A COMPARATIVE STUDY BETWEEN EXPRESSION OF A SYNTHETIC GENE OF HUMAN BASIC FIBROBLAST GROWTH FACTOR (hbFGF) AND ITS RELATED cDNA IN *ESCHERICHIA COLI*

H. Mirzahoseini*, F. Mehraein and F. Mozafari

Department of Biochemistry, Pasteur Institute of Iran, Tehran, Islamic Republic of Iran

Abstract

The gene encoding the human basic fibroblast growth factor (hbFGF) has been already chemically-synthesized and cloned in pET-3a expression vector (Pasteur Institute of Iran). In the present study, we compared the level of expression of this synthetic hbFGF and its related cDNA in *Escherichia coli*. The pBR322-cDNA of hbFGF supplied by Dr. Seno (from Molecular Biology Dept, Okaido prefectural university, Japan), was utilized to introduce two restriction sites (*Bgl*II and *Nde*I) in either ends of the gene by PCR. This PCR product was inserted into *Sma*I site of pUC18 plasmid (pUC-1003). This construct was inserted into *Bam*HI site of pET-3a expression vector (pET-1004). A final construct for expression of the gene was also made (pET-1005). The level of expression of the synthesized gene was compared to that of the cDNA using SDS-PAGE, western-blotting or ELISA. The results showed that expression of the hbFGF cDNA was much higher than that of the related synthetic gene, (16 mg/l and 0.025 mg/l respectively). Although in designing the synthetic gene the "codon usage" of *E. coli* was considered, it seems that "codon usage" did not improve the level of expression of the gene as was expected.

Introduction

Basic fibroblast growth factor (bFGF) or FGF-2 belongs to the FGF family, which consists of twenty-one (GenBank, 2001) structurally related polypeptides that promote the proliferation of cells of mesodermal and neuroectodermal origin [1]. It is an important factor during embryonic development, because it stimulates mitosis of different types of cells, such as fibroblast, and it also enhances angiogenesis. This protein makes profound interest in the pharmaceutical industry due to several potential applications, such as wound healing and redirection of blood vessels [2]. In contrast to the low concentration of hbFGF in natural sources such as

Keywords: hbFGF; Synthetic gene; cDNA; Codon usage

A:

ATGCCAGCTCTGCCGGAAGACGGTGGTTCTGGCGCCTTCCCGCCCGGCCAACTTCAAGGACCCCAAGCGTCTGTACTGCAAAAAC
GGTGGTTTCTTCCTGCGTATCCACCCGGACGGTCGTGTTGACGGTGTACGTGAAAAATCCGACCCGCACATCAAACTCCAACTGC
AGGCTGAAGAACGTGGTGTTGTTTCCATCAAAGGTGTTTCTGCTAACCGTTACCTGGCTATGAAAGAAGACGGTCGTCTGCTGGC

* *E-mail: mirzahoseini@yahoo.com*

```
TTCCAAATCTGTTACCGACGAATGCTTCTTCTTTGAACGTCTAGAATCCAACAACTACAACACCTACCGTTCCCGTAAATACACC
TCCTGGTACGTTGCTCTGAAACGTACCGGTTCAGTACAAACTGGGTTCCAAAACCGGTCCGGGTCAGAAAGCTATCCTGTTCCTG
CCGATGTCCGCTAAATCCTAA
                  ***
```

B:

```
        BglII NdeI
TAAATTAAGATCTCATATGCCCGCCTTGCCCAGGATGGCGGCAGCGGCGCCTTCCCGCCCGGCCACTTCAAGGACCCCAAGCGGC
TGTACTGCAAAAACGGGGGCTTCTTCCTGCGCATCCACCCCGACGGCCGAGTTGACGGGGTCCGGGAGAAGAGCGACCCTCACAT
CAAGCTACAACTTCAAGCAGAAGAGAGAGGAGTTGTGTCTATCAAAGGAGTGTGTGCTAACCGTTACCTGGCTATGAAGGAAGAT
GGAAGATTACTGGCTTCTAAATGTGTTACGGATGAGTGTTTCTTTTTTGAACGATTGGAATCTAATAACTACAATACTTACCGGT
CAAGGAAATACACCAGTTGGTATGTGGCACTGAAACGAACTGGGCAGTATAAACTTGGATCCAAAACAGGACCTGGGCAGAAAGC
TATACTTTTTCTTCCAATGTCTGCTAAGAGCTGATTTTAATGGCCAAGATCTAATAGTA
                               ***          BglII
```

Figure 1. Sequences of synthetic hbFGF gene in favor of codon usage of *E. coli* (A), and cDNA hbFGF after Site-directed mutagenesis with PCR (B).

primary endothelial cells, fibroblasts, glial cells and smooth muscle cells, the production of recombinant bFGF in *E. coli* was found to be between 2.3-3.4 g/l cell culture [2].

In the current work, the correlation between the level of expression of hbFGF and *E. coli* codon preference was studied. It was shown that although the codon preferences of *E. coli*, *BL21(DE3)* for expression of bFGF was considered, the level of expression was not higher than when cDNA of bFGF was used for expression.

Materials and Methods

Preparation of Synthetic hbFGF

The synthetic hbFGF, which was designed in favor of "codon usage" of *E. coli* (Fig. 1), was supplied by Dr. F. S. Mozafari [3]. Cysteines in positions 69 (#209-211) and 87 (#263-265) were replaced by serines, creating TCT codons instead of TGC codons for increasing the stabilization of the protein [4]. The synthetic gene was cloned into pET-3a expression vector (Novagene), to construct pET-1001.

Preparation of cDNA hbFGF

cDNA of the hbFGF in pBR322 (pBO124) was provided by Dr. Seno (from Japan). In this cDNA, cysteines in positions 69 and 87 were also replaced by serines.

Site-Directed Mutagensis with PCR

Unique *Nde*I and *Bgl*II sites were added to the either ends of cDNA by PCR (Fig. 1). PCR was carried out using a forward primer 25bp (5' TAAATTAAGATCTCATATGCCCGCC 3') and a reverse primer 30bp (5'TACTATTAGATCTTGGCCATTAAAATCAGC3') [made in Institute Bioorganic Chemistry, Moscow]. After 35 cycles of PCR amplification, the product was phenol/chloroform extracted, concentrated, and was run on 1% agarose gel.

Subcloning of cDNA in pUC18

The band which was corresponded to 450 bp, was purified from the gel and then subcloned into *Sma*I site of the plasmid pUC18, pUC-1003, (Fig. 2). DNA sequencing was performed on this construct to confirm the authenticity of its structure by dideoxy-chain-termination method [5].

Subcloning of cDNA in pET-3a

The cDNA sequence for hbFGF was digested with *Bgl*II and inserted into the compatible *BamH*I site of the plasmid pET-3a, downstream of a strong *T7* promoter to make pET-1004. pET-1004 was digested with *Nde*I and then relegated, the new construct was named pET-1005 in which 40 bp sequence between two *Nde*I sites in pET-1004 was deleted. Therefore, pET-1005 has a unique *Nde*I site (Fig. 2).

Expression of cDNA and synthetic hbFGF

BL21-DE3-plysS E. coli cells were transformed with pET-1005 and pET-1001 (synthetic gene). Strain *BL21(DE3)* is a lysogen of a λ phage containing *T7* gene 1 (coding T7 RNA polymerase) that is controlled by the *lacUV5* promoter-operator sequence and is inducible with isopropylthiogalactoside (IPTG) [6].The cells carrying pET-1001 or pET-1005 and the cells bearing pET-3a alone (as a negative control) were cultured in LB media containing 50 μg/ml ampicillin. When the A600 reached to 0.5, cultures was induced with 0.5 mM IPTG for approximately 3h at 37°C. The cells were collected by centrifugation and were lysed by sonication (3 times, 30 sec).

SDS-PAGE and Immunodetection

An equal volume of 2x protein loading buffer was added to 30 µl of each sample and boiled for 10 min. 20 µl of each sample was run on a 18% polyacrylamide gel, when the tracking dye reached to the bottom of the gel, electrophoresis was stopped and the gel was used for western blot analysis.

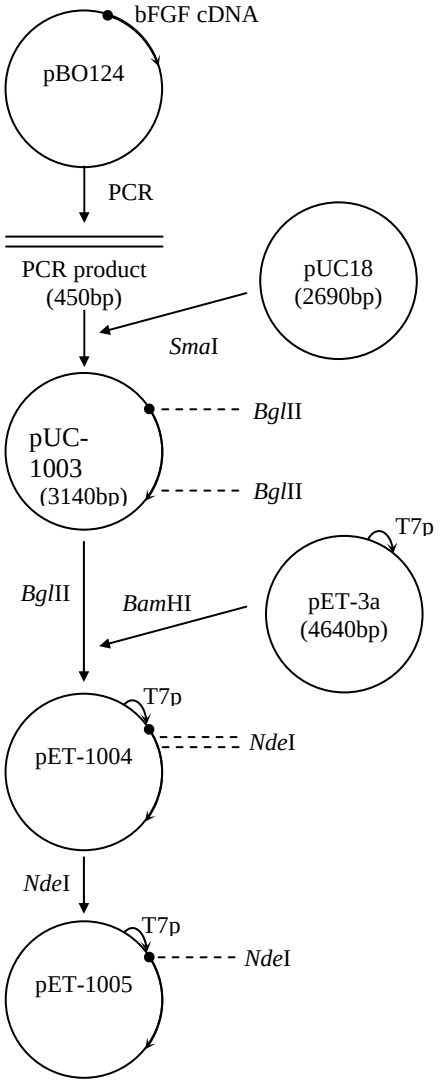

Figure 2. Schematic diagram for constructions of pUC-1003, pET-1004 and pET-1005 (final construction plasmid expressing hbFGF).

For western immunoblot analysis, the proteins which were separated on the gel was transformed onto a nitrocellulose membrane (0.2 µm pore size, *Schlecher & Schuell*). To block non specific binding, the filter was left for 1 h in PBS containing 0.5% bovine serum albumin (BSA). The first antibody (Ab), with 1/4000 dilution of the rabbit anti-hbFGF anti serum (SIGMA) in buffer A (0.5% BSA in PBS and 0.05% Tween 20) was added and left over night at 4°C. After 3 times of washing with buffer A, the second Ab, with 1/5000 dilution of Horse radish peroxidase conjugated goat-anti-rabbit Ab (SIGMA), was added and left at room temperature for 2 hours followed by three times washing with buffer A. The immunoreaction was visualized using 0.5 mg/ml diaminobenzidine HCl and 0.01% H2O2.

For quantitative measurement of hbFGF, an ELISA kit (R&D, USA) was used which includes a standard hbFGF for plotting a standard curve.

Results

Subcloning of cDNA of hbFGF

We received pBO124 (pBR322-cDNA of hbFGF) from Dr. Seno, (Japan). pUC-1003 , which is a derivative of this plasmid, was made by inserting the amplified product into *Sma*I site of pUC18 (as mentioned in Material and Methods). The authenticity of the cDNA sequence was confirmed by dideoxy-chain-termination method [5].

PET-1004 which is a derivative of pET-3a was also constructed in this study. In pET-1004, there were two recognition sites for *Nde*I (one of these sites was 40 bp upstream of the other). In order to remove one of these sites, pET-1005 was constructed. In this plasmid the 40bp distance between two *Nde*I sites was deleted. Therefore pET-1005 had a unique *Nde*I site.

Expression and Detection

Cells containing pET-1001, pET-1005 and pET-3a (as a negative control) were grown and induced with IPTG. The cells were then centrifuged, sonicated and were run on a 18% SDS-PAGE. The result from SDS-PAGE (Fig. 3) showed that there is an intense band corresponding to hbFGF (pET-1005-cDNA). However, the related band for pET-1001-synthetic gene was very faint. The results of western blotting also confirmed the result obtained from SDS-PAGE (Fig. 4).

We have also determined the amount of proteins in each system by ELISA. The results showed that the production of pET-1005-cDNA and pET-1001-synthetic gene were 16 mg/l and 0.025 mg/l, respectively.

Discussion

Since little amount of hbFGF can be purified from native sources by biochemical methods, it is desirable to produce sufficient quantities of recombinant protein in order to be able to study its biological function and for preclinical studies. Two approaches were adopted for expression of the protein in *E. coli*; chemical synthesis of the gene in favor of *E. coli* codon usage and its native

sequence. For highly expressed proteins there is a strong bias in favor of those codons for which the isoaccepting tRNA is in the greatest abundance [7,8]. In present study, we constructed two vectors for expression of hbFGF in *E. coli* (pET-1001 and pET-1005). The pET-1005 carries the sequence of the native protein and pET-1001 contains the modified sequence which was in favor of *E. coli* coding usage. The result of ELISA showed that the level of expression of the protein in pET-1005 *BL21 (DE3)* system was 16 mg/l. However this value was 0.025 mg/l for pET-1001 system.

Four criteria were considered for expression of this protein in *E. coli*: number of copies of the gene, rate of transcription, stability of the mRNA, and the efficiency of translation [9]. In both systems, pET-1001 and pET-1005, we have benefited from the existence of a very efficient *T7*/*lac* promoter for expression of the protein in *BL21(DE3)* cells. Therefore, the first and second criteria were the same in both systems, so the difference in level of expression could only be due to either the stability of the mRNA or the efficiency of translation.

A strong evidence has been presented for the involvement of mRNA secondary-structures as a major determinant of the translational initiation frequency in *E. coli* [10]. The efficiency of translation in *E. coli* is determined primarily at the stage of initiation, which includes binding of the 30S ribosome to 5' segments (from approximately nts –20 to +15) of mRNA spanning the Shine-Dalgarno sequence and the initiation codon [11]. Thus, the secondary structure of such mRNA segment can play an important role in modulating translation efficiency [12]. Although the extended mRNA secondary structure increases the half life of mRNA e.g. prevention of mRNA from degradation by 3'-5' exonucleases, the mRNAs with very high secondary structure are very poor substrates for translation system enzymes [13]. The presence of certain sequences at the 5'-end of mRNA may also influence mRNA stability [14]. Therefore primary sequence and/or structural changes influence the translational initiation frequency, efficiency of translational start region, and RBSs (ribosome binding sites) [15].

It was not possible to show whether it was the stability of the two mRNAs or the efficiency of translation in each system caused the great difference in the expression level of the protein. There are other possibilities which could affect the expression level of the protein e.g., the inaccessibility of the translational start region for the ribosomes (stem and loop structure) or in contrast to this, by changing the codons in favor of *E. coli* codon usage, new sites were created which could provide a better substrate for exonucleases.

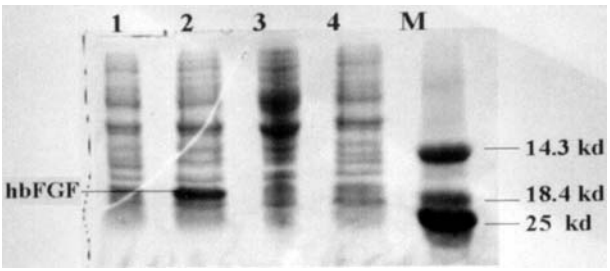

Figure 3. 18% SDS-PAGE. Lane 1: Induced cells continuing pET-3a (negative control); Lane 2,3: Induced and uninduced cells continuing cDNA; Lane 4: Induced cells continuing synthetic gene; Lane 5: Marker.

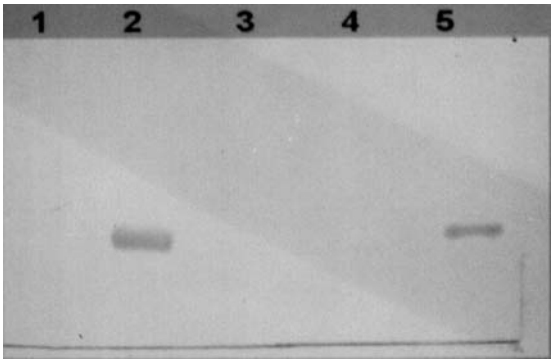

Figure 4. Western blotting analysis. Lane 1: Induced cells containing pET-3a (negative control); Lane 2,3: Induced and uninduced cells containing cDNA; Lane 4: Induced cells containing synthetic gene; Lane 5: Standard hbFGF.

References

1. Fax, G. M., Schiffer, S. G., Rohde, Tsai, L. B., Banks, A. R. and Arakawa, T. Production, biological activity, and structure of recombinant basic fibroblast growth factor and an analog with cysteine replaced by serine. *J. Biol. Chem*., **263**(34): 18452-58 (1988).
2. Birger Anspach, F. Purification of recombinant human basic fibroblast growth factor: stability of selective sorbents under cleaning in place conditions. *J. Chromatography A*, **711**: 129-39 (1995).
3. Mozafari, F. S. and Bakayev, V. V. Synthesis and expression of modified bFGF gene in *E. coli* cells. *Iranian Bio. J.*, **1**(1): 11-18 (1997).
4. Milev, P. V., Georgiev, O. I., Tzarnoretchki, P. O. and Hadjiolov, A. A. Cloning and high level expression of synthetic gene for human basic fibroblast growth factor. *J. Biotechnology*, **22**: 299-310 (1992).
5. Sanger, F., Nicklen, S. and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**: 5463-67 (1977).
6. Squires, C. H., Childs, J., Eisenberg, S. P., Polverini, P. J.

and Sommer, A. Production and characterization of human basic fibroblast growth factor from *E. coli. J. Biol. Chem.*, **263**, 31(5): 16297-302 (1988).

7. Ikamura, T. Correlation between the abundance of *E. coli* transfer RNAs and the occurrence of the respective codons in its protein genes. *J. Mol. Biol.*, **146**: 1-21 (1981).
8. Xuhua Xia. How optimized is the translational machinery in *E. col*i, salmonella typhimurium and saccharomyces cerevisiae? *Genetics*, **149**: 37-44 (1998).
9. Thomas, G., von Gabian, A., Nilsson, G., Andersson, M., Lundstrom, M., Lund, B. and Lundgren, E. Expression of an interferon-α gene variant in *E. coli* using tandemly repeated synthetic ribosomal binding sites. *DNA*, **6**(1): 41-46 (1987).
10. Spanjaard, R. A., van Dijk, M. C., Turion, A. J. and van Duin, J. Expression of the rat interferon-alpha 1 gene in *E. coli* controlled by the secondary structure of the translation–initiation region. *Gene*, **15**, 80(2): 345-51 (1989).
11. Morita, M., Kanemori, M., Yanagi, P. Heat-induced synthesis of Q32 in *E. coli*: Structural and functional dissection of rpoH mRNA secondary structure. *J. Bacteriology*, **181**(2): 401-10 (1999).
12. De Smit, M. H. and J. van Duin. Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. *Proc. Natl. Acad. Sci. USA*, **87**: 7668-72 (1990).
13. Gross, G. and Hollatz, I. Coliphage lambda to terminator lowers the stability of mRNA in *E. coli* hosts. *Gene*, **10**, 72(1-2): 119-28 (1988).
14. Groski, K., Roch, J. M., Prentki, P. and Krisch, H. M. The stability of bacteriophage T4 gene 32 mRNA: a 5' leader sequence that can stabilize mRNA transcripts. *Cell*, **43**(2 pt. 1): 461-69 (1985).
15. Gross, G., Mielke, C., Hollatz, I., Blocker, H. and Frank, R. RNA primary sequence or secondary structure in the translational initiation region controls expression of two variant interferon-β genes in *E. coli*. *J. Biol. Chem.*, **265**(29): 17627-36 (1990).