Short Communication: Duplication in the 5'-Flanking Region of the β -Lactoglobulin Gene is Linked to the *BLG* A Allele

M. H. Braunschweig

Institute of Genetics, University of Berne, CH-3001 Berne, Switzerland

ABSTRACT

β -Lactoglobulin (β -LG) is the major whey protein in the milk of cows and other ruminants. It is well established that the predominant genetic variants β -LG A and B are differentially expressed. Extensive investigation of the genetic variation in the promoter region of the *BLG* gene revealed the existence of specific haplotypes associated with the A and B variants. However, the genetic basis for the differentially expressed *BLG* A and B alleles is still elusive. In this study additional genetic variation further upstream in the 5'-flanking region of the *BLG* gene was identified, including 6 single nucleotide substitutions, a single nucleotide deletion, and a 7-bp duplication. Comparison of DNA sequences showed that the investigated 5'-flanking region is highly conserved between ruminants, and the duplication g.-1885_-1879dupCTCTCGC and the substitution g.-1888A>G are only found in the *BLG* A and D alleles in cattle. The cytosine at position g.-1957 and the thymine at positions g.-2008 and g.-2049 are only found in *BLG* B alleles of cattle. It is suggested that the described genetic variability contributes to the differential allelic expression of the *BLG* gene.

Key words: β -lactoglobulin, genetic polymorphism, milk protein, expression

β -Lactoglobulin (β -LG) is the major whey protein in the milk of ruminants. It is also present in the milk of most other mammals but is absent in rodents, lagomorphs, and in human milk. In a recent review it was concluded that the expression of β -LG in the mammary gland is primarily an important source of amino acids for the offspring and that this function arose by gene duplication from the physiologically essential glycode-lin (Kontopidis et al., 2004). Aschaffenburg and Drewry (1957) reported 50 yr ago that β -LG variants A and B are differentially expressed and they suggested a genetic background for this observation. In numerous

studies performed on different cattle breeds, the β -LG variant A was always synthesized in greater amounts (up to 50% more) than the variant B (Kim et al., 1996; Ehrmann et al., 1997; Lum et al., 1997; Robitaille et al., 2002). It was further demonstrated that these findings correlate with mRNA transcription, wherein more mRNA originates from the β -LG gene (*BLG*) A allele compared with the B allele (Wilkins et al., 1995). The expression of the *BLG* A and B alleles was investigated in several promoter studies (Wagner et al., 1994; Lum et al., 1997; Folch et al., 1999). Wagner et al. (1994) described 14 single nucleotide polymorphisms (SNP) in the 5'-flanking region and 2 SNP in the 5' untranslated region (5'-UTR) of the *BLG* gene. They found 3 frequent haplotypes consisting of a combination of 5 selected SNP, being in linkage disequilibrium to either the A or B protein-coding allele and concluded that this may be an explanation for the different synthesis of the β -LG protein variants A and B. In electromobility shift assay experiments, Lum et al. (1997) observed a 60% greater affinity to the activator protein-2 recognition site in the promoter of the *BLG* A allele compared with the corresponding site in the *BLG* B allele. In a reporter gene assay, the most frequent promoter variants linked to the respective A and B alleles showed relative expression levels of 57% for *BLG* A and 43% for *BLG* B (Folch et al., 1999).

Additional genetic variability in the *BLG* 5'-flanking region is presented in this study along with evidence of a newly identified cattle-specific duplication linked to the *BLG* A allele.

A standard phenol-chloroform extraction protocol was used to isolate DNA from blood. DNA was extracted from individuals representing a variety of breeds: 14 Swiss Brown, 3 Brown Swiss, 1 Angus, 1 Eringer, 2 Fleckvieh, 1 Jersey, 1 Hereford, 2 Holstein-Friesian, 2 N'Dama, 1 Simmental, and 1 gayal (mithun; *Bos frontalis*).

A PCR product of 287 bp spanned by the 28-nt primers *BLG*_678 and *BLG*_964 was amplified and sequenced using an ABI 3730 DNA Analyzer (ABI, Rotkreuz, Switzerland). The obtained sequences were analyzed with the Sequencer 4.6 software (Gene Codes

Received June 28, 2007.

Accepted August 30, 2007.

¹E-mail: martin.braunschweig@itz.unibe.ch

Corporation, Ann Arbor, MI). The numbering in the primer names corresponds to the nucleotide position of the 5' end of the primer with respect to the GenBank *BLG* sequence with the accession number Z48305. The PCR conditions were as follows: initial denaturation step at 95°C for 15 min followed by 30 cycles at 94°C for 30 s, 64°C for 30 s, 72°C for 60 s, and a final step at 72°C. Individuals were genotyped for duplication in the 5'-flanking region of *BLG* by separation of the PCR products obtained with the primer pair BLG_810 forward and BLG_964 reverse on a 3% agarose gel. The PCR conditions were the same as described above but with an annealing temperature of 60°C. The *BLG* A and B genotypes were determined by PCR-RFLP using *Hae*III as described by Medrano and Aquilar-Cordova (1990), but using the 22-nt primers BLG_5793 forward and BLG_5984 reverse. The annealing temperature of this primer pair was also 60°C and the other PCR conditions were the same as described above. The *BLG* D allele was determined as described by Braunschweig et al. (1998).

Bioinformatic analyses were performed using the GenBank *BLG* sequences of cattle (accession numbers Z48305 and X14710), yak (*Bos grunniens*; accession number AF194981), water buffalo (*Bubalus bubalis*; accession number AM238696), sheep (*Ovis aries*; accession numbers X68105, AY515301), goat (*Capra hircus*; accession numbers Z33881, DQ417346), and the sequence from gayal (mithun; *Bos frontalis*) as obtained in this study. The 5'-flanking region of the *BLG* A and B alleles were aligned to the corresponding sequences of the other ruminants by using the ClustalW program (<http://www.ebi.ac.uk/clustalw/>). The presence of potential regulatory elements was evaluated by searching the transcription factor binding sites database TRANSFAC (<http://www.gene-regulation.com>).

Numbering of polymorphisms was done with respect to the translation initiation codon where +1 corresponds to the adenine of the ATG. The reference sequence is the genomic sequence of the *BLG* B allele (accession number Z48305).

Sequence analysis of 287 bp from g.-2133 to g.-1847 in the 5'-flanking region of the *BLG* A allele revealed 1 duplication (g.-1885_-1879dupCTCTCGC), 1 insertion (g.-1891_-1892insG), and 6 substitutions (g.-1888A>G, g.-1903T>C, g.-1957C>T, g.-2008T>C, g.-2017C>T, and g.-2049T>C) compared with the *BLG* B allele (accession number Z48305; Figure 1A). Sequencing results from 4 *BLG* AA and 6 *BLG* BB animals of different breeds showed complete linkage disequilibrium between these novel polymorphisms and the respective genetic variants A and B. In Table 1, the genotypes for the polymorphisms g.-1885_-1879dupCTCTCGC and the substitution (c.401C>T) that discriminate the genetic variants A and

B are shown for individuals of different breeds. The genotypes were determined by means of PCR amplification and visualization on an agarose gel or by direct DNA sequencing. Furthermore, 1 Swiss Brown animal with the genotype *BLG* AD showed 2 *BLG* A haplotypes in the 5'-flanking region (Figure 1A) with one being linked to the D allele. Two individuals of the breed N'Dama (*Bos taurus*) were genotyped *BLG* BB and did not show any differences at the polymorphic sites in the 5'-flanking region compared with the corresponding sequence in the other *Bos taurus* cattle in the present study. The same result was found for one gayal (*Bos frontalis*), which was genotyped BB for the *BLG* alleles. However, this individual was additionally heterozygous C/T and A/G at position g.-1934 and g.-1976, where all *Bos taurus* animals were homozygous TT and GG, respectively. The results presented in Table 1 are in agreement with complete linkage disequilibrium between the g.-1885_-1879dupCTCTCGC and the c.401C>T substitution. It is worth mentioning that preliminary data indicate that the protein expression of the *BLG* D allele is similar to that of the A allele (my unpublished data).

The alignment of the *BLG* 5'-flanking regions of the A and B alleles to the corresponding regions in yak, gayal, water buffalo, sheep, and goat are shown in Figure 1B. The 5'-flanking region is highly conserved between these ruminants and the duplication (g.-1885_-1879dupCTCTCGC) and the guanine at position g.-1888 are only found in the *BLG* A (and D) allele in cattle. The cytosine at position g.-1957 and the thymine at positions g.-2008 and g.-2049 are only found in *BLG* B alleles of cattle. The thymine at position g.-1903 was also found in yak and gayal, and the cytosine (allele A) was present in water buffalo, sheep, and goat. The cytosine of *BLG* B at position g.-2017 was also found in water buffalo, sheep, and goat, whereas the thymine at position g.-2017 in *BLG* A was found in yak. These 2 *BLG* polymorphisms are not, therefore, cattle specific. The *BLG* A sequence from the 5'-flanking region of this and a previous study (Braunschweig and Leeb, 2006) was deposited in GenBank with the accession number EF694058. By standard in silico analyses, no transcription binding site is predicted to correspond to the region around the duplication (g.-1885_-1879dupCTCTCGC). In contrast, the g.-1888A>G substitution in the A and D alleles creates a DNA binding site for the upstream stimulatory factor as predicted by searching the TRANSFAC databases (Osborne et al., 1987; Figure 1C). The impact of the *BLG* B-specific nucleotides observed in cattle on potential transcription binding sites has yet to be determined.

The new genetic variability including the 7-bp duplication in the *BLG* A 5'-flanking region may contribute to the allele-specific differential expression of *BLG*. Re-

A **B**

Cattle <i>BLG B</i>	TGGCTGACCTCCCTCTGCTCTGCATCACCCAGTTCTGAGAGCAGAGTGGTGCTGGGGG-C	-1891
<i>BLG A</i>	TGGCTGACCTCCCTCTGCTCTGCATCACCCAGTTCTGAGAGCAGAGCGGTGCTGGGGGGC	
Gayal	TGGCTGACCTCCCTCCGCTCTGCATCACCCAGTTCTGAGAGCAGAGTGGTGCTGGGGG-C	
Yak	TGGCTGACCTCCCTCTGCTCTGCATCACCCAGTTCTGAGAGCAGAGTAGTGCTGGGGGGC	
Water buffalo	TGGCTGACCTCCCTCCGCCCTGCATCACCCAGTCCTGAGAGCAGAGCGGTGCTGGGGGGC	
Sheep	TGGCTGACCTCCCTCCTCCCTGCATCACCCAGTTCTGAAAGCAGAGCGGTGCTGGGGT-C	
Goat	TGGCTGACCTCCCTCCGCCCTGCATCACCCAGTTCTGAGAGCAGAGCGGTGCTGGGGGGC	
	***** * ***** * * * * * * * * * * * * * * *	
	g.-1885_1879dupCTCTCGC	
Cattle <i>BLG B</i>	ACAGCCTCTCGC-----ATCTGACACTTGTGTTCAAACCACCCATGCTGGTGTTTCGGG	-1838
<i>BLG A</i>	ACGGCCTCTCGCCTCTCGCATCTGACACTTGTGTTCAAACCACCCATGCTGCTGTTTCGGG	
Gayal	ACAGCCTCTCGC-----ATCTGACACTTGTGTTCAAACCACCCATGCTG	
Yak	ACAGCCTCTCGC-----ATCTGACACTTGTGTTCAAACCACCCGTGCTGGTGTTTCGGG	
Water buffalo	ACAGCCTCTCGC-----ATCTGACACTAGTGTTCAAACCACCCGTGCTGGTGTTTCGGG	
Sheep	ACAGCCTCTCGC-----ATCTAACGCCGGTGTTCAAACCACCCGTGCTGGTGTTTCGGG	
Goat	ACAGCCTCTCGC-----ATCTGACGCCGGTGTTCAAACCACCCGTGCTGGTGTTTCGGG	
	** ***** * * * * * * * * * * * * * * *	

C

USF (0.82) Sequence (+)-strand is shown
 USF gGCACGgcct

BLG A GGGGGGCACGGCCTCTCGCCTCTCGCATCT

g.-1888A>G
 g.-1885_1879dupCTCTCGC

Figure 1. A) Schematic representation of the β -lactoglobulin gene *BLG A* and *B* haplotypes in the 5'-flanking region from g.-1847 to g.-2133. The arrows indicate the primer positions of the PCR product that has been sequenced. The arrow in exon 1 indicates the position of the ATG translation start codon. B) Sequence alignment of cattle, gayal, yak, water buffalo, sheep, and goat 5'-flanking region of *BLG* encompassing the duplication of 7 nucleotides located between position -1885 and -1879 upstream of the ATG-translation initiation codon of the bovine *BLG* gene. The curly bracket indicates the duplication g.-1885_1879dupCTCTCGC. C) Potential binding site for the upstream stimulatory factor (USF) in the region around the duplication of *BLG A* and its orientation. The relative identities of the query sequence to the binding site consensus sequences are given in parentheses. The consensus sequences of the transcription factor USF is given on the right.

Table 1. β -Lactoglobulin gene (*BLG*) and g.-1885_-1879dupCTCTCGC genotypes in different breeds

Breed (number of animals)	(g.-1885_-1879dupCTCTCGC) ¹	BLG genotype (c.401C>T), (c.181G>C)
Swiss Brown (3)	22	AA
Swiss Brown (1)	22	AD ²
Holstein-Friesian (1)	22	AA
Simmental (1)	22	AA
Swiss Fleckvieh (1)	22	AA
Brown Swiss (2)	12	AB
Swiss Brown (2)	12	AB
Swiss Brown (1)	12	BD ²
Eringer (1)	12	AB
Hereford (1)	12	AB
Holstein-Friesian (1)	12	AB
Jersey (1)	12	AB
Swiss Fleckvieh (1)	12	AB
Angus (1)	11	BB
Brown Swiss (1)	11	BB
Swiss Brown (7)	11	BB
Mithun (1)	11	BB
N'Dama (2)	11	BB

¹The alleles are named 1 and 2 for 1 or 2 copies at the g.-1885_-1879dupCTCTCGC, respectively.

²The genetic variant *BLG* D is also linked to the duplication (g.-1885_-1879dupCTCTCGC).

cently, a *BLG* B* allele was described with an aberrant low level of expression associated with a g.-215C>A transversion (Braunschweig and Leeb, 2006). Sequencing in the *BLG* 5'-flanking region revealed no differences in the corresponding *BLG* B sequence. Detailed haplotype analysis of the 5'-flanking region and the 5'-UTR of *BLG* with respect to the content of β -LG variants A and B in the milk may elucidate the contribution of these haplotypes to *BLG* gene expression. If individuals exist that carry the recombinant haplotype with an adenine at position g.-1888 linked to the duplication (g.-1885_-1879dupCTCTCGC), this might shed light on the contribution of the g.-1888A>G substitution to the differential *BLG* allele expression. However, it is emphasized that such a haplotype study should consider all reported polymorphisms in the 5'-flanking region and 5'-UTR of the *BLG* gene.

ACKNOWLEDGMENTS

The author thanks Tosso Leeb for comments on the paper and revision of the text.

REFERENCES

- Aschaffenburg, R., and J. Drewry. 1957. Genetics of the β -lactoglobulins of cow's milk. *Nature* 180:376-378.
- Braunschweig, M., G. Stranzinger, and Z. Puhani. 1998. A PvuII PCR-RFLP test for the bovine β -lactoglobulin D allele. *Anim. Genet.* 30:76.
- Braunschweig, M. H., and T. Leeb. 2006. Aberrant low expression level of bovine β -lactoglobulin is associated with a C to A transversion in the *BLG* promoter region. *J. Dairy Sci.* 89:4414-4419.
- Ehrmann, S., H. Bartenschlager, and H. Geldermann. 1997. Polymorphism in the 5'-flanking region of the bovine-lactoglobulin-encoding gene and its association with β -lactoglobulin in the milk. *J. Anim. Breed. Genet.* 114:49-53.
- Folch, J. M., P. Dovc, and J. F. Medrano. 1999. Differential expression of bovine β -lactoglobulin A and B promoter variants in transiently transfected HC11 cells. *J. Dairy Res.* 66:537-544.
- Kim, J. S., M. Braunschweig, and Z. Puhani. 1996. Occurrence of extreme ratio of β -lactoglobulin variants A and B in Swiss Brown cattle quantified by capillary electrophoresis. *Milchwissenschaft* 51:435-438.
- Kontopidis, G., C. Holt, and L. Sawyer. 2004. Invited Review: β -Lactoglobulin: Binding properties, structure, and function. *J. Dairy Sci.* 87:785-796.
- Lum, L. S., P. Dovc, and J. F. Medrano. 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.
- Medrano, J. F., and E. Aquilar-Cordova. 1990. Polymerase chain reaction amplification of bovine β -lactoglobulin genomic sequences and identification of genetic variants by RFLP analysis. *Anim. Biotechnol.* 1:73-77.
- Osborne, T. F., G. Gil, M. S. Brown, R. C. Kowal, and J. L. Goldstein. 1987. Identification of promoter elements required for in vitro transcription of hamster 3-hydroxyl-3-methylglutaryl coenzyme A reductase gene. *Proc. Natl. Acad. Sci. USA* 84:3614-3618.
- Robitaille, G., M. Britten, J. Morisset, and D. Petitclerc. 2002. Quantitative analysis of β -lactoglobulin A and B genetic variants in milk of cows β -lactoglobulin AB throughout lactation. *J. Dairy Res.* 69:651-654.
- Wagner, V. A., T. A. Schild, and H. Geldermann. 1994. DNA variants within the 5'-flanking region of milk-protein-encoding genes. II. The β -lactoglobulin-encoding gene. *Theor. Appl. Genet.* 89:121-126.
- Wilkins, R. J., H. W. Davey, T. T. Wheeler, and C. A. Ford. 1995. Differential expression of β -lactoglobulin alleles A and B in dairy cattle. Pages 189-190 in *Intercellular Signalling in the Mammary Gland*. C. J. Wilde, M. Peaker, and C. H. Knight, ed. Plenum Press, New York, NY.