ABSTRACT

Lactoferrin (Lf) is a component of the natural protection system of animals. The present study was undertaken to characterize the buffalo lactoferrin gene by cloning and sequence analysis. Total RNA was isolated from the lactating buffalo mammary gland tissue and lactoferrin cDNA was synthesized by RT-PCR technique, then cloned and sequenced. Sequence obtained from the cloned product was analyzed and submitted to NCBI GenBank (Acc. No. JF825526). The buffalo Lf sequence revealed a 2127 nucleotide long ORF coding for 708 amino acids with a signal peptide of 19 amino acids. Comparison with other livestock species revealed buffalo lactoferrin having 71–97% homology at nucleotide level and 64–96% homology at amino acid level highest with cattle as compared to other species. Based on Lf c-DNA sequences, the phylogenetic analysis also indicated that buffalo Lf having a close relationship with that of cattle. The 3D structure of buffalo lactoferrin was generated by Swiss-model and it was verified by PDBeSum, PROCHECK, ProSA z-score and ProQ, which was proved to be satisfactory. Preliminary information generated will be helpful in utilizing this important molecule for further studies in buffalo.

Key words: Buffalo, Lactoferrin, Sequencing

MATERIALS AND METHODS

Sample collection and RNA isolation: Tissue sample was collected from lactating buffalo mammary gland at Delhi abattoir and immediately kept into liquid nitrogen. Total RNA was isolated from approximately 250 mg of homogenized tissue, by TRizol method and further purified and treated with DNaseI using a commercial kit as per manufacturer’s instructions and quantified by UV spectrophotometer.

cDNA synthesis and cloning of buffalo lactoferrin gene: Reverse transcription was performed using oligo (dT)_{18} primers from 1 μg of the total RNA of mammary gland.
tissue using the cDNA synthesis kit, following manufacturer’s instruction. Primers were designed for the amplification of complete ORF of buffalo lactoferrin gene using DNASTAR software, based on the reported bovine lactoferrin gene sequence (Accession No.NM_180998). Full-length buffalo lactoferrin gene PCR product was generated using the cycles: initial denaturation at 95°C for 2 min 30 sec, 32 cycles at 94°C for 30 sec, 55°C for 30 sec and extension at 72°C for 1.0 min, followed by final extension at 72°C for 10 min. Reactions performed in 20μl total volume and contained 2μl of buffalo cDNA, 10pmol of forward 5’-TGGATAAAGGACGCAGAAC-3’ and reverse 5’-GGAGGCAGGCTAGCTTCTTT -3′ primer, 10mM of dNTPs, one unit of Taq DNA polymerase and 2.0μl of 10X reaction buffer supplied with enzyme (containing 15mM MgCl2). PCR amplified product was visualized on 1.5% agarose gel and purified using the PCR purification kit. Purified PCR product was cloned into pTZ57R/T cloning vector following the manufacturer’s instructions. Recombinant clones were selected and plasmid DNA was extracted, using a kit and sequenced using primer walking.

Sequences analysis and comparison with livestock species: Different fragments of cloned sequences were assembled and analyzed using the Lasergene software and further submitted to GenBank. BLAST was used to match and compare buffalo lactoferrin nucleotide sequences (http://www.ncbi.nlm.nih.gov/) with other livestock species and multiple sequence alignment was performed to predict phylogeny using ClustalW2 program (http://www.ebi.ac.uk/Tools/clustalW2/). To predict signal peptide region of translated amino acids, buffalo lactoferrin gene sequence was submitted to online SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Further physiochemical properties such as, molecular weight, pI, and extinction coefficient estimated half-life, instability index and grand average of hydropathy (GRAVY) were studied using online Protparam program and compared. A homology modeling of buffalo lactoferrin gene was performed using template X-ray structure at 2.8 resolution of the bovine lactoferrin, PDB code file 1BLF. The 3 dimensional (3D) models of buffalo lactoferrin protein were created by Swiss-model. Further superimposition, alignment and root mean-square (RMSD) determination of query and template structure was performed in ASH SuperPosition online tool (http://sysimm.ifrec.osaka-u.ac.jp/ash). The resulting model was evaluated by PDBsum, ProSA z-score, PROCHECK and ProQ. The visualization of 3D structure was performed by PyMol software.

RESULTS AND DISCUSSION

The complete ORF of buffalo lactoferrin gene of 2.2 kb was amplified by RT-PCR from the mRNA of mammary gland tissue, and sequenced by primer walking after cloning in T/A cloning vector. Contig was constructed from the overlapping sequences generated to get full-length gene sequence, which was submitted to NCBI and is available under the accession no. JF825526. The complete ORF of buffalo lactoferrin gene was 2127 bases in length encoding 708 amino acids similar to other livestock species such as cattle (Shashidharan et al. 2011), yak (Dong and Zhang 2006), sheep and goat except pig and mouse (Table 1).

Based on the lactoferrin nucleotide and amino acids sequences, a phylogenetic tree was constructed among different species (Fig. 1), which revealed that buffalo, cattle and yak were in one cluster (Bovinae) with sheep and goat being in another cluster (Caprinae) close to them. In bovinae group, buffalo had the closest relationship with cattle and next with yak. Whereas, the rest of species including pig, horse, human and mouse, were placed in 4 separate lineages in the phylogenetic tree. This clustering based on lactoferrin cDNA clearly showed the evolutionary relationship among the closely related species, and was also in agreement with the known taxonomic relationship of these domesticated species.
species (Abdel-Rahman 2006).

In silico translation of buffalo lactoferrin gene revealed the presence of total 72 amino acids which were ruminants specific and out of these 26 amino acids were present in the N-terminal region, 28 in the C-terminal region, 1 in the signal peptide region and 1 in between the signal peptide and N-terminal region. Apart from these, 7 amino acid changes were found to be buffalo specific compared to

(Continued)

Fig. 2. Amino acid sequence alignment of buffalo lactoferrin with other livestock species.
cattle, sheep, goat and yak and compared to all livestock species 2 were buffalo specific with the amino acids leucine (L) and threonine (T) at positions 173 and 287 respectively. The rest of species were found to have phenylalanine (F) and isoleucine (I) on respective positions (Fig. 2).

The buffalo lactoferrin sequence obtained from this work had 71–97% homology at nucleotide level and 64–96% homology at amino acid level when it was compared with other livestock species. Moreover, buffalo lactoferrin had higher homology both at nucleotide and amino acid level with cattle. These results supported the previous research findings on lactoferrin having high sequence similarity across most of livestock species (Shashidharan et al. 2011, Tang et al. 2012). All the species (ruminants, non-ruminants...
and primates) were found to have a 1–19 nucleotides signal peptide sequence in the 5'-terminus and found to have 100% homology with other ruminant species. The secondary structure of different livestock species was predicted and the number of α-helices and β-sheet are given in Table 1. Buffalo, goat and horse had similar number of α-helices and β-sheets, whereas, closely related species cattle had 15 α-helices and 23 β-sheets compared to buffalo.

Further physiochemical properties of livestock lactoferrin protein were analyzed and compared using ProtParam tools is presented in Table 2. All proteins ranging from 77 to 78kDa were similar to previously identified bovine lactoferrin and human lactoferrin (Pierce et al. 1991). The computed isoelectric point will be useful for separating the protein on a polyacrylamide gel by isoelectric focusing. The extinction coefficient can be used to calculate the concentration of a protein in solution. The value of instability index was observed between 40 and 45 predicting that proteins are stable and estimated half-life is 30 h.

GRAVY index indicated the solubility of proteins: a negative GRAVY value for lactoferrin from all the compared sources designates it to be hydrophilic in nature. These characteristics of buffalo lactoferrin were highly similar to that of cattle (Pierce et al. 1991, Shashidharan et al. 2011).

The 3D model of buffalo lactoferrin protein provides us invaluable insights into the structural basis of its function. The model generated by Swiss-model was subjected to energy minimization and assessed for both geometric and energy aspects using different online program. Several structure assessment methods including Ramachandran plots, Z-score and RMSD were also used to check the reliability of the predicted 3D model. ProSA calculated the interaction energy per residue. In this analysis, the interaction energy of each residue with the remainder of the protein is computed to determine whether or not it fulfills certain energy criteria. In buffalo lactoferrin homology model, the resulting z-score was around –14.89, similar to that of the template i.e. cattle lactoferrin available in PDB (1blf), further confirming that the energy profile of the models is consistent with a reliable conformation based on similarity with that of the templates. ProQ is a method to predict the quality of a protein model that extracts structural features, such as frequency of atom-atom contacts, and predicts the quality of a model, as measured either by LG score or MaxSub. If LG/MaxSub score>1.5/0.1 fairly good model, LG/MaxSub score ≥2.5/0.5 very good model and LG/MaxSub score>4.0/0.8 extremely good model (Cristobal et al. 2001). Predicted buffalo lactoferrin model revealed LG/Maxsub score 6.6/0.495 indicating it to be, very good model. The root mean square deviation (RMSD) indicated the degree to which two 3D structures are similar; the lower the value the more similar the structures. Both

<table>
<thead>
<tr>
<th>Properties</th>
<th>Buffalo</th>
<th>Cattle</th>
<th>Yak</th>
<th>Goat</th>
<th>Sheep</th>
<th>Pig</th>
<th>Horse</th>
<th>Human</th>
<th>Mouse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>708</td>
<td>708</td>
<td>708</td>
<td>708</td>
<td>703</td>
<td>708</td>
<td>711</td>
<td>707</td>
<td></td>
</tr>
<tr>
<td>Mol. weight</td>
<td>77.6</td>
<td>78.1</td>
<td>78.0</td>
<td>77.3</td>
<td>77.5</td>
<td>77.3</td>
<td>78.3</td>
<td>77.8</td>
<td></td>
</tr>
<tr>
<td>Theoretical PI</td>
<td>8.3</td>
<td>8.7</td>
<td>8.6</td>
<td>8.3</td>
<td>8.4</td>
<td>8.6</td>
<td>8.4</td>
<td>8.5</td>
<td>8.8</td>
</tr>
<tr>
<td>α-helices</td>
<td>77</td>
<td>76</td>
<td>77</td>
<td>75</td>
<td>73</td>
<td>70</td>
<td>77</td>
<td>79</td>
<td>73</td>
</tr>
<tr>
<td>β-sheets</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Estimated t1/2 (h)</td>
<td>40.7</td>
<td>40.9</td>
<td>40.8</td>
<td>42.5</td>
<td>42.2</td>
<td>42.0</td>
<td>44.8</td>
<td>44.9</td>
<td></td>
</tr>
<tr>
<td>Instability index</td>
<td>-0.26</td>
<td>-0.29</td>
<td>-0.29</td>
<td>-0.24</td>
<td>-0.24</td>
<td>-0.29</td>
<td>-0.32</td>
<td>-0.34</td>
<td>-0.41</td>
</tr>
<tr>
<td>GRAVY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. (a) Homology model of buffalo lactoferrin protein using template bovine 1BLF; (b) Superimposition of the homology model and crystal structure of bovine template.
template and query structures were superimposed for the calculation of RMSD. The RMSD value obtained from superimposition of buffalo lactoferrin with known cattle lactoferrin (1blf), using ASH structural superposition, was found to be RMSD with template 0.082 and over a total of 685 aligned residues, suggesting reliable 3D structure (Fig. 3). So the representative structure verified by PDBsum, PROCHECK, ProSA z-score and ProQ was satisfactory and can thus be considered as a reliable source for further analyses.

Structure analysis of buffalo Lf revealed the presence of 2 lobes named N- and C- belongs to the transferring family at the positions 25–352 and 364–693 respectively. These lobes are connected covalently by a 3–turn $\alpha$-helix involving 11 amino acid residues (Fig. 3). Similar structure was also reported by Moore et al. (1997), these lobes are highly homologous with each other, and are connected by a hinge region between amino acids, containing parts of $\alpha$-helices, which provide additional flexibility to the molecule.

The information generated on buffalo lactoferrin gene will be helpful in utilizing this important molecule for further studies, including sub-cloning and expression in different prokaryotic and eukaryotic systems for the production of recombinant molecule, having therapeutic use.

ACKNOWLEDGEMENTS

The authors wish to thank Director, National Bureau of Animal Genetic Resources (ICAR) for providing necessary facilities to carry out the work. Technical assistance received from Mr Naresh Kumar, Technical Officer, Buffalo Genomics Lab, is gratefully acknowledged.

REFERENCES


