International Journal of Biological Macromolecules 45 (2009) 97–102

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Molecular cloning, expression in *Escherichia coli* and production of bioactive homogeneous recombinant human granulocyte and macrophage colony stimulating factor

Raquel Cristina Schwanke[a,b,c], Gaby Renard[c], Jocelei Maria Chies[c], Maria Martha Campos[d], Eraldo Luiz Batista Junior[a,d], Diógenes Santiago Santos[a,*], Luiz Augusto Basso[a,*]

[a] *Centro de Pesquisas em Biologia Molecular e Funcional, Instituto Nacional de Ciência e Tecnologia em Tuberculose, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre – RS 90619-900, Brazil*
[b] *Programa de Pós-Graduação em Biologia Celular e Molecular – PUCRS, Porto Alegre, Brazil*
[c] *Quatro G Pesquisa e Desenvolvimento LTDA – Tecnopuc, Porto Alegre - RS 90619-900, Brazil*
[d] *School of Dentistry, PUCRS, Porto Alegre - RS 90610-000, Brazil*

ARTICLE INFO

Article history:
Received 5 January 2009
Received in revised form 27 March 2009
Accepted 14 April 2009
Available online 21 April 2009

Keywords:
Human granulocyte-macrophage colony-stimulating factor
Inclusion bodies
Recombinant protein purification
Biosimilar

ABSTRACT

Human granulocyte and macrophage colony stimulating factor (hGM-CSF) is a glycoprotein that activates and enhances the differentiation and survival of neutrophils, eosinophils and macrophages, which play a key role in the innate immune response. Here we describe the construction of the hGM-CSF encoding gene, cloning, expression in *Escherichia coli*, purification of recombinant hGM-CSF, N-terminal amino acid sequencing, and biological activity assay using TF-1 cells. The results presented show that the combination of experimental strategies employed to obtain recombinant hGM-CSF can yield biologically active protein, and may be useful to scaling-up production of biosimilar protein.

1. Introduction

Biopharmaceuticals are medicinal products comprising biotechnology-derived recombinant proteins as active substances. As to all other medicines, they are regulated by the U.S. agency Food & Drugs Administration (FDA) [1] and the European Agency for the Evaluation of Medicinal Products (EMEA) [2], which create and establish standards and scientific mechanisms that ensure safety, efficacy and quality of biopharmaceutical drugs. EMEA has approved the first biosimilar (Sandoz's Omnitrope, somatropin, somatrophin) in April 2006 whereas the FDA has not approved any biosimilar yet [3]. Although the regulatory pathway for approval of biosimilars has not been completely finalized yet, these follow-on protein products should at least be shown to be pharmaceutically equivalent (that is, it contains the same active ingredient in the same strength, dosage form and route of administration) and to be bioequivalent [4]. Recombinant Granulocytes and Macrophages Colony Stimulating Factor (rGM-CSF) has been produced by Schering-Plough in *Escherichia coli* (Leucomax®, Molgramostim) and by Berlex in *Saccharomyces cerevisiae* (Leukine®, Sargramostim), the former had its patent expired in 2006. Leucomax® was co-developed by Novartis and Schering-Plough and has been co-marketed by the two companies in various countries since 1991 with the approval of EMEA. However, the FDA has only approved Leukine® for commercialization in USA. Sales in 2006 of hGM-CSF, in USA alone, were of US$ 73 million [5].

Colony-stimulating factors (CSFs) are a group of glycoproteins that regulate the proliferation and differentiation of hematopoietic cells. The classification is based on the stimulatory effects that they exert on various bone marrow progenitor cell lineages. GM-CSF is expressed in response to inflammatory stimuli by various cell types including T lymphocytes, macrophages, fibroblasts

Abbreviations: hGM-CSF, human granulocyte and macrophage colony stimulating factor; rhGM-CSF, recombinant hGM-CSF; FDA, Food & Drugs Administration; EMEA, European Agency for the Evaluation of Medicinal Products; CSFs, colony-stimulating factors; PCR, polymerase chain reaction; IPTG, isopropyl β-D-thiogalactopyranoside; IB, inclusion body; EDTA, ethylene diamine tetraacetic acid; PMSF, phenylmethylsulfonyl fluoride; RT, room temperature; SDS-PAGE, sodium dodecyl sulphate polyacrilamide gel electrophoresis; FBS, fetal bovine serum; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide.

* Corresponding authors at: Av. Ipiranga 6681 – Tecnopuc – Prédio 92A, 90619-900 Porto Alegre, RS, Brazil. Tel.: +55 51 33203629; fax: +55 51 33203629.
E-mail addresses: diogenes@pucrs.br (D.S. Santos), luiz.basso@pucrs.br (L.A. Basso).

0141-8130/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijbiomac.2009.04.005

and endothelial cells [6–8] and stimulates specifically the proliferation of cells of macrophage and granulocyte lineages [6,9]. GM-CSF enhances the production and survival of neutrophils, eosinophils and macrophages [10], which play a key role in the innate immune response. Accordingly, GM-CSF is used to restore hematopoietic dysfunctions, to stimulate the hyper-production of functionally primed effectors cells, and to augment host defense against infection and malignant diseases [11]. Endogenous myeloid CSFs, such as GM-CSF, have been used to enhance the clinical management of immunosuppressed patients with cancer. These agents are associated with significant decreases in chemotherapy-associated infections, antibiotic use, length of hospital stay and mortality [12].

The natural human GM-CSF is composed of 127 amino acids and shares 52% identity with murine GM-CSF [10]. Human GM-CSF has four cysteine residues which form two disulphide linkages but only the disulphide bond between Cys^{54} and Cys^{96} is required for biological activity of the protein [10]. Analysis of the three-dimensional structure of the non-glycosylated form of the protein [PDB: 1CSG] shows that GM-CSF has two-stranded antiparallel β-sheet with an open bundle of four α-helices [13,14]. The characterization of the structural elements of GM-CSF responsible for binding to its receptor and necessary for biological activity showed that residues 21–31 in helix A and 78–94 in helix C are essential for high affinity receptor binding and biological activity, whereas helix B does not appear to be essential [10]. In addition, Glu-21 of helix A has also been implicated in high affinity binding of the GM-CSF receptor [15,16].

Here we describe the synthesis of the coding DNA sequence of hGM-CSF without the signal peptide region, cloning into pET-30a(+) expression vector, and recombinant hGM-CSF (rhGM-CSF) protein expression in *E. coli* host cells. We also present an efficient downstream purification protocol for rhGM-CSF expressed in inclusion bodies. Homogeneous rhGM-CSF showed biological activity which was found to be similar to the control (commercially available product). The experimental results described here may contribute to improving rhGM-CSF production process and, hopefully, to lowering costs to healthcare payers and consumers.

2. Methods

2.1. Construction and cloning of hGM-CSF

Oligonucleotides were manually designed, in even number, and synthesized based on the nucleotide sequence of the GM-CSF gene (Accession number NM_000758 version 2). The synthesis of human GM-CSF coding DNA sequence was carried out as described elsewhere [17]. The DNA sequence that codes for the peptide signal was removed and NdeI (5′-end) and BamHI (3′-end) restriction sites were included in flanking primers. Briefly, the 400-bp coding sequence of human GM-CSF was divided into 12 sequences of approximately 50 bp each. The 12 designed sequences overlapped the ends of the immediately adjacent oligonucleotide (at least 10 bases over the adjacent ends). These fragments were then assembled by PCR (polymerase chain reaction) amplification in a step-by-step fashion. The final PCR product was gel-purified, cloned into pCR®-Blunt vector (Invitrogen) and subcloned into pET30a(+) expression vector (Novagen). Nucleotide sequence of rhGM-CSF was determined by automated sequencing to confirm the correct assembly of the coding sequence and ensure that no mutations were introduced by the PCR amplification.

2.2. Expression of human GM-CSF in E. coli

E. coli BL21(DE3) (Novagen) electrocompetent cells were transformed with recombinant pET 30a(+)::hGM-CSF plasmid. As negative control the cells were transformed with pET 30a(+) vector without insert. A single colony was inoculated into 100 mL of LB medium containing 30 μg mL^{-1} kanamycin, and grown overnight at 37 °C. This culture was used to inoculate (1:100) 5.5 L of 4YT medium (32 g bacto tryptone, 20 g yeast extract, and 5 g NaCl per liter, pH 7.2) and grown in shaker flasks at 180 rpm, 37 °C. After reaching OD_{600} 0.4–0.6 the cultures were grown for further 24 h with no IPTG induction. Cells were harvested by centrifugation at 15,900 × *g* for 30 min at 4 °C and stored at −20 °C.

2.3. Isolation of inclusion bodies

A strategy that involves inclusion body (IB) isolation and washing, solubilization of the aggregated protein, and refolding of the solubilized G-CSF protein [18] was employed to recover active rhGM-CSF from IBs. The frozen cell paste was resuspended in lysis buffer (50 mM Tris pH 8.0, 1 mM EDTA (Ethylene Diamine Tetraacetic Acid), and 1 mM phenylmethylsulfonyl fluoride (PMSF) 1:10, w/v). The cell suspension was submitted to a French press (Constant Systems LTD) under 137.9 MPa. The inclusion bodies were separated from the cell debris and intact cells by centrifugation at 15,900 × *g* for 45 min at 4 °C. The pellet was washed in three steps. In the first step, IB pellet was resuspended in 2% Triton X-100 in 50 mM Tris HCl, pH 8.0 and 5 mM EDTA at a pellet to buffer ratio of 1:40 (w/v). This solution was stirred at room temperature (RT) for 60 min and centrifuged at 15,900 × *g* for 30 min. This step was repeated once. In the second step of washing, the pellet was resuspended in 1% sodium deoxycholate in 50 mM Tris HCl pH 8.0 and 0.5 mM EDTA buffer (ratio of 1:40, w/v). This solution was stirred at RT for 1 h and centrifuged at 15,900 × *g* for 30 min. For the last step of washing the buffer 50 mM Tris HCl pH 8.0 containing 1 M NaCl and 0.5 mM EDTA was used to resuspend the pellet. The solution was stirred and centrifuged as before.

2.4. IB solubilization and refolding

IB pellet was resuspended in 2 M urea up to a final protein concentration of 2 mg mL^{-1}. The pH of the solution was adjusted to 11–12.5 with NaOH 1 M and stirred for 30 min at RT. The pH was then reduced with acetic acid to 8.0 and the protein solution diluted 10-fold with 0.1% polysorbate 20 for refolding. The solubilized solution was dialyzed twice against Tris HCl 50 mM pH 8.0 buffer and then twice against 25 mM sodium acetate buffer pH 4.5 (buffer A).

2.5. Purification of rhGM-CSF

The dialyzed solution was clarified by centrifugation at 15,900 × *g* for 30 min at 4 °C. The supernatant was loaded on a HiPrep Resource S column (GE Healthcare/Amersham Biosciences) equilibrated with buffer A and proteins were eluted with 0–1 M Tris HCl linear gradient (buffer B, 1 M Tris HCl, pH 8.0) at a flow rate of 1 mL min^{-1}. The eluted fractions were pooled and dialyzed against Tris HCl 50 mM, pH 8.0 (Buffer C) and loaded on a MonoQ HR10/10 column (GE Healthcare/Amersham Biosciences) equilibrated with buffer C. The bound proteins were eluted with 0–1 M NaCl linear gradient (Buffer D, Tris HCl 50 mM and 1 M NaCl, pH 8.0) at 1 mL min^{-1} flow rate. All purification steps were analyzed by SDS-PAGE 12% [19], and the protein concentration was determined by the Bradford's method [20], using a Bio-Rad Laboratories protein assay kit. The Owl Silver Stain kit (Thermo Scientific Owl Separation Systems, Inc, Portsmouth, UK) was used to silver stain protein gels.