*Pak. J. Biochem. Mol. Biol. 2010; 43(4):163-168*

# Sequence analysis, high level expression and one-step purification of recombinant caprine growth hormone

*M. Altaf Khan[1], Saima Sadaf[1,2] and M. Waheed Akhtar[1]**
*[1]School of Biological Sciences and [2]Institute of Biochemistry and Biotechnology,*
*University of the Punjab, Lahore, Pakistan*

**Abstract:** This study deals with the sequence analysis, high-level expression, single step purification and refolding of growth hormone of a local caprine breed "Beetal". Sequence analysis of caprine growth hormone (cGH) gene showed several nucleotide and two amino acid variations in cDNA of cGH with the previously reported caprine GH sequence. cGH was found to be more closely related phylogenetically to sheep, yak and cattle GHs as compared to the other mammalian GHs. cGH carrying a $His_6$-tag at the N-terminus was expressed in *E. coli* up to 50 % of the total cell protein. The expressed protein was in the form of inclusion bodies, which were solubilized with 6 M guanidinium hydrochloride and then purified by metal affinity chromatography. The purified cGH, refolded by a step-wise dialysis process, was over 92 % purified and this product was biologically active as determined by Nb2 rat lymphoma cell proliferation assay.

**Keywords:** Caprine growth hormone, sequence analysis, expression, purification, refolding.
**Received:** August 12, 2010 **Accepted:** November 10, 2010
***Author for Correspondence:** mwapu@brain.net.pk

## INTRODUCTION

Beetal goat, one of the most common caprine breed of South Asia, is a source of both milk and meat. Among many other genetic markers growth hormone (GH) gene is a key genetic trait playing crucial role in the overall growth of the organism and development of its mammary glands. GH is a member of a multi-gene family, which plays an essential role in the lactation process[1]. Most of the mammals and lower primates contain single copy of GH gene in their genome[2,3]. Administration of exogenous GH is known to increase milk yield and growth in dairy animals[4-5].

Introduction of histidine tag ($His_6$) to protein has been a popular strategy for the single-step purification of proteins by metal affinity chromatography. It is generally introduced at the C-terminus of the protein. However, in some reports, $His_6$ tag addition at the N-terminus has led to a high-level expression, which was due to destabilization of the secondary structures at 5'-end of the gene transcript[6]. The enhanced expression of cGH by introducing silent mutations t 5'-end of the cGH gene has been previously reported[7]. This study reports nucleotide sequence analysis of GH from the local caprine breed "Beetal", high-level expression of the cGH tagged with $His_6$ at the N-terminal, followed by single-step purification and refolding to a biologically active state.

## MATERIALS AND METHODS

### *Sample collection, bacterial strains, plasmids and reagents*

Blood samples of Beetal goat were collected from Bahadarnager farmhouse, Okara Pakistan. *E. coli* strain DH5α was used for vector propagation while *E. coli* BL21 CodonPlus was employed for expression analysis. *Taq* DNA polymerase and T4 DNA ligase were purchased from MBI Fermentas (Ontario, Canada). InsT/A cloning kit was also acquired from Fermentas. Qiaquick gel extraction and Qiaprep spin miniprep kits used were obtained from Qiagen Inc. USA. IPTG and other chemicals used in this study were of the highest purity. DTCS Quick start kit was obtained from Beckman Coulter. 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT), used as color development substrates, were procured from MBI Fermentas and US Biological. Rabbit anti-bovine growth hormone, used in western blot analysis, was obtained from US Biological, while goat anti-rabbit IgG conjugated with alkaline phosphatase was obtained from Bio-Rad, USA.

### *PCR amplification*

The genomic DNA was isolated from the blood of animal using a previously described method[8]. Five sets of PCR primers were designed on the basis of published nucleotide sequence (Accession No. D000476) to amplify the full-length gene of cGH including the promoter region. Pair of primers used to produce overlaps in the amplified DNA are shown below, while the relative positions of the sequences for which the primers were designed are given in Figure 1.

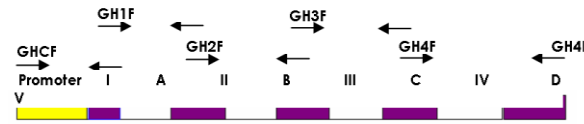


**Figure 1:** Schematic diagram of caprine GH gene showing the positions of exons (I-V), introns (A-D), and the primers used for amplifying overlapping segments

A typical PCR reaction contained 50 pmol of each primer pair, 100 ng of genomic DNA, 0.4 mM

of each dNTP, 5 µl of 10x PCR buffer with ammonium persulfate, 2 mM $MgCl_2$ and 2.5U of *Taq* DNA polymerase in a final volume of 50 µl. The thermocycling steps included an initial denaturation at 94ºC for 45 seconds, followed by 30 cycles of 94ºC for 45 seconds, 55ºC for 30 seconds and 72ºC for 1 minute and final extension at 72ºC for 30 minutes.

***Cloning and sequencing of cGH***

The PCR products were gel purified using Qiagen gel extraction kit and cloned into pTZ57R/T using InsT/A PCR product cloning kit according to manufacturer's recommendations. Positive colonies were confirmed by colony-pick PCR and the isolated plasmids were sequenced using M13 forward and reverse primers in a Beckman Coulter CEQ$^{TM}$ 8000 Genetic Analyzer. Sequencing reaction was prepared using DTCS Quick Start Kit. Each of the PCR products were sequenced individually and the complete sequence of Beetal GH was obtained by aligning the overlapping regions of the successive fragments. The positions of the intron-exon junctions were defined by aligning the sequence with that of the previously reported caprine GH sequence.

The nucleotide and amino acid sequences of other mammalian GHs were obtained by database homology searches using the BLASTN and BLASTP programs[9] available at ExPASy server. These were aligned using multiple sequence alignment program ClustalX[10,11], while the phylogenetic tree was constructed by neighbor joining method available on the same server.

***Expression, purification and refolding of His-tagged cGH***

For the construction of pHis-cGH expression plasmid, pET-cGH1[7] was used as a template. Hexa histidine tag was incorporated at the N-terminus of native sequence of cGH through PCR-based mutagenesis using HF/RP primer pair (HF 5'-GCAC*CATATG***CATCACCATCACCATCAC**GCCTTCCCAGCCATGTCCTTG-3',RP5'-GATAAGCTTGCAACTAGAAGGCACAGCTGGC-3'). The amplicon was cloned between *Nde*I and *Hin*dIII site of pET22b to generate pHis-cGH construct, which was thereafter used to transform *E. coli* expression host.

For expression analysis, transformed *E. coli* cells were grown in LB-ampicillin medium, induced with 0.5 mM IPTG to an $OD_{600}$ of 0.5-0.6 and harvested at 6 hours post induction. Inclusion bodies (IBs) were prepared from total cell lysate as described previously[12]. An aliquot of IBs was solubilized in 5 ml solubilization buffer [50 mM Tris-Cl (pH 8.0), 0.5 M NaCl, 5 mM imidazole, 6 M GdCl and 1 mM β-mercaptoethanol], stirred for 30-60 minutes at room temperature and then centrifuged at 20,000 rpm for 15 minutes at 4°C. The clear supernatant obtained after filteration through 0.45 µm was applied on a 5 ml Histrap chelating column (GE Healthcare, USA), which was equilibrated with solubilization buffer. After sample application, the column was washed with 25 ml equilibration buffer [50 mM Tris-Cl (pH 8.0), 0.5 M NaCl, 20 mM imidazole, 6 M GdCl and 1 mM β-mercaptoethanol] and then with 25 ml buffer containing 50 mM Tris-Cl (pH 8.0), 0.5 M NaCl, 20 mM imidazole, 6 M urea and 1 mM β-mercaptoethanol. Bound His-cGH was eluted using a linear gradient of imidazole [20 to 500 mM imidazole in 50 mM Tris-Cl (pH 8.0) containing 0.5 M NaCl]. 1 ml fractions containing the His-cGH were collected, pooled and subjected to dialysis against 250 ml of 50 mM Tris-Cl (pH 9.5) containing 4 M urea, 0.1 mM cysteine and 0.01 mM cystine overnight. After centrifugation at 20,000 rpm for 20 minutes, second dialysis was performed against 250 ml of 50 mM Tris-Cl (pH 9.5) containing 10 % glycerol and 5 % sucrose. The dialysate was again centrifuged at 20,000 rpm for 20 minutes and either lyophilized or used directly for biological activity assay.

***SDS-PAGE and immunoblot analysis***

Expression of His-cGH was analyzed by 13 % SDS-PAGE according to the method of Laemmli[13]. Resolved proteins were either visualized by staining with commassie brilliant blue R250 or subjected to immunoblot analysis according to previously described procedure[12]. After transfer of the protein to nitrocellulose membrane and blocking with TBS-T buffer [20 mM Tris (pH 7.6), 137 mM NaCl, 0.1 % (v/v) Tween 20] containing 2 % (v/v) gelatin and 0.4 % (v/v) sodium azide, the membrane was incubated with anti-bGH polyclonal antibodies for one hour. The incubation was followed by three washings with Tween 20, after which goat anti-rabbit IgG secondary antibody conjugated with alkaline phosphatase was applied. Detection of bound antibody was carried out by adding chromogenic substrate BCIP/ NBT.

Protein quantification in the samples was carried out by either UV absorbance or dye binding method[14] using bovine serum albumin (BSA) as standard.

***Biological activity assay***

The biological activity of His-cGH was assessed by Nb2 rat lymphoma cell proliferation assay as described earlier[12].

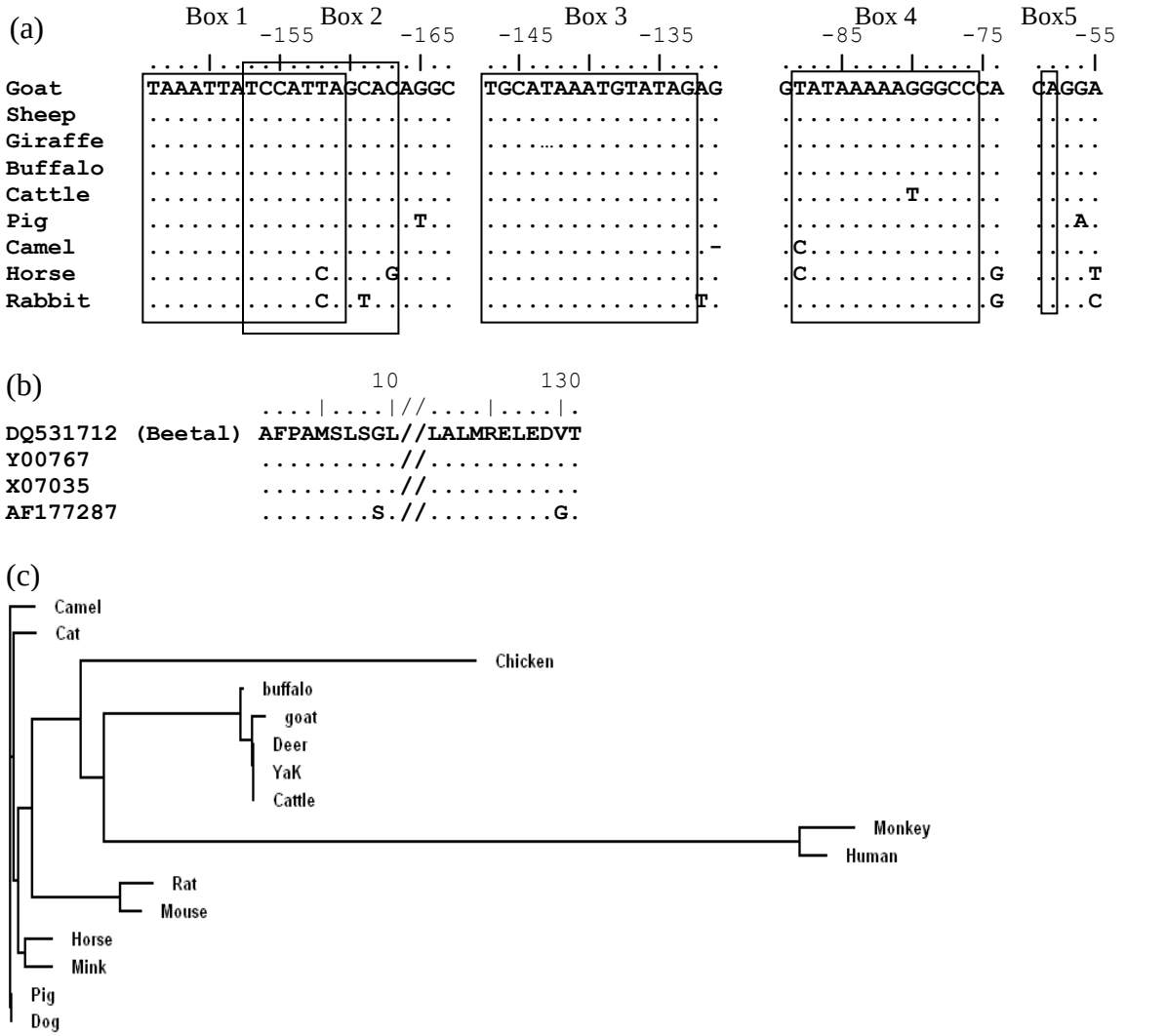


**Figure 2:** (a) Comparison of 5' untranslated region with other mammalian GHs viz., sheep (DQ461673), giraffe (AJ575421), water buffalo (AJ005116), cattle (M57764), pig (M17704), Arabian camel (AJ575419), horse (DQ845298), rabbit (Z38127). Conserved regions showing binding sites for transcription factors are shown in boxes. (b) Alignment of Beetal GH amino acid sequence with the reported sequence of other caprine GHs (Accession No. AF177287, Y00767 and X07035). (c) Phylogram based on amino acids sequences of GHs of different mammals i.e. buffalo (DQ307367), yak (EU344994), cattle (NM_180996), deer (AM049993), pig (X53325), camel (AJ575419), dog (AF069071), cat (NM_001009337), mink (X56120), chicken (M35609), rat (V01237), mouse (X02891), horse (U02929), human (BC090045), monkey (L16556).

The cells used in this study were a gift from Dr. Imran H. Khan (Center for Comparative Medicine, University of California Davis, USA)..

## RESULTS AND DISCUSSION

### *PCR amplification and sequence analysis*

The coding, non-coding and promoter regions of cGH gene were amplified from the genomic DNA of Beetal goat using five sets of overlapping primers as described under the Materials and Methods. Each set of primers yielded a PCR product of the expected size of 459-, 422-, 462-, 639-, and 640-bp, respectively. The amplicons were gel purified, cloned and sequenced both in the forward and reverse directions to ensure the correctness of sequence. The complete sequence of cGH gene, which was obtained by aligning the overlapping regions of the sequenced amplicons, was submitted to the GenBank database under the Accession No. DQ531712. Analyses revealed that the overall organization of cGH gene, consisting of five exons and four introns, is similar to the other mammalian GH genes. The exons I-V were 13-, 161-, 117-, 162- and 198- bps long, respectively. The four introns A-D contained 247-, 227-, 229- and 271 nucleotides,

respectively. Virtually around 60 % of the cGH gene is comprised of introns, which were identified in the coding sequence at positions identical to those in other mammalian GH genes and were of similar length except intron 4, which was slightly shorter than its counterparts.

The 5'-end untranslated region, which is of particular interest due to the sequences having role in expression regulation, was compared with the other mammalian GH genes. In the promoter region, several conserved sequences were identified, notably a TATAA sequence (Figure 2a, Box 4) located 24 bp upstream from the transcription start site (Box 5), two binding sites of pituitary transcription factors (Boxes 1 and 3, representing the distal and proximal binding sites for pituitary transcription factor 1, respectively) and the negative regulatory element (Box 2, overlapping the distal binding site in most mammals). No major difference could be seen during multiple sequence alignment, which reflects a strong conservation amongst the promoter sequences of related vertebrate GHs.

Furthermore, the cGH coding sequence largely agrees with the published sequence of Indian caprine GH (Accession No. AF177287), except that fifteen variations were observed at the nucleotide level. All the variations were silent except two, which encode glycine and valine at positions 9 and 130, respectively, instead of serine and glycine (Figure 2b). However, the amino acid sequence of cGH was found to be identical with another reported caprine GH sequence (Accession No. Y00767) reported previously[15]. Four cysteine residues, which are required to maintain structural integrity and biological activity, were also found almost at identical positions in all the mature GHs i.e., positions 53, 164, 182 and 189.

Beetal cGH showed higher degree of amino acid sequence similarity (90-99 %) with closely related species like buffalo, cattle, pig, deer, etc. but relatively lesser (64-77%) with monkey, chicken and human. The phylogenetic tree generated on the basis of GH amino acid sequences placed the goat, cattle, deer and yak on the same place/branching node and they all appear to have evolved by divergence from water buffalo, the most closely related ancestor of these species (Figure 2c).

**Expression of His-cGH**

*E. coli* cells harboring pETcGH1 carrying the native cGH sequence did not show any visible expression despite the presence of cGH in the correct reading frame[7]. It was observed that the 5'-end region of native cGH sequence and Shine Dalgarno (SD) sequence have a tendency to form hairpin loop which would provide hindrance to the accessibility of the ribosome for the formation of initiation complex and hence impede expression in the native form.

Addition of six histidine codons at the N-terminus of native cGH, however, helped in overcoming the expression barrier caused due to secondary structure formation. Upon induction with 0.5 mM IPTG, *E. coli* cells harboring pHis-cGH expressed the recombinant protein up to a level of ~50 % of the total cell protein, as analyzed by SDS-PAGE (Figure 3A). Due to the presence of $His_6$-tag, the number of amino acids was 197 amino acids instead of 191, a band of ~23 kDa could be analyzed on the gel. This agrees with a theoretical molecular weight of 22.67 kDa, calculated from the deduced amino acid sequence (ExPASy Protein Parameters Tools Analysis).

***Purification and refolding***

*E. coli* cells expressing His-cGH were harvested, resuspended in lysis buffer and then subjected to sonication. When analyzed on SDS-PAGE, almost all the His-cGH was found in the insoluble fraction with little or undetectable amount present in the soluble fraction (Figure 3A lanes 3 and 4). IBs were washed with triton X-100 and solubilized in 6 M GdCl before loading onto the Histrap column, which was preequilibrated with equilibration buffer. His-tagged recombinant protein can be purified in a single-step on an immobilized $Ni^{+2}$ affinity column, even in the presence of high concentration of denaturants like urea[16]. When the denatured solubilized proteins are applied onto the $Ni^{+2}$ column, protein containing His-tag gets adsorbed to the matrix while impurities and non-specific contaminants are washed away. The bound protein can then be eluted with increasing concentration of imidazole or by decreasing pH gradient[17]. In the present study, the bound His-cGH was eluted with a gradient of increasing concentration of imidazole. The product thus obtained was 92 % purified with a recovery of 54 % (Table 1).

**Table 1:** Purity and recovery of His-cGH at different stages of purification.

| Steps | Total Protein (mg)[a] | His-cGH (mg) | Purity (%) | Recovery (%) |
|---|---|---|---|---|
| Cell lysate | 48 | 24 | 50 | 100 |
| Inclusion bodies | 26 | 22 | 84 | 92 |
| Solubilization | 23 | 19 | 81 | 79 |
| $Ni^{2+}$ purification | 14 | 13 | 92 | 54 |

[a] 0.45 g wet weight cells were used as starting materials

For refolding of purified His-cGH, a stepwise dialysis procedure as previously described was adopted. The authenticity of the purified and refolded His-cGH was confirmed by western blot analysis, using rabbit anti-bovine GH, prior to biological activity assay (Figure 3A, lane 7).

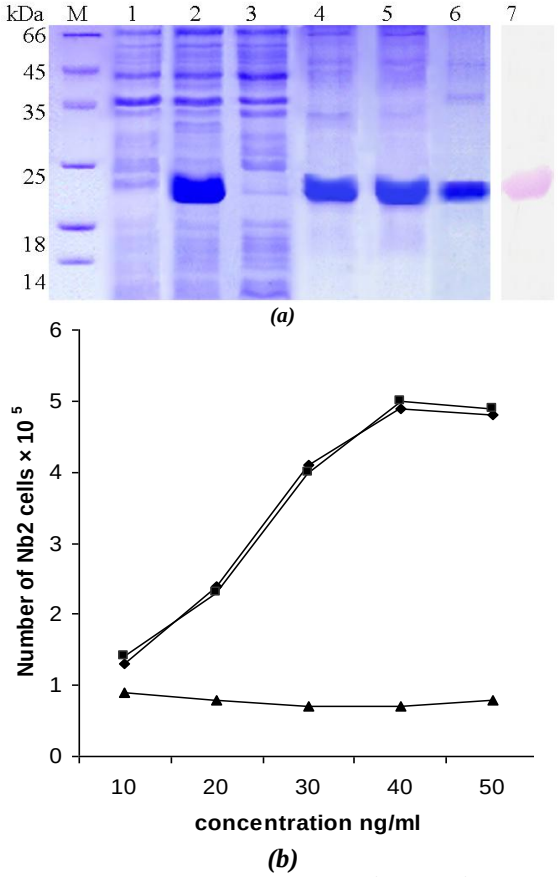


**Figure 3:** (a) SDS-PAGE analysis of proteins from *E. coli* cells transformed with recombinant pHis-cGH plasmids and the expressed His-cGH at different stages of purification and refolding. Lane M, size markers; lane 1, cells transformed with pETcGH1 with native sequence; lane 2, cells transformed with pHis-cGH; lanes 3, soluble cell lysate fraction; lane 4, insoluble protein from the cell lysate (IBs); lane 5, washed IBs; lane 6, IBs solublized and fractionated through a Ni-sepharose column; Lane 7, western blot analysis of the purified His-cGH. (b) Effect of different concentrations (10-50 ng/ml) of purified recombinant His-cGH on the proliferation of Nb2 rat lymphoma cells. (▲), Fischer's medium supplemented with BSA; (♦), refolded His-cGH; (■) and commercial bGH.

Refolded His-cGH showed the same rate of growth promoting activity in Nb2 lymphoma cell proliferation assay, as obtained with commercially available bovine GH (Figure 3b). As expected, the negative control did not show any activity.

This study extends the information on evolutionary relationships between the various caprine and other mammalian GH genes. Introduction of His-tag at N-terminal of the gene not only enhanced expression level but also allowed a simplified purification process. These procedures followed by stepwise removal of denaturant to renature the His-cGH to biologically active state should contribute towards its production at large scale.

## ACKNOWLEDGEMENT

We acknowledgement and thank Higher Education Commission of Pakistan for funding this research work.

## REFERENCES

1. Peel CJ and Bauman DE. Somatotropin and lactation. *J. Dairy Sci.*, 1987; 70: 474-486.
2. Adkins, RM, Nekrutenko A and Li WH. Bush baby growth hormone is more similar to nonprimate growth hormone than to rhesus monkey and human growth hormone. *Mol. Biol. Evol.*, 2001; 18: 55-60.
3. Wallis OC, Zhang YP and Wallis M. Molecular evolution of growth hormone (GH) in primates. Characterisation of the GH genes from slow loris and marmoset defines an episode of rapid evolutionary change. *J. Mol. Endocrinol.*, 2001; 26: 249-258.
4. Bauman DE. Bovine somatotropin and lactation: from basic science to commercial application. *Domest. Anim. Endocrinol.*, 1999; 17: 101-116.
5. Bonneau M and Laarveld B. Biotechnology in animal nutrition, physiology and health. *Livestock Prod. Sci*., 1999; 59: 223-241.
6. Sadaf S, Khan MA and Akhtar MW. Expression enhancement of bubaline somatotropin in E. coli through gene modifications in the 5'-end coding region. *J. Biotechnol*., 2008; 135: 134-139.
7. Khan MA, Sadaf S and Akhtar MW. Role of silent gene mutations in the expression of caprine growth hormone in *Escherchia coli. Biotechnol Prog*., 2007; 23: 1049-1052.
8. Blin N and Stafford DW. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.*, 1976; 3: 2303-2308.
9. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*., 1997; 25: 3389-3402.
10. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG and Thompson JD. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*., 2003; 31: 3497-3500.
11. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG. The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*., 1997; 25: 4876-4882.
12. Sadaf S, Khan MA and Akhtar MW. Production of bubaline somatotropin by auto-induction in *Escherichia coli*. *Biotechnol. Appl. Biochem*., 2007; 47: 21-26.
13. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 1970; 227: 680-685.
14. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of dye binding. *Anal. Biochem.*, 1976; 72: 248-254.

15. Yamano Y, Oyabayashi K, Okuno K, Miki Y, Noriyuki K, Manabe E, Hashi H, Sakai H, Komano T, Utsumi K and Iritani A. Cloning and sequencing of cDNA that encodes goat growth hormone. *FEBS Lett.*, 1988; 228: 301-304.
16. Li M, Su Z and Janson J. *In vitro* protein refolding by chromatographic procedures. *Protein Expr. Purif.,* 2004; 23: 1-10.
17. Arnau J, Lauritzen C, Petersen GE and John Pedersen J. Current strategies for the use of aYnity tags and tag removal for the purification of recombinant proteins. *Protein Expr. Purif.,* 2006; 48: 1-13.