

Full Length Research Paper

Genomic sequencing and recombinant expression of proinsulin isolated from cow and buffalo in Pakistan

Farheen Aslam^{1*}, Hooria Younas² and Saima Iftikhar Bajwa²

¹Department of Biotechnology, Lahore College for Women University, Lahore, Pakistan.

²School of Biological Sciences, University of the Punjab, Lahore, Pakistan.

Received 8 April, 2019; Accepted 16 May, 2019

This study was conducted to determine the heterologous expression of eukaryotic gene in *Escherichia coli* variable. The expression of Pakistani buffalo (*Bubalus bubalis*) and cow (*Bos taurus*) proinsulin genes in BL21 codon plus cells is 30% of total cellular protein. Total RNA was isolated from pancreatic tissues of local (Pakistani) breeds of buffalo (*B. bubalis*) and cow (*B. taurus*), converted to cDNA and cloned in a T/A cloning vector. There are five silent mutations: one in B-chain, two each in A-chain and C-peptide encoding regions of local Pakistani buffalo proinsulin DNA sequence as compared to the internationally reported cow proinsulin sequence. The DNA sequence of local (Pakistani) cow proinsulin shows there is one nucleotide encodes C-peptide that has mismatch with the reported cow proinsulin sequence but mismatched nucleotide is same between Pakistani cow and Pakistani buffalo proinsulin sequence. This indicates the genetic similarity of Pakistani cow with Pakistani buffalo. Both genes were expressed in BL21 Codon plus (DE3)-RIL cells with 0.2 mM IPTG at 37°C for 6 to 8 h. Proinsulin was expressed as 30% of total cellular proteins as insoluble inclusion bodies.

Key words: *Bubalus bubalis*, *Bos taurus*, cDNA, T/A cloning vector, BL21 Codon plus (DE3)-RIL cells.

INTRODUCTION

Insulin is synthesized as a larger precursor, preproinsulin (Chan et al., 1976), in pancreatic β -cells, which is rapidly cleaved to proinsulin (Patzelt et al., 1978). The proinsulin contains C-peptide connecting the C-terminus of B-chain with the N-terminus of A-chain. The insulin molecule served as a model for studies on the fundamental structure and properties of proteins. Originally, insulin used clinically was derived from cow and pigs because of their close resemblance with the human insulin. Knowing the sequence of bovine insulin (cow and buffalo) of the

local breed was important to know the evolutionary aspect behind the two species residing in the same area.

The sequence of cow (*Bos taurus*) preproinsulin was already reported in the literature (accession no. M54979). The sequence of buffalo (*Bubalus bubalis*) proinsulin was not available in the literature because of specific localization of this species in the Indo-Pak region. 97% of the world population of *B. bubalis* resides in Asia with half of the population in the Indian subcontinent (India and Pakistan). Sequencing of the buffalo proinsulin showed

*Corresponding author. E-mail: farheenpu@gmail.com. Tel: +92-42-99203801-9/277.

Table 1. Names, sequences, length, melting temperatures and restriction sites representation in the primers.

Primer name	Primer sequence (5' - 3')	Length	*TM (°C)	Restriction sites
WBI-N (Forward)	GAACATATGTTTCGTCACCAGCATCTGTGTGGC	33	66.6	NdeI
WBI-C (Reverse)	ATCTAGAAGCTTGTCTGGGGCAGGCCTAGTTACAG	34	69.0	HindIII

The restriction sites are shown with red color. *TM: Melting temperature.

that there are five nucleotides which cause five silent mutations in the buffalo DNA proinsulin sequence as compared to the sequence reported for cow proinsulin (Younas, 2009). In order to evaluate these differences, the cloning and sequencing of buffalo proinsulin was carried out along with that of the cow prohormone using animals from the same location, Lahore. For this study, total RNA was isolated from buffalo and cow pancreas and converted to cDNA, by using WBI-C as reverse primer. cDNA was then amplified by using WB1-N and WB1-C as forward and reverse primers, respectively. The proinsulin genes from the two species were successfully cloned in pTZ57R/T vector and sequenced.

MATERIALS AND METHODS

The strains used for the cloning and expression of the constructs were *Escherichia coli* DH5 α , BL21 Codon plus (DE3)-RIL cells of *E. coli*. All the reagents and chemical were from Sigma, Across and Fisher. Two primers WB1-N and WB1-C (Table 1) were designed with Amplify 1.0 (Bill Engels © 1992, University of Wisconsin, Genetics, Madison, WI 53706) and prepared by GeneLink™.

Isolation of total RNA

The pancreatic tissues of freshly slaughtered buffalo (*B. bubalis*) and cow (*B. taurus*) were cut to a size of approximately 1 cm² and total cellular RNA was extracted by acid phenol-guanidinium thiocyanate-chloroform extraction method developed by Sambrook and Russell (2001). Total RNA was visualized on 1.8% agarose gel containing formaldehyde.

RT-PCR of buffalo and cow proinsulin genes

The mRNA of buffalo and cow proinsulin genes were converted to cDNA by using reverse primer WB1-C. In a reaction mixture containing 4.2 μ g of total RNA of buffalo or 25.2 and 58.8 μ g of cow, 100 pmoles of reverse primer WB1-C was added. The reaction mixture was incubated at 75°C for 5 min, followed by rapid chilling on ice. To the later was added 4 μ l of 5X reverse transcriptase buffer, 0.5 mM dNTPs and the mixture was incubated at 37°C for 5 min. Following the addition of 200 units of RevertAid™ M-MuLV reverse transcriptase, the mixture was incubated at 42°C for 60 min and then heated at 70°C for 10 min and chilled on the ice. The cDNA of buffalo was amplified by PCR. The reaction mixture was incubated at 94°C for 3 min then denatured at 94°C for 30 s, annealing at 57°C for 30 s and polymerization at 72°C for 30 s with final incubation at 72°C for 5 min. The PCR products were extracted from the agarose gel by using DNA extraction kit of Fermentas.

Cloning of cow and buffalo proinsulin genes in pTZ57R/T vector

Ligation of the desired PCR product to pTZ57R/T vector was done by using InsTAclone™ PCR cloning kit. The ligated products were used to transform competent cells of *E. coli*. Transformed cells appeared as white colonies whereas non transformed cells appeared as blue colonies. Mini-preparation of plasmid was done by the alkaline lysis method (Bimboim and Doly, 1979). The miniprep samples were restricted with *NdeI* and *HindIII* to confirm the right sized clones.

Sequence analysis

The sequencing was done by dideoxy chain termination method on Beckman Coulter CEQ™ 8000 DNA sequencer and on applied biosystems 310 DNA sequencer. Gene of interest in pTZ57R/T vector was sequenced through M13 forward and reverse primers. Sequence was aligned on DNA data bank Japan (DDBJ) sequence analysis software ClustalW.

Cloning of cow and buffalo proinsulin genes in pET21a vector

The 264 bp fragment of cow and buffalo proinsulin genes after restriction with *NdeI* and *HindIII* was ligated to pET21a vector and used to transform BL21 Codon plus (DE3)-RIL cells. The culture was allowed to grow at 37°C for 3 h until the O.D. at 600 nm reached 0.6 to 0.7. Then, 0.2 mM IPTG was added and allowed to grow at 37°C for 6 to 8 h. 1 ml culture of each construct was centrifuged at 5000 rpm for 5 min, the pellet washed with 1 ml of 20 mM Tris pH 8.0 and resuspended in 100 μ l of 20 mM Tris pH 8.0. 10 to 20 μ l of the resuspended cells mixed with 20 μ l of 2X SDS gel loading buffer (100 mM Tris-Cl, pH 6.8, 200 mM DTT, 4% SDS, 0.2% bromophenol blue, 20% glycerol). The induced and uninduced samples were analyzed on 18% SDS-PAGE.

RESULTS AND DISCUSSION

Total cellular RNA was isolated from three different freshly slaughtered pancreatic tissues of buffalo (*B. bubalis*) and two from cow (*B. taurus*) by acid phenol-guanidinium thiocyanate-chloroform extraction method (Chomczynski and Sacchi, 1987), in which guanidinium thiocyanate homogenate is extracted with phenol: chloroform at low pH. The A_{260}/A_{280} ratio of the extracted RNA was 1.2 to 1.27. Of the total cellular RNA, 80 to 85% comprises ribosomal RNA. The analysis of the samples (Figure 1) on agarose gel containing 2.2 M formaldehyde clearly showed two bands of 28S rRNA and 18S rRNA.

Messenger RNA was converted to cDNA using

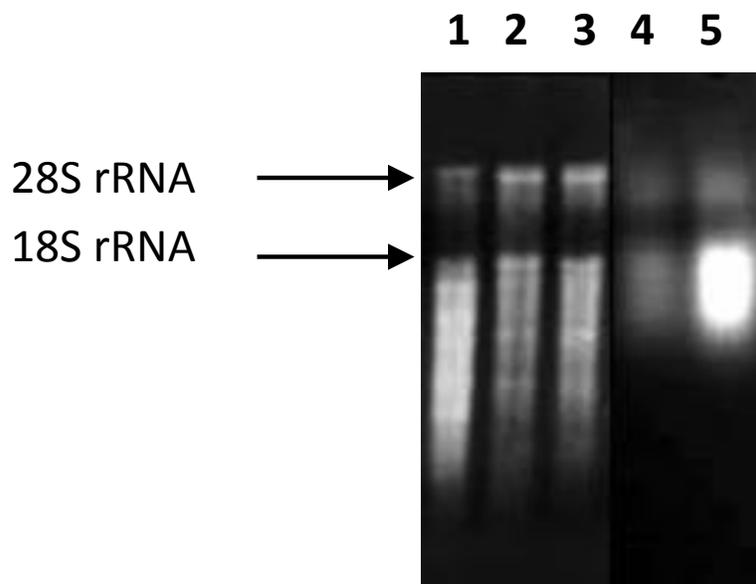


Figure 1. Agarose gel containing 2.2 M formaldehyde after 2 h, showing 28S rRNA (4.8 kb) and 18S rRNA (1.8 kb). Lanes 1, 2, 3: total RNA isolated from three different buffalo pancreatic tissues; Lanes 4, 5: total RNA isolated from cow pancreatic tissues.

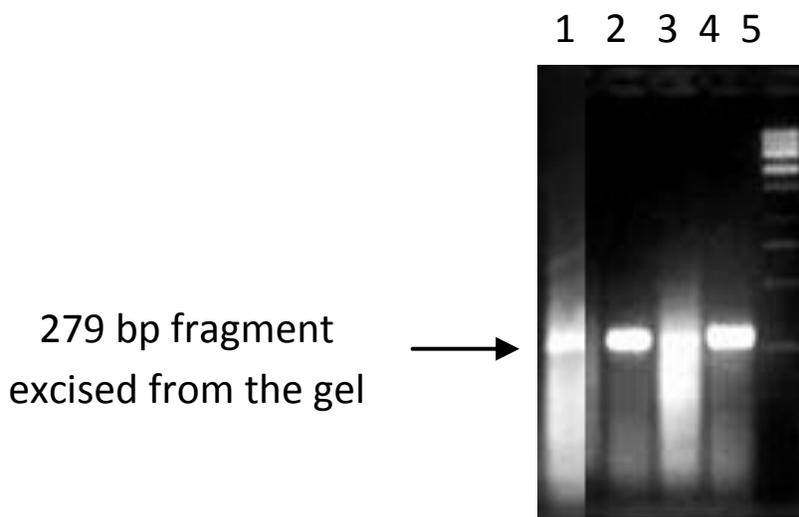


Figure 2. 2% agarose gel of RT-PCR of buffalo and cow proinsulin genes. Lane 1: RT-PCR of buffalo; Lanes 2, 3, 4: RT-PCR of cow; Lane 5: 1 kb DNA ladder (250-10000 bp).

Leukemia Virus reverse transcriptase and WB1-C primer having a *Hind*III site. cDNA was amplified by conventional PCR reaction using WB1-N as forward and WB1-C as reverse primer (Figure 2).

279 bp fragment after RT-PCR of buffalo and cow proinsulin gene was ligated to pTZ57R/T vector and used to transform DH5α cells. Figure 3 shows restriction of

plasmid with *Nde*I and *Hind*III.

Buffalo and cow proinsulin genes were sequenced with M13 forward as well as with M13 reverse primer. For the cow DNA sequence (Figure 5), it was found that in the C-peptide 7th amino acid, proline, is encoded by “CCC” in the case of Pakistani cow (*B. taurus*) which is the same codon “CCC” as present in Pakistani buffalo (*B. bubalis*).

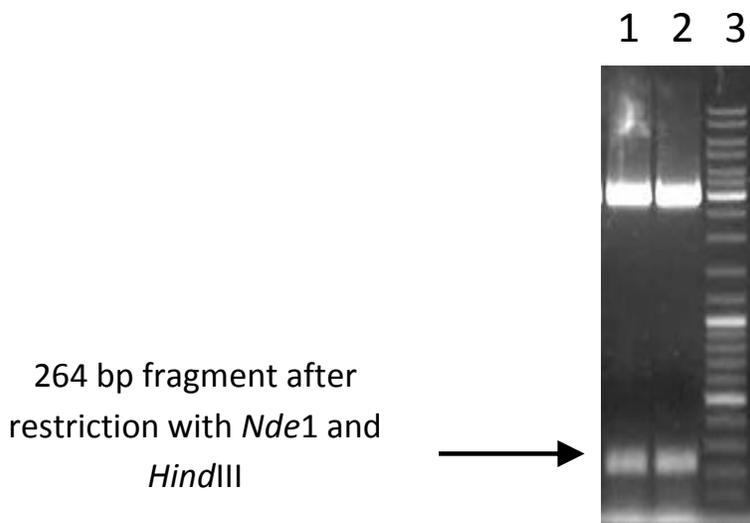


Figure 3. Agarose gel after *NdeI* and *HindIII* restriction of plasmid from three selected colonies. Lane 1: selected plasmid harboring buffalo proinsulin gene; Lane 2: selected plasmid harboring cow proinsulin gene; Lane 3: DNA ladder (100-10000 bp).

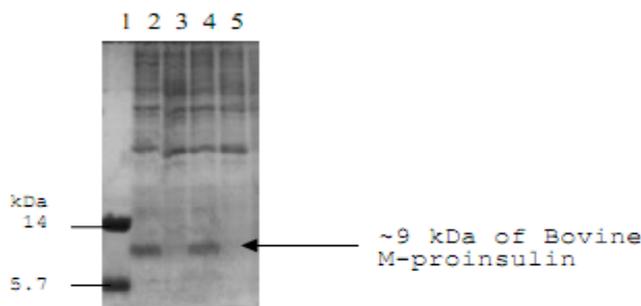


Figure 4. SDS-PAGE analysis of buffalo and cow M-proinsulin expressed in BL21 Codon plus (DE3)-RIL cells. Lanes 2 and 4: induced *E. coli* cells harboring cow proinsulin gene and buffalo proinsulin genes respectively; Lanes 3 and 5: their uninduced counterparts; Lane 1: protein marker, lysozyme (14 kDa) and bovine insulin (5.7 kDa).

However, it is “CCG” in the case of the sequence reported in the database for cow (*B. taurus*). This silent mutation in the proinsulin DNA sequence, indicates the genetic similarity of Pakistani cow with Pakistani buffalo. There is a study which suggests that cows and buffalos are derived from same ancestors (Abdel-Rahman, 2006).

It was found that in buffalo proinsulin sequence, compared to the published cow sequence, there was one mutation in the DNA encoding the B-chain, two in the C-peptide and two in the A-chain encoding regions. The sequence of buffalo proinsulin was submitted to GeneBank (DDBJ/EMBL) with an accession number of AB234871 (Olmos et al., 1994) and the sequence of Pakistani cow proinsulin with an accession number of

JX041514 (Aslam et al., 2013).

Many attempts have been made to express insulin and proinsulin in bacterial system using recombinant DNA technology (Olmos et al., 1994).

The buffalo M-proinsulin and cow M-proinsulin were expressed in BL21 Codon plus (DE3)-RIL cells, which were induced with 0.2 mM IPTG. Bacterial cells harboring buffalo and cow proinsulin genes expressed similar amounts of M-proinsulin, that is, approximately 30% of the total cellular proteins (Figure 4).

Insulin used clinically was derived from cow and pigs because of their close resemblance with the human insulin. Knowing the sequence of bovine insulin (cow and buffalo) of the local breed was important to know the

```

B. taurus (pak) TTC GTC AAC CAG CAT CTG TGT GGC TCC CAC CTG GTG GAG GCG CTG TAC CTG GTG TGC GGA GAG
B. taurus TTC GTC AAC CAG CAT CTG TGT GGC TCC CAC CTG GTG GAG GCG CTG TAC CTG GTG TGC GGA GAG
B. bubalis (pak) TTC GTC AAC CAG CAT CTG TGT GGC TCC CAC CTG GTG GAG GCG CTG TAT CTG GTG TGC GGA GAG
                   F V N Q H L C G S H L V E A L Y L V C G E

B. taurus (pak) CGC GGC TTC TTC TAC ACG CCC AAG GCC CGC CGG GAG GTG GAG GGC CCG CAG GTG GGG GCG CTG
B. taurus CGC GGC TTC TTC TAC ACG CCC AAG GCC CGC CGG GAG GTG GAG GGC CCG CAG GTG GGG GCG CTG
B. bubalis (pak) CGC GGC TTC TTC TAC ACG CCC AAG GCC CGC CGG GAG GTG GAG GGC CCG CAG GTG GGG GCG CTG
                   R G F F Y T P K A R R E V E G P Q V G A L

B. taurus (pak) GAG CTG GCC GGA GGC CCG GGC GCG GGC GGC CTG GAG GGG CCC CCG CAG AAG CGT GGC ATC GTG
B. taurus GAG CTG GCC GGA GGC CCG GGC GCG GGC GGC CTG GAG GGG CCC CCG CAG AAG CGT GGC ATC GTG
B. bubalis (pak) GAG CTG GCC GGA GGC CCG GGC GCG GGC GGC CTG GAG GGG CCC CCG CAG AAG CGC GGC ATC GTG
                   E L A G G P G A G G L E G P P Q K R G I V

B. taurus (pak) GAG CAG TGC TGT GCC AGC GTC TGC TCG CTC TAC CAG CTG GAG AAC TAC TGT AAC TAG
B. taurus GAG CAG TGC TGT GCC AGC GTC TGC TCG CTC TAC CAG CTG GAG AAC TAC TGT AAC TAG
B. bubalis (pak) GAG CAG TGC TGT GCC AGC GTC TGC TCT CTC TAC CAG CTG GAG AAC TAC TGT AAC TAG
                   E Q C C A S V C S L Y Q L E N Y C N Stop

```

Figure 5. Sequence alignment of Pakistani cow (*B. taurus* pak) with international cow (*B. taurus*) and pakistani buffalo (*B. bubalis* pak). Alignment represents five silent mutations between Pakistani buffalo and international cow, as highlighted in blue and green but Pakistani cow differs at only one nucleotide from international cow at proinsulin level, as highlighted in green.

evolutionary aspect behind the two species residing in the same area. Total RNA was isolated from pancreatic tissues of local (Pakistani) breeds of buffalo (*B. bubalis*) and cow (*B. taurus*), converted to cDNA and cloned in a T/A cloning vector. It was interesting to know that there are five silent mutations: one in B-chain, two each in A-chain and C-peptide encoding regions (Figure 5) of local Pakistani buffalo proinsulin DNA sequence as compared to the internationally reported cow proinsulin sequence. The sequence of *B. bubalis* proinsulin was reported for the first time and was submitted to GeneBank (DDBJ/EMBL) with an accession number of AB234871 (Olmos et al., 1994). The DNA sequence of local (Pakistani) cow proinsulin shows there is one nucleotide encodes C-peptide that has mismatch with the reported cow proinsulin sequence but mismatched nucleotide is same between Pakistani cow and Pakistani buffalo proinsulin sequence. This indicates the genetic similarity of Pakistani cow with Pakistani buffalo. There is a study which suggests that cows and buffalos are derived from same ancestors (Abdel-Rahman, 2006).

Many attempts have been made to express insulin and proinsulin in bacterial system using recombinant DNA technology (Olmos et al., 1994; Aslam et al., 2013). Buffalo and cow M-proinsulin were successfully expressed in BL21 Codon plus (DE3)-RIL cells and produced 30% of total cellular protein which is comparable with the protein expression of different

constructs of human proinsulin in BL21 Codon plus cells (Aslam et al., 2013)

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

- Abdel-Rahman S (2006). Evidences reveal that cattle and buffalo evolutionary derived from the same ancestor based on cytogenetic and molecular markers. *Biotechnology in Animal Husbandry* 22(3-4):1-9.
- Aslam F, Gardner Q-tAA, Zain H, Nadeem MS, Ali M, Rashid N, Akhtar M (2013). Studies on the expression and processing of human proinsulin derivatives encoded by different DNA constructs. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1834(10):2116-23.
- Bimboim H, Doly J (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids Research* 7(6):1513-23.
- Bolivar F (1994). Production in *Escherichia coli* of a rat chimeric proinsulin polypeptide carrying human A and B chains and its preparative chromatography. *Journal of biotechnology* 38(1):89-96.
- Chan SJ, Keim P, Steiner DF (1976). Cell-free synthesis of rat preproinsulins: characterization and partial amino acid sequence determination. *Proceedings of the National Academy of Sciences* 73(6):1964-8.
- Chomczynski P, Sacchi N (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical biochemistry* 162(1):156-159.

- Olmos J, Cruz N, Sánchez M, López M, Balbás P, Gosset G, Valle F, Patzelt C, Labrecque AD, Duguid JR, Carroll RJ, Keim PS, Heinrikson RL, Steiner DF (1978). Detection and kinetic behavior of preproinsulin in pancreatic islets. *Proceedings of the National Academy of Sciences* 75(3):1260-1264.
- Sambrook J, Russell D (2001). *Molecular Cloning: A Laboratory Manual*. 3rd Ed. New York: Cold Spring Harbor Laboratory Press.
- Younas H (2009). *Studies on recombinant buffalo (*Bubalus bubalis*) proinsulin and its derivatives* (Doctoral dissertation, University of Punjab, Lahore).