Simple and efficient expression of codon-optimized mouse leukemia inhibitory factor in Escherichia coli

Dewei Niu1, Qiuhong Wu1,2, Luyuan Huang3, Enze Yang1, Meiya Huang1, Yong Liu4 and Feng Wang1*

1Institute of Genomic Medicine, College of Pharmacy, Jinan University, Guangzhou 510632, 2Department of Pharmacy, Maternal and child health care Hospital of Zaozhuang, Zaozhuang 277100, 3Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, Key Laboratory of Regenerative Biology, South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, 4Guangdong Key Laboratory of Agro-Environment Integrated Control, Guangdong Institute of Eco-Environmental and Soil Sciences, Guangzhou 510650, China

*For correspondence: Email: jnubiopharm@126.com; Tel: 0086-15360022165

Received: 9 September 2016 Revised accepted: 10 March 2017

Abstract

Purpose: To obtain a higher yield of mouse leukemia inhibitory factor to maintain the proliferation potential of pluripotent stem cells at a low cost.

Methods: A method was designed to produce recombinant mLIF protein (rmLIF) in Escherichia coli. Through analysis of rmLIF sequence, it was found that rare codons were interspersed. After mutation from rare codons to Escherichia coli (E. coli) preferred ones were selected, the mutated gene mLIFm was cloned into pET15b vector. The pET15b-mLIFm was then transformed into Rosetta-gami strain and induced with optimal conditions at 18 °C for 16 h. Mass spectrometry was carried out to identify the peptides.

Results: After purification, the yield of the codon-optimized rmLIFm was 141 mg/L, compared with 110 mg/L for the original rmLIF. Mass spectral analysis showed the presence of four major peptides each with an intensity > 10 % at m/z 1031.57, 1539.82, 1412.01 and 2229.10 in mLIFm, respectively. His-tagged rmLIFm fusion protein displayed the potential to maintain the morphology of undifferentiated mouse embryonic stem cells (mESCs), which were positive for mESCs markers (Oct-4, Nanog, Sox-2, stage-specific embryonic antigen-1).

Conclusion: The findings provide a means to produce mLIF in a short, useful, cost-effective and environmentally-friendly manner, and thus lays a foundation for further studies of mLIF.

Keywords: Leukemia inhibitory factor, Mutated gene, Protein expression, Purification, Stem cells, Peptides, Escherichia coli

INTRODUCTION

Leukemia inhibitory factor (LIF) is a pleiotropic glycoprotein and a member of interleukin (IL)-6 family. It induces mouse myeloid leukemic M1 cells of terminal differentiation into macrophages [1]. A number of studies have shown that LIF is a multi-purpose cytokine with range of pleiotropic actions on different types of cells and tissues. Not only does it facilitate early embryogenesis and blastocyst implantation of mammals, it also induces the production of acute phase proteins by lipocyte formation and hepatocytes [2,3]. Furthermore, LIF is known as a significant growth factor for maintaining embryonic stem cells and inducing pluripotent stem cells in a pluripotent
status [4]. LIF is routinely added to the medium of pluripotent stem cells and without LIF leads to a fast differentiation of pluripotent stem cells. Therefore, it is extensively employed in the culture of pluripotent stem cells [5]. The cells play a primary role in cell therapies and tissue replacement therapies [6,7]. Recently, rhLIF was successfully produced and purified by several groups using different kinds of expression systems such as E. coli [8-10], yeast cells and even in rice [11]. An issue arising from this process is the soaring cost of LIF and other cytokines in pluripotent stem cells research. Numerous studies have suggested that the cost of commercially produced LIF accounts for about 90 % of pluripotent stem cells culture medium [8]. It is urgent to obtain various forms of LIF in high quantities and at a low cost and biological active.

E. coli is a commonly used host for expression of heterologous protein, but has a marked preference for certain codons. Differences in preference for codons between prokaryotes and eukaryotes can seriously affect protein production in E. coli. Rare codons exist in cloned genes that can affect many aspects such as protein expression and mRNA. Moreover, rare codons correlate with a small number of homologous tRNA in host cells.

In the present study, a cDNA encoding LIF from mice was cloned and sequenced. Subsequently, we developed an effective method to produce soluble rmLIF in E. coli after mutation from rare codons to E. coli preferred ones was performed. This method allowed rapid and effective purification of rmLIF in a soluble and biologically active form.

**EXPERIMENTAL**

**Reagents**

T₄ DNA ligase, Sma I and SYBR were purchased from Takara (Tokyo, Japan). FBS, L-glutamine and β-mercaptoethanol, minimum essential amino acids were purchased from Gibco (Karlsruhe, Germany). Goat anti-mouse IgG, Goat anti-mouse IgGs (containing horseradish peroxidase label), anti-Oct-4 and anti-Nanog were purchased from Santa Cruz Biotechnology (Santa Cruz, USA). E. coli cells were purchased from Novagen (New York, USA). The other reagents were commercially available.

**Analysis of codons and amino acid sequences of mLIF**

BLASTP program was used to search for homologs of mouse LIF from other species. Different LIF sequences were compared using the Clustal X program. The rare codons of mLIF were analyzed. In addition, the relationship between mouse LIF and other previously described mammal LIF were showed by a phylogenetic tree, which was constructed using the program Phylip-3.69.

**Cloning the coding sequence of mLIF into expression vector**

PCR was used to amplify the coding region of the mouse leukemia inhibitory factor with the cDNA of mouse embryonic fibroblast (MEF) cells as a template. The forward primers (5'-GGTCTGCACTGGAACACGG-3') and the reverse (5'-CTAGAAGGCTGGACCACCG-3') were designed based on the mRNA sequence of mouse leukemia inhibitory factor (NM_008501).

The mLIF sequence which mutated from rare codons to E. coli preferred ones, was directly synthesized by Invitrogen Company. The mLIF and mLIF were ligated into the modified expression vector pET15b (Sma I) using T4 DNA ligase.

**Expression, purification and western blot analysis of rmLIF**

E.coli Rosetta-gami cells harboring expression plasmids were incubated at 37 °C until OD₆₀₀ reached 0.6. IPTG was then added at a concentration of 0.5 mM. Subsequently, the culture was shaken constantly at 180 rpm for 16 h at 18 °C. The supernatant and precipitate were harvested by centrifugation (12,000 rpm for 10 min). The supernatant was loaded onto a 5 mL Ni-NTA-His binding column equilibrated with 10 mM imidazole (20 mM sodium phosphate pH7.4, 500 mM NaCl). The protein bound matrix was washed with 20 mM imidazole (20 mM sodium phosphate pH 7.4, 500 mM NaCl). The protein was eluted with 250 mM imidazole (20 mM sodium phosphate pH 7.4, 500 mM NaCl). The concentrated sample was tested by BCA assay.

Protein sample of the rmLIF was separated on SDS-PAGE, followed by transfer onto a PVDF membrane. The PVDF membrane was blocked for 1.5 h at 25 °C. It was then incubated with mouse anti-His IgG antibody (1:1000) at room temperature. After 2 h, the PVDF membrane was incubated with goat anti-mouse IgG conjugated with horseradish peroxidase (1:1000) for 1.5 h at room temperature. After incubation, the specific protein bands were developed using the diaminobenzidine (DAB) system.
Mass spectrometry analysis

Sequencing grade modified trypsin (Promega, Madison, USA) was used to digest samples, matrix-assisted laser desorption/ionization time-of-flight mass (MALDI-TOF-MS) was used to analyze samples by an ABI 4800 Plus MALDI TOF/TOF Analyzer (Invitrogen, Carlsbad, USA). IPI mouse database was used to analyze the result (V3.36) by GPS Explore software (V3.6, Invitrogen, Carlsbad, USA).

Maintenance of mouse embryonic stem cells (mESCs) in culture

We expanded the mouse embryonic stem cells line OG2 which was cultivated on mitomycin C handled mouse embryonic fibroblast (MEF) feeders. The culture medium was composed of DMEM (HyClone, Bonn, Germany) supplemented with 200 mM L-glutamine, 0.1 mM sodium bicarbonate, 0.1 mM β-mercaptoethanol, 0.1 mM nonessential amino acids, 10 % FBS (all from Gibco, Karlsruhe, Germany) and 100 U/mL penicillin/streptomycin (Invitrogen, Camarillo, USA). The purified rmLIF was added to the culture medium at a final concentration of 10 ng/mL. 1000 U/mL of commercially available mouse leukemia inhibitory factor (Millipore, Billerica, USA) was added as positive control to compare function and efficiency [12].

Real-time PCR

RNA was extracted from 5 passages of mESCs with each mLIF supplementation. The expression of embryonic stem cell markers was investigated by Real-time PCR and primers were designed as: Oct-4 (FP- TAGGTGAGCCGCTTTCCAC, RP- GCTTACCAGTTTGAGGAT), Sox-2 (FP- AGGGCTGGGAGAAAGAAG, RP- CCGGAT TTGTGTGATTAGT), Nanog (FP- CTCAGTGCC AGGCTGACA, RP- CCGGATTTGAGGAT), GAPDH (FP- AACTTTGCGCATTTGGA AGGGCTCA, RP- TTGGCAGCCACAGTGATG CAGGGA). For data normalization, the GAPDH was used as a reference gene to analyze the gene quantitatively. The reaction efficiency of primers was estimated by a standard curve. The 2^ΔΔCt method was used to calculate the expression levels of relative gene [13].

Immunofluorescence

At the tenth passage of mESC culture (30 - 40 days), cell morphology and the expression of ESC markers by immunofluorescence were used to ensure rmLIF functionality. The undifferentiated state of mESC was determined by immunofluorescence staining with Oct-4 (1:250; Santa Cruz Biotechnology, Santa Cruz, USA), Nanog (1:50; Santa Cruz Biotechnology, Santa Cruz, USA), Sox-2 (1:100; Stem Cell Technologies, Vancouver, Canada), and stage-specific embryonic antigen-1 (1:50; Chemicon, Billerica, USA) 4 % paraformaldehyde was used to fixed the cells. After fixed for 10 min, the cells were washed with PBS 3 times, incubated with blocking buffer containing 0.1% Triton X-100 and 10 % FBS in a humidity chamber at 37 ºC. After 1 h, primary antibody was diluted in blocking buffer and applied for 2 h at 37 ºC; then washed three times. After incubation with secondary antibodies conjugated with TRITC (1:200, Sigma, Santa Cruz, USA) for 1 h at 37 ºC, they were washed twice in PBS, followed by incubation with PBS containing DAPI (10 μg/mL) for 5 min. DAPI stain solution was washed, fluorescence microscope was used to obtain pictures.

Statistical analysis

Data were expressed as mean ± standard deviation. The relative expression of Nanog, Oct-4, and Sox-2 relative expressions were analyzed using Student’s t-test, only if p < 0.05 the differences were considered statistically significant. SPSS for Windows software (version 13.0; SPSS Inc, Chicago, IL, USA) was used to analysis the data.

RESULTS

Sequence of mLIF

Bioinformatic analysis showed that the mLIF gene contains a 612-bp full open reading frame which encoded a 22.3 kDa protein with a pI of 9.39.

A signal peptide cleavage site was found between positions 23 and 24 of the mLIF sequence based on SignalP database. Therefore, it encoded a supposed 180 amino-acid mature mLIF of 19.9 kDa with a pI of 9.2.

The results revealed that full mLIF amino acid sequences showed 75%-85 % identities to those from Rattus norvegicus, Bos taurus, Homo sapiens, Pteropus alecto, Sorex araneus, Macaca mulatta (Figure 1A). The phylogenic tree of LIF was established according to the amino acid sequence of selected LIF from different species (Figure 1B), mLIF was initially clustered with Rattus norvegicus LIF.

Rare codons

Rare codons in mLIF sequence were searched in Rare Codons Calculator (Figure 2A). There were
four rare Pro codons at positions 13, 62, 115 and 157, three rare Ile codons at positions 23, 128 and 182, two rare Arg codons at positions 132 and 167 and one rare Leu codon at position 68 in the 570 bp mLIF sequence encoding mature mLIF. All 10 rare codons in the mLIF sequence were replaced by the preferred codons of *E. coli*, generating the mLIF<sup>m</sup> sequence (Figure 2B).

**Figure 1:** (A) Alignment of amino acid sequences of LIF for mouse and other species. (B) A phylogenetic tree of LIF from different species. It was constructed by Phylip-3.69 program. Numbers represent the values given by bootstrap analysis. The species names (in parenthesis) of different LIF are shown on the right.

**Figure 2:** (A) Rare codons in mLIF sequence (B) Rare codon replaced by preferred codons of *E. coli* in mLIF<sup>m</sup> sequence. The codons marked in red were the modified rare codons.
Expression, purity and characteristics of rmLIF<sup>m</sup>

The best optimized conditions for expression of recombinant protein in soluble form was acquired by induction with 0.5 mM IPTG at 18 °C for 16 h. Induced culture was harvested which produced 3.2 g wet weight cells per 500 ml culture. After sonication, the supernatant and pellet was analyzed by SDS-PAGE (Figure 3A). Both rmLIF and rmLIF<sup>m</sup> showed an apparent band about 23.55 kDa after induction. After reforming the rare codons soluble form (Lane 2, Figure 3A) accounts for 35.2 % of total expressed soluble rmLIF<sup>m</sup> in the supernatant of the lysate, compared with rmLIF accounted for only 29.1 % (Lane 3, Figure 3A).

The supernatant was purified by Ni-NTA affinity chromatography. The results showed that the protein had been purified with a purification yield of 97.5 % according to its migration (Figure 3A). The presence of the protein was confirmed with anti-His-tag by Western blot (Figure 3B). As noted in the present study, a tag-based protein purification method was used. The expression level of rmLIF<sup>m</sup> approximated 141 mg/L. Its production increased by one third when compared with the original rmLIF.

Mass spectra

Figure 4 shows the mass spectrum of trypsin digest of the purified rmLIF<sup>m</sup>. There were four major peptides at 1031.57, 1539.82, 1412.01 and 2229.10 m/z. The strongest peak was appeared on 1412.01, which fits to the predicted molecular weight of HGAGSPLPITPVNATCAI peptide fragment. This result was a further confirmation that rmLIF<sup>m</sup> protein was successfully expressed.

Figure 4: MALDI-TOF mass spectrum of purified rmLIF<sup>m</sup>. The purified rmLIF<sup>m</sup> was determined on SDS-PAGE gel and the rmLIF<sup>m</sup> in a piece of gel was subjected to in-gel digestion with trypsin. The resulting polypeptide was analyzed by MALDI-TOF MS.
Biological activity

The biological activity of rmLIF<sup>m</sup> could be identified on account of its ability to maintain the proliferation potential of pluripotent stem cell after numerous passages while retaining their pluripotency [12,14]. The results in Figure 5A indicated that the group supplied with rmLIF<sup>m</sup> expressing the ESCs markers were similar to the positive control group. Importantly, no ESCs markers expression was detected in MEF which were used as a negative control. In addition, after 10 passages in culture, compared with the mESC colonies not treated with rmLIF<sup>m</sup>, the group treated with rmLIF<sup>m</sup> fusion protein maintained cell structure integrity, the shape is circular with an intact nucleus, mESC colonies not treated with rmLIF<sup>m</sup> were varied and most presented nested or island shapes (Figure 5B). The mESC colonies added to rmLIF<sup>m</sup> fusion protein were further detected by immunofluorescence staining with nuclear markers, along with the stage-specific embryonic antigen-1 (SSEA-1), and the results showed that the four stem cell factors were all induced by rmLIFm (Figure 5C).

Figure 5: (A) Realtime-PCR analysis of the pluripotent markers, Nanog, Oct-4, and Sox-2 relative expression over GAPDH in mESCs after 5 passages in culture with commercial mLIF, rmLIF<sup>m</sup> supplemented media; (B) Morphology of mESC colonies (phase contrast) after 10 passages in medium containing 10 ng/mL rmLIF<sup>m</sup> (left) and not treated with rmLIFm (right); (C) Immunofluorescence imaging of stem cell markers Sox-2 (upper left), Oct-4 (upper right), Nanog (below left) and SSEA-1 (lower right), respectively. DAPI was used to counterstain cell nucleus, Scale bar, 10 μm
DISCUSSION

To date, a large number of expression systems were used to express LIFs. Among them, majority is hLIF. Low expression yields, misfolds and aggregates, however, prevented laboratories from applying LIF.

Sequence analysis of mLIF demonstrated that there were 10 rare codons in the 570 bp mLIF sequence encoding mature mLIF which would lower mLIF protein expression levels. In order to overcome the problem of rare codons, we optimized it from rare codons to E. coli preferred ones. Although E. coli is a helpful host for expression of heterologous protein, differences in codon usage preference between the original organism and E. coli could lead to significant reduction in expression level [15]. Similar reports showed that replacing rare codon increased the expression of beta-G glucosidase by 2- and 3.6-fold [16].

A major drawback would be that expression of heterologous protein in large quantities in E. coli often leads to the formation of inclusion bodies [17]. Most of the protein were expressed as inclusion bodies at 37 ºC. Along with the form of inclusion bodies, there are usually complicated and unstable refolding procedures which tend to bring about low purification yields [18]. To solve this problem, we used the optimal induction temperature at 18 ºC, the protein was expressed mostly as a soluble form. Lower temperature makes it easy to form correct protein folding and increase protein solubility [19]. On the other hand, mLIFm protein contains three disulfide bonds formed by six conserved cysteine residues [20]. The Rosetta-gami host strain also carried the trxB/gor mutations, which could enhance disulfide bond formation. Our study also demonstrates that the rmLIFm produced had correct folding structure. Furthermore, this approach took advantage of the small flexible N-terminal six His affinity tag system for easy purification of fusion protein tagged with a polyhistidine sequence. Compared with other tags, Glutathione S-transferase (GST) and thioredoxin (Trx) which can improve solubility and yields, it is small enough to not alter functional properties of the tagged proteins and maintain biological activity of recombinant proteins.

As noted in the present study, we used an affinity tag-based purification method to make the purification easy. Meanwhile, soluble mLIFm protein can easily be obtained and its expression level rose to approximately 141 mg/L per 1 L medium. In contrast with previous study, the yield of rmLIFm protein was higher. 1.3 mg hLIF were got from 3.8 g cells (wet weight) of 1 L culture by Jung et al [21]. About 30 mg untagged hLIF per 1L EnBso Flo culture was achieved after one-step eXact™ affinity chromatographic purification by Imaizumi et al [10]. Furthermore, Western blot and mass spectrometry were used to identify rmLIFm fusion protein.

As demonstrated, purified rmLIFm was helpful in maintaining the multiplication potential of mESCs throughout numerous passages. These results indicated that mESCs with rmLIFm protein treatment showed common multipotent markers. The undifferentiation potential of pluripotent stem cells can be maintained after long-term culture with rmLIFm. More importantly, His-tag portion showed no significant effect on the bioactivity of rmLIFm. Production is cost-effective in bacteria culture. In addition, the protein purification methodology is simple, practical and cost-effective which may decrease rmLIFm costs to a large extent.

CONCLUSION

By adopting optimizing codons, using small molecular weight His tags and prokaryotic host Rosetta-gami, a straightforward and effective method to produce large amounts of soluble mLIFm is presented. The strategy not only provides an effective and fast way to produce rmLIFm which is not only helpful in maintaining the multiplication potential of mESCs and retaining their pluripotency, but also might be useful for the production of other growth factors.

DECLARATIONS

Acknowledgement

This research was supported by Science and Technology Planning Project (no. 2012KJCX0017).

Conflict of Interest

No conflict of interest associated with this work.

Contribution of Authors

The authors declare that this work was done by the authors named in this article and all liabilities pertaining to claims relating to the content of this article will be borne by them.

Open Access

This is an Open Access article that uses a fund-

REFERENCES


