Optimization and construction of human insulin-like growth factor 1 gene related to human health

Hamidreza Iranpoor[1], Venus Vatankhah[2], Masoud Lotfizadeh[2*]

[1]*Medical Biotechnology Dept., Golestan University of Medical Sciences, Gorgan, I.R. Iran;*
[2]*Social Health Determinants Research Center, Shahrekord University of Medical Sciences, Shahrekord, I.R. Iran.*

Original article

ABSTRACT

Background and aims: Laron syndrome is a disease that treated by Insulin-like Growth Factor 1 (IGF-1). This protein is a single chain and has three disulfide bonds. People with Laron syndrome have low rates of cancer and diabetes, although they appear to be at increased risk of casual death due to their stature. IGF-1 is synthesized by many tissues and is secreted from liver as an endocrine hormone which is transmitted to other tissues. IGF-1 is responsible for cell differentiation, transformation, suppression of apoptosis, cell cycle progression, cell proliferation and differentiation. Expression system, the igf-1 gene, codon adaptation index (CAI) and GC contents are very critical for the large scale production of this protein. Current study was aimed to Optimization and Construction of human IGF-1 gene in E. coli DH5α.

Methods: we used DNA2 and ProtParam softwares for designing the best form to produce IGF-1. First, the coding sequence was verified and then synthesized. For confirmation of the pUC18-IGF-1, sequencing test was carried out using M13 reverse primer. Finally it was inserted into the cloning site of pUC18.

Results: After coding optimization, the CAI rate was increased from 84 % to 90% and GC content from 55.07 % to 56.62%. The presence of the band near 225bp resulted from enzymatic digestion with 2 restriction enzymes demonstrates the correct cloning of the recombinant vectors in the cloning site of pUC18 cloning vector.

Conclusion: According to software and experimental analysis, the designed sequence probably in the best form could be used for production of recombinant protein.

Keywords: Human Insulin-like Growth Factor 1, Cloning, Vector, Optimization, Sequencing, Construct.

INTRODUCTION

In 1957, a growth hormone-dependent factor was discovered, which could affect the increase of growth of the epiphyseal cartilage. A more general expression "somatomedine" was presented from various metabolic effects and the insulin-like nature view.[1] They have a high structural homology with insulin. They were renamed as Insulin-like growth factor 1 and 2.[2] IGF-1 is known as Somatomedin C, is produced in the liver and is responsible for cell differentiation, transformation, suppression of apoptosis, cell cycle progression, cell proliferation and differentiation, nervous system and muscular system and internal secretion.[3] IGF-1 recombinant protein plays a role in the treatment of many diseases. The

*Corresponding author: Masoud Lotfizadeh, Social Health Determinants Research Center, Shahrekord University of Medical Sciences, Shahrekord, I.R. Iran. Tel: 0098383334754, E-mail: masoud_lotfizadeh@yahoo.com

use of its system in diseases related to such as severe shortage of primary IGF, like Laron syndrome, reduced effect of growth hormone, idiopathic short stature, wasting syndrome.[4] The most of reported cases of Laron syndrome have been in people with Semitic origins. There is also a disproportionate number of sufferers found in remote villages in the South American country of Ecuador who are descended from colonial-era Jewish-origin New Christian conversos who had covertly migrated to Ecuador during the Spanish Conquest despite the Spanish Crown's prohibition of their immigration to its colonies and territories as a result of the Inquisition.[5,6] Other patients include people of other Semitic non-Jewish origins, including from Saudi Arabia. Recent publications have proposed that Homo floresiensis showed a population with prevalent Laron syndrome.[7] For large-scale production of these proteins, engineered bacteria can be used.[8] To do this, to increase the efficiency of protein expression, IGF-1 gene sequence should be optimized in the used host bacteria.[9] E. coli DH5α Bacteria is one of the engineered strains which is compatible with pUC cloning vectors.[10] In this descriptive study, we obtain GC amount and CAI (codon adaptation index) using available software to design this protein. Then after designing and its synthesis, after cloning in the E. coli DH5α bacteria we sequence it.

METHODS

At first, we obtained the main sequence of amino acids in IGF-1 protein from the NCBI website. Then, functional section which is the CDS can be achieved by removal of the signalling sequence and the protein sections that does not have function. This section has 70 functional amino acids. After that, we design 2 limiting position (restriction site) of the NdeI and BamHI-HF enzymes at the N-terminal and C-terminal. We optimized nucleotide sequence using DNA2 software. This optimizing is done to select the codons

that are used more in the bacteria. The chromatogram sequencing was done also to that. We also determine the GC% content. ProtParam software was used for the stability of the recombinant protein. Optimized gene synthesis was done. Synthesized fragment length was 225 bp. This gene was received in the pUC18 cloning plasmid. The sequence of the gene was sequenced by contig method. Then, after receiving the plasmid carries the gene, it was transformed by $CaCl_2$ and heat shock to the cloning bacteria. Since this bacterium is ampicillin plasmid resistant, colonies that were grown in the antibiotics culture can be selected. Then, reproduced plasmids were extracted using Qiagen Company plasmid extraction kit. Then, after confirming of the gene amplification on 1% agarose gel, with two NdeI and BamHI-HF enzymes of Biolab Company, enzyme digestion reaction was performed on this plasmid. Enzyme digestion was performed to confirm the correct cloning in the cloning bacteria. After cutting the designed fragment, pUC18 plasmid made linear using the mentioned enzymes. Then, the cut fragment and the linear vector were isolated on 1% agarose gel.

RESULTS

Gene optimization was carried in the E. coli host and the CAI and GC content were obtained as follows (Figures 1,2,3,4).
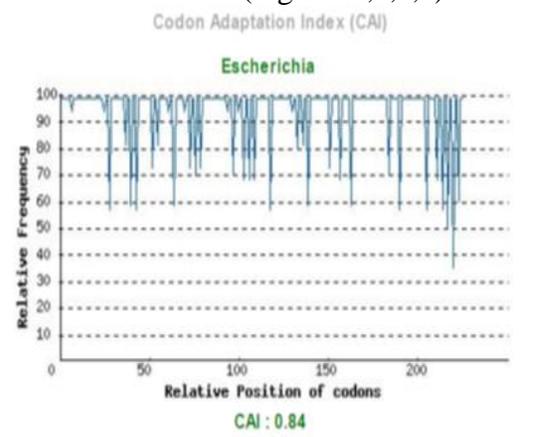
Figure 1: CAI before optimization
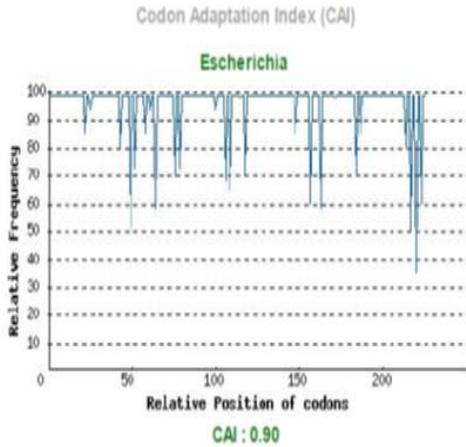
Figure 2: CAI after optimization
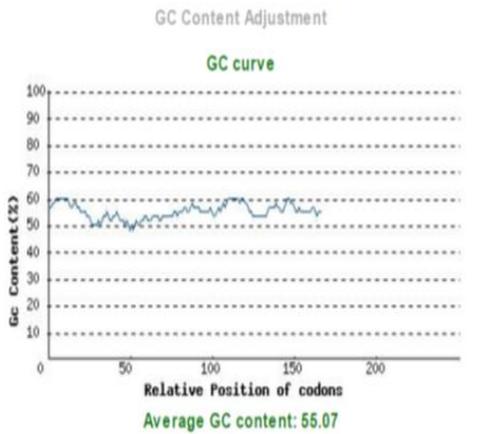
Figure 3: GC content before optimization optimization
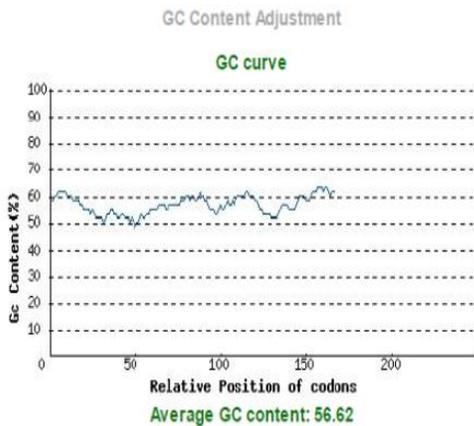
Figure 4: GC content after
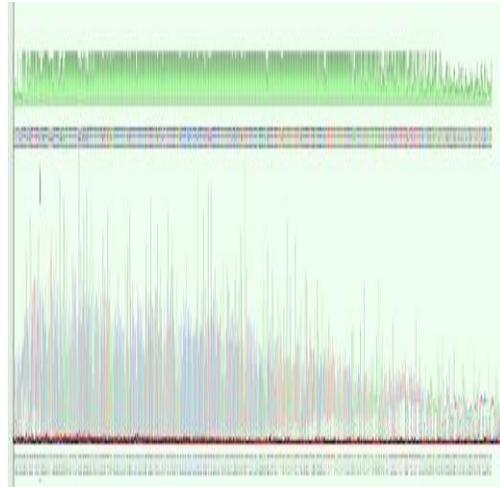
Figure 5: Sequence chromatogram

Then after gene optimization it was transformed into the cloning E. coli DH5α bacteria.

After White-blue colonies screening and Plasmids extraction, electrophoresis was done on 1% agarose gel, a band with a molecular weight of 2659 indicates the cloning plasmid with the considered gene (Figure 6).

Figure 6: pUC18-IGF-1 with 2659 bp on 1% agarose gel

```
                    10        20        30        40        50        60
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCT│
C04-C14305-419493-3-M13-88.abl(1>695)← │CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCT│

                    70        80        90        100       110       120       130
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │TTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAG│
C04-C14305-419493-3-M13-88.abl(1>695)← │TTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCtGACACATGCAGCTCCCGGAGACGGTCACAG│

                    140       150       160       170       180       190       200
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │CTTGTCTGTAAGCGGATGCCGGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTC│
C04-C14305-419493-3-M13-88.abl(1>695)← │CTTGTCTGTAAGCGGATGCCGGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTC│

                    210       220       230       240       250       260       270
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │GGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGCCCCGGCTCGCT│
C04-C14305-419493-3-M13-88.abl(1>695)← │GGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGGCCCCGGCTCGCT│
419493-3.seq(1>397)                  →                                          CATATGGCCCCGGCTCGCT│

                    280       290       300       310       320       330       340
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │CGCCGTCACCGTCTACGCAACCGTGGGAACACGTCAATGCCATCCAAGAAGCTCGTCGTCTGCTGAATC│
C04-C14305-419493-3-M13-88.abl(1>695)← │CGCCGTCACCGTCTACGCAACCGTGGGAACACGTCAATGCCATCCAAGAAGCTCGTCGTCTGCTGAATC│
419493-3.seq(1>397)                  → │CGCCGTCACCGTCTACGCAACCGTGGGAACACGTCAATGCCATCCAAGAAGCTCGTCGTCTGCTGAATC│

                    350       360       370       380       390       400       410
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │TGAGTCGTGATACGGCGGCCGAAATGAACGAAACCGTGGAAGTTATTAGCGAAATGTTTGACCTGCAGG│
C04-C14305-419493-3-M13-88.abl(1>695)← │TGAGTCGTGATACGGCGGCCGAAATGAACGAAACCGTGGAAGTTATTAGCGAAATGTTTGACCTGCAGG│
419493-3.seq(1>397)                  → │TGAGTCGTGATACGGCGGCCGAAATGAACGAAACCGTGGAAGTTATTAGCGAAATGTTTGACCTGCAGG│

                    420       430       440       450       460       470       480
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │AACCGACGTGCCTGCAAACCCGTCTGGAACTGTATAAACAGGGCCTGCGCGCGGTAGTCTGACGAAACTGA│
C04-C14305-419493-3-M13-88.abl(1>695)← │AACCGACGTGCCTGCAAACCCGTCTGGAACTGTATAAACAGGGCCTGCGCGCGGTAGTCTGACGAAACTGA│
419493-3.seq(1>397)                  → │AACCGACGTGCCTGCAAACCCGTCTGGAACTGTATAAACAGGGCCTGCGCGCGGTAGTCTGACGAAACTGA│

                    490       500       510       520       530       540       550
         ┌───┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴────┴──┐
         │AAGGCCCGCTGACCATGATGGCATCCCATTACAAACAGCACTGCCCGCCGACGCCGGAAACCAGCTGTG│
C04-C14305-419493-3-M13-88.abl(1>695)← │AAGGCCCGCTGACCATGATGGCATCCCATTACAAACAGCACTGCCCGCCGACGCCGGAAACCAGCTGTG│
```
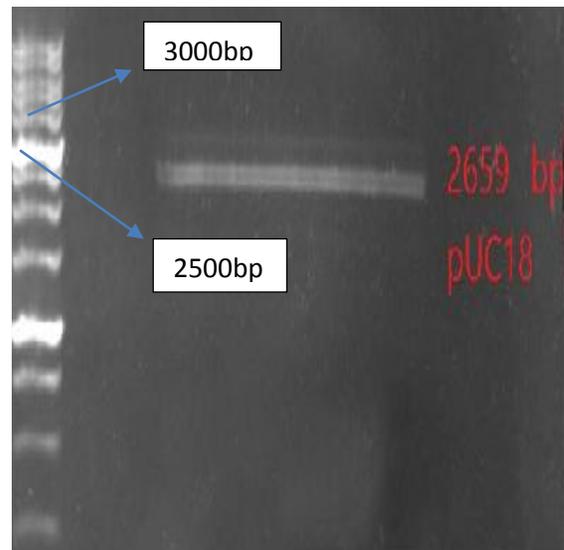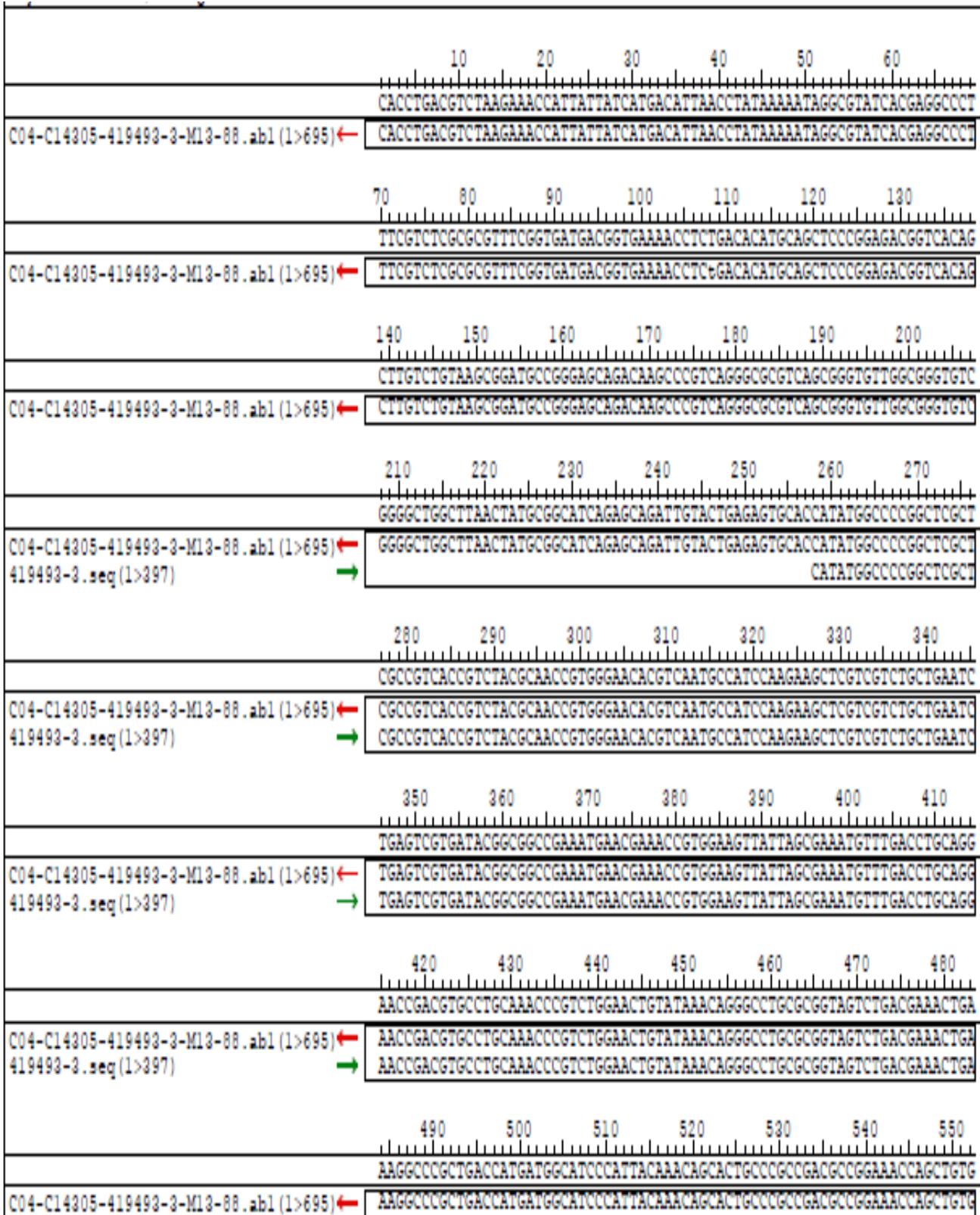
Figure 7: sequencing of IGF-1

Table 1: IGF-1 gene before and after optimization

IGF-1 gene (225 *bp*)	
Before optimization	CATATGGGTCCGGAAACCCTGTGCGGTGCTGA ACTGGTTGACGCTCTGCAGTTCGTTTGCGGTGA CCGTGGTTTCTACTTCAACAAACCGACCGGTTA CGGTTCTTCTTCTCGTCGTGCTCCGCAGACCGG TATCGTTGACGAATGCTGCTTCCGTTCTTGCGA CCTGCGTCGTCTGGAAATGTACTGCGCTCCGCT GAAACCGGCTAAATCTGCTTGAGGATCC
After optimization	CATATGGGCCCGGAAACCCTGTGTGGTGCGGA ACTGGTGGATGCCCTGCAATTCGTGTGTGGTGA CCGTGGCTTTTACTTCAACAAACCGACCGGCTA TGGTAGCTCTAGTCGTCGCGCACCGCAGACCG GCATTGTGGATGAATGCTGTTTTCGTTCCTGCG ACCTGCGTCGCCTGGAAATGTACTGTGCGCCGC TGAAACCGGCGAAAGCGCCTGAGGATCC

Then, sequencing was done by contig method in order to the accuracy of gene cloning by reverse M13 primer (Figure 7).

DISCUSSION

Human insulin-like growth factor type 1 is responsible for the increase of growth in the body.[11] People lacking this growth factor have disease with defects in IGF-1 like Laron.[12] In addition, this growth factor also is used for wound healing and the treatment of diabetes type 1 and 2.[13] E. coli DH5α bacteria is one of the engineered strains to reproduce the gene. Beta galactosidase incomplete form (lacZ ΔM15) is designed in the genome of this bacteria that using α-complentation can be screened the pUC cloning vector using white-blue colonies. pUC cloning vector has an ampicillin resistance gene for easy screening.[14] In previous studies, it has been shown that cloning of the gene in E. coli bacteria optimization is needed. Since the objective of this project is amplification of this gene in bacteria with maximum efficiency to be performed, optimization of this gene using appropriate Bioinformatics software as well as studying protein expression in them is needed. At the end of this gene cutting site was designed for 2 NdeI and BamHI-HF cutting enzyme. TGA sequence was selected for the terminal codon. Nucleotide sequence of the recombinant protein was optimized by DNA2 software. Using this software Codon Adaptation Index (CAI) and GC content of the coding were evaluated. If CAI results be more close to 100 percent it will be more suitable for expression in the host. The second case is GC content of the sequence. The more GC percent, the more energy for opening 2 DNA strings. So, it is better that GC content be 70 percent. Recombinant sequence CAI was 84% prior to optimization in the normal condition of the gene which was improved to the reasonable rate of 90% and GC increased from 55.07 to 56.62 percent. ProtParam Software was used to determine protein stability. Also identification of gene cloning was confirmed by electrophoresis on 1% agarose gel.

CONCLUSION

This study is related to design, optimization and production of recombinant vector containing encoding fragment of human type 1 insulin-like growth factor. Having CAI= 90%, we could achieve the highest expression rate of this gene in E. coli bacteria.

REFERENCES

1. Laron Z. Insulin-like growth factor 1 (IGF-1): a growth hormone. Mol Pathol. 2001; 54(5): 311.

2. Yuqing Y, Minghui X, Hongmei S, Hongjie Z. Cloning and higher expression of recombinant human insulin-like growth factor-1. Life Sci J. 2009; 6(3).

3. Rinderknecht E, Humbel RE. The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin. J Biol Chem. 1978; 253(8): 2769-76.

4. MB R. Introduction of new recombinant insulin-like growth factor-1-current and future perspectives. Touch Brief. 2008; 4(1): 6-52.

5. Guevara-Aguirre J, Balasubramanian P, Guevara-Aguirre M, Wei M, Madia F, Cheng CW, et al. Growth hormone receptor deficiency is associated with a major reduction in pro-aging signaling, cancer, and diabetes in humans. Sci Transl Med. 2011; 3(70): 70ra13.

6. LexisNexis. The New York times on the Web 1857. Available from: http://www.nexis.com.emu.londonmet.ac.uk/api/version1/sf?sfi=GB00NBGenSrch&csi=6742&shr=t.

7. Hershkovitz I, Kornreich L, Laron Z. Comparative skeletal features between Homo floresiensis and patients with primary growth hormone insensitivity (Laron Syndrome). Am J Phys Anthropol. 2007; 134(2): 198-208.

8. Builder S, Hart R, Lester P, Reifsnyder D. Refolding of misfolded insulin-like growth factor-I. Google Patents; 1997.

9. Sweet CR. Expression of recombinant proteins from lac promoters. Methods Mol Biol. 2003; 235: 277-88.

10. Preston A. Choosing a cloning vector. Methods Mol Biol. 2003; 235: 19-26.

11. Hede MS, Salimova E, Piszczek A, Perlas E, Winn N, Nastasi T, et al. E-peptides control bioavailability of IGF-1. PLoS One. 2012; 7(12): e51152.

12. Horney MJ, Evangelista CA, Rosenzweig SA. Synthesis and characterization of insulin-like growth factor (IGF)-1 photoprobes selective for the IGF-binding proteins (IGFBPS). photoaffinity labeling of the IGF-binding domain on IGFBP-2. J Biol Chem. 2001; 276(4): 2880-9.

13. Bizzarri C, Benevento D, Giannone G, Bongiovanni M, Anziano M, Patera IP, et al. Sexual dimorphism in growth and insulin-like growth factor-I in children with type 1 diabetes mellitus. Growth Horm IGF Res. 2014; 24(6): 256-9.

14. Casali N. Escherichia coli host strains. E coli Plasmid Vectors: USA: Humana Press; 2003; 235: 27-48.