

Soluble Prokaryotic Overexpression and Purification of Bioactive Human Granulocyte Colony-Stimulating Factor by Maltose Binding Protein and Protein Disulfide Isomerase

Bich Hang Do¹, Han-Bong Ryu¹, Phuong Hoang, Bon-Kyung Koo, Han Choe*

Department of Physiology and Biomedical Institute of Technology, University of Ulsan College of Medicine, Seoul, Korea

Abstract

Human granulocyte colony-stimulating factor (hGCSF), a neutrophil-promoting cytokine, is an effective therapeutic agent for neutropenia patients who have undergone several cancer treatments. Efficient production of hGCSF using *E. coli* is challenging because the hormone tends to aggregate and forms inclusion bodies. This study examined the ability of seven different N-terminal fusion tags to increase expression of soluble hGCSF in *E. coli*. Four tag proteins, namely maltose-binding protein (MBP), N-utilization substance protein A, protein disulfide isomerase (PDI), and the b'a' domain of PDI (PDlb'a'), increased the solubility of hGCSF under normal conditions. Lowering the expression temperature from 30°C to 18°C also increased the solubility of thioredoxin-tagged and glutathione S-transferase-tagged hGCSF. By contrast, hexahistidine-tagged hGCSF was insoluble at both temperatures. Simple conventional chromatographic methods were used to purify hGCSF from the overexpressed PDlb'a'-hGCSF and MBP-hGCSF proteins. In total, 11.3 mg or 10.2 mg of pure hGCSF were obtained from 500 mL cultures of *E. coli* expressing PDlb'a'-hGCSF or MBP-hGCSF, respectively. SDS-PAGE analysis and silver staining confirmed high purity of the isolated hGCSF proteins, and the endotoxin levels were less than 0.05 EU/μg of protein. Subsequently, the bioactivity of the purified hGCSF proteins similar to that of the commercially available hGCSF was confirmed using the mouse M-NFS-60 myelogenous leukemia cell line. The EC₅₀s of the cell proliferation dose-response curves for hGCSF proteins purified from MBP-hGCSF and PDlb'a'-hGCSF were 2.83±0.31 pM, and 3.38±0.41 pM, respectively. In summary, this study describes an efficient method for the soluble overexpression and purification of bioactive hGCSF in *E. coli*.

Citation: Do BH, Ryu H-B, Hoang P, Koo B-K, Choe H (2014) Soluble Prokaryotic Overexpression and Purification of Bioactive Human Granulocyte Colony-Stimulating Factor by Maltose Binding Protein and Protein Disulfide Isomerase. PLoS ONE 9(3): e89906. doi:10.1371/journal.pone.0089906

Editor: Paul D. Riggs, New England Biolabs, United States of America

Received: October 10, 2013; **Accepted:** January 24, 2014; **Published:** March 3, 2014

Copyright: © 2014 Do et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by MRC grant (2008-0062286) and Priority Research Center Program (2009-0094054) of National Research Foundation funded by the Ministry of Education, Science and Technology of Korea. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have applied for a Korean patent on the method of producing hGCSF in *E. coli*: The patent titled 'Soluble expression and purification method of active recombinant human GCSF' has been filed at the Korea Patent Office on Sept. 26, 2013. The patent application number is '10-2013-0114432'. This patent application does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials. Therefore all materials presented in the manuscript will be freely available to the scientific community although not for commercial purposes.

* E-mail: hchoe@ulsan.ac.kr

† These authors contributed equally to this work.

Introduction

Granulocyte colony-stimulating factor (GCSF), also known as pluripoiectin, controls the production, differentiation, and function of granulocytes, which account for 70% of white blood cells [1,2]. The recruitment of two monomers of GCSF triggers dimerization of the GCSF receptor and initiates a signaling cascade [3–5]. Production of GCSF, which is secreted predominantly by macrophages, fibroblasts and endothelial cells, is stimulated by several inflammatory stimuli, including interleukin-1β, tumor necrosis factor-α, and lipopolysaccharide [6–8]. Human GCSF (hGCSF) has been approved for the treatment of neutropenia, a common disorder in cancer patients following radiotherapy or chemotherapy treatments, characterized by an extremely low number of neutrophils in the blood [9,10]. GCSF also has neuroprotective properties [11]; accordingly, the protein

has been used as a protective agent in mouse models of various neurodegenerative diseases, including amyotrophic lateral sclerosis [12,13].

Human GCSF was initially purified from a tumor cell line that continuously secreted the protein [14]. When expressed in the methylotrophic yeast *Pichia pastoris*, hGCSF is secreted in a soluble form; however, the secreted protein is highly aggregated and must be solubilized using high concentrations of denaturants such as guanidine hydrochloride or urea. Consequently, purification of the biologically active form of hGCSF from yeast requires the removal of these denaturants and refolding of the protein [15]. *Escherichia coli* also produces aggregated hGCSF in inclusion bodies (IBs) [16–22]; however, the overall yield of biologically active protein from these structures is usually low [23]. Alternatively, hGCSF can be secreted into the periplasm of *E. coli* [24,25], although low yields are also usually obtained using this method. Maltose-binding

protein (MBP), and stress-responsive proteins such as peptidyl-prolyl cis-trans isomerase B, bacterioferritin, and glutathione synthase, have previously been tested as fusion partners to increase the production of solubilized hGCSF in *E. coli* [26,27].

In this study, several new methods of overexpressing soluble hGCSF in the cytoplasm of *E. coli* were investigated, enabling efficient production of biologically active protein. The following seven N-terminal fusion tags were used: hexahistidine (His6), thioredoxin (Trx), glutathione S-transferase (GST), MBP, N-utilization substance protein A (NusA), protein disulfide bond isomerase (PDI), and the b'a' domain of PDI (PDib'a'). The MBP, NusA, PDI, and PDib'a' tags increased the solubility of hGCSF markedly at 30°C. Lowering the expression temperature to 18°C also increased the solubility of Trx- and GST-tagged hGCSF, whereas His6-hGCSF was insoluble at both temperatures. The expression level and the solubility of the tag-fused hGCSFs were also tested in the *E. coli* Origami 2(DE3) strain that have mutations in both the thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes, which may assist the disulfide bond formation in the cytoplasm of *E. coli* [28–30]. Simple methods of purifying hGCSF from the PDib'a' or MBP tagged proteins were developed using conventional chromatographic techniques. In total, 11.3 mg of biologically active hGCSF was obtained from 500 mL of culture. Silver staining indicated that the extracted hGCSF was highly pure and the endotoxin level was very low. The activity of the purified protein was measured using a bioassay with mouse M-NFS-60 myelogenous leukemia cells.

Materials and Methods

Construction of plasmids and expression in *E. coli*

The hGCSF gene (Uniprot identifier: P09919-2) encodes a protein comprising 204 amino acids, the first 29 of which form the signal peptide. To enable the expression and purification of hGCSF in *E. coli*, a tobacco etch virus (TEV) protease recognition site (TEVrs; ENLYFQ_G) was appended to the N-terminus of mature hGCSF (175 amino acids), and two site-specific recombination sequences, *att*B₁ (5'-GGGGACAAGTTTGTACAAAAAGCAGGCTTC-3') and *att*B₂ (5'-ACCCAGCTTTCTGTACAAAAGTGGTCCCC-3'), were added to each end of the gene sequence (Figures 1A and 1B). The hGCSF DNA sequence which is substituted Met¹ to Ala¹ was synthesized and subcloned into plasmid pUC57 (Genscript, Piscataway, NJ), which was then recombined with the pDONOR207 vector (Invitrogen, Carlsbad, CA) to produce the entry vector pENTR-hGCSF (Figure 1A). LR recombination cloning between pENTR-hGCSF and seven destination vectors containing the relevant fusion tags (pDEST-HGWA, pDEST-HXGWA, pDEST-HGGWA, pDEST-HMGWA, pDEST-HNGWA, pDEST-PDI, and pDEST-PDib'a') [31,32] was performed to produce expression vectors containing tagged hGCSF. The expression plasmids were confirmed by DNA sequencing (Macrogen, Daejeon, Korea) and then transformed into *E. coli* BL21(DE3) and Origami 2(DE3).

To overexpress hGCSF, the transformed BL21(DE3) cells were grown at 37°C in 200 rpm of shaking incubator in 2 mL of Luria-Bertani (LB) broth containing 50 µg/mL ampicillin. For the culture of the transformed Origami 2(DE3), 12.5 µg/mL tetracycline was also added. One mM isopropyl-β-D-thiogalactoside (IPTG) was added at 0.4~0.6 OD₆₀₀ to induce the expression of the hGCSF fusion proteins. The cells were harvested after incubation for 5 h at 30°C or 12 h at 18°C.

Purification of hGCSF from the PDib'a'-hGCSF fusion protein

E. coli BL21(DE3) cells transformed with the PDib'a'-hGCSF expression vector were cultured for 12 h at 18°C in 500 mL of LB medium. When OD₆₀₀ was reached to 0.4~0.6, 1 mM IPTG was added to induce the expression of the fusion protein. The collected cells were resuspended in 50 mL of immobilized metal ion affinity chromatography (IMAC) binding buffer comprising 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, and 5% glycerol (v/v). The solution was sonicated until completely transparent and then centrifuged for 20 min at 27,000 g to generate the supernatant. After equilibrating with binding buffer, the pre-packed 3×5 mL HisTrap HP column (GE Healthcare, Piscataway, NJ) was fed with the lysate solution and non-specific proteins were then removed by washing with IMAC buffer containing 100 mM imidazole. The PDib'a'-hGCSF fusion protein was eluted in IMAC buffer containing 500 mM imidazole. To support TEV protease cleavage, the buffer was then exchanged to NaCl-free IMAC buffer (50 mM Tris-HCl, pH 8.0, 5% glycerol (v/v)) using a dialysis membrane (Viskase, Darien, Illinois). For digestion, the fusion protein was incubated with TEV protease at a ratio of 1:20 for 12 h at 18°C. For IMAC, the digested sample was loaded onto a pre-packed 2×5 mL HisTrap HP column filled with IMAC buffer. Unlike other proteins in solution, hGCSF had a low affinity to the Ni resin and was easily eluted from the HisTrap column using IMAC buffer containing 50 mM imidazole. Based on the chromatogram, the collected hGCSF was analyzed by 10% Tris-tricine SDS-PAGE.

Purification of hGCSF from the MBP-hGCSF fusion protein

E. coli BL21(DE3) cells transformed with the MBP-hGCSF expression vector were cultured for 12 h at 18°C in 500 mL of LB medium and induced by 1 mM IPTG when OD₆₀₀ was 0.4~0.6. Due to the high affinity of MBP-hGCSF to the MBP column, a 2×5 mL MBPTrap HP column (GE Healthcare) was used as the first purification step. The cells were resuspended in 50 mL of MBP-binding buffer comprising 50 mM Tris-HCl (pH 8.0), 0.5 mM EDTA, 200 mM NaCl, and 5% glycerol (v/v), and then sonicated to form a soluble solution. The supernatant was loaded onto a 2×5 mL MBPTrap HP column equilibrated with MBP-binding buffer. Non-specific bound proteins were removed by washing with binding buffer and MBP-hGCSF was eluted with binding buffer containing 10 mM maltose monohydrate. The eluted sample was diluted until the final concentration of NaCl was 50 mM and then cleaved with TEV protease under the same conditions as described for PDib'a'-hGCSF. Cleaved hGCSF was then purified using the same method of hGCSF cleavage from PDib'a'-hGCSF.

SDS-PAGE and silver staining

Proteins were separated and visualized on a 10% Tris-tricine gel stained with Coomassie Brilliant Blue R-250 (AMRESCO, Solon, OH). The expression, solubility, and purity were quantified using ImageJ software (<http://imagej.nih.gov/ij>). For silver staining, the polyacrylamide gel was placed into Fixative Enhancer Solution (Bio-Rad Laboratories, Hercules, CA) for 20 min and then rinsed with distilled water to increase the sensitivity and contrast of the staining. Staining and developing were performed using a mixture of silver complex solution, reduction moderator solution, and image development reagent (Bio-Rad Laboratories, Hercules, CA). The reaction was stopped by the addition of 5% acetic acid.

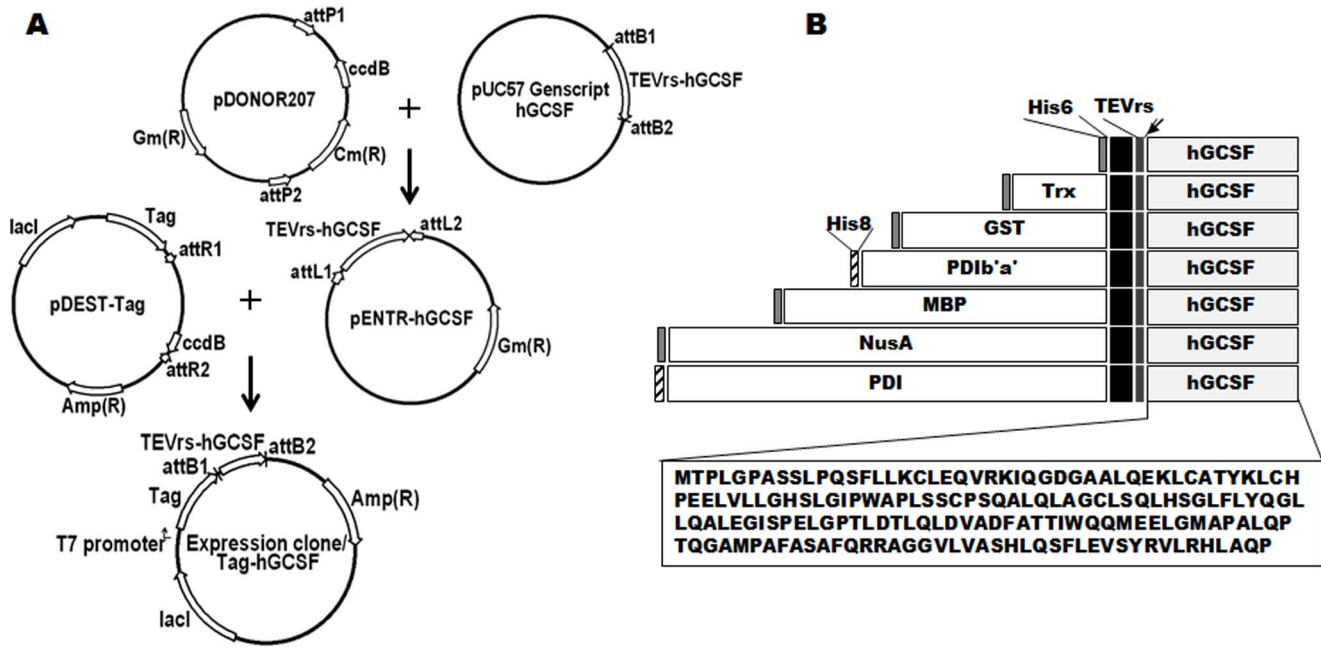


Figure 1. Construction of the hGCSF expression vectors and schematic representations of the domain structures. A. The method of construction and vector map of the tag-hGCSF construct. All fusion constructs were generated in the same way via LR recombination cloning. Expression of the fusion proteins in *E. coli* was controlled by the IPTG-inducible T7 promoter, and ampicillin was used as the selection marker. B. Schematic representation of the seven hGCSF fusion proteins used in this study (His6-, Trx-, GST-, PDib'a'-, MBP-, PDI-, and NusA-hGCSF). The arrow indicates the TEV protease cleavage site. Black is extra sequences from BP and LR recombinations. The amino acid sequence of mature hGCSF is also shown.

doi:10.1371/journal.pone.0089906.g001

Endotoxin assay

To remove endotoxins from purified hGCSF, the solution was incubated with 1% Triton X-114 (Sigma-Aldrich, St. Louis, MO) at 4°C for 30 min. Triton X-114 was accumulated after incubating the sample at room temperature and removed by centrifugation at 9,000 g for 10 min [33]. The Endpoint Chromogenic Limulus Amebocyte Lysate test (Lonza, Basel, Switzerland) was used to quantify the remaining endotoxin in the target solution. Briefly, Limulus Amebocyte Lysate was incubated with the hGCSF sample at 37°C for 10 min before the substrate was added. Stop agent (25% v/v glacial acetic acid) was then added to the mixture and the released p-nitroaniline was evaluated by photometric measurement at 405–410 nm.

Cell proliferation assay

The M-NFS-60 mouse myelogenous leukemia cell line [34,35], kindly provided by Dr. Kyung-Woon Kim (Rural Development Administration, Suwon, Korea), was grown in RPMI-1640 medium (Invitrogen, Carlsbad, CA) containing 10% fetal bovine serum, 1X penicillin and streptomycin (Invitrogen), and 0.05 mM β-mercaptoethanol [36]. The cells were maintained at 37°C in a humidified atmosphere containing 5% CO₂. The bioassay of purified hGCSF using M-NFS-60 cells was based on the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (Invitrogen). The cultured cells were seeded at a density of 3 × 10⁴ cells/well into 96-well plates containing growth medium. To determine its effect on proliferation of the cells, different concentrations (0.1, 1, 10, 100, 1,000, 10,000 and 100,000 pg/mL) of commercially available hGCSF purified from IB (Genscript, Piscataway, NJ) and hGCSF produced from the PDib'a' and MBP fusion proteins were added to each well in a final volume of 100 μL. After 72 h of incubation, 15 μL of 5 mg/mL MTT was

added to each well and the cells were incubated in the dark at 37°C for a further 4 h. After removing all solutions from the cells, 100 μL of dimethyl sulfoxide was added to each well to completely solubilize the formed aggregates. The optical density of the solution was measured at 570 nm using an ELISA plate reader (Molecular Devices, Sunnyvale, CA).

Data analysis

A non-linear regression analysis was used to determine the M-NFS-60 cell proliferation dose-response to hGCSF. The data were fitted using the following equation and Microsoft Excel software, where *Re* is response of the cells, *Bl* is the baseline at low concentration, *Max* is the maximum response, *conc* is the concentration of the protein, and *Hs* is the Hill coefficient of stimulation, *Bh* is the baseline at high concentration, and *Hi* is the Hill coefficient of inhibition:

$$Re = Bl + \frac{Max - Bl}{1 + \left(\frac{EC_{50}}{conc}\right)^{Hs}} - \frac{Max - Bh}{1 + \left(\frac{IC_{50}}{conc}\right)^{Hi}} \quad (1)$$

All data are presented as the mean ± standard error (SE) of n ≥ 3 of 2 independent experiments. To determine the statistical significance of the responses of cells to hGCSF, group means were compared using a Student's *t*-test or a one-way analysis of variance followed by Bonferroni's multiple comparisons test. Graphpad Prism 5 software (GraphPad, San Diego, CA) was used for statistical analyses and *P* < 0.05 was considered significant.

Results

Construction of plasmids and expression of tagged hGCSF in *E. coli*

To enable soluble expression of hGCSF in the cytoplasm of *E. coli*, the following seven tags were fused to the N-terminus of the protein via LR recombination cloning: His6, Trx, GST, PDI b'a', MBP, PDI, and NusA (Figure 1). A TEVrs was also inserted between each tag and hGCSF to facilitate removal of the tags during purification, and the sequence was codon-optimized for *E. coli* expression (Figure 1B). Vectors containing the fusion tags were recombined with the hGCSF plasmid, then the resulting plasmids were sequence-verified and transformed into the BL21(DE3) *E. coli* strain, which lacks protease expression.

Expression of the hGCSF fusion genes in *E. coli* was controlled by a T7 promoter and induced with 1 mM IPTG at two different expression temperatures of 30°C and 18°C. The expression levels of all tagged hGCSF proteins were 33–68%, and the expression levels of all proteins were higher at 18°C than 30°C (Figure 2 and Table 1). The solubilities of the proteins varied depending on both the type of fusion tag used and the expression temperature. The solubility of hGCSF at 30°C was markedly enhanced by the addition of the MBP, NusA, PDI, and PDIB'a' tags (Figure 2B and Table 1). Lowering the expression temperature to 18°C additionally increased the solubility of the Trx-hGCSF and GST-hGCSF proteins to similar levels (Figure 2A and Table 1); however, His6-hGCSF was insoluble at both expression temperatures. We also tested *E. coli* Origami 2(DE3), a strain that may promote disulfide bond formation in the cytoplasm of *E. coli*, as an expression host. The expression levels of the fusion proteins in Origami 2(DE3) were lower than those in BL21(DE3), and the solubilities were similar at both 18°C and 30°C (Figure S1). Based on the expression level, solubilities and sizes of the tagged proteins, PDIB'a'-hGCSF and MBP-hGCSF in BL21(DE3) were selected for further study.

Purification of hGCSF from the PDIB'a'-hGCSF fusion protein

Separation of hGCSF from the PDIB'a'-hGCSF fusion protein was performed by two rounds of IMAC, with an intervening TEV protease digestion step (Figure 3A). IMAC was possible because all of the tags used in the study contained an additional His6 or His8 tag at their N-terminal end (Figure 1B). Cells transformed with the plasmid containing PDIB'a'-hGCSF were induced with IPTG and then collected (Figure 3B, lane 2). The cells were lysed and centrifuged to harvest the supernatant (Figure 3B, lane 3), which was then loaded onto a Ni column and the binding protein was eluted after a washing step (Figure 3B, lane 4). Most of the non-specific proteins were removed at this step; however, some minor contaminant bands were observed. Despite the presence of these additional proteins, TEV protease digestion was performed. After optimizing the digestion conditions (data not shown), the majority of the PDIB'a'-hGCSF protein was cleaved by TEV protease (Figure 3B, lane 5). A second HisTrap HP column was then used to remove the PDIB'a' tag, undigested PDIB'a'-hGCSF, and TEV protease, which also contained a His6-tag. Cleaved hGCSF weakly bound to the Ni column and was eluted by 50 mM imidazole (Figure 3C). An SDS-PAGE analysis revealed the absence of any contaminating proteins after this step (Figure 3B, lane 6). Silver staining of the SDS-PAGE gel under reducing and non-reducing conditions showed that the purified hGCSF protein was highly pure and mostly monomeric (Figure 3D). Typically, 11.3 mg of hGCSF was obtained from a 500 mL culture of *E. coli* expressing PDIB'a'-hGCSF, with a yield of 36.7% (Table 2). After treatment

with Triton X-114, the endotoxin level of hGCSF purified from the PDIB'a'-hGCSF fusion protein was 0.05 EU/μg.

Purification of hGCSF from the MBP-hGCSF fusion protein

Figure 4A shows an outline of the process used to purify hGCSF from MBP-hGCSF in the cell lysate. MBP chromatography isolated the MBP-hGCSF fusion protein from the total protein mixture with a purity of approximately 80% (Figure 4B, lane 4). After cleavage of the fusion protein with TEV protease (Figure 4B, lane 5), the sample was applied to a Ni-NTA column and purified hGCSF was obtained by eluting with 50 mM imidazole (Figure 4B, lane 6; Figure 4C). Similar to the highly pure hGCSF (approximately 99%) obtained from PDIB'a'-hGCSF, silver staining of the SDS-PAGE gel under reducing and non-reducing conditions revealed the presence of highly pure hGCSF isolated from MBP-hGCSF (Figure 4D). Most of the purified protein was monomeric; although a small amount of hGCSF dimer was observed under non-reducing conditions (Figure 4D). Typically, 10.2 mg of purified hGCSF was obtained from a 500 mL culture of MBP-hGCSF. This total yield of 38.3% (Table 2) was lower than from PDIB'a'-hGCSF. After treatment with Triton X-114, the level of endotoxin in the purified hGCSF sample was 0.013 EU/μg. The endotoxin level of bio-products is typically less than 1 EU/μg.

Biological activity of hGCSF

The bioactivities of the purified hGCSF proteins were measured using an MTT assay and the mouse M-NFS-60 myelogenous leukemia cell line. The number of M-NFS-60 cells increased dramatically after incubation with commercially available hGCSF or hGCSF purified from the PDIB'a'-hGCSF or MBP-hGCSF fusion proteins (Figures 5A–E). At concentrations below 1 nM, the dose-response curves were sigmoidal for all three forms of hGCSF (Figure 5E); however, higher concentrations produced mild inhibition, resulting in a bell-shaped curve (Figure 5E). The EC₅₀s of commercial hGCSF, hGCSF from MBP-hGCSF, and hGCSF from PDIB'a'-hGCSF were 10.69±2.62 pM, 2.83±0.31 pM, and 3.38±0.41 pM, respectively, with Hill coefficients of 1.06±0.29, 1.00±0.05, and 1.06±0.11, respectively. The differences between the EC₅₀s and Hill coefficients were not statistically significant, suggesting that the hGCSF proteins purified from MBP-hGCSF and PDIB'a'-hGCSF are as slightly better effective as commercially available hGCSF.

Discussion

Many human proteins expressed in prokaryotes such as *E. coli* are prone to accumulation in IBs. Consequently, time-consuming solubilization and refolding are necessary to generate the purified proteins; processes that are also hampered by low yields, poor reproducibility, and the generation of proteins with low biological activity [23,37]. When expressed in *E. coli*, hGCSF is also insoluble, and so to address this problem, this study examined the effect of seven different fusion tags that function as chaperones, as well as the effect of a low expression temperature, on the solubility of hGCSF.

The MBP, PDI, PDIB'a', and NusA tags solubilized greater than 70% of the hGCSF fusion protein at 30°C, whereas the solubilities of the Trx-, GST-, and His6-tagged proteins were low at this temperature (Table 1, Figure 2). MBP is thought to act as a general molecular chaperone [38] by binding to hydrophobic residues present on protein surfaces [39]. MBP-tagged proteins can be easily purified with commercially available MBP-binding columns. PDI forms and breaks disulfide bonds of proteins in the lumen of the endoplasmic reticulum. The cytoplasm is usually a

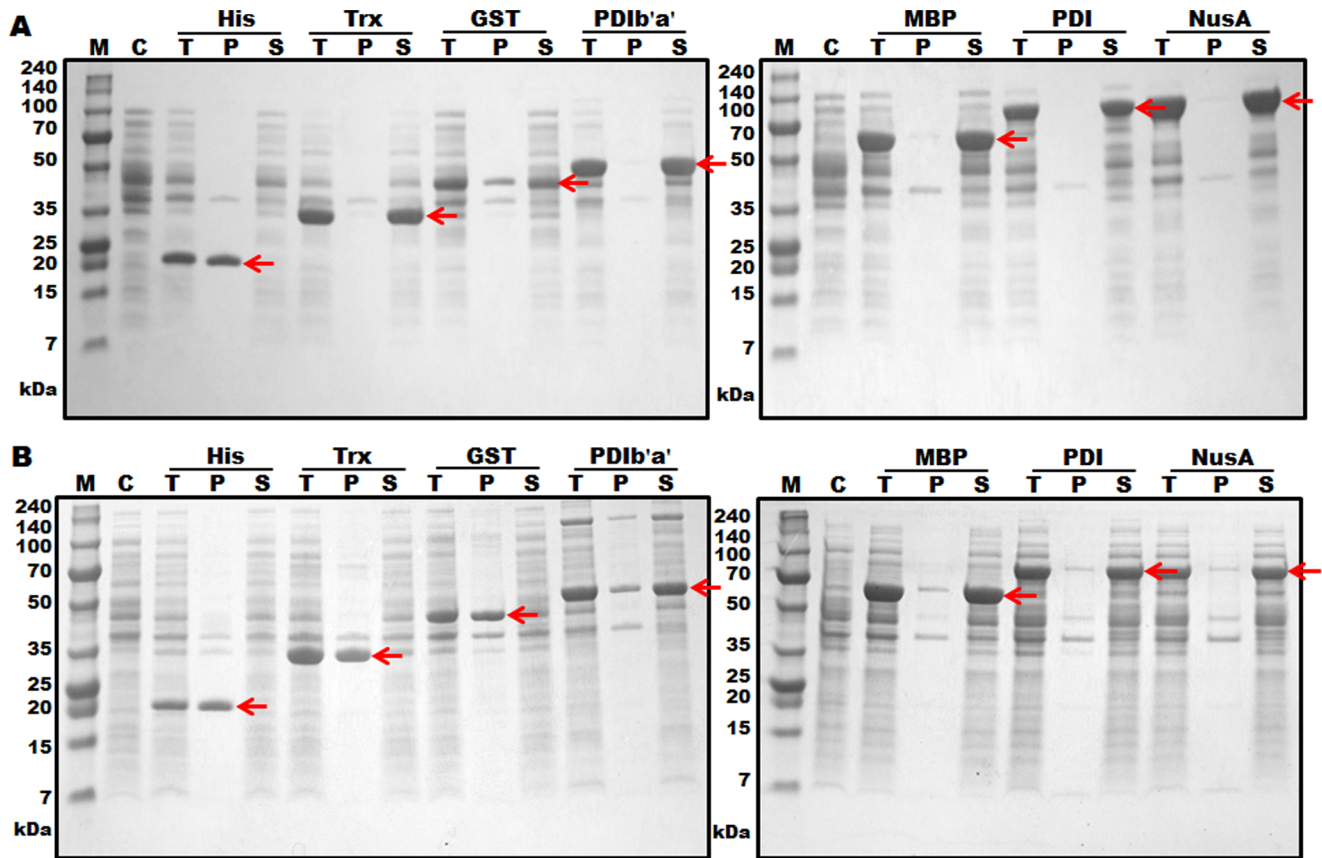


Figure 2. Expression levels of hGCSF fused with seven different tags in *E. coli* BL21(DE3). Protein expression was induced with 1 mM IPTG at either 18°C (A) or 30°C (B). After sonication, 20 µg of each total protein was loaded onto a 10% Tris-tricine gel. The arrows indicate the hGCSF fusion proteins. M, molecular weight size marker; C, total protein before IPTG induction (control); T, total protein after IPTG induction; P, protein in the cell pellet after sonication; S, protein in the supernatant after sonication.
doi:10.1371/journal.pone.0089906.g002

reducing environment that prevents proper disulfide bond formation, but PDI increases the production of soluble proteins in both the cytoplasm [40] and periplasm of *E. coli* [41]. PDI is composed of four thioredoxin-like domains, named a, b, b', and a'. The a and a' domains display redox-active catalytic and chaperone activities, whereas the b and b' domains only demonstrate some chaperone functions [42]. Previous experiments in our laboratory have shown that PD1b'a' increases the solubility of several proteins to the same degree as PDI (data not shown);

however, the data presented here show that PD1b'a' was less effective than PDI at solubilizing hGCSF. NusA was suggested as a solubilizing tag protein based on the revised Wilkinson-Harrison solubility model [43,44], which predicted NusA to be 95% soluble and to improve the solubility of several proteins. PDI and PD1b'a' were also predicted to be good solubilizing agents according to this model (data not shown). The revised Wilkinson-Harrison solubility model considers the number of four turn-forming residues (Asn, Gly, Pro, and Ser) and determines the net charge by subtracting

Table 1. Expression levels and solubilities of hGCSF fused with seven different N-terminal tags.

Tag	Tag size (kDa)	Fusion protein size (kDa)	Expression (%)		Solubility (%)		
			18°C	30°C	18°C	30°C	
hGCSF (18.8 kDa)	His6	0.8	23.5	43.8	33.6	-	-
	Trx	11.8	35.3	61.4	48.8	98.3	5.0
	GST	25.7	49.2	41.3	40.0	78.4	3.2
	PD1b'a'	35.6	59.1	66.3	42.2	96.0	73.5
	MBP	40.3	63.8	61.4	58.4	96.5	88.1
	PDI	55.1	78.7	55.6	43.8	98.1	89.3
	NusA	54.9	78.4	68.0	44.8	97.5	89.5

doi:10.1371/journal.pone.0089906.t001

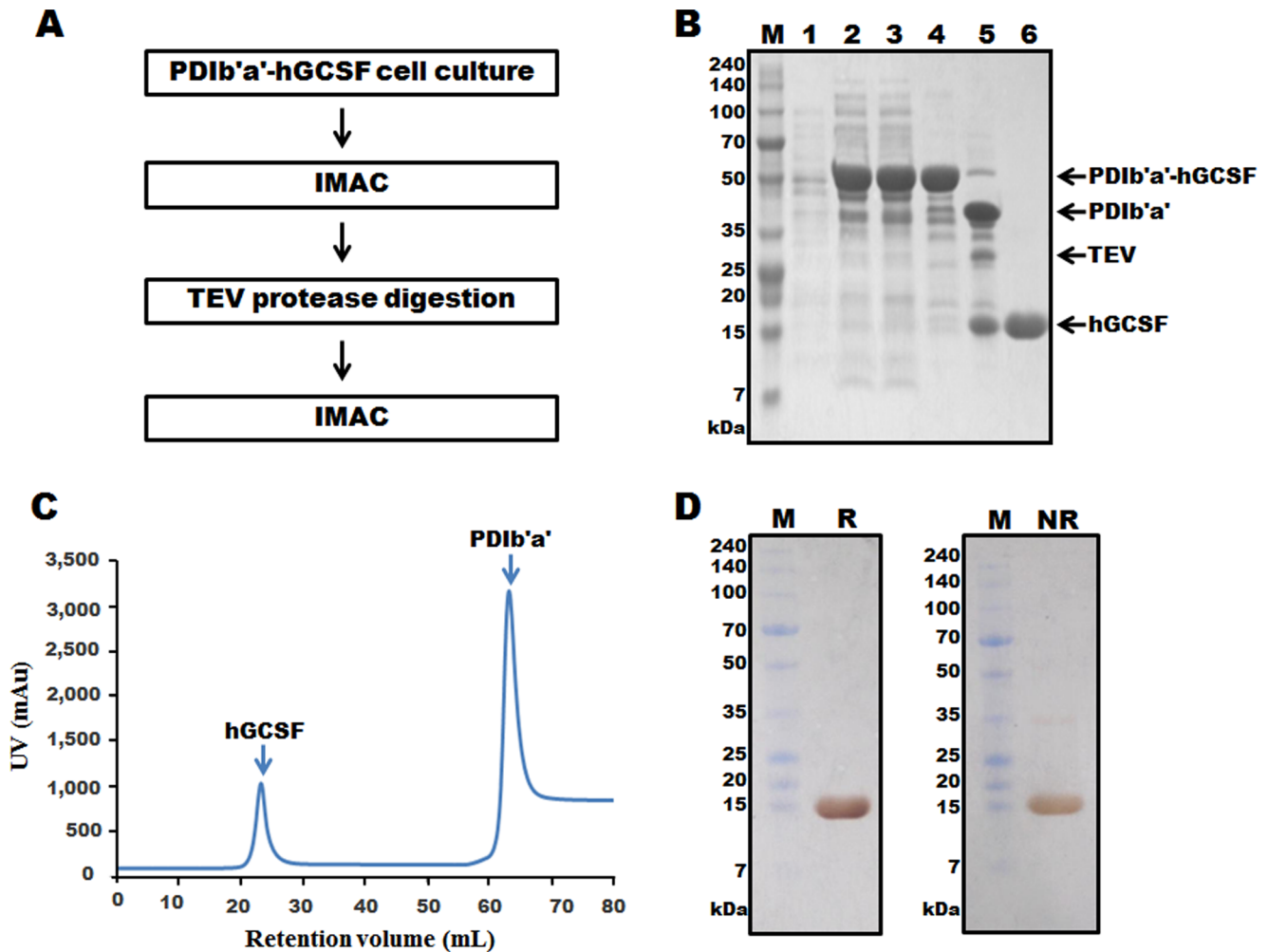


Figure 3. Purification of hGCSF from the PDib'a'-hGCSF fusion protein expressed in *E. coli* BL21(DE3). A. Overview of the purification process using IMAC chromatography and TEV protease digestion. B. SDS-PAGE analysis of hGCSF at various stages of the purification process. Ten micrograms of total protein was loaded onto a 10% Tris-tricine gel. M, molecular weight marker; lane 1, total protein before IPTG induction; lane 2, total protein after IPTG induction; lane 3, soluble fraction after cell sonication; lane 4, PDib'a'-hGCSF fusion protein purified by the first round of IMAC (59.1 kDa); lane 5, PDib'a'-hGCSF fusion protein following TEV protease cleavage showing separated PDib'a' (35.6 kDa) and hGCSF (18.8 kDa); lane 6, purified hGCSF after the second round of IMAC (18.8 kDa). C. IMAC chromatogram of the TEV protease cleaved PDib'a'-hGCSF fusion protein showing clear separation of the PDib'a' tag and hGCSF. D. Silver staining of 7.5 µg of purified hGCSF. M, molecular weight marker; R, reduced hGCSF; NR, non-reduced hGCSF.

doi:10.1371/journal.pone.0089906.g003

the number of acidic residues from the number of basic residues. However, this model may have some limitations because it predicted relatively low solubility for the MBP, Trx, and GST tags

(data not shown), despite the fact that hGCSF fused with these tags showed good solubility.

With the exception of His6-hGCSF, lowering the expression temperature from 30°C to 18°C increased the solubility of all

Table 2. The characteristics of hGCSF purified from PDib'a'-hGCSF and MBP-hGCSF fusion proteins expressed in *E. coli*.

Purification step	hGCSF purified from PDib'a'-hGCSF				hGCSF purified from MBP-hGCSF			
	Total protein(mg)	Purity (%)	hGCSF(mg)	Yield (%)	Total protein(mg)	Purity (%)	hGCSF (mg)	Yield (%)
Cell weight	1500	-	-	-	1500	-	-	-
Supernatant	140	69.1	30.8	100	118.8	75.9	26.6	100
1 st Chromatography (IMAC/MBP)	71.5	73.3	16.7	54	79.8	88	20.7	77.8
2 nd Chromatography	11.4	99	11.3	36.7	10.3	99	10.2	38.3

doi:10.1371/journal.pone.0089906.t002

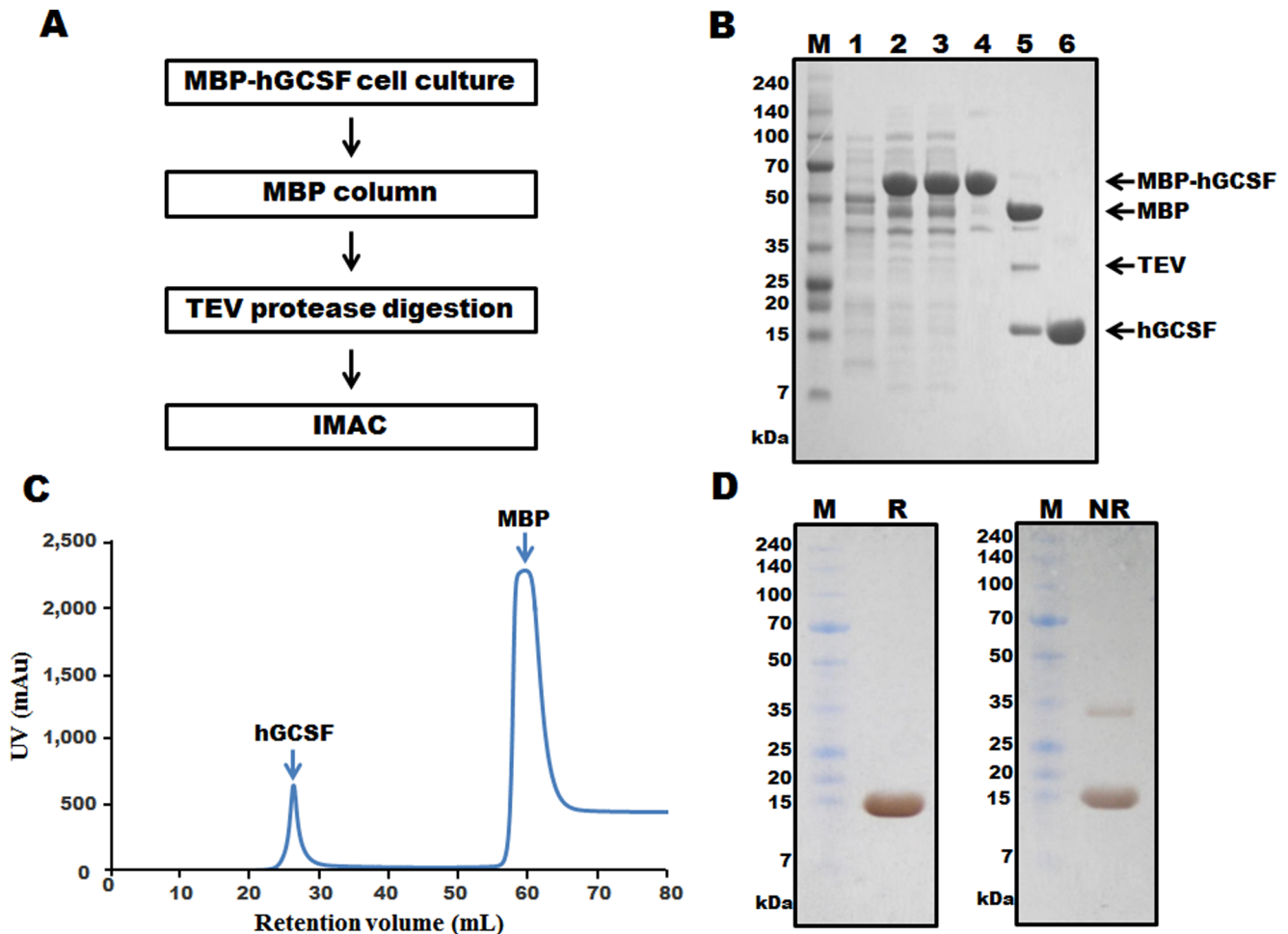


Figure 4. Purification of hGCSF from the MBP-hGCSF fusion protein expressed in *E. coli* BL21(DE3). A. Overview of the purification process using an MBP affinity column, TEV protease digestion, and IMAC. B. SDS-PAGE analysis of hGCSF at various stages of the purification process. M, molecular weight marker; lane 1, total protein before IPTG induction; lane 2, total protein after IPTG induction; lane 3, soluble fraction after cell sonication; lane 4, MBP-hGCSF fusion protein purified using an MBP affinity column (63.3 kDa); lane 5, MBP-hGCSF fusion protein after cleavage with TEV protease showing separated MBP (40.3 kDa) and hGCSF (18.8 kDa); lane 6, purified hGCSF after IMAC (18.8 kDa). C. IMAC chromatogram of the TEV protease cleaved MBP-hGCSF fusion protein showing clear separation of the MBP tag and hGCSF. D. Silver staining of 7.5 µg of purified hGCSF. M, molecular weight marker; R, reduced hGCSF; NR, non-reduced hGCSF. doi:10.1371/journal.pone.0089906.g004

tagged hGCSF proteins (Table 1, Figure 2). Low expression temperatures have been successfully used in the past to increase the solubility of many proteins expressed in *E. coli* [45–50]; however, the molecular mechanisms responsible for this effect are not fully understood at present. The cold temperature protein chaperones are induced at low temperatures [51]; peptidyl-prolyl isomerase is a known cold temperature protein chaperone that catalyzes cis/trans isomerization of the peptide bonds found in proline residues [52]. In addition, several ATP-consuming heat shock proteins may also play a role in improving protein solubility at low expression temperatures [53]. Although highly inducible by heat shock treatment, these proteins are expressed at normal temperatures and have chaperone functions. However, the effects of lowering the expression temperature on protein solubility cannot be generalized because His6-tagged hGCSF was not soluble at all at 18°C.

The effects of hGCSF purified from MBP-hGCSF or PDIB'a'-hGCSF on the proliferation of M-NFS-60 cells were slightly higher than that of commercially available hGCSF (Figure 5E). The EC₅₀ values for hGCSF purified from MBP-hGCSF (2.83 pM) and

PDIB'a'-hGCSF (3.38 pM) were consistent with a previous study that reported an EC₅₀ value in the range of 0.8–6 pM for hGCSF [25,54,55]. At high concentrations, the purified hGCSF proteins induced mild inhibition of cell proliferation, resulting in a bell-shaped biphasic dose-response curve (Figure 5E). This is consistent with a previous report that other cytokines also show a biphasic dose-response curve [56].

There are three splicing variants of hGCSF. The short isoform (b) used in this study is reportedly more active than the longer isoform (a) [57], and the third isoform lacks the region spanning amino acids 37 to 73. In this study, we substituted the first amino acid (Ala) with Met, and this mutation increased binding of hGCSF to its receptor [58] and facilitated PEGylation of the N-terminus of the protein, which increased the half-life of GCSF in blood [59].

Mature hGCSF contains five cysteine residues, four of which form two native intramolecular disulfide bonds, Cys³⁷-Cys⁴³ and Cys⁶⁵-Cys⁷⁵. A previous study in which Cys¹⁸ was mutated to Ser demonstrated that Cys¹⁸ is not required for bioactivity of hGCSF [60]. However, during folding of hGCSF, intermolecular disulfide

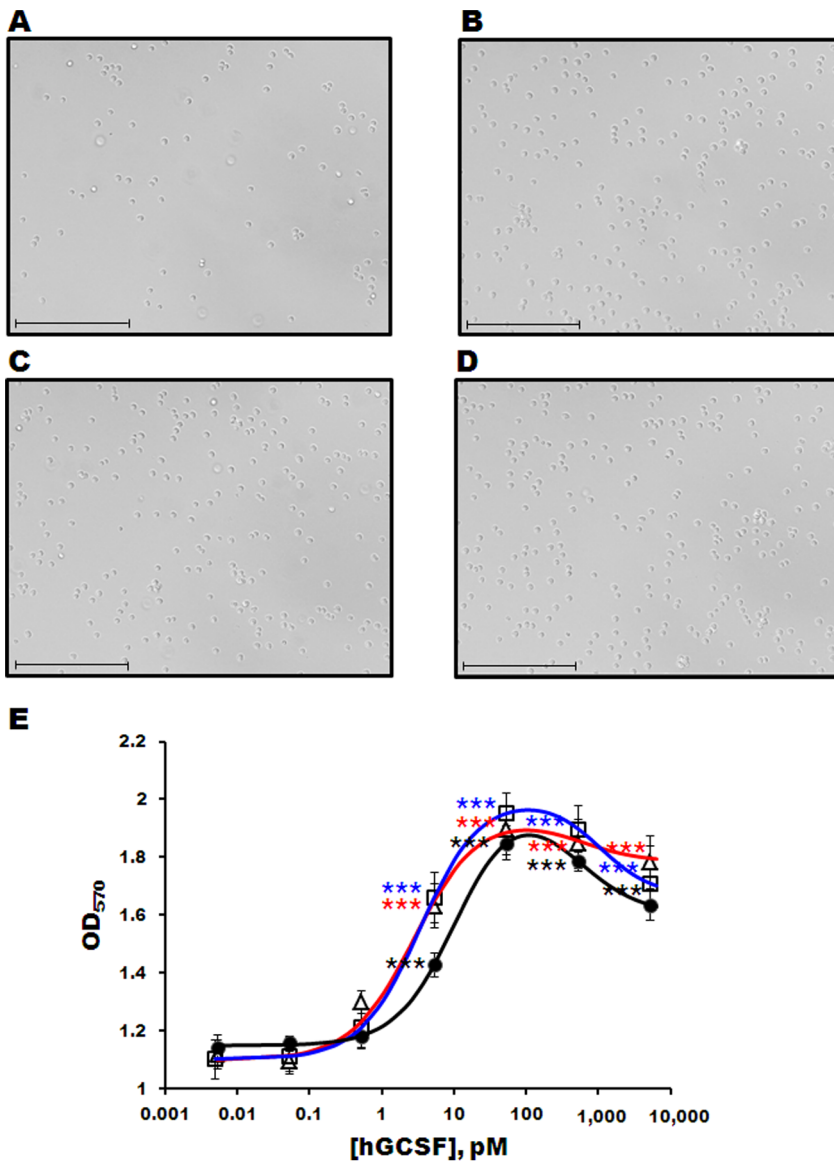


Figure 5. Cell proliferation assay of purified hGCSF using the M-NFS-60 cell line. Light microscopy images of M-NFS-60 cells incubated without (A) or with (B–D) 1 ng/mL hGCSF for 48 h. Commercially available hGCSF (B), hGCSF purified from MBP-hGCSF (C), and GCSF purified from PDIb'a'-hGCSF (D) were used. The scale bar represents 100 μm. E. Dose-response curve of M-NFS-60 cells following exposure to different concentrations of purified hGCSF and commercial hGCSF. The number of cells was measured as the OD₅₇₀. ●, commercially available hGCSF; △, hGCSF purified from MBP-hGCSF; hGCSF purified from PDIb'a'-hGCSF. Data are represented as the mean ± SE of n≥3 of 2 independent experiments. Statistical significance compared to no hGCSF treatment group: *, p<0.05; **, p<0.01; ***, p<0.001. doi:10.1371/journal.pone.0089906.g005

bonds between two Cys¹⁸ residues or Cys¹⁸ and another Cys residue can occur in aggregates [61]. The formation of subsequent dimers or multimers can render hGCSF insoluble in *E. coli* cytoplasm. As a result of the non-optimal spatial orientation of the molecules, the activity of the GCSF dimer is much lower than that of the GCSF monomer *in vitro* [62]. Some effective solutions, such as the mutation of Cys¹⁸ [21,36] or the addition of a specific secretory signal peptide that directs the secretion of hGCSF into the periplasmic space [24], have been used to overcome this obstacle in *E. coli*. Here, soluble monomeric hGCSF with bioactivity similar to that of hGCSF purified from HEK cells was obtained using a fusion protein strategy and a low expression temperature.

Mature hGCSF is glycosylated at Thr¹³⁴. One limitation of using *E. coli* to produce hGCSF is the lack of glycosylation machinery in the bacterial cells; therefore, overexpressed hGCSF obtained from *E. coli* is non-glycosylated. Glycosylation prevents protein aggregation and increases the half-life of circulating proteins in the blood by protecting proteins from protease cleavage; however, it does not affect the binding of proteins to receptors. Indeed, the clinical effects of glycosylated and non-glycosylated hGCSF on chemotherapy-induced neutropenia were not significantly different statistically in a clinical trial [63].

Conclusion

This study demonstrates that fusion proteins and a low expression temperature can be used to successfully express soluble hGCSF in the cytoplasm of *E. coli*. Using simple chromatographic techniques and TEV protease digestion, >10 mg of highly bioactive hGCSF was purified from 500 mL cultures of cells expressing MBP-hGCSF or PD1b'a'-hGCSF.

Supporting Information

Figure S1 Expression levels and solubilities of hGCSF fused with seven different tags in *E. coli* Origami 2(DE3). Protein expression was induced with 1 mM IPTG at

References

- Bath PM, Sprigg N (2006) Colony stimulating factors (including erythropoietin, granulocyte colony stimulating factor and analogues) for stroke. *Cochrane Database Syst Rev*: CD005207.
- Metcalf D (2010) The colony-stimulating factors and cancer. *Nat Rev Cancer* 10: 425–434.
- Demetri GD, Griffin JD (1991) Granulocyte colony-stimulating factor and its receptor. *Blood* 78: 2791–2808.
- Young DC, Zhan H, Cheng QL, Hou J, Matthews DJ (1997) Characterization of the receptor binding determinants of granulocyte colony stimulating factor. *Protein Sci* 6: 1228–1236.
- Horan T, Wen J, Narhi L, Parker V, Garcia A, et al. (1996) Dimerization of the extracellular domain of granulocyte colony stimulating factor receptor by ligand binding: a monovalent ligand induces 2:2 complexes. *Biochemistry* 35: 4886–4896.
- Zsebo KM, Cohen AM, Murdock DC, Boone TC, Inoue H, et al. (1986) Recombinant human granulocyte colony stimulating factor: molecular and biological characterization. *Immunobiology* 172: 175–184.
- Nicola NA, Metcalf D, Matsumoto M, Johnson GR (1983) Purification of a factor inducing differentiation in murine myelomonocytic leukemia cells. Identification as granulocyte colony-stimulating factor. *J Biol Chem* 258: 9017–9023.
- Roberts AW (2005) G-CSF: a key regulator of neutrophil production, but that's not all! *Growth Factors* 23: 33–41.
- Hollingshead LM, Goa KL (1991) Recombinant granulocyte colony-stimulating factor (rG-CSF). A review of its pharmacological properties and prospective role in neutropenic conditions. *Drugs* 42: 300–330.
- Ghalaut PS, Sen R, Dixit G (2008) Role of granulocyte colony stimulating factor (G-CSF) in chemotherapy induced neutropenia. *J Assoc Physicians India* 56: 942–944.
- Meurer K, Pitzer C, Teismann P, Kruger C, Goricke B, et al. (2006) Granulocyte-colony stimulating factor is neuroprotective in a model of Parkinson's disease. *J Neurochem* 97: 675–686.
- Pollari E, Savchenko E, Jaronen M, Kanninen K, Malm T, et al. (2011) Granulocyte colony stimulating factor attenuates inflammation in a mouse model of amyotrophic lateral sclerosis. *J Neuroinflammation* 8: 74.
- Schneider A, Kruger C, Steigleder T, Weber D, Pitzer C, et al. (2005) The hematopoietic factor G-CSF is a neuronal ligand that counteracts programmed cell death and drives neurogenesis. *J Clin Invest* 115: 2083–2098.
- Nomura H, Imazeki I, Oheda M, Kubota N, Tamura M, et al. (1986) Purification and characterization of human granulocyte colony-stimulating factor (G-CSF). *EMBO Journal* 5: 871–876.
- Lasnik MA, Porekar VG, Stalc A (2001) Human granulocyte colony stimulating factor (hG-CSF) expressed by methylotrophic yeast *Pichia pastoris*. *Pfluegers Archiv European Journal of Physiology* 442: R184–R186.
- Yamamoto A, Iwata A, Saitoh T, Tuchiya K, Kanai T, et al. (2002) Expression in *Escherichia coli* and purification of the functional feline granulocyte colony-stimulating factor. *Veterinary Immunology and Immunopathology* 90: 169–177.
- Rao DV, Narasu ML, Rao AK (2008) A purification method for improving the process yield and quality of recombinant human granulocyte colony-stimulating factor expressed in *Escherichia coli* and its characterization. *Biotechnology and Applied Biochemistry* 50: 77–87.
- Vanz AL, Renard G, Palma MS, Chies JM, Dalmora SL, et al. (2008) Human granulocyte colony stimulating factor (hG-CSF): cloning, overexpression, purification and characterization. *Microb Cell Fact* 7: 13.
- Wang C, Wang L, Geng X (2008) High recovery refolding of rhG-CSF from *Escherichia coli*, using urea gradient size exclusion chromatography. *Biotechnology Progress* 24: 209–213.
- Arokayan S (2011) Chemical synthesis and improved expression of recombinant human granulocyte colony-stimulating factor cDNA. *Genet Mol Res* 10: 2671–2678.
- Jiang Y, Jiang W, Qiu Y, Dai W (2011) Effect of a structurally modified human granulocyte colony stimulating factor, G-CSFa, on leukopenia in mice and monkeys. *J Hematol Oncol* 4: 28.

either 18°C (A) or 30°C (B). After sonication, 20 µg of each total protein was loaded onto a 10% Tris-tricine gel. The arrows indicate the hGCSF fusion proteins. M, molecular weight size marker; C, total protein before IPTG induction (control); T, total protein after IPTG induction; P, protein in the cell pellet after sonication; S, protein in the supernatant after sonication. (TIF)

Author Contributions

Conceived and designed the experiments: BHD HBR HC. Performed the experiments: BHD HBR PH. Analyzed the data: BHD HBR PH BKK HC. Wrote the paper: BHD HBR BKK HC.

- Li M, Fan H, Liu J, Wang M, Wang L, et al. (2012) High pH solubilization and chromatography-based renaturation and purification of recombinant human granulocyte colony-stimulating factor from inclusion bodies. *Applied Biochemistry and Biotechnology* 166: 1264–1274.
- Fahnert B, Lilie H, Neubauer P (2004) Inclusion bodies: formation and utilisation. *Advances in Biochemical Engineering/Biotechnology* 89: 93–142.
- Jeong KJ, Lee SY (2001) Secretory production of human granulocyte colony-stimulating factor in *Escherichia coli*. *Protein Expression and Purification* 23: 311–318.
- Jin H, Cantin GT, Maki S, Chew LC, Resnick SM, et al. (2011) Soluble periplasmic production of human granulocyte colony-stimulating factor (G-CSF) in *Pseudomonas fluorescens*. *Protein Expression and Purification* 78: 69–77.
- Kim CW, Han KS, Ryu KS, Kim BH, Kim KH, et al. (2007) N-terminal domains of native multidomain proteins have the potential to assist de novo folding of their downstream domains in vivo by acting as solubility enhancers. *Protein Science* 16: 635–643.
- Song JA, Han KY, Park JS, Seo HS, Ahn KY, et al. (2009) Human G-CSF synthesis using stress-responsive bacterial proteins. *FEMS Microbiology Letters* 296: 60–66.
- Wang Y, Wei A, Wang M, Wei X, Zhang C, et al. (2012) [Prokaryotic expression, purification and enzymatic properties of nuclease P1]. *Sheng Wu Gong Cheng Xue Bao* 28: 1388–1397.
- Lefebvre J, Boileau G, Manjunath P (2009) Recombinant expression and affinity purification of a novel epididymal human sperm-binding protein, BSPH1. *Mol Hum Reprod* 15: 105–114.
- Chen X, Wu J, Liu H, He Z, Gu M, et al. (2010) Approaches to efficient production of recombinant angiogenesis inhibitor rhVEG1-192 and characterization of its structure and antiangiogenic function. *Protein Sci* 19: 449–457.
- Busso D, Delagoutte-Busso B, Moras D (2005) Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in *Escherichia coli*. *Analytical Biochemistry* 343: 313–321.
- Song JA, Koo BK, Chong SH, Kwak J, Ryu HB, et al. (2013) Expression and purification of biologically active human FGF2 containing the b'a' domains of human PDI in *Escherichia coli*. *Applied Biochemistry and Biotechnology* 170: 67–80.
- Liu S, Tobias R, McClure S, Styba G, Shi Q, et al. (1997) Removal of endotoxin from recombinant protein preparations. *Clinical Biochemistry* 30: 455–463.
- Weinstein Y, Ihle JN, Lavu S, Reddy EP (1986) Truncation of the c-myc gene by a retroviral integration in an interleukin 3-dependent myeloid leukemia cell line. *Proc Natl Acad Sci U S A* 83: 5010–5014.
- Hara K, Suda T, Suda J, Eguchi M, Ihle JN, et al. (1988) Bipotential murine hematopoietic cell line (NFS-60) that is responsive to IL-3, GM-CSF, G-CSF, and erythropoietin. *Exp Hematol* 16: 256–261.
- Chung HK, Kim SW, Byun SJ, Ko EM, Chung HJ, et al. (2011) Enhanced biological effects of Phe140Asn, a novel human granulocyte colony-stimulating factor mutant, on HL60 cells. *BMB Rep* 44: 686–691.
- Villaverde A, Carrio MM (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 25: 1385–1395.
- Bach H, Mazor Y, Shaky S, Shoham-Lev A, Berdichevsky Y, et al. (2001) *Escherichia coli* maltose-binding protein as a molecular chaperone for recombinant intracellular cytoplasmic single-chain antibodies. *Journal of Molecular Biology* 312: 79–93.
- Kapust RB, Waugh DS (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8: 1668–1674.
- Liu Y, Zhao TJ, Yan YB, Zhou HM (2005) Increase of soluble expression in *Escherichia coli* cytoplasm by a protein disulfide isomerase gene fusion system. *Protein Expression and Purification* 44: 155–161.
- Kurokawa Y, Yanagi H, Yura T (2001) Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in *Escherichia coli*. *Journal of Biological Chemistry* 276: 14393–14399.

42. Appenzeller-Herzog C, Ellgaard L (2008) The human PDI family: versatility packed into a single fold. *Biochimica et Biophysica Acta* 1783: 535–548.
43. Davis GD, Elisee C, Newham DM, Harrison RG (1999) New fusion protein systems designed to give soluble expression in *Escherichia coli*. *Biotechnology and Bioengineering* 65: 382–388.
44. Diaz AA, Tomba E, Lennarson R, Richard R, Bagajewicz MJ, et al. (2010) Prediction of protein solubility in *Escherichia coli* using logistic regression. *Biotechnology and Bioengineering* 105: 374–383.
45. Shirano Y, Shibata D (1990) Low temperature cultivation of *Escherichia coli* carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level. *FEBS Letters* 271: 128–130.
46. Ferrer M, Chernikova TN, Timmis KN, Golyshin PN (2004) Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Applied and Environmental Microbiology* 70: 4499–4504.
47. Vera A, Gonzalez-Montalban N, Aris A, Villaverde A (2007) The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering* 96: 1101–1106.
48. Kim EK, Moon JC, Lee JM, Jeong MS, Oh C, et al. (2012) Large-scale production of soluble recombinant amyloid-beta peptide 1–42 using cold-inducible expression system. *Protein Expression and Purification* 86: 53–57.
49. Yi AR, Lee SR, Jang MU, Park JM, Eom HJ, et al. (2009) Cloning of dextranase gene from *Leuconostoc citreum* HJ-P4 and its high-level expression in *E. coli* by low temperature induction. *J Microbiol Biotechnol* 19: 829–835.
50. Imsoonthornruksa S, Noisa P, Parnpai R, Ketudat-Cairns M (2011) A simple method for production and purification of soluble and biologically active recombinant human leukemia inhibitory factor (hLIF) fusion protein in *Escherichia coli*. *Journal of Biotechnology* 151: 295–302.
51. Phadtare S (2004) Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* 6: 125–136.
52. Kandror O, Goldberg AL (1997) Trigger factor is induced upon cold shock and enhances viability of *Escherichia coli* at low temperatures. *Proceedings of the National Academy of Sciences of the United States of America* 94: 4978–4981.
53. Baneyx F, Mujacic M (2004) Recombinant protein folding and misfolding in *Escherichia coli*. *Nature Biotechnology* 22: 1399–1408.
54. Conzelmann N, Schneider A (2011) A screen for peptide agonists of the G-CSF receptor. *BMC Res Notes* 4: 194.
55. Bai Y, Ann DK, Shen WC (2005) Recombinant granulocyte colony-stimulating factor-transferrin fusion protein as an oral myelopoietic agent. *Proceedings of the National Academy of Sciences of the United States of America* 102: 7292–7296.
56. Magistrelli G, Gueneau F, Mushmani M, Ravn U, Kosco-Vilbois M, et al. (2005) Chemokines derived from soluble fusion proteins expressed in *Escherichia coli* are biologically active. *Biochemical and Biophysical Research Communications* 334: 370–375.
57. Gascon P (2012) Presently available biosimilars in hematology-oncology: G-CSF. *Target Oncol* 7 Suppl 1: S29–34.
58. Haniu M, Horan T, Arakawa T, Le J, Katta V, et al. (1995) Extracellular domain of granulocyte-colony stimulating factor receptor. Interaction with its ligand and identification of a domain in close proximity of ligand-binding region. *Archives of Biochemistry and Biophysics* 324: 344–356.
59. Cox GN, Smith DJ, Carlson SJ, Bendele AM, Chlipala EA, et al. (2004) Enhanced circulating half-life and hematopoietic properties of a human granulocyte colony-stimulating factor/immunoglobulin fusion protein. *Experimental Hematology* 32: 441–449.
60. Wingfield P, Benedict R, Turcatti G, Allet B, Mermod JJ, et al. (1988) Characterization of recombinant-derived granulocyte-colony stimulating factor (G-CSF). *Biochemical Journal* 256: 213–218.
61. Roessl U, Wiesbauer J, Leitgeb S, Birner-Gruenberger R, Nidetzky B (2012) Non-native aggregation of recombinant human granulocyte-colony stimulating factor under simulated process stress conditions. *Biotechnol J* 7: 1014–1024.
62. Young DC, Zhan H, Cheng QL, Hou J, Matthews DJ (1997) Characterization of the receptor binding determinants of granulocyte colony stimulating factor. *Protein Science* 6: 1228–1236.
63. Bonig H, Silbermann S, Weller S, Kirschke R, Korholz D, et al. (2001) Glycosylated vs non-glycosylated granulocyte colony-stimulating factor (G-CSF)—results of a prospective randomised monocentre study. *Bone Marrow Transplantation* 28: 259–264.