Molecular cloning, promoter analysis and SNP identification of Italian Nicastrese and Saanen lactoferrin gene

Molecular cloning, promoter analysis and SNP identification of Italian Nicastrese and Saanen lactoferrin gene

Pauciullo A.[a], Cosenza G.[a], Nicodemo D.[a], Gallo D.[a], Mancusi A.[a], Crepaldi P.[b], Di Berardino D.[a], Ramunno L.[a*]

[a]*Department of Soil, Plant, Environment and Animal Production Science, University of Naples "Federico II", Napoli, Italy*

[b]*Department of Animal Science, University of Milano, Milano, Italy*

**Corresponding author: Ramunno Luigi. Department of Soil, Plant, Environment and Animal Production Science, University of Naples "Federico II", Via Università 133, 80055 Portici (Na), Italy. Phone: ++39/0812539004; fax: ++39/0817762886; e-mail: ramunno@unina.it*

ABSTRACT

Lactoferrin (Lf) is an iron-binding glycoprotein found in the exocrine secretions including milk. High levels of lactoferrin may have a role in the prevention of microbial infection of the mammary gland. Our preliminary studies on mastitis incidence in Italian goats led us to consider a different active role of lactoferrin in the mammary gland defence mechanisms. In this report we sequenced and characterized the goat lactoferrin cDNA and its promoter region in two different breeds of goat. The complete cDNA comprised 2356 nucleotides, including 38 bp at the 5'-UTR and 194 bp at the 3'-UTR. The open reading frame is 2127 bp long and it encodes a mature protein of 689 aminoacids. A total of 19 nucleotide differences, 11 of them being responsible for 8 aminoacid changes, were identified through the comparison with French, Korean and Tibetan goat lactoferrin cDNAs. About 1700 bp of the lactoferrin gene promoter were sequenced. Sequence analysis revealed a noncanonical TATA box, multiple SP1/GC elements, and other putative binding sites for transcription factors, such as NF-κB, STAT3 and AP2. Two SNPs were identified, one of

which would seem to create a new putative AP2 consensus sequence. The presence of an additional AP2 binding site could be associated to quantitative differences of such protein fraction.

KEYWORDS: Lactoferrin, Iron-binding protein, Gene promoter, Transcription factor, AP-2, Goat.

Lactoferrin is a non-haem bioactive iron-binding glycoprotein of about 80-KDa which has been found in various concentrations in milk and other exocrine secretions. Many functions have been attributed to lactoferrin, including iron homeostasis, immuno-modulatory, anti-inflammatory activity, anti-viral, anti-fungal and anti-parasitic activity, enzymatic and anti-oxidant action (for a review, González-Chávez et al. 2009); furthermore, it can induce apoptosis and arrest tumor growth in vitro and it has the ability to modulate the production of cytokines in cancer (Öztaş Yeşim & Özgüneş, 2005). However, the best known role of lactoferrin is the primary defence against microbial infection, mainly through iron sequestration required for microbial growth, direct interaction with bacterial surface (for a review, Legrand et al. 2008) and direct bactericidal activity due to the lactoferricin peptide (Bellamy et al. 1992; van der Kraan et al. 2006).

The well documented antimicrobial action of lactoferrin and its essential role in the natural defence mechanism of the mammary gland make it a candidate gene for increasing resistance against bacterial infections in farm animals. In dairy cattle, its concentration dramatically increases during the dry period and a mastitis infection (Kutila et al. 2003). This suggests that this protein plays important physiological roles and it can reduce the incidence of clinical and subclinical mastitis (Hagiwara et al. 2003). In piglets, the lactoferrin content can be an indicator of enterobacterial infection (Tore et al. 1988); in the goat, the same concentration of lactoferrin secreted from two different genetic types (Korean native and Saanen goat) showed different antibacterial activity. In particular, Saanen goat lactoferrin exhibited no activity against *Escherichia coli* O111, even at 7.5 mg/ml, whereas Korean native goat lactoferrin was already effective at 5 mg/ml. These observations suggested that the different antibacterial activities may be the result of

differences in the protein conformation caused by the polymorphisms in the goat lactoferrin gene (Lee et al. 1997). This gene has been mapped on chromosome 22 (Le Provost et al. 1994); its main feature is the extremely split architecture, consisting in 17 exons ranging in size from 48 bp (exon 11) to 226 bp (exon 17). Although the functions of this protein have received adequate attention, the molecular mechanisms of gene expression and regulation remain relatively unknown.

Studying mastitis incidence in Southern Italy goat breeds, we noticed that its occurrence in Nicastrese goat was lower compared to the observed cases for other breeds as, for example, Saanen goat (unpublished data). Our interest in these data prompted us to speculate that lactoferrin could play a more active role in mammary gland innate defence mechanisms against infections. The identification of genetic polymorphisms responsible for improved resistance to mammary infections could be conveniently used in breeding programmes in order to reduce mastitis incidence in farm animals. Aim of this work was, therefore, to characterize the lactoferrin cDNA, the regulatory regions of gene promoter in the Italian Nicastrese and Saanen breeds and to identify mutations which could potentially affect milk lactoferrin concentration or produce aminoacid changes.

Genomic DNA was isolated from leukocytes obtained from individual blood samples, using conventional phenol-chloroform extraction method followed by DNA precipitation with absolute ethanol. The isolated DNA was then resuspended in 100 µl TE buffer pH 7.6 (10 mM Tris, 1mM EDTA) (Gossens & Kan, 1981). The DNA was quantified spectrophotometrically and the integrity assessed via gel electrophoresis (1.0%).

According to Chomczynsky and Sacchi (1987), total RNA was extracted from milk somatic cells of five goats per breed, at the end of the lactation stage. The goats were comparable for age, type of feed, diet, and feeding level. RNA was quantified by spectrophotometer and its integrity tested by gel electrophoresis (1.5%). cDNA synthesis was performed using Improm-II™ Reverse Transcriptase (Promega, Madison WI, USA) as follows: 1 µg of total RNA was incubated with 10 µM of the reverse primer LF17R: 5'-AGGGAATGAAAATCAACAGCA-3' in a volume of 10 µl at 70 °C for 5min to denature RNA secondary structures. The 10-µl RT mix were prepared

according to the manufacturer's guidelines, and reactions were incubated at 42 °C for 1 h, followed by 70°C for 15 min to inactivate the RT enzyme.

The whole cDNA and the promoter region of the goat lactoferrin gene were amplified using an iCycler IQ (Bio-Rad, Hercules, California). Primers for amplifications and sequencing were designed by DNASIS-Pro (Hitachi Software, San Bruno, CA) using, as templates, the complete sequence of the goat lactoferrin cDNA (EMBL Acc. No.: X78902) for the couple LF1F 5'-CGGAGTCGCCCCAGG-3' and LF17R (amplicon A) and the sequence of the homologous cattle lactoferrin gene and promoter region (EMBL Acc. No.: AY319306) for the following primers LFPRF 5'-TCCTTTTCATTGGCAAATGAG-3' and LF1R 5'-GGCGGGGACGAAGAG-3' (amplicon B); LF5'F 5'-AGATACAAAGATGCTTCA-3' coupled with a primer designed on newly determined goat promoter sequence LFPRR 5'-TGGCAGAGGCAATAT-3' (amplicon C).

The PCR reaction mix (50 μl) comprised: 100 ng of genomic DNA or cDNA, 1X PCR Buffer (Promega), 2.5 mM $MgCl_2$, 5 pmol of each primer, dNTPs each at 400 μM, 2.5 U of *Taq* DNA Polymerase (Promega). PCR was performed under the following conditions: the first cycle involved a denaturation step at 95°C for 4 min, followed by 35 cycles at 95°C for 60 s, 58°C (amplicons A and B) – 54°C (amplicon C) for 45 s, 72°C for 90 s, the final extension step was carried out at 72°C for 10 min. PCR products were purified with QIAquick columns and cloned into pDrive vector (QIAGEN S.p.A., Milano, Italy). The ligation products were transformed into QIAGEN EZ Competent cells following the manufacturers' guidelines (QIAGEN PCR Cloning plus Kit). Screening of the recombinant clones was performed by PCR using the plasmid primers M 13. Nucleotide sequencing was carried out according to the dideoxynucleotide chain-termination technique (Sanger et al., 1977) by using a BigDye™ Terminator cycle sequencing kit (Applied Biosystems, Warrington, UK) and an ABI PRISM 377-18 (Applied Biosystems, Foster City CA, USA) nucleotide sequencer.

More than 10 clones of each PCR product on both strands were sequenced. Chromatograms were edited and trimmed to remove the vector sequences using DNASIS-Pro (Hitachi). Homology

searches, comparison among sequences, and multiple alignments were accomplished using ClustalW2 (Larkin et al., 2007). The putative transcription factor binding sites were searched by AliBaba 2.1 software (Grabe, 2002).

The complete cDNA coding for the goat lactoferrin was sequenced for the two investigated breeds. It comprised 2356 nucleotides (EMBL Acc. No. FM875929). Analysis of the sequence revealed one open reading frame 2127 bp long (positions 39–2162), plus 38 nucleotides and 194 nucleotides in the 5'- and 3'-UTR, respectively. The signal peptide (19 aminoacids) is encoded by the nucleotides 39-95 (last 43 nucleotides of the exon 1 and first 14 bp of the exon 2), the translation stop codon (TAA) is realized between the nucleotides 2163-2165, whereas the polyadenylation signal sequence is in positions 2322-2327. No nucleotide differences were found in the two investigated breeds, whereas the comparison with the lactoferrin cDNA of French (EMBL Acc. No.: X78902), Korean native (U53857) and the Tibetan goat (DQ387456), showed a total of 19 nucleotide differences (respectively 9, 8 and 5), 11 of them being responsible for 8 aminoacid changes (Table 1). As a consequence, the Italian goat lactoferrin cDNA appears to be a new and particular allelic combination. In fact, the first part of the Italian lactoferrin cDNA (ex 1 – ex 8) appears to be more similar to the Korean and Tibetan cDNAs, conversely the second fraction (ex 9 – ex 17) more homologue to the French counterpart. Neverthless, the different nucleotide combination does not determine remarkable difference in the estimated molecular weight of the deduced mature protein (689 aminoacids) (Table 1).

By comparing the deduced lactoferrin aminoacid sequence with the known Lf sequences available in GenBank, we found that the degree of conservation of Lf among species varies from 56.4% *vs* mouse to 97.7% *vs* sheep (Fig. 1). As for the other species, even the goat lactoferrin is a simple polypeptide chain folded into two symmetrical lobes (N lobe: aminoacids 1-332 and C lobe: aminoacid 344-689 of the mature protein), connected by a hinge region containing an α-helix (aminoacids 333-343), which provides flexibility to the molecule (Fig. 1). The analysis of the primary structure of the goat lactoferrin showed: a) the signal peptide, at the amino terminus, which

is fully conserved among the ruminants with a cleavage site at positions 19 and 20 (*LGLCLA*↓AP); and b) mature protein at the carboxyl terminus. Other features include: a) Three potential bioactive peptides: one lactoferricin (KLGAPSITCVRR), which is one of the main peptides responsible for antibacterial properties of lactoferrin in cattle (Bellamy et al., 1992) and appears to be strongly conserved among ruminants; and two lactoferrampins (NLIWELLRKAQEKFGKNKSQ and WELLRKAQEKFGKNKSQ), which in cattle show strong bactericidal activities against Gram-positive and Gram-negative bacteria as, for instance, *Bacillus subtilis* and *Echerichia coli* (van der Kraan et al. 2006); b) Five potential glicosilation sites (Asn-X-Ser/Thr) one of which (NQT) at the Asn495 residue is fully conserved among the species (Fig. 1).

In order to understand whether lactoferrin could affect the different mastitis incidence between these two breeds even in absence of polymorphisms in the coding regions, we decided to extend the sequencing to the 5' flanking region and to analyze the putative regulatory regions. A fragment of about 1700 bp of the lactoferrin gene was amplified and sequenced (EMBL acc. no. AJ784283). The lactoferrin promoter contains a noncanonic TATA box located, with reference to the first nucleotide of the first exon, at nucleotides -28/-24; at least three GC box (-1369/-1364, -948/-934, -164/-159) and two SP1 binding sites (-200/-187, -69/-55). The high GC content (64.9%) is a typical feature of housekeeping genes promoters and the presence of these last sites could contribute to the constitutive expression of the lactoferrin in various tissues. In particular, a noncanonical TATA box (ATAAA) and a SP1 binding site at position -69/-55 are almost perfectly conserved in the region immediately upstream the exon 1 in other 7 species (cattle, human, pig, mouse, rat, camel and buffalo), suggesting that these two sites might play a fundamental role in the regulation of lactoferrin gene *in vivo* (Zheng et al. 2005). The region spanning from nucleotides -1009 to -841 contains three potential LPS-responsive elements, in particular a STAT3 binding site (-1009/-991), and two NF–κB sites (-960/-949, -852/-838). STAT 3 is known as a response factor to acute inflammatory phases, promoting the growth and cellular recovery, whereas NF-κB plays a

central role in immunitary response and it can be activated by Toll-like receptors (Muzio et al. 2000).

Compared to the Saanen conterpart, the promoter sequence of the Nicastrese lactoferrin gene, is characterized by two SNPs: a transition A→G (nt -333) and a cytosine insertion realized between the nucleotides -87/-86 (Fig. 2). The first one, apparently does not interest any known regulatory site, whereas the cytosine insertion would seem to create a new putative AP2 binding site (GCCTGGGG) (Fig. 2), located 53 bp upstream the TATA box, and between two SP1 binding sites, motif already known to participate in the recruitment of the general transcription machinery and to play an important role in the basal activity of lactoferrin promoter (Zheng et al. 2005). The estrogen receptor AP2 is known to be an important transcription factor regulating the differential expression of lactoferrin gene in pigs and mice (Wang et al. 1998; Liu & Teng 1991). Furthermore, the new AP2 falls into the CpG island of the gene promoter. Such region is known to facilitate a rapid response to microbial infection by inducing the production of IL-6 in milk, prompting the release of the TNF-α, reducing *E. coli* counts in milk, attenuating the damage of inflammatory mediatators with a reduction in symptoms of inflammation (Zhu et al., 2007). The insertion of a cytosine and the creation of an additional binding site for the activator protein transcription factor-2 in such CpG region could enhance the protein production and its correlated effects. An example of how an AP2 can enhance the transcription is shown in bovine lactating mammary gland, where the different binding affinity of the AP2 transcription factor to the gene promoter of another whey-protein (β-lactoglobulin) brings to the creation of a more or less efficient transcriptional complex, which is responsible for differences in gene expression (Lum et al. 1997) even associated with total milk protein content (Kuss et al., 2003). Thus, as observed for such whey-protein, it would be reasonable to think that the presence of an additional AP2 binding site for the goat lactoferrin could increase the transcriptional activity and be associated to quantitative differences in gene expression, which could enhance all the activities related to such protein, improve mammary gland defence and, thereby, have a beneficial effects against mastitis infection in goat.

In conclusion, we found 19 nucleotide differences in the Italian goat lactoferrin cDNA which appears to be a new and particular allelic combination compared to the French, Korean native and the Tibetan goats. These nucleotide differences, responsible for 8 aminoacid changes, could be translated in a different antibacterial activity as result of a different lactoferrin conformation caused by the polymorphisms at gene level. Furthermore, we identify two SNPs in the 5' region, one of which affects an AP2 transcription factor binding site, which could increase the gene expression and enhance all the activities of the protein, including the mastitis defence. These findings serve as a first orientation for further studies, focused on the evaluation of lactoferrin activity in the goat mammary gland. Such activity can be influenced by the polymorphisms identified at cDNA level, which could modify the functionality of the protein itself, or by the effect of possible differences in lactoferrin concentration due to the additional presence of an AP2 consensus sequence. As already suggested in water buffalo for the *Nramp1* gene (Capparelli et al., 2007), it would be useful to investigate possible associations between the activity of the lactoferrin and the identified SNPs; this could represent, in the next future, a useful tool for 'marker' assisted selection programmes, in order to reduce mastitis incidence in goat farms and -consequently- to increase the productive efficiency of the dairy industry.

ACKNOWLEDGEMENTS

This research was financially supported by the Italian Ministry for Agriculture and Forestry Policy - MiPAF (SELMOL project).

REFERENCES

Bellamy, W., Takase, M., Wakabayashi, H., Kawase, K., Tomita, M., 1992b. Antibacterical spectrum of lactoferricin B, a potent bactericidal peptide derived from the N-terminal region of bovine lactoferrin. J. Appl. Bacteriol. 73, 472-479.

Capparelli, R., Alfano, F., Amoroso, M.G., Borriello, G., Fenizia, D., Bianco, A., Roperto, S., Roperto, Iannelli, D. 2007. Protective Effect of the Nramp1 BB Genotype against *Brucella abortus* in the Water Buffalo (*Bubalus bubalis*). Infect. Immun. 75 (2), 988-996.

Chomczynsky, P., Sacchi, N., 1987. Single step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. Anal. Biochem. 162, 156–159.

González-Chávez, S.A., Arévalo-Gallegos, S., Rascón-Cruz, Q., 2009. Lactoferrin: structure, function and applications. Int J Antimicrob Agents 33 (4), 301.e1-e8.

Gossens, M., Kan, Y.W., 1981. DNA analysis in the diagnosis of haemoglobin disorders. Methods Enzymol. 76, 805–817.

Grabe, N., 2002. AliBaba2: Context specific identification of transcription factor binding site. In Silico Biol. 2, S1–S11.

Hagiwara, S., Kawai, K., Anri, A., Nagahata, H., 2003. Lactoferrin concentration in milk from normal and subclinical mastitic cows. J. Vet. Med. Sci. 65, 319-323.

Kuss, A.W., Gogol, J. and Geldermann, H., 2003. Associations of a polymorphic AP-2 binding site in the 5'-flanking region of the bovine β-lactoglobulin gene with milk proteins. J. Dairy Sci. 86, 2213-2218.

Kutila, T., Pyorala, S., Saloniemi, H., Kaartinen, L., 2003. Antibacterial effect of bovine lactoferrin against udder pathogens. Acta Vet. Scand. 44, 35-42.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. ClustalW and ClustalX version 2. Bioinformatics 23, 2947–2948.

Lee, T.H., Shimazaki, K., Yu, S.L., Nam, M.S., Lee, K.K., Yu, D.Y., 1997. Polymorphic sequence of korean native goat lactoferrin exbhiting greater antibacterial activity. Anim. Gen. 28, 367-369.

Legrand, D., Pierce, A., Elass, E., Carpentier, M., Mariller, C., Mazurier, J. 2008. Lactoferrin structure and functions. Adv. Exp. Med. Biol. 606, 163-194.

Le Provost, F., Nocart, M., Guerin, G., Martin, P., 1994. Characterization of the goat lactoferrin cDNA: assignment of the relevant locus to bovine U12 synteny group. Biochem. Biophys. Res. Commun. 203, 1324-1332.

Liu, Y., Teng, C.T., 1991. Characterization of estrogen-responsive mouse lactoferrin promoter. J. Biol. Chem. 266 (32), 21880-21885.

Lum, L.S., Dovc, P., Medrano, J.F., 1997. Polymorphisms of bovine β-Lactoglobulin promoter and differences in the binding affinity of activator protein-2 transcription factor. J. Dairy Sci. 80, 1389-1397.

Muzio, M., Polentarutti, N., Bosisio, D., Manoj Kumar, P.P., Mantovani, A., 2000. Toll-like receptor family and signalling pathway. Biochem. Soc. T. 28, 563-566.

Öztaş Yeşim, E.R., and Özgüneş, N., 2005. Lactoferrin: a multifunctional protein. Adv. Mol. Med. 1, 149-154.

Sanger, F., Nicklen, S., and Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U S A. 74 (12), 5463–5467.

Tore, J.G., Ola, R., Trond, J., Ove, A., 1988. Lactoferrin as inhibitor of septicaemia and endotoxemia in pigs. Scand. J. Infect. Dis. 20, 659-666.

van der Kraan, M.I.A., Nazmi, K., van't Hof, W., Amerongen, A.V., Veerman, E.C., Bolscher, J.G., 2006. Distinct bactericidal activities of bovine lactoferrin peptides LFampin 268–284 and LFampin 265–284; Asp-Leu-Ile makes a difference. Biochem. Cell Biol. 84, 358–362.

Wang, S.R., Lin, J., Cheng I.C., Lin T.Y., 1998. Characterization and functional analysis of the porcine lactoferrin gene promoter. Gene 215, 203-212.

Zheng, J., Ather, J.L., Sonstegard, T.S., Kerr, D.E., 2005. Characterization of the infection-responsive bovine lactoferrin promoter. Gene 353, 107-117.

Zhu, Y.M., Miao, J.F., Zhang, Y.S., Zhen, L., Zou, S.X., Deng, Y.E., 2007. CpG-ODN enhances mammary gland defense during mastitis induced by *Escherichia coli* infection in goats. Vet. Immunol. Immunopathol. 120, 168-176.