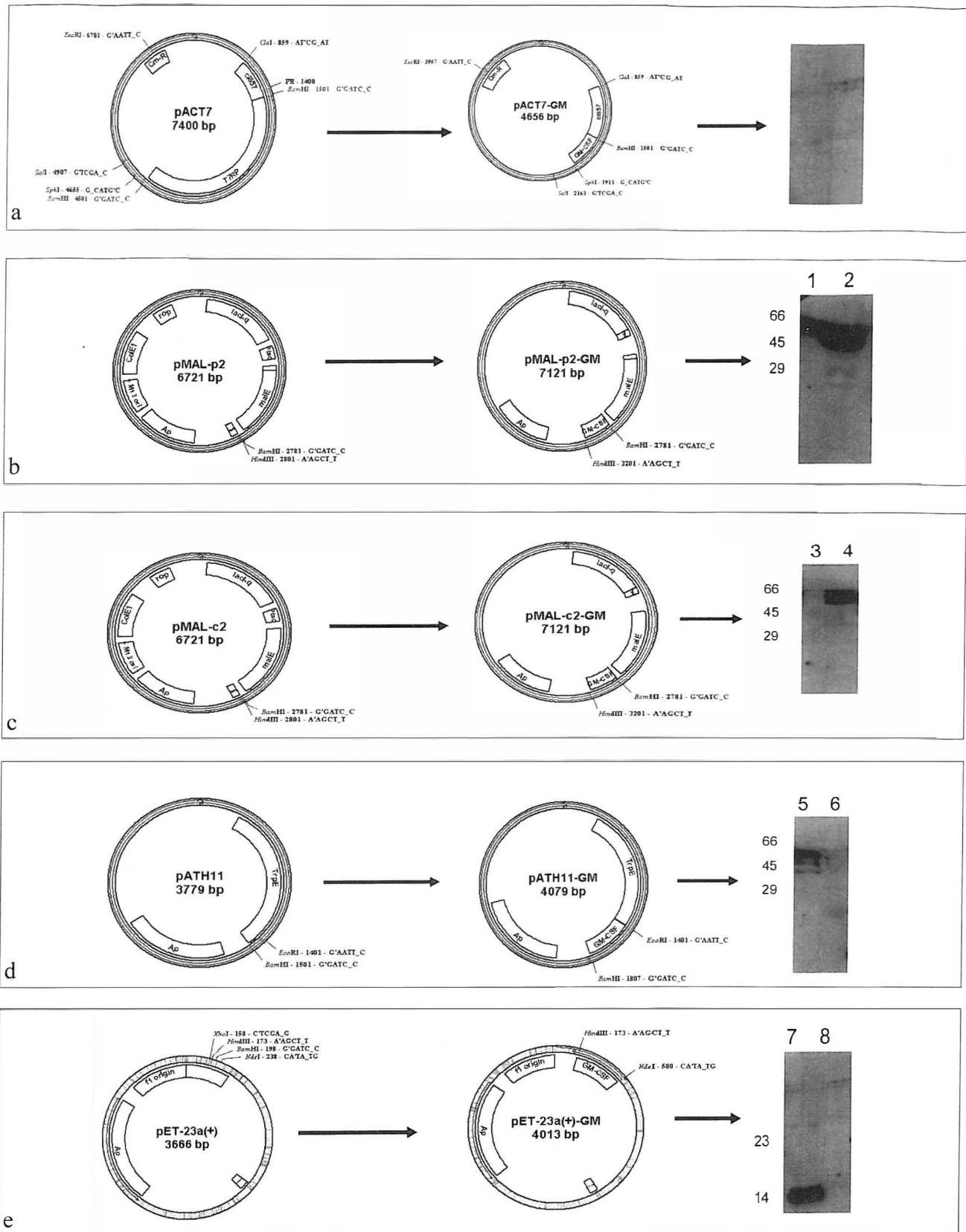


Fig. 1. Schematics of constructs used for the protein expression.

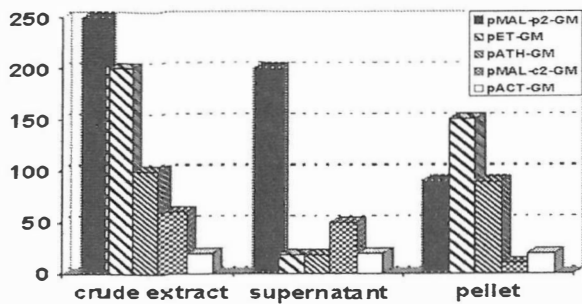


Fig. 2. Expression level before and after sonication in supernatant and pellet of cell lysate measured by ELISA.

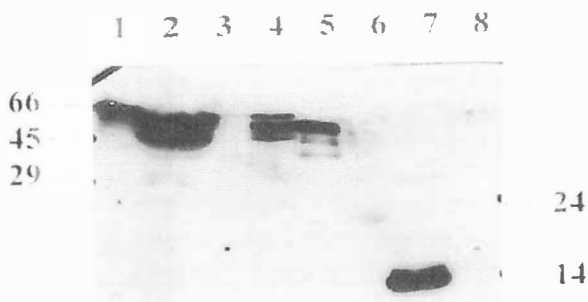


Fig. 3. Western Blot analysis. Lanes 1 to 8, pellet and supernate fractions of pMAL-p2, pMAL-c2, pATH11, and pET23a respectively.

segregates within inclusion bodies. This problem can be addressed by using fusion proteins or promoters that are activated by temperature downshift.<sup>10</sup>

The expressed protein in pMAL-p2-GM (sup) and in pET23a-GM (pellet) was purified by affinity chromatography. The yield of purified protein was found to be 5% of total cell protein for pMAL-p2 and 4% for pET and the expressed protein was found to be biologically active.

In our experiment, the expression level in pMAL-p2 was the highest and it was not only related to the promoter (because expression with the same promoter in pMAL-c2 has not shown the same expression level). This data suggests that the expression level is rather dependent on the nature of protein produced (soluble, insoluble, cytoplasmic, periplasmic) than the type of promoter used.

It was shown that gene fusion expression systems have been used to enhance protein solubility within recombinant hosts. The widely used fusion systems include maltose-binding protein (MBP), glutathione S-transferase (GST), Ubiquitin, and thioredoxin (TrxA). These have been shown to facilitate the soluble expression of several cytokines, peptides and a diverse range

of other proteins.<sup>20</sup> Fusions have been useful for increasing protein solubility in many occasions.<sup>21</sup>

In the present study two fusion systems were used, Trp E and maltose-binding protein. Trp E in pATH11 didn't affect the solubility of the expressed protein and the expressed protein was in insoluble form. Expression of protein with maltose-binding protein in pMAL-c2 makes the protein soluble but the expression was low, however in case of pMAL-p2 high expression was obtained, despite the fact that the two vectors use the same promoter i.e. tac (Fig. 1). In this system a signal sequence directs protein to the periplasmic space that can lead to overexpression and solubilization of the protein. The results showed that periplasmic expression or exporting the expressed protein to outer space could be a way to increase the level of expression in case of hGM-CSF. Unlike cytoplasm, the periplasmic space is an oxidizing compartment where disulfide bond formation can occur rapidly.<sup>22</sup> However, it has been shown that inclusion body formation can also occur in the periplasm.<sup>23</sup> The reason for the high level of expression in case of soluble fusion maltose binding protein (MBP) when it is exported to periplasmic space rather than pMAL-c2, could be due to localization of protein in the latter case in the cytoplasm. Therefore the possibility of degradation with proteolytic enzymes may exist. However in case of pMAL-p2 since protein is transported to the periplasmic space there is no accumulation and no sense of stress for bacteria, therefore the result is better expression and higher expression of hGM-CSF can be achieved by this system.

Secretion into the periplasm allows one to separate the gene product from proteases and modification enzymes in the cytosol to avoid potentially toxic effects of the product in the cytosol and to yield efficient disulfide bond formation in the oxidizing environment of the periplasm.

Other factors that can affect expression rather than secretion directly, such as correct protein folding aided by molecular chaperones, mRNA stability, codon usage, and cytosolic proteases should also be taken into account. Thus the full potential of protein secretion and recovery may be attained by using not only the efficient expression and secretion system described in this article, but also additional factors could be considered.

## REFERENCES

1. Gribskov M, Burgess RR: Overexpression and purification of the sigma subunit of *Escherichia coli* RNA polymerase. *Gene* 26: 109-118, 1983.
2. Lilie H, Schwarz E, Rudolph R: Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology* 9: 497-501, 1998.
3. Kane JF, Hartley DL: Formation of recombinant protein

## Improving hGM-CSF Expression

- inclusion bodies in *Escherichia coli*. Trends Biotech 6: 95-101, 1998.
- Pezzuto JM, Jognson M, E. Manasse HR: Biotechnology and Pharmacy, Chapman & Hall Inc. Ch.3, 1993.
  - Cantrell MA, Anderson D, Cerretti DP, et al: Cloning, sequence and expression of a human granulocyte/ macrophage colony stimulating factor. Proc Natl Acad Sci USA 82: 6250-6254, 1985.
  - Lee F, Yokota T, Otsuka T, et al: Isolation of cDNA for a human granulocyte-macrophage colony stimulating factor by functional expression in mammalian cells. Proc Natl Acad Sci USA 82: 4360-4364, 1985.
  - Schrimsher JL, Rose K, Simona MG, Wingfield P: Characterization of human and mouse granulocyte-macrophage colony stimulating factors derived from *E.coli*. Biochem J 247: 195-199, 1987.
  - Libby RT, Braedt G, Kronheim SR, et al: Expression and purification of native human granulocyte-macrophage colony stimulating factor from an *Escherichia coli* secretion vector. DNA 6: 221-229, 1987.
  - Greenberg R, Lundell D, Alory Y, et al: Expression of biologically active, mature human granulocyte/ macrophage colony stimulating factor using an *E.coli* secretory expression system. Curr Microbiol 17: 321-332, 1988.
  - Baneyx F: Recombinant protein expression in *Escherichia coli*. Curr Opin Biotechnol 10: 411-421, 1999.
  - Oloomi M, Bouzari S, Rechinsky V: PCR-mediated expression of the human GM-CSF gene in *Escherichia coli*. Iranian Biomedical J 2 (1): 21-25, 1998.
  - Oloomi M, Bouzari S, Rechinsky V: Purification and characterization of the cloned human GM-CSF gene expressed in *Escherichia coli*. MJIRI 12 (4): 353-357, 1999.
  - Laemmli UK: Cleavage of structural proteins during the assembly of the head of Bacteriophage T4. Nature 227: 680-685, 1970.
  - Lowry OH, Rosebrough NJ, Farr AL, Randall RJ: Protein measurement with the folin phenol reagent. J Biol Chem 193: 265-275, 1951.
  - Neu HC, Heppel LA: The release of enzymes from *Escherichia coli* by osmotic shock during the formation of spheroplasts. J Biol Chem 240: 3685-3692, 1965.
  - Guise AD, West SM, Chaudhuri JB: Protein folding *in vivo* and renaturation of recombinant proteins from inclusion bodies. Mol Biotech 6: 53-64, 1996.
  - Sawers G, Jarsch M: Alternative regulation principles for the production of recombinant proteins in *Escherichia coli*. Appl Microbiol Biotechnol 46: 1-9, 1996.
  - Kapust RB, Waugh DS: *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci 8: 1668-1674, 1999.
  - Miroux B, Walker JE: Over-production of protein in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260: 289-298, 1996.
  - LaVallie ER, Diblasio EA, Kovacic S, et al: A thioredoxin gene fusion expression that circumvents inclusion body formation in the *E. coli* cytoplasm. Biotechnology 11: 187-193, 1993.
  - Samuelsson E, Moks T, Nilsson B, Uhlen M: Enhanced *in vitro* refolding of insulin-like growth factor I using a solubilizing fusion partner. Biochemistry 33: 4207-4211, 1994.
  - Ostermeier M, Georgiou G: The folding of bovine pancreatic trypsin inhibitor in the *Escherichia coli* periplasm. J Biol Chem 269: 21072-21077, 1994.
  - Georgiou G, Telford JN, Shuler ML, Wilson DB: Localization of inclusion bodies in *Escherichia coli* overproducing  $\beta$ -lactamase or alkaline phosphatase. Appl Environ Microbiol 52: 1157-1161, 1986.