Short Communication

Site-directed mutagenesis in human granulocyte-colony stimulating factor, cloning and expression in *Escherichia coli*

Hamed Naghoosi[1], Farida Behzadian[*2], Alireza Saeedinia[2], Seyed Ali Ghorashi[3]

[1]Department of Microbiology, Faculty of Science, Islamic Azad University, Karaj Branch, P.O. Box 31485-313, Karaj, I.R. Iran [2]Department of Genetic Engineering, Research Center for Science and Biotechnology, P.O. Box 19395-1949, Tehran, I.R. Iran [3]Department of Animal Biotechnology, National Institute of Genetic Engineering and Biotechnology, P.O. Box 14965/161, Tehran, I.R. Iran

Abstract
Human granulocyte colony stimulating factor (hG-CSF) induces proliferation and differentiation of granulocyte progenitor cells. This glycoprotein is currently being used for treatment of neutropenia, in patients who have undergone bone marrow transplantation. So far, different researchers have tried to enhance hG-CSF biological activity and stability. In this study, Polymerase Chain Reaction (PCR) based site-directed mutagenesis was performed on hG-CSF cDNA. The final amplified DNA fragment was cloned into the pBluescript SK(-) plasmid and after verification of the desired mutations by sequencing, it was subcloned into the pET-21a(+) vector and expressed in *Escherichia coli* BL21. The mutant G-CSF product was analyzed by SDS-PAGE and Western-blot analyses. The results show that the recombinant mutant G-CSF has been cloned and expressed successfully in prokaryotic system. This research aimed to produce a new recombinant hG-CSF expected to show enhanced biological characteristics in contrast to those of the native hG-CSF. The analysis of its function and biological characteristics remain to be examined.
Keywords: Granulocyte colony stimulating factor (G-CSF); Site-directed mutagenesis; expression

Correspondence to:* **Farida Behzadian, Ph.D.
Tel: +98 21 22945141-6 Ext. 2588; Fax: +98 21 22959491
E-mail: Fbehzadian@yahoo.com

The hematopoietic system is a complex structure with multipotent stem cells which produce mature hematopoietic cells that circulate in peripheral blood. Hematopoiesis is a highly regulated process and all of its stages operate under the control of specific factors called cytokines. Majority of the cytokines are glycoproteins that regulate survival, proliferation and differentiation of hematopoietic progenitor cells and also affect the function of mature blood cells. Colony stimulating factors (CSFs) are a family of cytokines which consist of five types of glycoproteins consisting of M (macrophage)-CSF, G (granulocyte)-CSF, E (eosinophil)-CSF, GM (granulocyte macrophage)-CSF and multi-CSF or Il-3 (Creighton, 1999).

Granulocyte colony stimulating factor (G-CSF) is a major member of this family that is produced by bone marrow stromal cells, endothelial cells, macrophages and fibroblasts. This factor induces proliferation and differentiation of neutrophil progenitor cells as well as activation of mature granulocytes for more efficient immune responses. hG-CSF is a glycoprotein consisting of 174 amino acid residues (hG-CSFb, 18.8 KD) or 177 amino acid residues (hG-CSFa, 19.6 KD) with one O-linked glycosyl group on Threonine (Thr) at position 133. These two different G-CSFs are encoded by two different mRNAs produced by alternative splicing from a single precursor RNA. Even though these different forms have similar biological activities, hG-CSFb is 20 times more active than hG-CSFa (Nagata

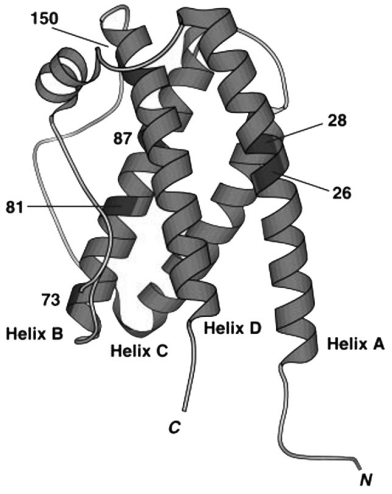

Figure 1. Ribbon representation of the structure of hG-CSF (from Bishop *et al.*, 2002, with permission from corresponding author). The numbers mentioned in figure are just for understanding the position of amino acids in the 3D structure, and they are not all related to this work.

1989; Nagata *et al*., 1986 a, b).

The molecular structure of this cytokine is an up-up-down-down, antiparallel, left-handed four α-helical bundle without any β-sheets (Fig. 1). There are two disulfide bonds residing between Cysteine (Cys) 36/42 and Cys 64/74 of this molecule (Werner *et al*., 1994; Hill *et al*., 1993).

The two forms of hG-CSF cDNA, hG-CSFa and hG-CSFb, have already been cloned by Nagata *et al*. (1986 a) and Souza *et al*. (1986), respectively. Currently, hG-CSFb is synthesized by biotechnological methods and used for treatment of neutropenia arising from chemotherapy and radiotherapy and in patients who have undergone bone marrow transplantation (Fernández-Varón and Villamayor, 2007; Klingemann, 1989). So far, several researchers have attempted to improve G-CSF biological activity, stability and shelf life by using different mutagenesis techniques (Luo *et al*., 2002; Bishop *et al*., 2001; Lu *et al*., 1999; Ishikawa *et al*., 1993 and 1992; Devlin *et al*., 1988).

The aim of this study was to design and produce an engineered recombinant G-CSF, using site-directed mutagenesis, in order to improve its biological characteristics in comparison to those of native hG-CSF.

PCR based site-directed mutagenesis by the overlap extension method (Sambrook and Russel, 2001) was performed in three stages. Sequence, polarity and position of the primer pairs as well as restriction enzyme cutting sites are presented in Table 1. The primers were designed according to the codon usage frequency of *E. coli* by using the GeneRunner software version 3.05 (Hasting software, Inc.) *Nde* I and *Xho* I specific sites and additional nucleotides on either sides (on 3′ end of F and 5′ end of R primer) were created in the F and R designated primers, respectively. The His hexamer (His_6-tag) sequence as well as the ATTA stop sequence was designed at the R primer 5′ overhang, all conforming to the cloning and expression strategies of this study.

PCR reactions were carried out in a total reaction mixture volume of 25 µl containing *Pfu* buffer 1X (20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 0.1% (v/v) Triton X-100, 0.1 mg/ml of BSA, 2 mM $MgSO_4$) 0.2 mM of each dNTP (Cinnagen, Iran), 10 pmol of each forward and reverse primers, 50-150 ng of template DNA and 1 U of *Pfu* DNA polymerase (Fermentas, Lithuania).

In the first and second PCR stages, a clone containing native hG-CSF cDNA (Saeedinia *et al.*, 2003) was used as template. The first round was performed using F and R1 primers as shown schematically in Figure 2 under the following conditions: initial denaturation at

Table 1. Sequence, position and polarity of the primers used in PCR mutagenesis and amino acid substitutions directed by them. Underlined: restriction sites; double underlined: codon changes in contrast to the wild type sequence. Thr: threonine, Ala: alanine, Leu: leucine, Gly: glycine, Tyr: tyrosine, Pro: proline, Arg: arginine, His: histidine, Cys: cysteine.

Primer designation	Sequence (5' to 3')	Position & Polarity	Amino acid substitutions
F	AACATATGGCGCCCACCTACCGTGCTAGCTCCCTGCCCCAGAG	-8-35 Sense	Thr2/Ala Leu4/Thr Gly4/Tyr Pro5/Arg
R	ACTCGAGATTAATGATGATGATGATGATGGGGCTGGGCAAGGTGGCGTAG	502-552 Antisense	His6-tag addition
F1	GCGTTAGAGCAAGTGAGGAAGATCCAGGGTGATGCAGCAGCGCTCCAG	48-96 Sense	Cys17/Ala Gly28/Ala
R1	TGCATCGCCCTGGATCTTACTCACTTGGTCTAACGCCTTGAGCAGGAAGC	32-82 Antisense	Cys17/Ala Gly28/Ala

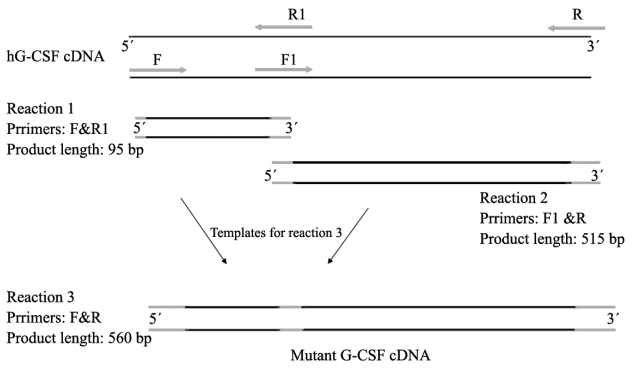

Figure 2. Schematic diagram of PCR based site-directed mutagenesis in 3 stages. shaded arrows: primers and their orientation and position on the template (hG-CSF cDNA).

94ºC for 3 min followed by 35 cycles of denaturation at 94ºC for 45 sec, primer annealing at 60ºC for 45 sec and extension at 72ºC for 1 min followed by a final extension step at 72ºC for 8 min. During the second PCR round, F1 and R primers were used in the above-mentioned PCR mixture and thermocycle profile. In the third PCR stage, the amplicons obtained through the first and second rounds were diluted by 1/100, and used together as template. F and R primers were used to produce and amplify the full length mutant G-CSF (muG-CSF) cDNA using the following thermal cycle profile: initial denaturation at 94ºC for 3 min followed by 35 cycles of denaturation at 94ºC for 45 sec, primer annealing at 60ºC for 45 sec and extension at 72ºC for 2.5 min followed by a final extension step at 72ºC for 10 min.

The PCR product, mutant G-CSF (muG-CSF) amplicon, was gel purified and cloned into the *Eco*R V linearized pBluescript SK(-) plasmid. Restriction analysis of this intermediate plasmid was carried out to confirm cloning fidelity. In order to confirm the desired mutations, the recombinant vector was sequenced in a bidirectional manner, using T_3 and T_7 universal primers.

The muG-CSF fragment was exposed to *Nde* I/*Xho* I double digestion and subcloned into the pET-21a(+) vector (Novagen, USA) that was linearized with the same restriction endonucleases. All DNA manipulations including restriction digestion, T_4 ligation and agarose gel electrophoresis techniques were carried out as described by Sambrook and Russel (2001).

The recombinant construct was used to transform *E. coli* BL21 (DE3) competent cells. After selection and verification of recombinant colonies, cells harboring pET21-muG-CSF were cultivated in LB broth medium (Scharlau, Spain). The expression of muG-CSF was induced by addition of Isopropyl β-D-1-thio-galactopyranoside (IPTG) at a final concentration of 1 mM when the cells had reached an optical density of 0.7 at 600 nm (OD_{600}). The cells were then incubated for further 5 h. Bacterial cells were harvested at 0, 0.5, 1, 2, 3, 4 and 5 h after induction. To get the total cell protein extract, cells from 1 ml culture samples were separated by centrifugation at 6000 ×*g* for 10 min at 4°C, and then lysed directly by adding 100 µl of sample buffer (80 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 10% (v/v) glycerol, 0.001% (w/v) bromophenol blue and boiled for 10 min. 25 µl of the prepared protein samples were electrophoresed on a 12.5% (w/v) SDS-polyacrylamide gel. Commercial recombinant hG-CSF (Filgrastim-neupogen, Roche, Germany) was used as a positive control. Electrophoresis was carried out at 20 mA, for 2.5 h.

Protein bands on the SDS-polyacrylamide gel were transferred to a polyvinylidene difluorid (PVDF) filter. The filter was blocked with 3% (w/v) bovine serum albumin (BSA) (Merck, Germany) and incubated with 0.5 µg/ml of murine monoclonal anti-human G-CSF (Sigma, USA) for 1 h. After washing, it was incubated with 1/2000 dilution of anti-mouse IgG conjugated with Horse Radish Peroxidase (HRP) (Sigma, USA) as secondary antibody for 1h. 0.5 mg/ml 3,3-Diaminobenzidine reagent and 0.1% (v/v) H_2O_2 in tris buffer saline (TBS) were used as the HRP substrate for color development.

During the first round of PCR, a 95 bp DNA fragment encoding the first 30 N-terminus amino acids of the mutant G-CSF was amplified. In the second round, a 515 bp DNA fragment was amplified which covers nucleotides 44 to 560 of the full length mutant G-CSF cDNA, encoding amino acids 14 to 180 of muG-CSF (including His_6-tag). In the third reaction, these amplicons joined each other to produce a 560 bp full length muG-CSF cDNA (Fig. 3a). This final fragment was flanked by the *Nde* I and *Xho* I restriction sites at its 5′ and 3′ termini, respectively. The fragment was then cloned into the pET-21a(+) via the same restriction sites in the correct orientation and under the control of the T_7 promoter. Sequencing results confirmed the desired mutations in muG-CSF cDNA and cloning fidelity. Restriction analysis was also carried out to demonstrate the latter (Fig. 3b).

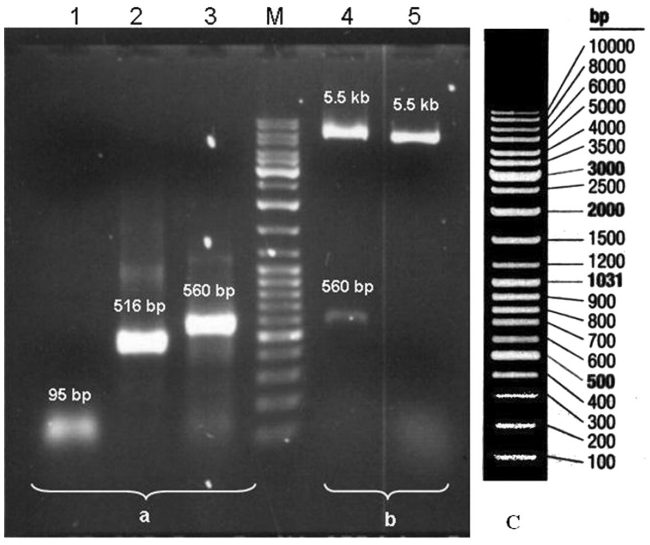

Figure 3. Agarose gel electrophoresis of PCR products and restriction analysis. Panel a: lanes 1-3 are PCR products obtained from stages 1-3, respectively. Panel b: Nde I/Xho I double digestion of pET21-muG-CSF (lane 4) and noncloned pET-21a (as a control), the pattern shows cloning fidelity. M) Molecular weight marker: Generuler DNA ladder mix #SM0333 (Fermentas, Lithuania). Panel C: The map of DNA ladder.

As demonstrated in Figure 4, the expressed muG-CSF with a size of approximately 19 kD was detected in total cell protein samples obtained from induced cells harboring pET21-muG-CSF. The intensity of the target band gradually grown increased over times, during which most of the recombinant muG-CSF was

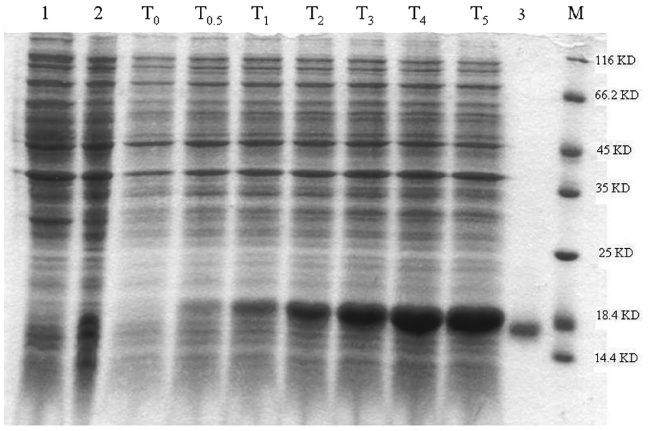

Figure 4. Time course SDS-PAGE analysis of muG-CSF expression. 25 µl of total cell protein (TCP) samples: Lane 1: *E. coli* BL21 (DE3) transformed with pET-21a vector. Lane 2: Nontransformed *E. coli* BL21 (DE3). Lane T0: TCP samples from cells harboring pET21-muG-CSF at the time of induction Lanes $T_{0.5}$ to T_5: TCP samples from cells harboring pET21-muG-CSF 0.5 to 5 h after induction, respectively. Lane 3: Commercial hG-CSF (filgrastim-neupogen, Roche, Germany) applied as positive control. Lane M: Protein molecular weight marker (#SM0431, Fermentas, Lithuania).

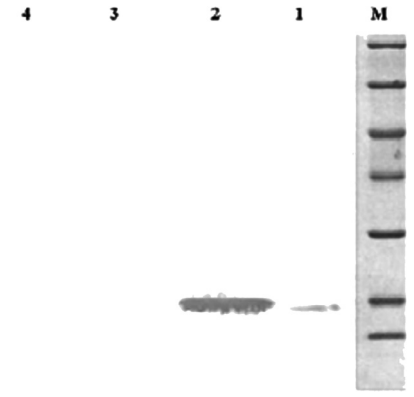

Figure 5. Detection of recombinant muG-CSF by Western blot analysis using murine monoclonal anti-hG-CSF. Lane 1: commercial hG-CSF (filgrastim-neupogen) used as a positive control. Lane 2: recombinant muG-CSF produced 4 h after induction of cells harboring pET21-muG-CSF. Lane 3: sample of transformed cells before induction. Lane 4: sample of non-transformed cells.

observed 5 h after induction.

Western-blot analysis using monoclonal antibody against hG-CSF proved the specificity of the detected band in SDS-PAGE. Figure 5 illustrates the result of Western-blot analysis carried out with monoclonal antibody.

Mutational analysis studies of hG-CSF have revealed that internal and C-terminal regions of this protein play an essential role in interaction with its cell surface receptor and manipulation of these evolutionally-conserved sequences mostly results in significant loss of biological activity (Young *et al*., 1997). Conversely, mutagenesis studies on N-terminal regions have shown a delicate relationship between the non-essential N-terminal structure and the G-CSF biological activity via a mechanism which dose not directly change the backbone of the molecular structure. In this regard, Kuga *et al*. (1989) and Okabe *at al*. (1990) have shown that substitution of Thr 1, leucine (Leu) 3, glysine (Gly) 4, proline (Pro) 5 and Cys 17 with alanine (Ala), Thr, tyrosine (Tyr), arginine (Arg) and serine (Ser) will result in 2-4 times more biological activity than that of the wild type protein. Cys 17 is not involved in structural disulfide bonds, and its substitution with Ser or Ala does not have any impact on its biological activity. Therefore many researchers have considered these substitutions in amino acid sequence to avoid the formation of unwanted disulfide bonds through protein folding process. Moreover this alteration results in a more thermodynamically stable G-CSF structure. (Yamasaki *et al*., 1998; Reidhaar-Olson *et al*., 1996; Ishikawa *et al*., 1993; Lu *et al*., 1992;

Wingfield *et al*., 1988).

Other mutagenesis studies have proved that substitution of each Gly residue at positions 26, 28, 149 and 150 singly or in combination, with Ala will result in proteins with dramatically enhanced stability while retaining wild type levels of biological activity (Bishop *et al*., 2001). Cys 17 and Gly 28 substitution with Ala, has also led to a 5-fold improvement in G-CSF storage stability and shelf life (Luo *et al*., 2002).

Imaginary structural analysis using the Swiss-Pdb Viewer (v3.7) software of the crystal structure of hG-CSF (protein data bank (PDB) record 1GNC) (Zink *et al*., 1994), demonstrated that all the abovementioned substitutions on the same oligopeptide do not affect the 3-dimensional molecular structure, which may still be left functional.

Here we combined these favorable mutations in the same coding sequence of G-CSF by performing a two step site-directed mutagenesis (Table 1). The main objective was a combination of the results of abovementioned studies to create a new mutant G-CSF, developed both physically and biologically, with enhanced efficacy as a recombinant drug. In summary, the pET system, a well known prokaryotic expression vector, was used to express the muG-CSF recombinant protein. The cloning strategies employed in this study were: i) Using the ATG triplets within the *Nde* I recognition site to encode the N-terminal Met base-paired to the AUG start codon in the transcript. This restriction site was engineered into flanking reverse primer (Sambrook and Russel, 2001). ii) To avoid adding two unwanted amino acid residues at the C-terminus of muG-CSF, the pET carrying His_6-tag sequence was not used, instead, the six His triplet codons followed by a stop sequence was engineered into the flanking reverse primer (Sambrook and Russel, 2001). Such a fusion facilitates detection of the expressed muG-CSF and will be very useful in the purification procedures of later experiments. It has also been shown that conjugating the unrelated small oligopeptides to the C-terminus of hG-CSF does not have any effect upon its biological activity (Oshima *et al*., 2000).

With respect to SDS-PAGE and western-blot analysis, the bands corresponding to muG-CSF were observed at a level slightly higher than that of the commercial G-CSF (Figs. 4 and 5) because the addition of the His hexamer to the polypeptide increases molecular weight by approximately 0.7 KD.

In this research, a new formulation of mutant G-CSF has been designed and expressed in prokaryotic system that would be an improved version of native hG-CSF and an appropriate candidate for pharmaceutical purposes after verifying its function, biological activity and safety.

Acknowledgments

The authors would like to thank Mohamad R. Masoumian, Hesam Barjesteh, Hossein A. Sami, Kaveh Baghai and Mehdi Zeinoddini for their excellent technical assistance.

References

Bishop B, Koay DC, Sartorelli AC, Regan L (2001). Reengineering granulocyte colony-stimulating factor for enhanced stability. *J Biol Chem*. 276: 33465-33470.

Creighton Thomas E (1999). *Encyclopedia of Molecular Biology*, John Wily & Sons Inc. New York, USA.

Devlin Patricia E, Drummond Robert J, Toy Pam, Mark David F, Watt Kenneth WK, Devlin James J (1988). Alteration of amino-terminal codons of human granulocyte-colony-stimulating factor increases expression levels and allows efficient processing by methionine aminopeptidase in *Escherichia coli*. *Gene* 65: 13-22.

Fernández-Varón E, Villamayor L (2007). Granulocyte and granulocyte macrophage colony-stimulating factors as therapy in human and veterinary medicine. *Vet J*. 174: 33-41.

Hill Christopher P, Osslund Timothy D, Eisenberg D (1993). The structure of granulocyte-colony-stimulating factor and its relationship to other growth factors. *Proc Natl Acad Sci USA*. 90: 5167-5171.

Ishikawa M, Iijima H, Satalce-Isikawa R, Tsumura H, Iwamatsu A, Kadoya T, Shimad Y, Fukamachi H, Kobayashi K, Matsuki S, Asuno K (1992). The substitution of Cysteine 17 of recombinant human G-CSF with Alanine greatly enhanced its stability. *Cell Struct Funct.* 17: 61-65.

Ishikawa M, Okada Y, Ishikawa R, Tsumura H, Matsuki S, Asano K (1993). Protein tailoring of human granulocyte colony-stimulating factor. *Biotechnol Lett*. 15: 673-678.

Klingemann Hans-G (1989). Clinical applications of recombinant human colony-stimulating factors. *CMAJ*. 140: 137-142.

Kuga T, Komatsu Y, Yamasaki M, Sekine S, Miyaji H, Nishi T, Sato M, Yokoo Y, Asano M, Okabe M, Morimoto M, Itoh S (1989). Mutagenesis of human granulocyte colony stimulating factor, *Biochem Biophys Res Commun*. 159: 103-111.

Lu HS, Clogston CL, Narhi LO, Merewether LA, Pearl WR, Boone TC (1992). Folding and oxidation of recombinant human granulocyte colony stimulating factor produced in *Escherichia coli*, *J Biol Chem*. 267: 8770-8777.

Lu HS, Fausset PR, Narhi LO, Horan T, Shinagawa K, Shimamoto G, Boone TC (1999). Chemical modification and site-directed mutagenesis of methionine residues in recombinant human granulocyte colony-stimulating factor: Effect on stability and biological activity. *Arch Biochem Biophys*. 362: 1-11.

Luo Peizhi, Heyes Robert J, Chan Cheryl, Stark Dian M, Hwang Marian Y, Jacinto Jonathan M, Juvvadi Padmaja, Chung

Helen S, Kundu A, Ary ML, Dahiyat BI (2002). Development of a cytokine analog with enhanced stability using computational ultrahigh throughput screening, *Protein Sci*. 11: 1218-1226.

Nagata S, Tsuchiya M, Asano S, Kaziro Y, Yamazaki T, Yamamoto O, Hirata Y, Kubota N, Oheda M, Nomura H, Ono M (1986a). Molecular cloning and expression of cDNA for human granulocyte colony-stimulating factor. *Nature* 319: 415-418.

Nagata S (1989). Gene structure and function of granulocyte colony-stimulating factor. *Bio Essays* 10: 113-117.

Nagata S, Tsuchiya M, Asano S, Yamamoto O, Hirata Y, Kubata N, Oheda M, Nomura H, Yamazaki T (1986b). The chromosomal gene structure and two mRNAs for human granulocyte colony-stimulating factor. *EMBO J*. 5: 575-581.

Novagen Inc. (2005). pET system manual, 11th edition.

Okabe M, Asano M, Kuga T, Komatsu Y, Yamasaki M, Yokoo Y, Itoh S, Morimoto M, Oka T (1990). *In vitro* and *in vivo* hematopoietic effect of mutant human granulocyte colony-stimulating factor. *Blood* 75: 1788-1793.

Oshima Yasuo, Tojo Arinobu, Niho Yoshiyuki, Asano Shigetaka, (2000), Biological Activity of Human Granulocyte Colony Stimulating Factor with a Modified C-Terminus. *Biochem Biophys Res Commun*. 267: 924-927.

Reidhaar-Olson John F, De Souza-Hart Janet A, Selick Harold E (1996). Identification of residues critical to the activity of human granulocyte colony-stimulating factor. *Biochemistry* 35: 9034-9041.

Saeedinia A, Sadeghizadeh M, Maghsoudi N, Fallah Mehrabadi J, Karimi M, Akbari B (2003). Construction and cloning of human granulocyte colony stimulating factor (hG-CSF) cDNA. *Modarres JMS* (in Persian). 5: 1.

Sambrook J, Russel DW (2001). *Molecular Cloning*: a laboratory manual, Cold Spring Harbor Laboratory Press. New York, USA.

Souza LM, Boone TC, Gabrilove J, Lai PH, Zsebo KM, Murdock DC, Chazin VR, Bruszewski J, Lu H, Chen KK (1986). Recombinant human granulocyte colony-stimulating factor: effects on normal and leukemic myeloid cells. *Science* 232: 61-65.

Yamasaki M, Konishi N, Yamaguchi K, Itoh S, Yokoo Y (1998). Purification and characterization of recombinant human granulocyte colony-stimulating factor (rhG-CSF) derivatives: KW-2228 and other derivatives. *Biosci Biotechnol Biochem*. 62: 1528-1534.

Werner JM, Breeze AF, Kara B, Rosenbrock G, Boyd J, Soffe N, Campbell ID (1994). Secondary structure and backbone dynamics of human granulocyte colony-stimulating factor in solution. *Biochemistry* 33: 7184-7192.

Wingfield Paul, Benedict Robert, Turcatti Gerardo, Allet Bernard, Mermod Jean-Jacques, Delamarter John, Simona Marco G, Rose Keith, (1988) Characterization of recombinant derived granulocyte-colony stimulating factor (G-CSF). *Biochem J*. 256: 213-218.

Young Dennis C, Zhan Hangjun, Cheng Qi-lin, Hou Jinzhao, Matthews David J (1997). Characterization of the receptor binding determinants of granulocyte colony stimulating factor. *Protein Sci*. 6: 1228-1236.

Zink T, Ross A, Luers K, Cieslar C, Rudolph R, Holak TA (1994). Structure and dynamics of the human granulocyte colony-stimulating factor determined by NMR spectroscopy. Loop mobility in a four-helix-bundle protein. *Biochemistry* 33: 8453-8463.