Gene Technology
ISSN: 2329-6682
Aslam et al., Gene Technol 2019, 8:1
DOI: 10.4172/2329-6682.1000147
Research Article
Open Access

# Recombinant Human Proinsulin Expression in *E. Coli* by Altering 5′ Untranslated and Translated Region

**Aslam F[1]*, Latif K[1], Waseem R[1], Naz S[1] and Iftikhar S[2]**

[1]*Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan*
[2]*School of Biological Sciences, University of the Punjab, Lahore, Pakistan*

## Abstract

Messenger RNA initiate the process of translation, by transferring the code of DNA to transfer RNA. A poly purine rich sequence called Shine-Dalgarno sequence help to determine the position of the start codon, the Shine-Dalgarno sequence is different because it allows the ribosome to be built at an interior position on the mRNA through direct binding to this sequence. Ribosomal binding site at the 5'end of translation initiation site used to bind mRNA secondary structure, distance between RBS and start codon effects translational efficiency of a gene. In this study we change the distance between ribosome binding site and start codon to get the different ratios of translational expression. For this purpose, proinsulin gene cloned in pET21a vector, the distance between the binding site and starting codon has been retained to 8 nucleotides. At this distance between ribosome binding site (RBS) and start codon, the expression of proinsulin was 30% of total cellular proteins. When there are 10 nucleotides between, expression decreased up to 2-4%. With 12 nucleotides between RBS and start codon, expression further decreased up to 1-2%. As these attempts are made random, so by checking the mRNA secondary structures by M-fold showed a binding between ribosome binding site and start of proinsulin gene. Another attempt was made by incorporating ten different nucleotides at the start of proinsulin gene to make the secondary structure of mRNA less stable (ΔG=-5.5) does not significantly alter the expression of proinsulin in BL21 codon plus cells. There are other controlling factors of protein expression i.e., metabolic instability, rapid degradation of mRNA or accumulation of protein may downregulate expression of mRNA.

**Keywords:** Shine dalgarno sequence; Ribosome binding site; Proinsulin

## Introduction

The transcription start site is the location where transcription starts at the 5'-end of a gene sequence [1]. In prokaryotic protein production can be controlled, with three steps:

1) Protein transcription (code information in form of mRNA).

2) Availability of messenger RNA to continue the translation.

3) Degradation of mRNA.

Availability of mRNA controls the rate of translation. The 3' terminal sequence of 16S RNA, recognize shine Dilgarno sequence and initiation codon (AUG) [2]. Availability as well as primary structure of this region, to ribosomal particles, contributes towards translation initiation and strength [3]. Distance between RBS and start codon effects binding with 16S RNA. Greater than 12 and less than 5 nucleotides decrease the initiation rate of translation.

Proinsulin is expressed as recombinant protein which is processed enzymatically to insulin [4]. An expression vector has transcriptional promoter, translational initiation region, initiation codon termination. Secondary structure of RBS is very important for translation initiation and presence of thyamnine and adenine increases the efficiency of translation [5].

Recombinant protein expression in expression host can be improved by mRNA stability, codon bias, and controlling inclusion bodies formation and changing genetic tools in different combinations [6]. Expression system of *E. coli* is used and insulin is made from its precursor proinsulin, expressed as inclusion bodies and made catalytically active by refolding [7].

## Materials and Methods

PCR kit, restriction enzymes *Xba*I and *Hind*III and T/A cloning kit from Fermentas, Gel extraction kit and plasmid extraction kit was from QIAGEN, DH5 alpha and BL21 codon plus DE3-RIL were from Novagen, pET21a and pTZ57R/T was from Novagen. All the other chemicals used were analytical grade from Sigma and Invitrogen. The template was synthetic human proinsulin with appending methionine-lysine-arginine gene sequence at 5' end of gene (Invitrogen), with *E. coli* preferred codon, in pET21a vector (Figure 1).

Six primers were designed to incorporate different number of nucleotides between RBS and start codon to amplify human proinsulin gene with 8-13 nucleotides between RBS and start codon (Figures 2-4). There used living system to examine the proinsulin gene expression is pET21a vector, six derivatives of human proinsulin gene. Free energy of mRNA secondary structure of six derivatives were calculated by Mfold as Figure 5 [8]. To amplify six derivatives of human proinsulin gene, 100 pmoles of forward and reverse primer, 4 µg of synthetic human proinsulin gene as template was used. The conditions for PCR were initial denaturation at 94°C for 30 sec, annealing at 65°C for 30 sec and polymerization at 72°C for 0.30 sec and final extension at 72°C for 5 min. The PCR product was analyzed on 1% agarose gel. PCR products of the derivatives of proinsulin were ligated to pTZ57R/T vector. The ligation mixture was transformed to competent cells of DH5α and plated on LB plates containing 1.5% agar, 100 µg/ml of ampicillin, 1.33 µg/ml 5-bromo-4-chloro-3-indolyl- β-D-galactoside (X-gal) and 35µg/ml of 0.2 M isopropylthio-β-D-galactoside (IPTG). Different vectors were restricted with *Xba*1 and *Hind*III. The fragment of ~300 bp, confirmed the presence of proinsulin gene in respective vector. ~300 bp of human proinsulin gene obtained by restriction with *Xba*1 and

***Corresponding author:** Aslam F, Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan, Tel: +92-42-99203801-9; E-mail: farheenpu@gmail.com

**Received** January 09, 2019; **Accepted** January 22, 2019; **Published** January 31, 2019

**Citation:** Aslam F, Latif K, Waseem R, Naz S, Iftikhar S (2019) Recombinant Human Proinsulin Expression in *E. Coli* by Altering 5' Untranslated and Translated Region. Gene Technol 8: 147. doi: 10.4172/2329-6682.1000147

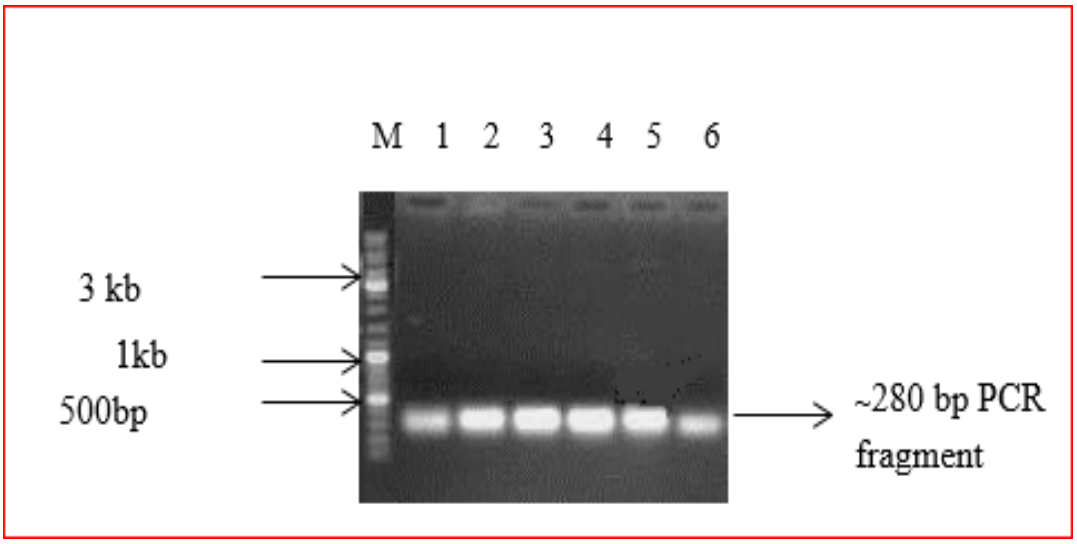


**Figure 1:** 1% agarose gel of PCR amplified fragment of different proinsulin derivatives. Lane M: DNA Ladder mix (100-10000 bp), Lanes 1, 2, 3, 4, 5, 6: PCR fragment of 280 bp by using primers k-1, k-2, k-3, k-4, k-5, k-6 respectively.

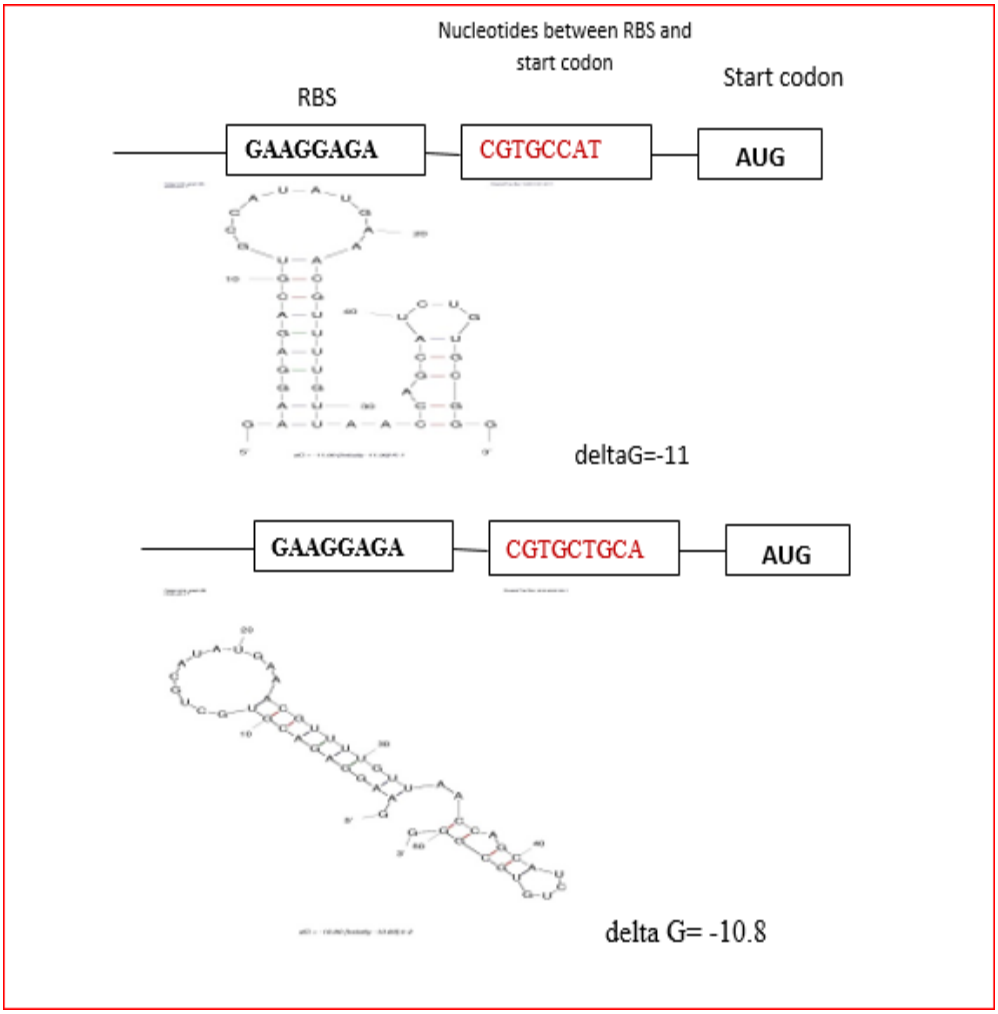


**Figure 2:** Varying number of nucleotides between RBS and start codon along with the secondary structure of mRNA predicted from m-fold [8].

*Hind*III, is gene cleaned and ligated to pET21a vector. The ligation mixture was transformed to DH5α cells. The presence of insert in Hpi-8nt/pET, Hpi-10nt/pET, Hpi-12nt/pET, Hpi-13a/pET, Hpi-13b/pET and Hpi-n/pET were confirmed by double restriction with *Xba*1 and *Hind*III.

## Results

Six derivatives of human proinsulin gene were assembled by varying the distance between RBS and start codon i.e., 8-13 nucleotides between the ribosome binding site and start codon. 280 bp PCR fragment confirmed the different derivatives of proinsulin. The ΔG value for mRNA secondary structure of these derivatives varies between -10.60 to -10.80. ΔG shows presence of strong secondary structure because of bonding between RBS and start of the proinsulin gene. The fragment of ~300 bp, confirmed the presence of proinsulin gene in respective vector. For expression analysis, the different constructs were transformed to BL21 codon plus cells and induced with 0.2 mM IPTG (Figures 6-8).

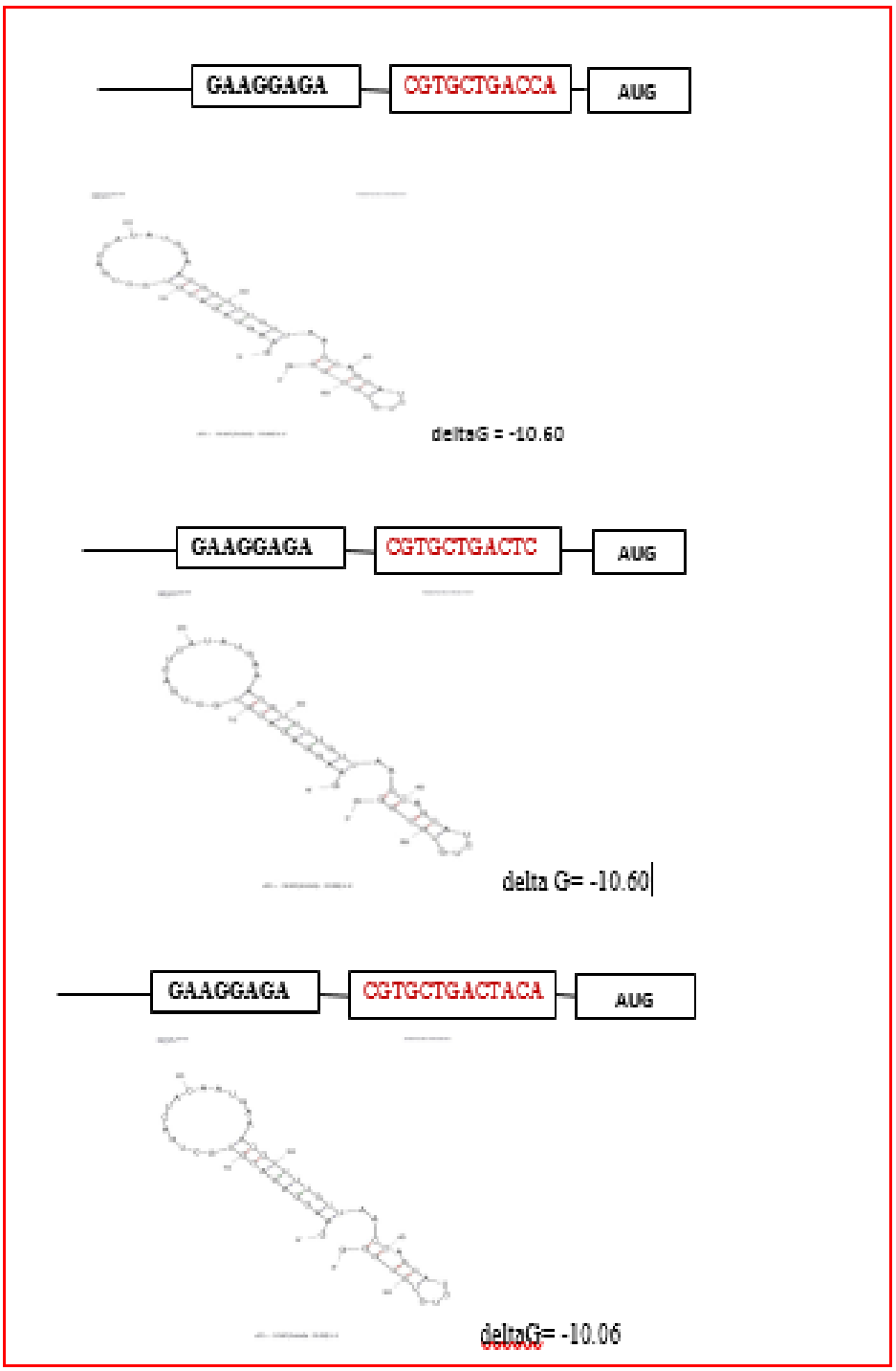


**Figure 3:** Varying number of nucleotides between RBS and start codon along with the secondary structure of mRNA predicted from m-fold [8].

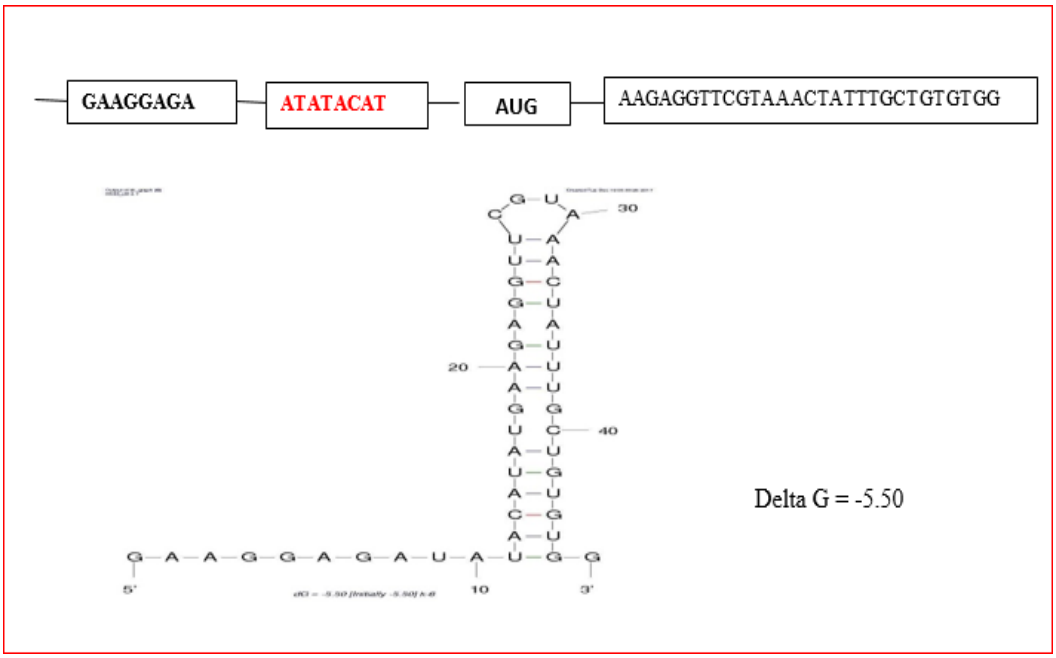


**Figure 4:** Varying number of nucleotides between RBS and start codon along with the secondary structure of mRNA predicted from m-fold [8].

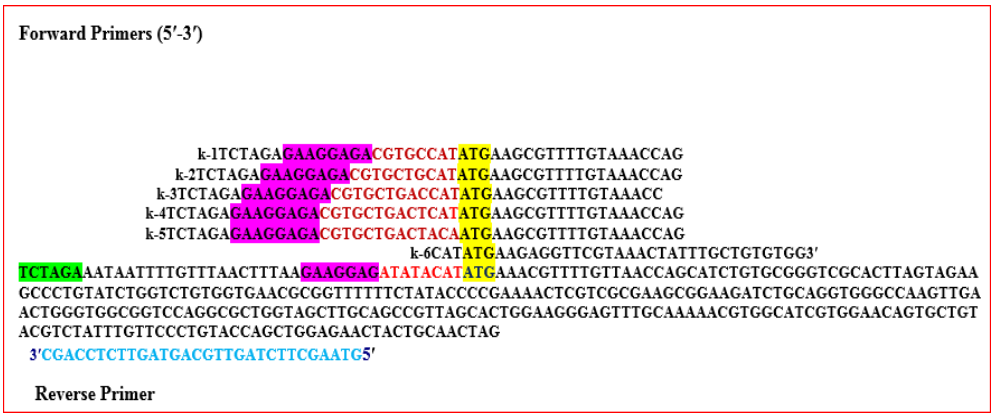


**Figure 5:** Sequence of human proinsulin gene with six designed primers. Distance between ribosome binding site and start codon are shown in red font. Reverse primer shown in blue font. Ribosome binding site (RBS) is highlighted with pink color. Note: *Xba*I site is highlighted with green color, Start codon is highlighted with yellow color.

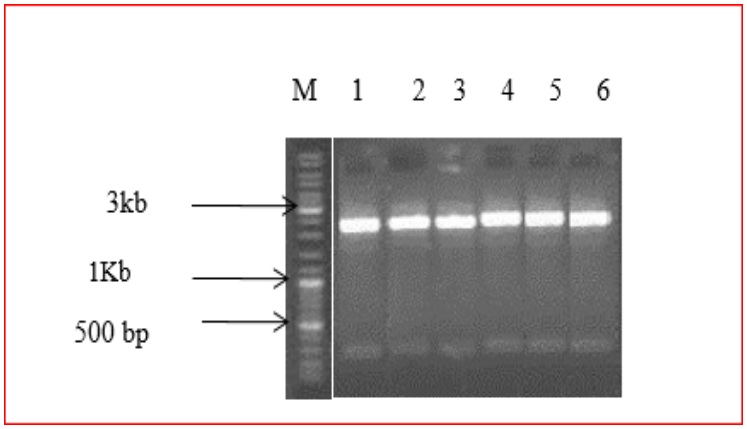


**Figure 6:** 1% agarose gel showing different constructs of human proinsulin gene restricted with *Xba*1 and *Hind*111. Lane M: DNA Ladder mix (100-10000 bp); Lane 1, 2, 3, 4, 5 and 6 (Hpi-8nt/pTZ, Hpi-10nt/pTZ, Hpi-12nt/pTZ, Hpi-13a/pTZ and Hpi-13b/pTZ and Hpi-n/pTZ) were restricted with *Xba*1 and *Hind*111 respectively.

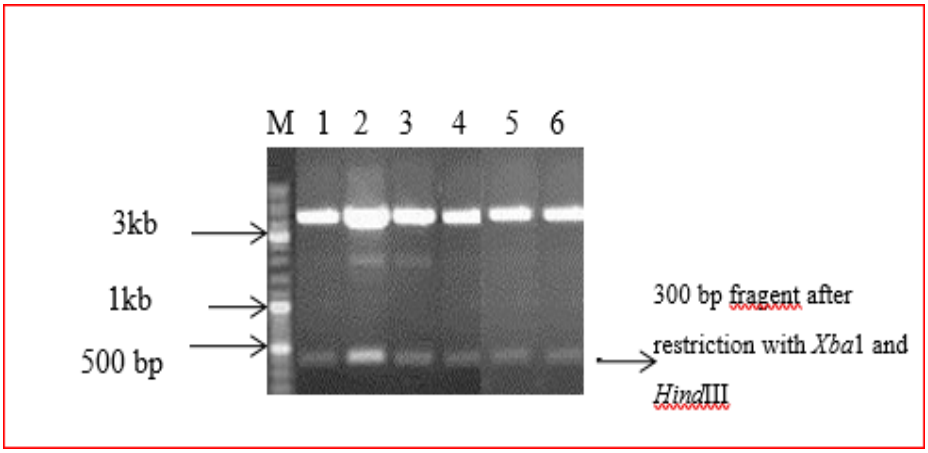


**Figure 7:** Hpi-8nt/pET, Hpi-10nt/pET, Hpi-12nt/pET, Hpi-13a/pET, Hpi-13b/pET and Hpi-n/pET confirmed by double restriction with *Xba*1 and *Hind*111.

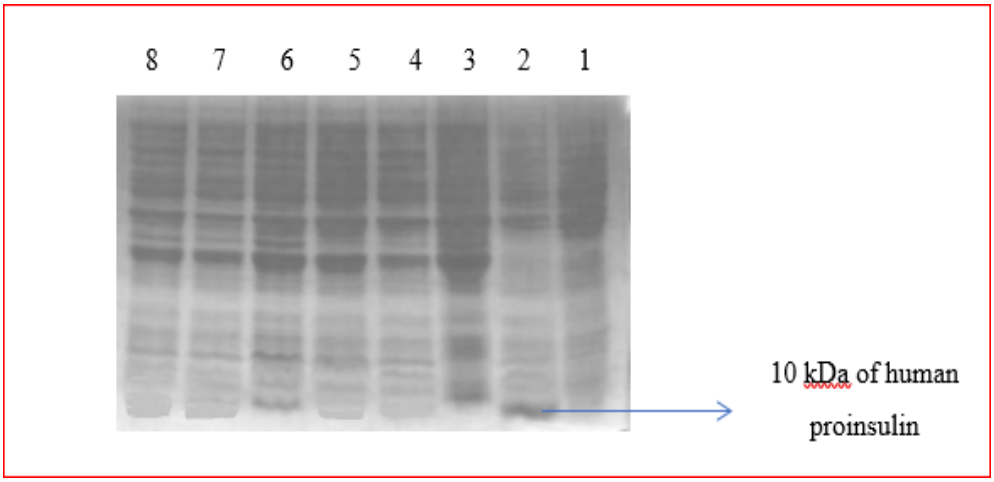


**Figure 8**: SDS-PAGE analysis of Human proinsulin gene. Lane 1: uninduced MKR; Lane 2: induced MKR with 0.2 mM IPTG; Lane 3, 4, 5, 6, 7 and 8: Hpi-8nt/pET, Hpi-10nt/pET, Hpi-12nt/pET, Hpi-13a/pET, Hpi-13b/pET and Hpi-n/pET induced with 0.2 mM IPTG.

## Discussion and Conclusion

Translation initiation region (TIR) comprises of thirty nucleotides upstream of the start codon. TIR has the most important region of ribosome binding site, located 5-11 nucleotides upstream of the start codon. TIR and sequence of the gene is important in controlling the transcriptional as well as translational efficiency of specific gene [9].

The synthesis of protein in *E. coli* begins when the 30S ribosomal subunit detects Shine-Dalgarno sequence present on mRNA. Shine-Dalgarno sequence (S-D sequence) forms Watson-Crick base pairs with the 30S ribosomal subunit, it is present upstream of the start codon of mRNAs [10]. Presence of secondary structures in the TIR region influence the binding of 30S ribosomal RNA complex and decreases the translation initiation rate [11].

Changing nucleotides at the 5' end UTR can enhance the expression of recombinant protein up to 600 times [12]. In the present research the sequence and number of nucleotides were changed between RBS and start codon. With 8 nucleotides between RBS and start codon, expression was the same as by appending M-K-R sequence at the N-terminus of proinsulin gene i.e., 30% of total cellular proteins [13].

In the case of 10 nucleotides between RBS and start codon, the expression decreases up to 2-4% of total cellular proteins. With 12 nucleotides between RBS and start codon showed expression was 1-2% of total cellular proteins. With 13 nucleotides between RBS and start codon, expression was 30% of total cellular protein. With 13 nucleotides between RBS and start codon only three different nucleotides, expression was 2-4% of total cellular protein although ΔG of mRNA secondary structure for both the constructs with 13 nucleotides between RBS and start codon are same i.e., -10.6. It is reported that expression of certain genes enhanced up to 1000-fold by varying sequence between initiation codon and pET vector [14]. It showed that distance as well as the sequence of nucleotides between RBS and start codon is very important for expression. Changing three nucleotides between RBS and start codon decrease the expression. By changing N-terminal sequence of human proinsulin gene. It has 2 amino acid changed as compared to the other constructs. Construct Hpi-nt/pET has the $B^7$ (Glutamine Q) and $B^8$ (Histidine H) replaced with $B^7$ (Tyrosine) and $B^8$ (Leucine), so by changing the amino acids of the N- terminal of the proinsulin, decrease its expression as compared to MKR-Hpi. mRNA secondary structure of all derivatives of proinsulin were checked by Mfold [8] that showed stable structures of K-1,K-2, K-3, K-4, K-5 ΔG ranging from -10.60 to -11 because of binding between RBS site and start of proinsulin gene, so K-6 was designed by mutating 5' end of

proinsulin gene to make mRNA secondary structure less stable (ΔG=-5.50), but there was not effect on expression of proinsulin gene.

Incorporation of poly adenylate in the spacer region between RBS and start codon moderately increases the expression of laccase CotA from *B. subtilis* [15].

### Acknowledgement

This research was supported by Higher Education Commission of Pakistan by granting Rs. 0.5million.

### Conflict of Interest

There is no conflict of interest amongst authors.

### References

1. Zvelebil MJ, Baum JO (2007) Understanding bioinformatics, Garland Science.
2. Spirin AS (1999) Translational control in prokaryotes. Ribosomes pp: 309-338.
3. Smit MHD, Duin JV (1990) Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. Proc Natl Acad Sci USA 87: 7668-7672.
4. Steiner DF, Cunningham D, Spigelman L, Aten B (1967) Insulin biosynthesis: Evidence for a precursor. Science 157: 697-700.
5. Laursen BS, Steffensen SA, Hedegaard J, Moreno JM, Mortensen KK, et al. (2002) Structural requirements of the mRNA for intracistronic translation initiation of the enterobacterial in fB gene. Genes Cells 7: 901-910.
6. Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol 115: 113-128.
7. Nilsson J, Jonasson P, Samuelsson E, Stahl S, Uhlén M (1996) Integrated production of human insulin and its C-peptide. J Biotechnol 48: 241-250.
8. Zuker M (2003) M-fold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406-3415.
9. McCarthy JE, Gualerzi C (1990) Translational control of prokaryotic gene expression. Trends Genet 6: 78-85.
10. Shine J, Dalgarno L (1974) The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71: 1342-1346.
11. Gold L, Pribnow D, Schneider T, Shinedling S, Singer BS, et al. (1981) Translational initiation in prokaryotes. Annu Rev Microbiol 35: 365-403.
12. Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, et al. (2013) Precise and reliable gene expression *via* standard transcription and translation initiation elements. Nat Methods 10: 354-360.
13. Aslam F, Gardner QT, Zain H, Nadeem MS, Ali M, et al. (2013) Studies on the expression and processing of human proinsulin derivatives encoded by different DNA constructs. Biochem Biophys Acta 1834: 2116-2123.
14. Mirzadeh K, Martínez V, Toddo S, Guntur S, Herrgård MJ, et al. (2015) Enhanced protein production in *Escherichia coli* by optimization of cloning scars at the vector–coding sequence junction. ACS Synth Biol 4: 959-965.
15. Liebeton K, Lengefeld J, Eck J (2014) The nucleotide composition of the spacer sequence influences the expression yield of heterologously expressed genes in Bacillus subtilis. J Biotechnol 191: 214-220.