

## The impact of genetic polymorphisms on the protein composition of ruminant milks

Patrice MARTIN<sup>a,\*</sup>, Małgorzata SZYMANOWSKA<sup>b</sup>,  
Lech ZWIERZCHOWSKI<sup>b</sup>, Christine LEROUX<sup>c</sup>

<sup>a</sup> Institut National de la Recherche Agronomique, Laboratoire de Génétique Biochimique et de Cytogénétique, bâtiment Jacques Poly, Centre de Jouy, 78352 Jouy-en-Josas, France

<sup>b</sup> Institute of Genetics and Animal Breeding, Polish Academy of Sciences, Jastrzębiec, 05-552 Wólka Kosowska, Poland

<sup>c</sup> Institut National de la Recherche Agronomique, Unité de Recherches sur les Herbivores, Équipe TALL, Centre de Theix, 63122 Saint-Genès-Champanelle, France

**Abstract** — The purpose of this review is to give an overview of our current knowledge on the polymorphisms occurring in genes coding for milk proteins and responsible for quantitative variability in their expression, thus influencing the protein composition of livestock ruminant milk. The overall genomic organisation of the 6 main ruminant milk protein genes:  $\alpha$ -lactalbumin,  $\beta$ -lactoglobulin and the four caseins ( $\alpha_{s1}$ ,  $\alpha_{s2}$ ,  $\beta$  and  $\kappa$ ), their chromosomal location and their expression pattern are first summarised before presenting general mechanisms controlling gene expression both at the transcriptional and the post-transcriptional levels. Polymorphisms found in *cis*-regulatory elements, mainly within the 5'-flanking region of the genes encoding  $\beta$ -lactoglobulin and  $\alpha_{s1}$ - and  $\alpha_{s2}$ -caseins, have been found, in cattle, to influence their transcription rate. In addition, polymorphisms found in the transcription unit, within intron as well as exon sequences, have been shown to be responsible for defects in the processing of primary transcripts and/or the export of messenger RNA to the cytoplasm. Mutations responsible for the occurrence of premature stop codons in  $\alpha_{s1}$ - and  $\beta$ -casein mRNAs have been shown to be associated both with a decrease in the level of the relevant transcripts and the existence of multiple forms of messengers due to alternative splicing (exon skipping, usage of cryptic splice sites). Such a situation, well-exemplified by the gene encoding  $\alpha_{s1}$ -casein in the goat, may have dramatic biological consequences (secretion pathway, casein micelle structure, fat content, etc.) by modifying the message and accordingly the primary structure of the protein as well as its expression. Since some of these polymorphisms dramatically affect technological properties of milk, including cheese yields and organoleptic characteristics, methods mainly based on the PCR technique have been designed and applied in selection and breeding programmes to improve milk protein quality.

**gene expression / milk protein / genetic polymorphism / ruminants**

---

\* Correspondence and reprints  
E-mail: martin@diamant.jouy.inra.fr

## 1. INTRODUCTION

Gene expression in higher eukaryotes is regulated at multiple levels, including the accessibility of genes to regulatory factors in opened chromatin, modulation on accessible promoters of the transcription rate, RNA processing (splicing and polyadenylation), export of matured transcripts to the cytoplasm, stability as well as translation of messengers. Regulation at the transcriptional level essentially influences the amount of transcripts whereas regulation at the post-transcriptional level may also qualitatively affect the protein finally synthesised. Such a general pattern of regulation is obviously also true for the milk protein genes. In the last several years it has become clear that post transcriptional mechanisms are also critical control points and this has been particularly well documented in the goat species, as far as the caseins are concerned. The aim of the present review is not to provide the reader with an exhaustive overview of the mechanisms involved in the expression of milk protein genes nor to cover the large literature dealing with polymorphisms occurring in their transcription units and/or promoter and regulatory regions, in different mammalian species. Our goal was rather to draw attention on some significant and well documented examples showing how mutations responsible for polymorphisms at the genomic level and exploitable in ruminant selection programmes can influence milk protein composition at the quantitative as well as at the qualitative level.

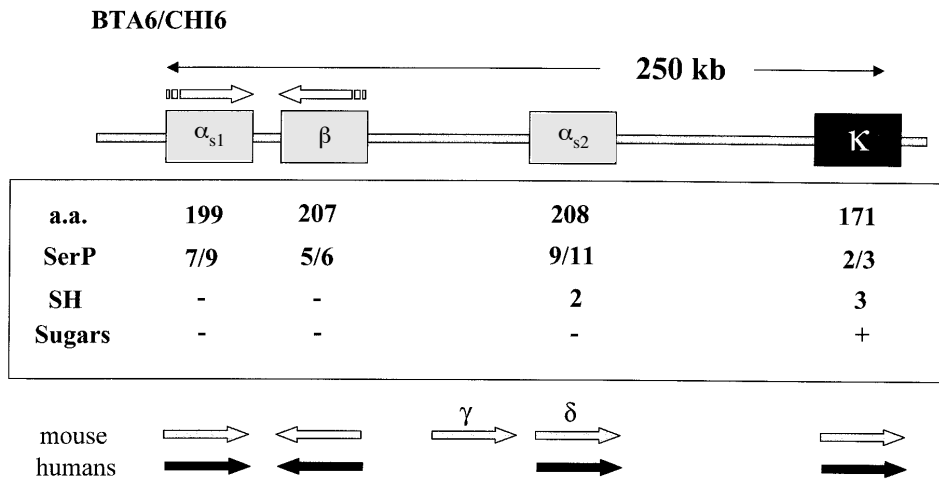
## 2. STRUCTURAL ORGANISATION OF RUMINANT MILK PROTEIN GENES

More than 95% of the proteins contained in ruminants' milk (mainly goats, sheep and cattle) are synthesised from 6 structural genes encoding proteins, now well-characterised:  $\alpha$ -lactalbumin and  $\beta$ -lactoglobulin, the two main whey proteins in ruminants,

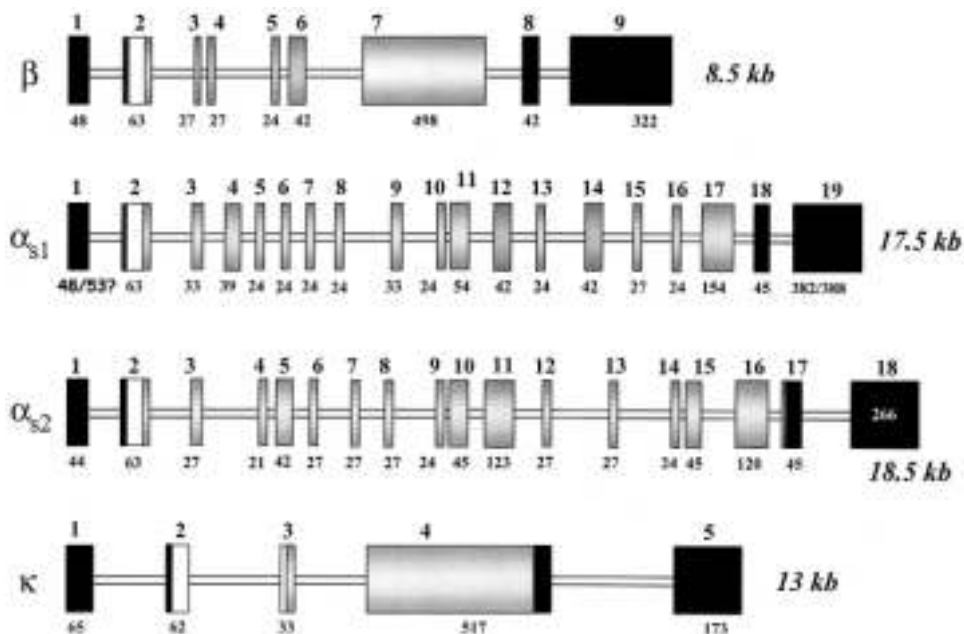
and the four caseins  $\alpha_{s1}$ ,  $\beta$ ,  $\alpha_{s2}$  and  $\kappa$  which are encoded by four tightly linked and clustered genes [27, 87], in that order, in a 250-kb genomic DNA fragment (Fig. 1). The four casein genes have been mapped on chromosome 6 in cattle and goats [38, 87], whereas the genes encoding  $\alpha$ -lactalbumin and  $\beta$ -lactoglobulin have been mapped on chromosomes 5 [38] and 11 [37], respectively.

The structural organisation of the 6 main milk protein genes has been determined in cattle (Figs. 2 and 3), as well as in sheep and goats. To ensure a specific expression during lactation in the mammary epithelial cells (MEC), activation of these genes requires, besides TATA and CAAT boxes signalling the start of transcription to RNA polymerase II, recognition sites for miscellaneous effectors involved in stage-tissue specificity, induction and modulation of expression. Induction of milk protein genes is under the control of a complex multihormonal process essentially involving prolactin, glucocorticoids, as well as insulin and several other lactogenic factors including growth hormone [80, 88, 93, 101]. Interactions between MEC and the extracellular matrix (ECM) seem to play a crucial role in the expression of milk protein genes [1, 42].

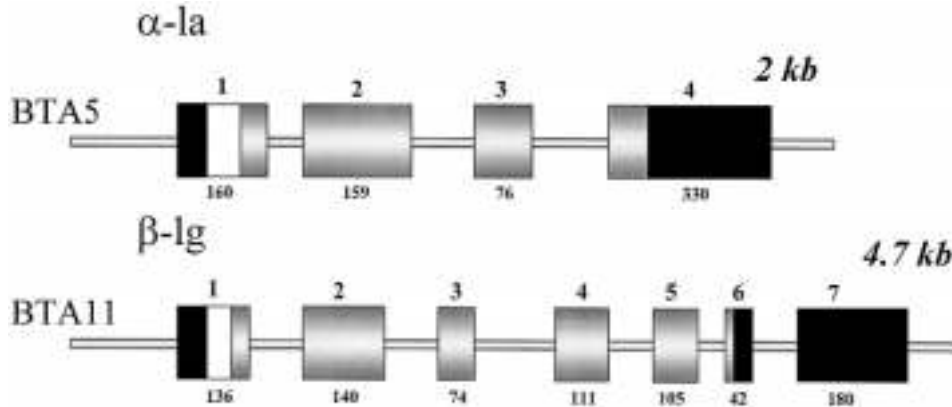
In the goat, loci encoding  $\alpha_{s1}$ - and  $\beta$ -caseins have been shown to be *ca.* 12 kb apart and convergently transcribed [50]. These results have been further confirmed in cattle for which the genomic organisation of the casein gene locus was determined [78]. Despite some differences in the distance separating casein genes and their numbers (with an extra  $\alpha_{s2}$ -like gene, the mouse locus comprises five transcription units:  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\epsilon$  and  $\kappa$ ), the overall genomic organisation of the locus is fairly conserved between mammalian species and the presence of dominant *cis*-acting regulatory elements, required for the high-level coordinate expression of the casein genes, is suspected in the  $\alpha_{s1}$ / $\beta$ -region [78].



**Figure 1.** Genomic organisation of the bovine/goat casein locus. The genes encoding the four caseins are depicted by grey (calcium sensitive caseins) and black ( $\kappa$ -casein) boxes and the major biochemical properties (length of the peptide chain, number of amino acid residues, number of phosphoserine residues, number of cysteine residues and glycosylation status) of the goat proteins synthesized starting from the corresponding genes are given under each gene. The overall structural organisation of the mouse and human loci are also given.



**Figure 2.** Structural organisation of the four bovine casein transcription units. Open bars represent introns, and exons are depicted by large, black (5' and 3' untranslated regions), white (part of exon encoding the signal peptide) and grey (exons and part of exons encoding matured proteins) boxes. Sizes of exons are given, in base pairs, under each exon whose numbering is indicated on the top.



**Figure 3.** Structural organisation of the transcription unit encoding the two main bovine whey proteins. See legend of Figure 2 for a description.

All four genes are coordinately expressed at high levels in a tissue- and stage-specific fashion. The three genes encoding the calcium-sensitive caseins ( $\alpha_{s1}$ ,  $\alpha_{s2}$  and  $\beta$ ), that are evolutionarily related, share common regulatory motifs in the proximal 5' flanking region [32]. Although the organisation of the 5' flanking region of the  $\kappa$ -casein gene is different [19], its expression pattern seems to be similar to that of the other casein genes. There are more and more evidence demonstrating that a common set of transcription factors is required in most mammalian species, for the expression of milk protein genes. The mechanisms controlling milk protein gene expression, especially those pertaining to the STAT5 behaviour (although they are significantly different between mammalian species, probably due to a different organisation of the hormone response regions of the casein genes), apparently does not result in fundamental differences in their responsiveness to lactogenic hormones, at least in transfected cell lines. The species-specific arrangement of the transcription factor binding sites in the  $\beta$ -casein gene appears to be crucial for the strength and stage at which this gene is expressed in different species, including humans, rodents and ruminants [97]. For example, the bovine, but not the mouse,  $\beta$ -casein gene is strongly

induced shortly before parturition. This difference in stage-specific expression was recapitulated in the expression of a bovine  $\beta$ -casein transgene (including 16 kb of 5' - and 8 kb of 3' -flanking regions) in transgenic mice, thus indicating that *cis*-acting sequences might be, at least in part, responsible for species-specific expression patterns [77], demonstrating their involvement in the control of their expression.

As mentioned before, transcription is not, however, the only level at which regulation of gene expression occurs. In the following paragraphs, we shall see that there are many other factors acting at the post-transcriptional level, including messenger RNA stability and processing. The process of intron removal (excision), and exon joining (splicing), is a major function ensured in the nucleus, by a large multi-component complex: the spliceosome. This supramolecular structure made of four large RNA-protein complexes (U1, U2, U4/U6 and U5 small nuclear ribonucleoprotein particles), or snRNPs, and some non-snRNP proteins (more than 50 proteins) assemble around three short conserved sequences within the intron in an ordered manner to form the active spliceosome. Such an accurate mechanism is controlled by a set of strict rules to achieve high fidelity and efficiency in

splicing. A dysfunction of this machinery may have dramatic biological consequences by modifying the message and accordingly the primary structure of the protein as well as its expression.

### 3. POLYMORPHISMS IN *CIS*-REGULATORY ELEMENTS OF MILK PROTEIN GENES

Computer aided analysis of the 5'-flanking regions of the milk protein genes [55, 56] allowed the identification of several putative *cis*-regulatory sequences and potential sites of binding of trans-acting regulatory factors. Milk protein genes of cattle contain, in their 5'-flanking regions, putative binding sites for transcription factors: MGF/STAT5A<sup>1</sup>, MAF<sup>2</sup>, MPBF<sup>3</sup>, PMF<sup>4</sup>, first identified in the mammary tissue, as well as for ubiquitous transcription factors: AP1, CREB, CTF/NF1, OCT1, SP1, NF- $\kappa$ B, GR, TR, YY1, HSE. The function of some of these factors in regulating milk protein gene expression has been confirmed experimentally [32, 57, 79]. Most of the regulatory sequences are located in the proximal regions of milk protein gene promoters, within 200–250 bp adjacent to the initiation of the transcription site. However, in the bovine  $\beta$ -casein gene a distal regulatory element *bceI* has been localised approximately 1 500 bp upstream of the transcription starting point [79].

Recently, genetic polymorphisms within the 5'-flanking regions of milk protein genes have been reported by several authors [9, 23, 30, 43, 68, 81]. A comprehensive review of these reports is given in Table I. Since nucleotide substitutions, deletions or insertions in gene promoters may be located in transcription factor-binding sites or *cis* reg-

ulatory sequences, they can increase or decrease the efficiency of the transcription of a relevant structural gene, and thus influence gene expression and the physiological traits of animals.

#### 3.1. $\alpha$ -lactalbumin

$\alpha$ -lactalbumin is essential for the biosynthesis of lactose in the mammary gland. It enhances the substrate-affinity of  $\beta$ -1,4-galactosyltransferase which catalyses the formation of lactose from glucose and UDP-galactose [12]. Thus, due to its prominent role in milk synthesis,  $\alpha$ -lactalbumin is considered a valuable genetic marker for milk production traits in cattle. However,  $\alpha$ -lactalbumin appears rather monomorphic in European cattle breeds, after screening on the protein as well as on the nucleotide level within the coding sequence of the gene.

Bleck and Bremel [8] cloned and sequenced the 5'-flanking region of the  $\alpha$ -lactalbumin gene derived from the genome of Holstein cattle. The sequenced clone contains 1952 bp of the 5'-flanking region and 66-bp of the protein-coding region. Three single-bp polymorphisms were identified within this region. These polymorphisms occur at positions +15, +21 and +54 relative to the mRNA transcription starting point. The +15 and +21 variations occur in the region of the gene encoding the 5'-untranslated region (5'-UTR) of the mRNA.

A single base difference (A/G) at position +15 in the 5'-region of the  $\alpha$ -lactalbumin mRNA has been shown [7] as a useful marker associated with an increased milk yield in Holstein cows. Cows with allele A of the gene have higher milk yield, protein yield, and fat yield; the B allele is associated with

<sup>1</sup> Mammary Gland Factor/Signal Transducer and Activator of Transcription.

<sup>2</sup> Mammary cell-Activating Factor.

<sup>3</sup> Milk Protein-Binding Factor, identified as STAT5A.

<sup>4</sup> Pregnancy-specific Mammary nuclear Factor.

**Table 1.** Polymorphism within 5'-flanking regions of milk protein genes.

Gene	Species	Chromosomal location	Number of exons	Gene size (length of sequenced 5'-region); GenBank accession number	Type/site of polymorphism	Method of analysis	References
$\alpha_{s2}$ -casein	Cattle	6	18	18483 (2509); M94327	10 nucleotide substitutions (-1150 to +118 region)	PCR-RFLP; Sequencing	Schild and Geldermann [81]
$\alpha_{s1}$ -casein	Cattle	6	19	17508 (2131; X59856)	15 substitutions, 2 deletions/insertions (-1145 to +100 region)	PCR-RFLP; Sequencing	Schild and Geldermann [81]
$\alpha_{s1}$ -casein					218-bp region (+3 to -215)	SSCP	Kamiński and Józefowicz [46]
$\alpha_{s1}$ -casein					A/G substitution at pos. -175	RFLP	Koczan et al. [48]
$\beta$ -casein	Cattle	6	9	8498 (1722; X14711)	10 nucleotide substitutions, 1 deletion/insertion (-1049 to +88 region)	PCR-RFLP; Sequencing	Schild and Geldermann [81]
$\beta$ -casein					T deletion at pos. -516	Sequencing	Bleck and Bremel [8]
$\beta$ -casein	Goat	6	9	9006 (1001; AJ011018)	A/G substitution at pos. -161	Sequencing; RFLP	Pappalardo et al. [68]
$\kappa$ -casein	Cattle	6	5	13000 (2167; M75887)	15 nucleotide substitutions; -1028 to +59 region	Sequencing; RFLP	Schild et al. [82]
$\kappa$ -casein					214-bp 5'-flanking region	RFLP	Kamiński et al. [43]

Table I. (Continued).

Gene	Species	Chromosomal location	Number of exons	Gene size (length of sequenced 5'-region); GenBank accession number	Type/site of polymorphism	Method of analysis	References
$\beta$ -lactoglobulin	Cattle	11	7	4723 (795; X6313)	6 nucleotide substitutions; 3 haplotypes	PCR-RFLP	Ehrmann et al. [23]
$\beta$ -lactoglobulin					13 nucleotide substitutions in the -607 to +34 region	Sequencing; SSCP	Lum et al. [52]
$\beta$ -lactoglobulin					14 nucleotide substitutions in the -765 to -22 region and 2 at pos. +27 and +31	Sequencing; RFLP	Wagner et al. [94]
$\beta$ -lactoglobulin					209-bp region (-501 to -293)	SSCP	Kamiński and Zabołewicz [47]
$\beta$ -lactoglobulin	Goat	11	7	4697 (2148; Z33881)	C/T substitution at pos. -60	PCR-RFLP	Yahyaoui et al. [98]
$\alpha$ -lactalbumin	Cattle	5	4	2023 (1951; M90645)	3 substitutions, pos. +15, +21 and +54	Sequencing; RFLP	Bleck and Brenel [8]
$\alpha$ -lactalbumin					A/G substitution at pos. -1 689 and +15	Sequencing; RFLP	Voelker et al. [92] Kamiński [44]

a higher percentage of protein and fat. It is not known, however, whether the polymorphism of  $\alpha$ -lactalbumin itself has a functional effect or it is linked to some other gene polymorphism that has a functional influence on milk production. The +15 mutation in  $\alpha$ -lactalbumin gene, although not located in the gene promoter, might influence the expression of  $\alpha$ -lactalbumin at the level of mRNA translation, but this remains to be demonstrated.

Voelker et al. [92] sequenced 230 bp of the  $\alpha$ -lactalbumin 5'-flanking region in order to examine this region for potential sequence variations. In addition to the +15 A/G mutation, they identified a second single base pair difference, located at position -1689 from the transcription starting point. This polymorphism is also an A  $\rightarrow$  G transition. The allele showing an adenine at position -1689 was designated as allele A, and that with a guanine at this position was designated as allele B. The -1689 and +15 polymorphisms were compared within the Holstein population to determine their linkage relationship. In this study, the +15 A variant was always linked to the variant A at -1689. Taken together, these results suggest the existence of haplotype A (+15A and -1689A) associated with higher milk, protein and fat yields.

### 3.2. $\beta$ -lactoglobulin

$\beta$ -lactoglobulin is a protein present in the milk of many species, mainly in ruminants. It is absent from the milk of rodents, rabbits and camels in which instead another major whey protein – the whey acidic protein or WAP – is found.  $\beta$ -lactoglobulin is also absent from human milk. However, the presence of WAP in human milk has still to be demonstrated. Recently, besides  $\beta$ -lactoglobulin, the WAP has also been identified in the milk of swine [83].

Twelve polymorphic variants of  $\beta$ -lactoglobulin are known in cattle, but the two

most frequent, A and B, were shown to be associated with differences in milk protein yield and composition. These variants differ by two a.a. substitutions in the polypeptide chain arising from two single nucleotide substitutions in the  $\beta$ -lactoglobulin gene: Asp 64 (GAT)  $\rightarrow$  Gly (GGT) and Val 118 (GTC)  $\rightarrow$  Ala (GCC). The latter T  $\rightarrow$  C substitution creates an *Hae*III restriction site, thus enabling an RFLP analysis at the  $\beta$ -lactoglobulin locus [64].

Quantitative effects of these common variants on milk composition and cheese-making properties have been reported [2]. Allele B of  $\beta$ -lactoglobulin is associated with high casein and fat contents in cow's milk, while, milk of Holstein cows with the  $\beta$ -lactoglobulin AA genotype were shown to contain more whey proteins and total protein than those of the other genotypes [2, 6, 20, 53, 58, 62, 66, 67]. In addition, a higher expression of allele A has been described in heterozygous (AB) animals [31]. This differential allelic expression has been explained by nucleotide differences in the promoter regions associated with these two alleles.

In their study, Lum et al. [52] evaluated 733 bp of the  $\beta$ -lactoglobulin promoter region and 92 bp of the first exon for possible polymorphisms using denaturing gradient gel electrophoresis (DGGE) and nucleotide sequence analysis. Within this region, 13 single nucleotide substitution polymorphisms were detected. Several potential binding sites for transcription factors were found within the promoter sequence. The role of the G to C transversion within a consensus binding site for activator protein-2 (AP2) at position -430 was investigated. Using the DNase-I footprint assay, the functional importance of this point mutation was confirmed and different binding affinities of AP2 were shown for both alleles – G and C. A possible regulatory role of AP2 in the transcriptional regulation of the  $\beta$ -lactoglobulin gene has been proposed. Several other transcription factors are known to bind to recognition



sequences of the  $\beta$ -lactoglobulin gene promoter [28, 95], but the polymorphisms reported in the 5'-flanking region of the bovine  $\beta$ -lactoglobulin gene were not located in these binding sites.

By sequence analysis of the 5'-flanking regions of the milk protein encoding genes, altogether 65 variable sites have been revealed by Geldermann et al. [30]. Sixty of these sites are base substitutions and 5 are deletions/insertions. About 50% of the variable sites are located in potential protein binding sites, identified so far by computer-aided analysis. In cell culture tests, the investigated promoter variants lead to different reporter gene expression. The most divergent variants in the 5' flanking regions of the milk protein encoding genes were cloned in the pGL2-Basic expression vector, transfected into mammary gland cells, and were measured for reporter gene expression. In the case of the  $\beta$ -lactoglobulin encoding gene, the promoter variant of the AA genotype produced up to 3.5 times greater expression of a reporter gene than the BB genotype. Also Folch et al. [29] showed differential expression of a reporter gene fused to bovine  $\beta$ -lactoglobulin A or B promoter variants in transiently transfected HC11 cells, the A variant having a more efficient expression of the reporter than the B variant (57% vs. 43%).

Wagner et al. [94] identified 14 single-bp substitutions within the 5'-flanking region and two in the 5'-untranslated region (5'-UTR) of exon I of the bovine  $\beta$ -lactoglobulin gene. Some of them are located in potential binding sites for trans-acting factors or in the 5'-UTR. Sequence data of the 5'-flanking regions and the protein genotypes indicate that the A allele of the  $\beta$ -lactoglobulin coding region mostly associated with the B allele within the promoter sequence, whereas the B allele of the coding region is often combined with the A allele of the 5'-flanking region. This observation suggests that alleles A or B in the coding regions were connected with distinct pro-

motor variants. Graml et al. [31] suggested that such intragenic haplotype associations may explain the observed differences in the effects of A or B variants of  $\beta$ -lactoglobulin on milk production traits particularly on  $\beta$ -lactoglobulin synthesis ( $A > B$ ) in heterozygous cows.

Polymorphism within the 5'-flanking region of the bovine  $\beta$ -lactoglobulin gene ( $\beta$ -lgPTU) was studied by Kamiński and Zabolewicz [46] in the Polish Black-and-White dairy cattle, in which the two predominant  $\beta$ -lactoglobulin variants are also A and B. Moreover, within the promoter region ( $\beta$ -lgP) of the gene, several point mutations were found by SSCP analysis. A statistical analysis revealed significant associations between the transcription unit ( $\beta$ -lgTU) and  $\beta$ -lgP genotypes as well as intragenic ( $\beta$ -lgP/ $\beta$ -lgTU) haplotypes and milk protein content during the first complete lactation. Cows with the AA genotype at the  $\beta$ -lgTU locus, R3  $\beta$ -lgP SSCP pattern and AA/R3 haplotype had the highest protein content, suggesting that the nucleotide sequence variation within the promoter of the  $\beta$ -lactoglobulin gene is a marker associated to differences in the protein content in milk.

### 3.3. $\kappa$ -casein

Two major genetic variants of  $\kappa$ -casein, A and B, have been identified in cattle. Variant A shows Thr (ACC) and Asp (GAT) at positions 136 and 148, respectively, whereas the B variant shows Ile (ATC) and Ala (GCT) at the same positions. These differences are the result of single base mutations in the  $\kappa$ -casein gene [22] and the two alleles (A and B) can be distinguished by the presence or absence of a *Hind*III restriction endonuclease recognition site [22, 73]. In addition, the change in amino acid position 148 abolishes a *Hinf*I site in the  $\kappa$ -casein B allele [63].

In general, the B variant of  $\kappa$ -casein has been recognised as superior for milk quality

in European cattle breeds; the B allele of  $\kappa$ -casein is associated with shorter renneting time of the milk [53]. Cheeses produced from  $\kappa$ -casein BB milks have also been shown to have higher protein content, higher yield and better quality than those produced from  $\kappa$ -casein AA or AB milk. Milks of Holstein cows with the  $\kappa$ -casein BB genotype have been shown to contain, on average, 0.13% more protein than those of the other genotypes [66].

Significant differences have been found in the content of both variants A and B of  $\kappa$ -casein in milks of heterozygous AB cows [90]. Usually there is more of the protein variant encoded by allele A than that encoded by allele B. Reasons for such a variability were searched for by analysing sequences within the 5'-upstream regulatory regions of the  $\kappa$ -casein gene. In order to identify DNA polymorphisms, Schild et al. [82] analysed the 5'-flanking region of the bovine  $\kappa$ -casein-encoding gene from 13 cows belonging to seven breeds. Fifteen nucleotide variations have been identified, some of which are located within potential regulatory sites possibly involved in the expression of the  $\kappa$ -casein gene. However, further investigations are required to show a possible influence of these allelic variants on  $\kappa$ -casein gene expression.

Kamiński [45] identified a *DdeI* RFLP polymorphism within a 214-bp fragment of the 5'-upstream region of the  $\kappa$ -casein gene, containing 5 potential consensus sequences for different transcription factors. In a population of 112 *Bos taurus* (86 cows and 26 bulls of Polish Black and White cattle) and 7 *Bison bonasus* individuals, 7 had no recognition sites for *DdeI*, 23 were heterozygous, and the DNA of 89 were cut completely into two fragments. One hundred and twenty four cows were genotyped for two polymorphisms in the  $\kappa$ -casein gene – CASK (exon IV – A and B alleles) and CASK-R ( $\kappa$ -casein promoter, *DdeI* restriction site at position –385 – P and M alleles) [43]. Associations between individual CASK and CASK-R genotypes as well as intragenic

haplotypes CASK/CASK-R and milk performance traits were investigated. Significant differences ( $P \leq 0.01$ ) were found only for the milk protein percentage. Cows with the PM CASK-R genotype and AA/PM intragenic haplotype showed the highest protein percentage.

Studies on the differential expression of different alleles at the  $\kappa$ -casein locus were carried out by Debeljak et al. [21]. On average, 13.5% more transcripts of the A allele were found as compared with allele B mRNA, in the mammary gland of cows of the AB genotype. No allele specific polymorphism has been found in the proximal promoter region of the A, B and E alleles of the  $\kappa$ -casein gene. Electromobility-shift assay (EMSA) and DNaseI footprinting analyses showed binding sites for AP2, NF1 and MGF/STAT transcription factors but no differences in DNA-protein binding have been reported for the different alleles. Moreover, seven polymorphic sites have been identified within the 3'-flanking region of the  $\kappa$ -casein gene and the differences were shown in the length of mRNAs derived from A and B alleles of the gene. The regulation on the post-transcriptional stages of  $\kappa$ -casein expression has been suggested as a possible mechanism of differential allelic expression.

### 3.4. Other caseins

Bleck et al. [9] sequenced the 5'-flanking region of the  $\beta$ -casein gene for a single Holstein bull, a single Brown Swiss bull, and a single Jersey bull. The only difference observed among the three animals was a single base pair deletion that occurred at position –516 from the transcription starting point. The Holstein was homozygous for the form A of the allele, and the Brown Swiss and Jersey were both heterozygous. Four additional sequence differences (deletions) were found when the sequenced fragment was compared with the previously published sequence of the bovine  $\beta$ -casein gene [11].

Schild and Geldermann [81] identified a total of 34 variable sites in approx. 1.2 kb 5'-flanking regions of genes encoding bovine calcium-sensitive caseins: 17, 10 and 7 for  $\alpha_{s1}$ ,  $\alpha_{s2}$  and  $\beta$ -casein, respectively. The sequenced 5'-flanking regions of the genes were computer-analysed for homologies with consensus sequences for mammary gland-specific factors, hormone receptor sites and ubiquitous transcription factors. As shown, 17 of the 34 mutations (50%) were located within known potential regulatory sequences, suggesting an influence of these mutations on the binding of transcription factors. Thus, these variants may affect the expression of the Ca-sensitive casein-encoding genes. We applied the cell culture and transfection techniques to directly determine the effect of the mutations on gene expression. The results indicated that differences in expression of the casein genes could not be attributed to the influence of the variability at single sites, but rather to the additive effect of intragenic haplotypes within the 5'-flanking regions.

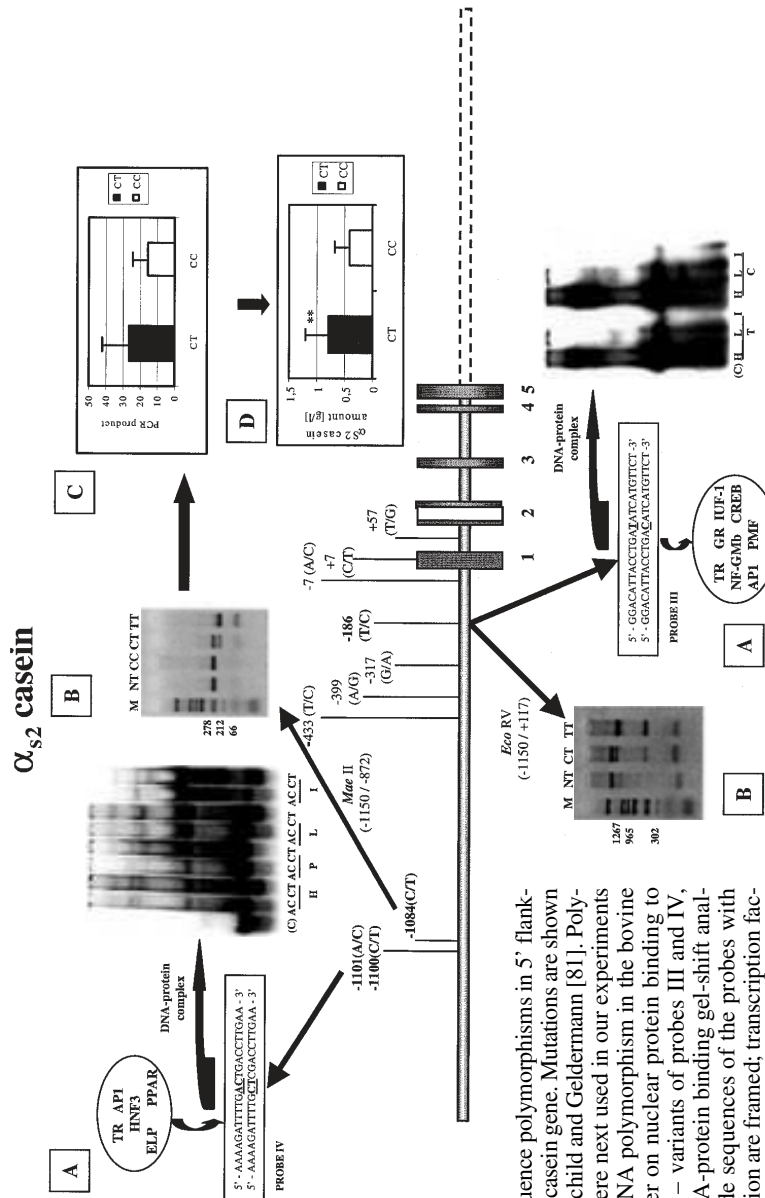
Sequence variability in the 5'-flanking regions of milk protein genes of species other than cattle was studied only sporadically. Pappalardo et al. [68], while sequencing the goat  $\beta$ -casein gene detected a polymorphic *MseI* restriction site at the 5'-flanking end of the gene in region containing the putative regulatory elements. Two alleles were observed: allele A, showing three fragments (197 bp, 124 bp and 63 bp) and the B allele showing only two fragments (321 bp and 63 bp). Nucleotide sequences showed that in position -161 the A allele is characterised by an adenine and allele B by a guanine.

### 3.5. The association between sequence polymorphisms in the 5'-flanking region of bovine casein genes, transcription factor binding, and milk composition in Polish breeds

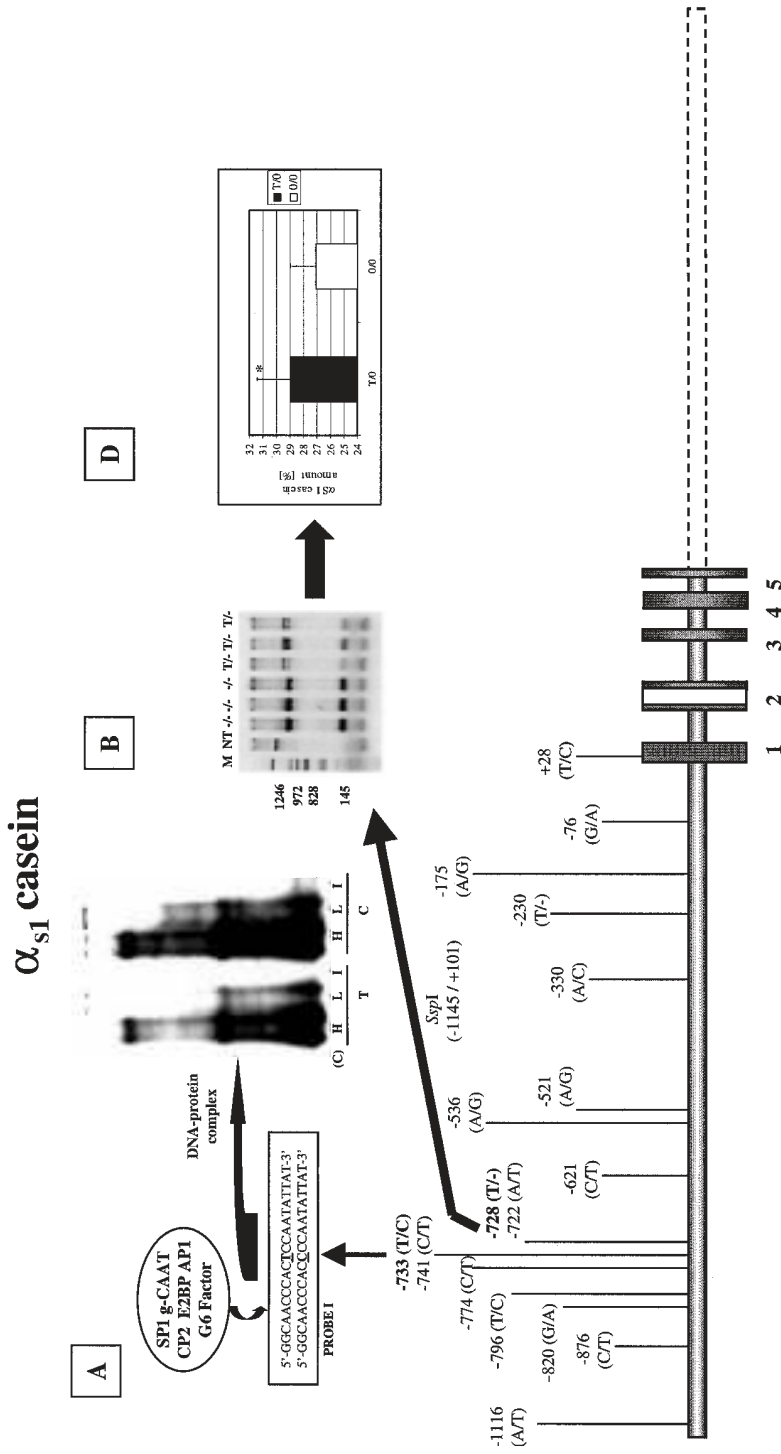
Figures 4, 5 and 6 show mutations in the 5' flanking regions of bovine casein

genes previously identified by Schild and Geldermann [81], Koczan et al. [48], and Bleck et al. [8]. These mutations are also listed in Table I. In our experiments nine of these mutations (marked in bold) were screened to see whether they might change affinity to nuclear proteins – transcription factors, affect expression of relevant genes and quantitatively influence composition of milk proteins. Results concerning four of these mutations are shown in this paper. PCR-RFLP tests were developed [85, 86] using primer sequences given by Schild and Geldermann [81] to genotype within promoter regions of the  $\alpha_{s2}$ - (positions -186 and -1084),  $\alpha_{s1}$ - (position -728) and  $\beta$ - (position -109) casein genes (Figs. 4B, 5B and 6B). Eighty-one Polish Black-and-White (BW) and 195 Polish Red (PR) cows were screened for polymorphism using this approach. Both breeds were found to differ significantly in the distribution of various genotypes in the 5'-flanking regions of  $\alpha_{s1}$ - and  $\alpha_{s2}$ -casein genes. No polymorphism was found in the 5'-flanking region of the  $\beta$ -casein gene. However, preferential associations were found between individual promoter genotypes and protein variants of  $\alpha_{s1}$ - and  $\beta$ -caseins. In all PR individuals with the CT promoter variant of the  $\alpha_{s2}$ -casein gene (position -186) only variant A of the  $\alpha_{s1}$ -casein was found (statistically significant at  $P \leq 0.001$ ). On the contrary, all BW cows with the CC genotype at position -1084 had the A<sub>1</sub> allele of  $\beta$ -casein ( $P \leq 0.02$ ). Such associations are in agreement with the notion of a tight linkage between casein genes, first demonstrated by Grosclaude et al. [35, 36] and since confirmed at the molecular level through genomic analyses. It provides strong evidences for the existence of specific haplotype combinations, including both coding and regulatory sequences at the casein *locus* in the Polish cattle breeds.

The proteins of milk derived from cows of various genotypes at  $\alpha_{s2}$ - and  $\alpha_{s1}$ -casein gene promoters were analysed quantitatively

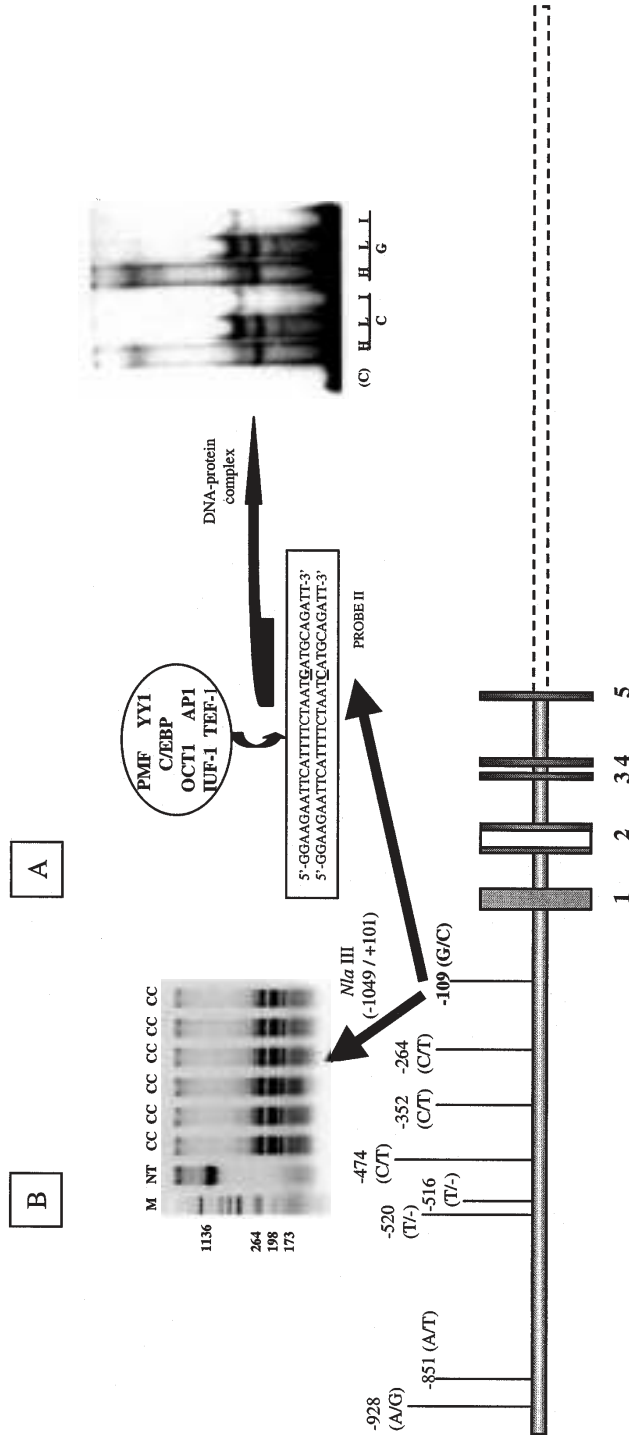


**Figure 4.** Nucleotide sequence polymorphisms in 5' flanking regions of bovine  $\alpha_{s2}$ -casein gene. Mutations are shown previously identified by Schild and Geldermann [81]. Polymorphic sites, in bold, were next used in our experiments [85, 86]. (A) Effect of DNA polymorphism in the bovine  $\alpha_{s2}$ -casein gene promoter on nuclear protein binding to DNA. T, C and AC, CT – variants of probes III and IV, respectively, used in DNA-protein binding gel-shift analyses (EMSA). Nucleotide sequences of the probes with marked position of mutation are framed; transcription factors binding to the probes III and IV are circled. Proteins were isolated from cell nuclei of heifers (H), pregnancy (P), lactating (L) and involuting (I) cow's mammary gland. (B) RFLP analyses of the bovine  $\alpha_{s2}$ -casein gene in the 5'-flanking region (positions -1084 and -186). Genotypes were identified through *Mae*II and *Eco*RV; restriction patterns; M: DNA standard; NT: undigested product; CC, CT and TT: cow genotypes. (C) Detection of  $\alpha_{s2}$ -casein transcripts from mammary glands of various genotypes in  $\alpha_{s2}$ -casein gene promoter at position -1084. RNA was isolated from the mammary glands of cows with CT and CC genotypes. Differences in casein transcripts after scanning and densitometry are expressed in arbitrary units. (D) HPLC analysis of  $\alpha_{s2}$ -casein contents in milk of cows with CC and CT genotypes in the  $\alpha_{s2}$ -casein gene promoter. Casein contents are expressed in g·L<sup>-1</sup>. Values represent the average of three independent experiments ( $\bar{X} \pm SD$ ).



**Figure 5.** Nucleotide sequence polymorphisms in 5' flanking regions of bovine  $\alpha_{s1}$ -casein gene. Mutations are shown previously identified by Koczan et al. [48] and Schild and Geldermann [81]. Polymorphic sites, in bold, were next used in our experiments. **(A)** Effect of DNA polymorphism in the bovine  $\alpha_{s1}$ -casein gene promoter on nuclear protein binding to DNA. T and C; variants of probe I used in EMSA analysis. **(B)** RFLP analysis of the bovine  $\alpha_{s1}$ -casein gene in the 5'-flanking region (position -728) with *SspI* restriction enzyme **(D)** HPLC analysis of  $\alpha_{s1}$ -casein content in milks of cows with -/- and T/- genotypes. Other descriptions are as in the legend to Figure 4.

# β casein



**Figure 6.** Nucleotide sequence polymorphisms in the 5' flanking regions of bovine β-casein gene. Mutations are shown previously identified by Schild and Geldermann [81] and Bleck et al. [8]. Polymorphic site, in bold, was next used in our experiments. (A) Effect of DNA polymorphism in bovine casein gene promoter on nuclear protein binding to DNA. Variants C and G of probe II were used in EMSA analysis. (B) RFLP analysis of the bovine β-casein gene polymorphism in the 5' flanking region (position -109) with *Nla*III. Other descriptions are as in the legend to Figure 4.



using SDS-PAGE and HPLC techniques. Differences were shown in casein content in milk obtained from cows carrying specific mutations at position -1084 and at position -728 of the  $\alpha_{s2}$ - and  $\alpha_{s1}$ -casein genes, respectively. Milk of BW cows with the T/- genotype of  $\alpha_{s1}$ -casein gene promoter contained on average 11% more  $\alpha_{s1}$ -casein than cows with the -/- genotype promoter ( $P \leq 0.01$ ), (Fig. 5D). Also, the polymorphism within the  $\alpha_{s2}$ -casein gene promoter was shown to affect the content of the relevant protein in milk; cows of the BW breed with the CT genotype of the  $\alpha_{s2}$ -casein gene at position -1084 contained twice as much  $\alpha_{s2}$ -casein as the CC animals ( $P \leq 0.002$ ), (Fig. 4D). Moreover, differences were found in the expression rates of the different  $\alpha_{s2}$ -casein alleles using semi-quantitative RT-PCR analysis performed on RNAs isolated from biopsies of mammary glands or from somatic cells derived from cow's milk (Fig. 4C). The mammary glands of the BW cows of the CT genotype (position -1084) accumulated approximately 25% more  $\alpha_{s2}$ -casein mRNA than those with the CC genotype.

Binding of nuclear proteins to variable 5'-noncoding regions of bovine  $\alpha_{s1}$ -,  $\alpha_{s2}$ -, and  $\beta$ -casein genes was observed, using mobility-shifts assays (EMSA), as described previously by Schild and Geldermann [81], (Figs. 4A, 5A and 6A). These sequences were first analysed in detail using the TESS program (Transcription Element Search System; [www.cbil.upenn.edu/tess](http://www.cbil.upenn.edu/tess)) for the presence of putative transcription factor-binding sites. Such an analysis not only allowed the identification of binding sites but also allowed to predict how nucleotide substitutions can influence binding affinity of transcription factors to polymorphic sequences.

Based on this data the following oligonucleotide probes were designed for EMSA: probe I, covering the -743/-723 bp region of  $\alpha_{s1}$ -casein gene promoter with single substitution C/T at position -733 and binding

sites for AP1, SP1, g-CAT, CP2, E2BP and G6 factors, (Fig. 5A); probe II, a -129/-100 bp region of the  $\beta$ -casein gene promoter, with the G/C mutation at position -109 and binding sites for AP1, PMF, YY1, Oct1, IUF-1, C/EBP and TEF1, (Fig. 6A); probe III, covering the -199/-176 bp region of the  $\alpha_{s2}$ -gene promoter, with a single substitution C/T at position -186 and binding sites for TR, GR, AP1 and CREB, (Fig. 4A); probe IV, a -1113/-1090 bp region of the  $\alpha_{s2}$ -casein gene promoter, including two substitutions A/C and C/T at positions -1101 and -1100 and binding sites for TR, AP1, GCN4 and HNF3, (Fig. 4A). Proteins extracted from cell nuclei derived from the mammary glands of cows at different physiological stages: non-lactating (heifers), pregnant, lactating, involuting, were used in the EMSA analyses.

We showed (Figs. 4A and 5A) that proteins contained in the mammary glands of heifers and lactating cows bind more efficiently to allele C of probes I (mutation C/T at position -733 in the  $\alpha_{s1}$ -casein promoter region) and III (mutation C/T at position -186 in the  $\alpha_{s2}$ -casein promoter region), while extracts obtained from lactating and drying cows showed (Fig. 4A) greater affinity to allele CT of probe IV (mutation AC/TC at position -1101/1100 in the  $\alpha_{s2}$ -casein promoter region). Differences in the binding of nuclear proteins to polymorphic sequences were statistically significant ( $P \leq 0.05$ ). EMSA competition experiments showed that transcription factors AP1, CREB, TR, YY1, SP1 and NF $\kappa$ B might contribute to the formation of the DNA-protein complexes.

Taken together, these results showed that nucleotide sequence variations in the promoter regions of bovine casein genes might change the affinity of these regions to nuclear proteins – transcription factors – and thus affect the expression of relevant genes that quantitatively influence the composition of milk proteins.

#### 4. POST-TRANSCRIPTIONAL REGULATION: THE IMPACT OF GENETIC POLYMORPHISM ON THE PROCESSING AND STABILITY OF CASEIN MESSENGER RNAs

##### 4.1. Defect in the processing of primary transcripts: alternative splicing

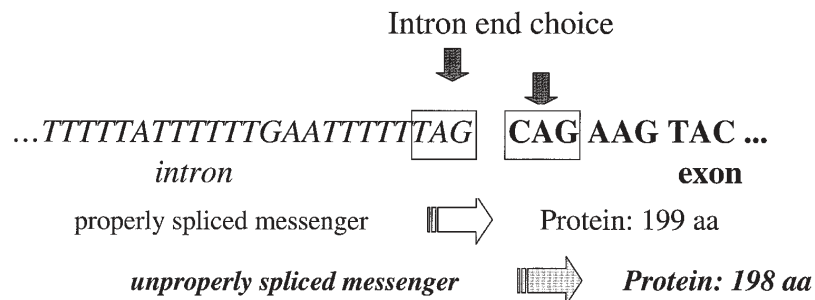
Two kinds of events may intermittently arise during the course of pre-mRNA splicing, both leading to a shorter peptide chain. The first one, referred to as cryptic splice site usage, is possibly due to a sort of slippage of the spliceosome, induced by an “error-prone” junction sequence. As far as caseins are concerned, this defect in accuracy leads casually to the loss of the first codon (usually a CAG) of the 3' exon. The second intermittent event, which is particularly well-exemplified in small ruminants, gives rise to an alternative exon-skipping (sometimes referred to as non allelic exon-skipping). It is thought to be provoked by weaknesses in the consensus sequences, either at the 5' and/or 3' splice junctions or at the branch point, or both.

##### 4.1.1 – Casual usage of cryptic splice sites

The casual deletion of a glutaminy residue (Q<sub>78</sub>), first detected in the goat  $\alpha_{s1}$ -casein [51], seems to be a rather frequent

phenomenon, occurring in most of the species examined so far. This “codon skipping”, which is likely due to an erroneous 3' cryptic splice site usage when exons 10 and 11 are spliced (Fig. 7), has been found, at the protein level, in the four major ruminant species: caprine [25], ovine [24] bovine and water buffalo [26]. Interestingly, such a cryptic splice site slippage does not only occur in ruminants, since a similar event, affecting another glutaminy residue (Q<sub>37</sub> in the mature protein) encoded by the first codon of exon 6', has been reported in the human  $\alpha_{s1}$ -casein [41, 61].

This loss in accuracy of the splicing machinery could be a general phenomenon promoted by the nucleotide sequence at the intron-exon junction. The mechanism by which the AG defining the 3' splice site is accurately and efficiently recognised involves, indeed, a 5'-to-3' scanning process. The first AG downstream of the branch point-polypyrimidine tract is selected preferentially (Fig. 7). The occurrence of a competitive AG, downstream from the proximal one, can be used alternatively [84]. Starting an exon sequence with a CAG (coding for a glutaminy residue) would be a facilitating feature, since the intron 3' splice site consensus sequence ends by XAG, following a pyrimidyl stretch. Both the short size of the polypyrimidyl stretch (interrupted in the goat intron 10 of the  $\alpha_{s1}$ -casein gene by a GAA trinucleotide, Fig. 7) and of the intron might be an enhancing factor. Indeed,



**Figure 7.** Nucleotide sequence at the intron10/exon11 junction to illustrate a “codon skipping” event due to a loss of accuracy in the splicing process. Intron 10 sequence is in italics.

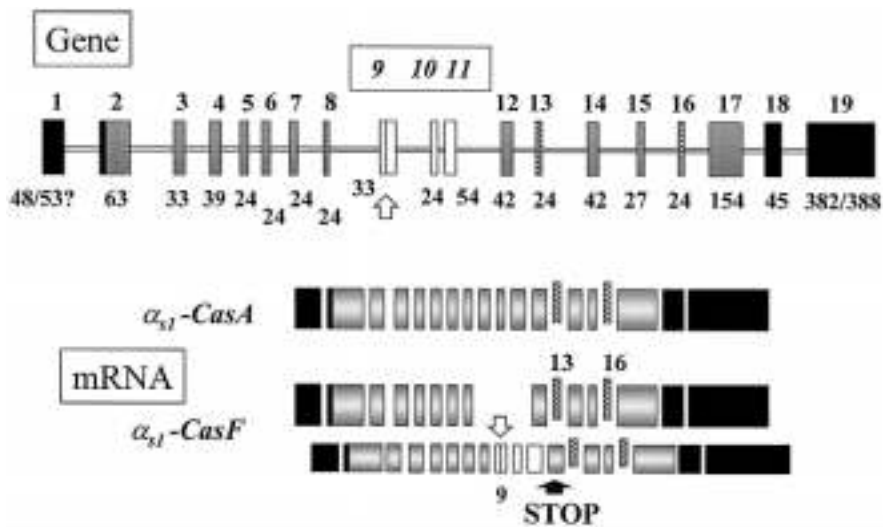


introns 6 and 10, involved in the human and ruminant “codon skipping” events, are 150-bp [61] and 100-bp long [48, 51], respectively.

#### 4.1.2. Casual exon-skipping

Structural characterisation of the casein fraction and/or relevant mRNA analyses have enabled to provide objective elements, in the four major ruminant species, for the occurrence of multiple forms of  $\alpha_{s1}$ -casein. However, the extent of this heterogeneity depends on the species. While  $\alpha_{s1}$ -casein consists of a mixture of two forms (199 and 198 amino acid residues) in cattle and water buffalo, due to the alternative deletion of Q<sub>78</sub>, there are, in sheep and goats, at least seven molecular forms which differ in their peptide chain length, regardless of the genetic polymorphism [25]. The main com-

ponent corresponds to the 199-residue form initially described in goat milk [13]. The others, in lower amounts, are shorter forms of  $\alpha_{s1}$ -casein and differ by the deleted sequences 110–117 and/or 141–148. Genomic and mRNA analysis of  $\alpha_{s1}$ -casein demonstrated that these forms originated from exon-skipping events at the level of exon 13 (encoding peptide 110–117) and/or 16 (encoding peptide 141–148) during the processing of the primary transcripts (Fig. 8). Combinations of such deletions have also been observed. Deletion of peptide 110–117, which contains the charged residues S<sup>P</sup><sub>115</sub>, E<sub>110</sub>, E<sub>117</sub> and K<sub>114</sub>, and of peptide 141–148, which contains E<sub>141</sub>, produces proteins with a different net charge. Only the protein lacking the sequence 141–148 has to date been localised in the gel isoelectric focusing patterns of ovine  $\alpha_{s1}$ -casein [18], whereas the other short forms still remain to be localised.



**Figure 8.** Schematic representation of the inton/exon structure organisation of the goat  $\alpha_{s1}$ -casein gene and of the transcripts arising from alleles A and F. Exons are depicted by large black (3' and 5' untranslated regions) white, hatched, and grey boxes. Matured transcripts (mRNA) are given (bottom). Exons 13 and 16 that are casually skipped are depicted by hatched boxes within the transcript exon assemblies. Open vertical arrows indicate the single nucleotide deletion thought to be responsible for exon-skipping events recorded. Exon 9, 10 and 11 skipped in bloc in short sized transcripts of the F allele are depicted in white boxes in the long-sized transcripts, and the stop codon due to the frame shift mutation (arising in the 9th exon) is indicated by a black vertical arrow. Exon numbering (upper numbers) is given in the gene (top) as well as exon sizes which are indicated in base pairs (lower number).

In caprine and ovine milk, the presence of such deleted  $\alpha_{s1}$ -casein species has been shown with highly expressed genetic variants, as well as with weakly expressed alleles. This finding suggests that it is a frequently encountered event, not related to allelic variability which might be due to the weakness of common sequences either at the 3' and/or 5' splice sites or at the branch point. However, it is worth noting that the alternative splicing of exon 13 and/or 16 reported for small ruminant species has not been detected in cattle and water buffalo [26]. Such differences in the processing of  $\alpha_{s1}$ -casein pre-messengers in closely related species has to be mentioned. Sequencing of a 400-bp genomic DNA fragment, spanning exon 16 in ovine, caprine and bovine species has identified 15 nucleotide substitutions and one nucleotide deletion at the 3' end of intron 15, between the bovine and the small ruminant species, and only 6 substitutions within the 5' sequence of intron 16. None of these substitutions affects the consensus splice sites of exon 16. In the same way, using the Zucker algorithm to predict the secondary structure of pre-messengers, no difference has been detected which could explain the casual alternative splicing of this exon in ovine and caprine species but not in cattle [49]. However, according to Passey et al. [69], a substitution within the donor splice site could be responsible for the casual outsplicing, estimated to affect 20% of the total ovine  $\alpha_{s1}$ -casein mRNA, through the formation of an inhibitory RNA secondary structure.

Once again, this phenomenon is not restricted to ruminants. The existence of three types of  $\alpha_{s1}$ -casein transcripts has been reported in human mammary tissue [41, 61]. This heterogeneity is due to an alternative splicing, involving exon 7 (using the bovine gene numbering), and to the usage of a cryptic splice site. Likewise, porcine  $\alpha_{s1}$ -casein also shows such a heterogeneity, since multiple forms of porcine  $\alpha_{s1}$ -casein cDNA

have been described [3]. These forms differ by internal deletion hypothesised to be due to an altered RNA processing, affecting other exons (12 and 13', using bovine gene numbering).

Multiple forms arising from casual alternative splicing have also been found in ovine  $\alpha_{s2}$ -casein. Two non-allelic forms produced by different mRNAs [10] have been shown to occur in all the individuals analysed, apparently indicating a stabilised mechanism for the production of protein isoforms of different lengths. Two non-allelic forms of  $\alpha_{s2}$ -casein occurring in ovine milk differ by an internal deletion of nine amino acid residues at positions 34–42 of the  $\alpha_{s2}$ -casein chain. Analysis of the products obtained by reverse transcription of mRNA has shown a greater heterogeneity of  $\alpha_{s2}$ -casein transcripts. In addition to the expected deletion of codons (34 to 42) affecting 30–40% of mRNA, another structural difference involving an internal stretch of 44 nucleotides in the 5' untranslated region has been reported. The presence of these different  $\alpha_{s2}$ -casein mRNA has been explained by the authors to be caused by partial exon-skipping.

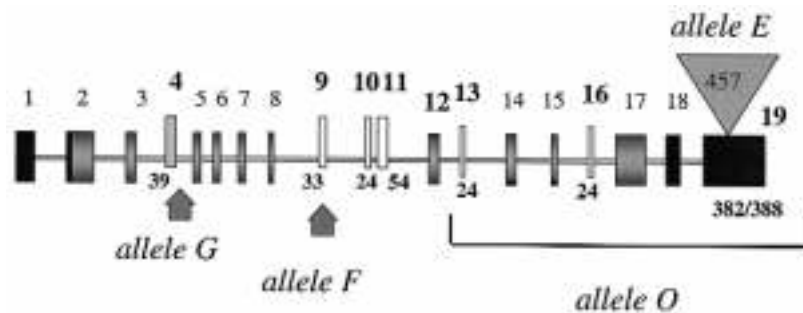
One can therefore propose that exon skipping has to be considered as a frequent event, mainly in the case of  $\alpha_{s1}$ - and  $\alpha_{s2}$ -casein genes, for which the coding region is divided into so many short exons. The maturation of long cognate primary transcripts, such as the one encoding ovine and caprine  $\alpha_{s1}$ -casein scattered by a score of introns [51], appears to be an intricate process requiring many successive steps. Do those known deletions in  $\alpha_{s1}$ -casein and  $\alpha_{s2}$ -casein simply reflect the lack of accuracy of an intricate processing mechanism whenever mutations induce conformational modifications of pre-mRNA, preventing the normal progress of events? Notwithstanding, these phenomena are mainly responsible for the great complexity of casein composition arising from the presence of both full-length and deleted  $\alpha_{s1}$ - and  $\alpha_{s2}$ -casein forms.

#### 4.1.3. Genetic polymorphisms also increase casein heterogeneity and quantitatively influence the milk protein composition

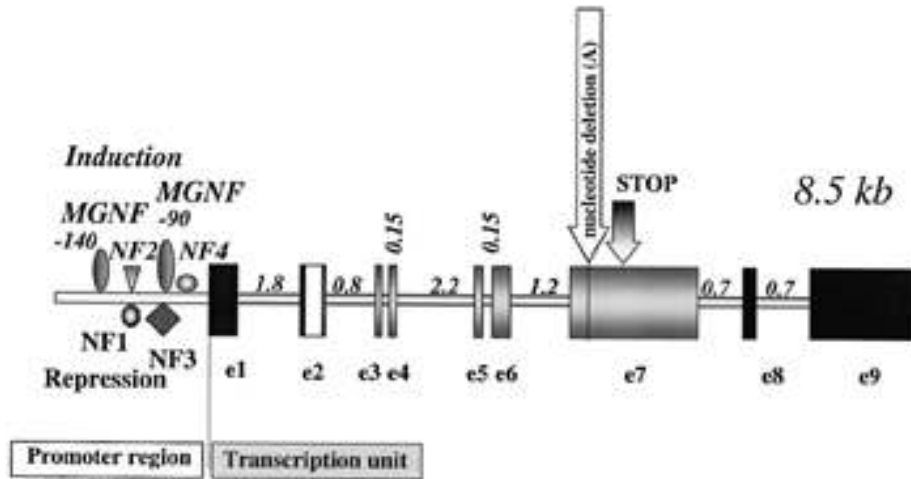
To add further to this casual exon skipping, genetic polymorphism of milk proteins which has been intensively studied especially in cattle [33] and in goat [59], may sharply increase the heterogeneity of caseins in milk. It may also affect the amount of casein in milk either in inducing a loss in the stability and/or translatability of the messengers or by triggering defects in their processing. This is particularly well exemplified in small ruminants. Studies performed on goat milk have reported extensive genetic polymorphism of  $\alpha_{s1}$ -casein with at least 14 autosomal alleles at the goat  $\alpha_{s1}$ -casein locus, distributed in seven different classes of protein variants associated with four levels of expression [34].  $\alpha_{s1}$ -Casein A, B, C and E variants differ from each other in amino acid substitutions, while  $\alpha_{s1}$ -casein F and G variants, associated with a low level of protein synthesis, are internally deleted [14, 60]. The establishment of the overall organisation of the goat  $\alpha_{s1}$ -casein gene (Fig. 9), the characterisation of allele F [51] at the genomic level, as well as the analysis of their transcription

products, demonstrated that the internal deletion of 37 amino acid residues, occurring within variant F, arises from the out splicing of three consecutive exons (9, 10, and 11), skipped *en bloc* during the processing of the primary transcripts (Fig. 8). Furthermore, the  $\alpha_{s1}$ -CasF allele was shown to yield multiple alternatively spliced transcripts, among which were transcripts lacking 24 nucleotide-long sequences encoded by exons 13 and 16 [51]. By comparison with the non-defective  $\alpha_{s1}$ -CasA, B and C alleles, a reduction in the amount of mRNA accounting for a lower  $\alpha_{s1}$ -casein content in milk was observed.

Likewise, a single nucleotide deletion resulting in a premature stop codon is associated with a marked reduction in the amount and an extensive heterogeneity of transcripts from the goat  $\beta$ -casein null allele in Créole and Pyrenean breeds [71, 72]. These authors have shown the occurrence of multiple and shorter  $\beta$ -Cas0 transcripts which differ from their full-length counterparts in large nucleotide stretches that were missing within exon 7. Four in frame nonsense codons (the first one at codon 58), due to a single nucleotide change, were found in the  $\beta$ -Cas0 allele (Fig. 10) and the cognate mRNA. Another  $\beta$ -casein null allele has been identified in a Napolitan goat breed [17]. It was



**Figure 9.** Mutations characterised in defective alleles and shown to be responsible for the “quantitative polymorphisms”. For the description of the gene, see the legend of Figure 2. Boxes in white represent exons that are possibly skipped during the course of the splicing process, whereas exon 4 (stippled) is constitutively skipped due to a mutation (G/A) at the first nucleotide of the 5' splice site consensus sequence. The insertion of a line element (457 nt) in the last exon and the large genomic deletion affecting the 3' end of the transcription unit giving rise to the null allele are also indicated.



**Figure 10.** Schematic representation of the protein binding sites located in the promoter region, upstream from the transcription unit of the  $\beta$ -casein gene, where ubiquitous Nuclear Factors (NF) and Mammary Gland-“specific” Nuclear Factors (MGNF) have been shown to bind. For the description of the gene, see the legend of Figure 2. White and grey vertical arrows indicate: the site of the frame shift mutation responsible for the absence of  $\beta$ -casein synthesis associated with the null allele and the position of the first stop codon arising from this mutation, respectively.

shown to differ from the wild-type by a C  $\rightarrow$  T transition, affecting codon 157 in the seventh exon [74]. The resulting premature termination codon is associated with a 10-fold decrease in  $\beta$ -casein mRNAs. However, data are lacking about the possible occurrence of multiple mRNAs for the Italian allele.

As far as both  $\alpha_{s1}$ -*CasF* and  $\beta$ -*Cas0* are concerned, mutations are responsible for the existence of premature stop codons, associated with a decrease in the relevant level of transcripts and responsible for the presence of multiple forms of messengers, due to alternative splicing. Many reports (reviewed by Valentine [89]) have drawn attention to a possible relationship between nonsense codons and exon skipping. Indeed, some genes containing premature codons express alternatively spliced mRNA in which the exon containing the nonsense codon has been skipped. Among the hypotheses proposed to explain such a safeguard mechanism, one can mention “nuclear scanning”

which recognises nonsense codons and then has an effect on exon definition [99, 100]. This raises the question: How is a normal termination codon (which does not usually mediate a reduction in the abundance of mRNA) distinguished from a premature stop codon? A rule has been proposed for the position of the termination codon. According to Nagy and Maquat [65], it must be located less than 50–55 nucleotides upstream from the 3'-most exon-exon junction. The normal termination codon in  $\alpha_{s1}$ -casein, as well as in  $\beta$ -casein transcripts, is in part or fully encoded by the penultimate exon (exons 18 and 8, respectively) at 43 and 36 nucleotides upstream from the last exon-exon junction, respectively, thus conforming perfectly with the stated rule. Conversely, stop codons identified both with  $\alpha_{s1}$ -*CasF* and the French and Italian  $\beta$ -*Cas0* alleles, are located well beyond the 55 nucleotide limit. Therefore, they could be suspected to mediate mRNA decay and promote the occurrence of multiple forms of transcripts.

#### 4.2. Genetic polymorphism and stability/translatability of messengers

The insertion of a truncated LINE element, highly repeated in the goat genome, in the 3'UTR of the  $\alpha_{s1}$ -CasE mRNA (Fig. 9), was shown to be responsible for a three-fold reduction in the amount of messenger as compared with its A counterpart [70]. Such an insertion in a region known to be crucial for the stability of the messengers [40], might be involved in this phenomenon, probably owing to the occurrence of AU rich elements (ARE) in the inserted LINE sequence. ARE are known to reduce the stability of mRNA [39]. In addition, the presence of a polyU tail at the 5' end of the LINE insertion was expected to induce a sequestering through base pairing interactions of the messenger polyA tail, that could impair the modulation of translation efficiency which is under the control of the 3'UTR and the poly A tail. The apparent absence of the polyA tail might also induce a fast decay of the mRNA molecules as first proposed for *c-fos* by Wilson and Treisman [96]. A similar result was reported a few years later by Rando et al. [75] with the G variant of  $\alpha_{s1}$ -casein in the bovine species.

#### 5. PHYSIOLOGICAL AND TECHNOLOGICAL CONSEQUENCES

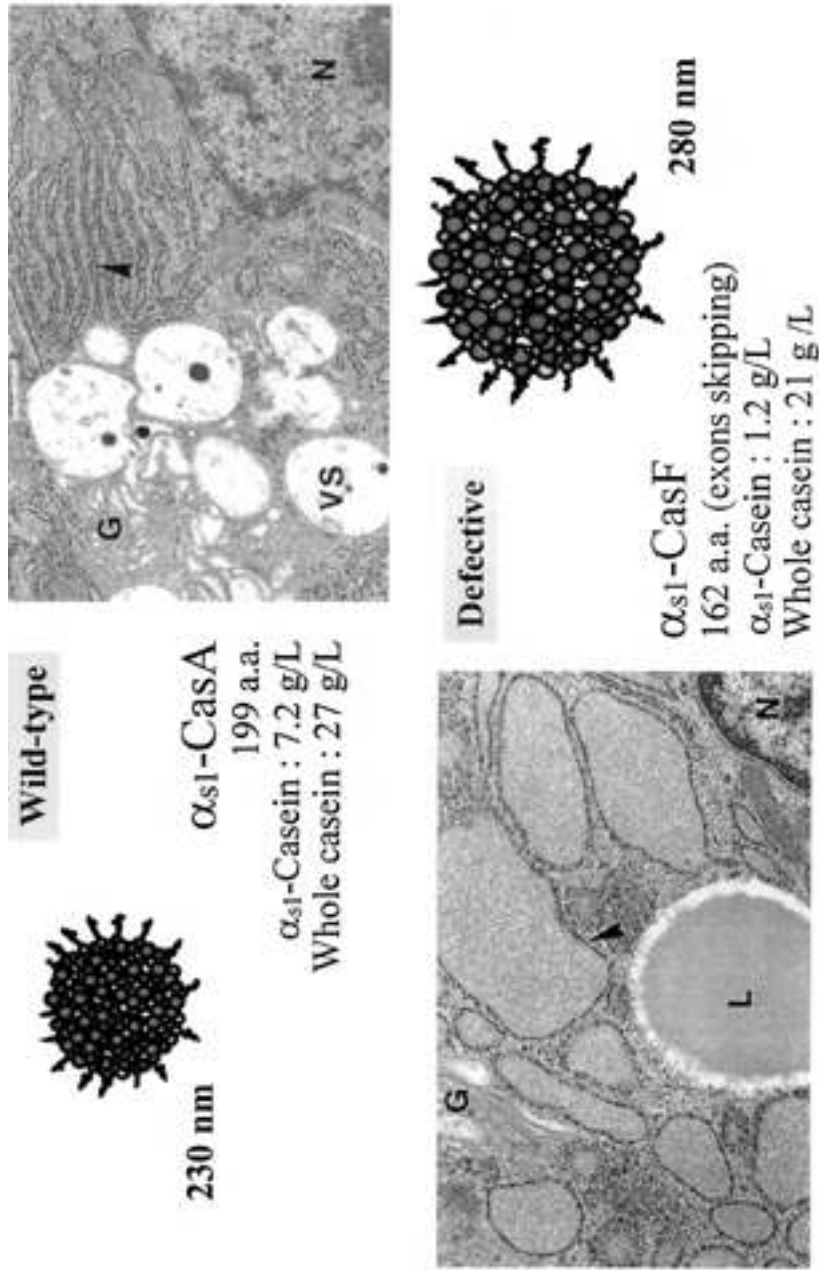
Since defective alleles at the  $\alpha_{s1}$ -casein locus (mainly E and F) were largely dominant in French dairy goat breeds at the end of the nineteen-eighties, then there was an opportunity to increase the casein content of goat milk by committing a selection program in favour of highly expressed alleles (A, B and C). To assess the possible effects on production traits and technological properties of milk, comparative analyses were carried out over 8 years on an experimental farm (Station caprine de Moissac, 48110 Sainte-Croix-Vallée Française) and in breeders.

The main results of the "technological" study, were reported by Remeuf [76] and Vassal et al. [91]. Briefly, the differences in  $\alpha_{s1}$ -casein and protein contents were confirmed and unexpectedly, the genotype at the  $\alpha_{s1}$ -casein locus was shown to influence fat content [5]. The genotype showed a significant effect on the diameter of the casein micelles and on their calcium content which are lower in the A-type milk. Significant effects have also been observed on coagulation parameters, cheese yields and organoleptic properties.

On-farm performance records of the progeny of five bucks heterozygous at the  $\alpha_{s1}$ -casein locus has provided data indicating that  $\alpha_{s1}$ -casein genetic polymorphism is without any significant effect on milk yield and confirmed the effect on fat content which is significantly higher with strong alleles [54]. It is suggested that the reduction in fat content following the reduction in casein might be due to a disruption in the intracellular transport and secretion mechanisms within the mammary epithelial cell (MEC).

Electron microscopy of MECs from lactating goats homozygous for alleles A, E, F and O at the  $\alpha_{s1}$ -Cas locus revealed a very different general morphology. A-type MEC showed typical images with a basal spherical nucleus surrounded by parallel lamellae of rough endoplasmic reticulum (RER), a Golgi apparatus with more or less distended saccules surrounded by numerous small and secretory vesicles containing ramified electron-dense proteins and casein micelles in the supra nuclear region. By contrast, F- and O-types (and to a lesser extent the E-type) MEC were noticeably different, with a dramatically dilated and distended RER (Fig. 11) in which a granular electron-dense material accumulated [15]. Kinetics of intracellular transport of newly synthesised proteins in mammary tissue biopsies revealed that the rate of progression of the other caseins towards the Golgi apparatus is highly reduced, when  $\alpha_{s1}$ -casein is lacking [16].





**Figure 11.** Electron micrographies showing the morphology of the mammary epithelial cell from lactating goats homozygous A (wild type) or F (defective) at the  $\alpha_{s1}$ -casein locus [16]. Note the remarkable enlargement of the cisternae of the Rough Endoplasmic Reticulum (arrow heads) observed with the F defective allele. The micelle size and the features of each  $\alpha_{s1}$ -casein variant as well as its impact on the casein content of goat milk is given. L: lumen; VS: secretory vesicles; G: Golgi; N: nucleus.

This leads to the accumulation of immature caseins in distended RER cisternae. In contrast, no accumulation was found in MECs which lack  $\beta$ -casein, suggesting that efficient export of the caseins from the RER to the Golgi is dependent upon the interaction between  $\alpha_{s1}$ -casein and the other caseins.

## 5. CONCLUSIONS

Since the first report in 1955, by Aschaffenburg and Drewry [4] of the polymorphic status of milk proteins there has been a growing interest in determining genotypes of breeding cattle. This opened the door to extensive research aiming at characterising genetic variants of the six main milk proteins. To make a more efficient production possible and to achieve a tailored genetic improvement of milk protein quality using established genetic approaches (selection and breeding programmes), it was also necessary at the DNA level, to increase our understanding of the regulation of milk protein gene expression and to characterise the factors responsible for the individual variability in the structure of milk proteins.

The identification of transcription factors involved in the control of milk protein gene expression was a major breakthrough in the understanding of the mechanisms that modulates the biosynthesis of milk components. The presence of mutations within potential binding sites for these nuclear factors, found in the promoter regions of milk protein genes, have been shown to be actually responsible for different binding affinities and result in differences in gene expression. Likewise, the description of allele specific splicing patterns leading to both internal peptide deletion and significant reductions in the amount of protein synthesised has to be considered as a relevant progress in the perception that we now have of the way gene expression is modulated.

The present knowledge of the structure and function of the major milk protein genes

is already being applied to the selection of animals according to their genotypes. Several PCR-based methods for analysing cattle DNA sequence polymorphisms such as AS-PCR (Allele Specific-Polymerase Chain Reaction), PCR-RFLP (Polymerase Chain Reaction-Restriction Fragment Length Polymorphism), ACRS (Amplification-Created Restriction Site), SSCP (Single-Strand Conformation Polymorphism), have been proposed for genotyping milk protein loci in both sexes (cows and bulls, goats and bucks, ewes and rams), in very young animals, even in embryos. Thus, molecular genetic markers could play an important role in the pre-selection of embryos prior to transfer, young bulls to progeny tests, or heifers to rear when there is a large "surplus" of embryos, heifer or bull calves.

## ACKNOWLEDGEMENTS

The authors wish to thank Ms Eulalia Siadkowska for excellent technical assistance. We also wish to thank Dr T. Malewski for his scientific and laboratory support. This study was funded by the State Committee for Scientific Research (KBN) grant No. 6 P04B 019 17.

## REFERENCES

- [1] Aggeler J., Park C.S., Bissell M.J., Regulation of milk protein and basement membrane gene expression: the influence of the extracellular matrix, *J. Dairy Sci.* 71 (1988) 2830–2842.
- [2] Aleandri R., Buttazzoni L.G., Schneider J.C., Caroli A., Davoli R., The effects of milk protein polymorphisms on milk components and cheese-producing ability, *J. Dairy Sci.* 73 (1990) 241–255.
- [3] Alexander L.J., Das Gupta N.A., Beattie C.W., The sequence of porcine  $\alpha_{s1}$ -casein cDNA, *Anim. Genet.* 23 (1992) 365–367.
- [4] Aschaffenburg R., Drewry J., Occurrence of different  $\beta$ -lg in cow's milk, *Nature* 176 (1955) 218–219.
- [5] Barbieri M., Manfredi E., Elsen J.-M., Ricordeau G., Bouillon J., Grosclaude F., Mahé M.F., Bibé B., Influence du locus de la caséine  $\alpha_{s1}$  sur les performances laitières et les paramètres génétiques des chèvres de race Alpine, *Genet. Sel. Evol.* 27 (1995) 437–450.

- [6] Bech A.M., Kristiansen K.R., Milk protein polymorphism in Danish dairy cattle and influence of genetic variants on milk yield, *J. Dairy Res.* 57 (1990) 53–62.
- [7] Bleck G.T., Bremel R.D., Correlation of the  $\alpha$ -lactalbumin (+15) polymorphism to milk production and milk composition of Holsteins, *J. Dairy Sci.* 76 (1993) 2292–2298.
- [8] Bleck G.T., Bremel R.D., Sequence and single-base polymorphisms of the bovine  $\alpha$ -lactalbumin 5' flanking region, *Gene* 126 (1993) 213–218.
- [9] Bleck G.T., Conroy J.C., Wheeler M.B., Polymorphisms in the bovine  $\beta$ -casein 5' flanking region, *J. Dairy Sci.* 79 (1996) 347–349.
- [10] Boissard M., Hue D., Bouniol C., Mercier J.C., Gaye P., Multiple mRNA species code for two non-allelic forms of ovine  $\alpha_{s2}$ -casein, *Eur. J. Biochem.* 201 (1991) 633–641.
- [11] Bonsing J., Ring J.M., Stewart A.F., Mackinlay A.G., Complete nucleotide sequence of the bovine  $\beta$ -casein gene, *Aust. J. Biol. Sci.* 41 (1988) 527.
- [12] Brew K., Hill R.L., Lactose biosynthesis, *Rev. Physiol. Biochem. Pharmacol.* 72 (1975) 105–157.
- [13] Brignon G., Mahé M.F., Grosclaude F., Ribadeau Dumas B., Sequence of caprine  $\alpha_{s1}$ -casein and characterization of those of its genetic variants which are synthesized at a high level,  $\alpha_{s1}$ -CN A, B and C, *Protein Seq. Data Anal.* 2 (1989) 181–188.
- [14] Brignon G., Mahé M.F., Ribadeau Dumas B., Mercier J.C., Grosclaude F., Two of the three genetic variants of goat  $\alpha_{s1}$ -casein which are synthesized at a reduced level have an internal deletion possibly due to altered RNA splicing, *Eur. J. Biochem.* 193 (1990) 237–241.
- [15] Chanut E., Leroux C., Martin P., Ollivier-Bousquet M.,  $\alpha_{s1}$ -casein polymorphism in goat: perturbation of the intracellular transport of milk proteins, in: Blum J.W., Bruckmaier R.M., Proceedings of the Symposium on milk synthesis, secretion and removal in ruminants, Berne, Switzerland, 1996, p. 31.
- [16] Chanut E., Martin P., Ollivier-Bousquet M.,  $\alpha_{s1}$ -casein is required for the efficient transport of  $\beta$ - and  $\kappa$ -casein from the endoplasmic reticulum to the Golgi apparatus of mammary epithelial cells, *J. Cell. Sci.* 112 (1999) 3399–3412.
- [17] Chianese L., Garro G., Nicola M.A., Mauriello R., Ferranti P., Pizzano R., Cappuccio U., Laezza P., Addeo F., Ramunno L., Rando A., Rubino R., The nature of  $\beta$ -casein heterogeneity in caprine milk, *Lait* 73 (1993) 533–547.
- [18] Chianese L., Garro G., Mauriello R., Laezza P., Ferranti P., Addeo F., Occurrence of five  $\alpha_{s1}$ -casein variants in ovine milk, *J. Dairy Res.* 63 (1996) 49–59.
- [19] Coll A., Folch J.M., Sanchez A., Structural features of the 5' flanking region of the caprine kappa-casein gene, *J. Dairy Sci.* 78 (1995) 973–977.
- [20] De Lange D.J., Meyer E.H.H., Van Rensburg L., Swanepoel A., The polymorphism protein system in cow's milk. I. The association with the production of milk protein and fat during lactation, *South Africa J. Dairy Sci.* 22 (1990) 67–71.
- [21] Debeljak M., Susnik S., Marinsek-Logar R., Medrano J.F., Dovc P., Allelic differences in bovine kappa-CN gene which may regulate gene expression, *Pfligers Arch.* 439 (2000) R4–R6.
- [22] Eggen A., Fries R., Die Untersuchung von Kaseingenen mittels DNA-Analyse, *ETH Landwirtschaft Schweiz Band 2* (1989) 231–235.
- [23] Ehrmann S., Bartenschlager H., Geldermann H., Polymorphism in the 5' flanking region of the bovine  $\beta$ -lactoglobulin-encoding gene and its association with  $\beta$ -lactoglobulin in the milk, *J. Anim. Breed. Genet.* 114 (1997) 49–53.
- [24] Ferranti P., Malorni A., Nitti G., Laezza P., Pizzano R., Chianese L., Addeo F., Primary structure of ovine  $\alpha_{s1}$ -casein: localization of phosphorylation sites and characterization of genetic variants, *J. Dairy Res.* 62 (1995) 281–296.
- [25] Ferranti P., Addeo F., Malorni A., Chianese L., Leroux C., Martin P., Differential splicing of pre-messenger RNA produces multiple forms of goat  $\alpha_{s1}$ -casein, *Eur. J. Biochem.* 249 (1997) 1–7.
- [26] Ferranti P., Lilla S., Chianese L., Addeo F., Alternative nonallelic deletion is constitutive of ruminant  $\alpha_{s1}$ -casein, *J. Protein Chem.* 18 (1999) 595–602.
- [27] Ferretti L., Leone P., Sgaramella V., Long range restriction analysis of the bovine casein genes, *Nucleic Acids Res.* 18 (1990) 6829–6833.
- [28] Folch J.M., Coll A., Sanchez A., Complete sequence of the caprine  $\beta$ -lactoglobulin gene, *J. Dairy Sci.* 77 (1994) 3493–3497.
- [29] Folch J.M., Dovc P., Medrano J.F., Differential expression of bovine  $\beta$ -lactoglobulin A and B promoter variants in transiently transfected HC11 cells, *J. Dairy Res.* 66 (1999) 537–544.
- [30] Geldermann H., Gogol J., Kock M., Tacea G., DNA variants within the 5' flanking region of bovine milk protein encoding genes, *J. Anim. Breed. Genet.* 113 (1996) 261–267.
- [31] Graml R., Weiss G., Buchberger J., Pirchner F., Different rates of synthesis of whey protein and casein by alleles of the  $\beta$ -lactoglobulin and  $\alpha_{s1}$ -casein locus in cattle, *Genet. Sel. Evol.* 21 (1989) 547–554.
- [32] Groenen M.A.M., Van de Poel J.J., Regulation of expression of milk protein genes: a review, *Livestock Prod. Sci.* 38 (1994) 61–78.
- [33] Grosclaude F., Le polymorphisme génétique des principales lactoprotéines bovines, *Prod. Anim.* 1 (1988) 5–17.



- [34] Grosclaude F., Martin P., Casein polymorphisms in the goat. IDF Seminar "Milk Protein Polymorphism II", Int. Dairy Fed. ed., Palmerston North, New Zealand, February, 1997, pp. 241–253.
- [35] Grosclaude F., Garnier J., Ribadeau Dumas B., Jeunet R., Étroite dépendance des loci contrôlant le polymorphisme des caséines  $\alpha$  et  $\beta$ , C.R. Hebd. Scéances Acad. Sci. 259 (1964) 1569–1571.
- [36] Grosclaude F., Joudrier P., Mahé M.F., Polymorphisme de la caséine  $\alpha_{s2}$  bovine : étroite liaison du locus  $\alpha_{s2}$ -Cn avec les loci  $\alpha_{s1}$ -Cn,  $\beta$ -Cn et  $\kappa$ -Cn, mise en évidence d'une délétion dans le variant  $\alpha_{s2}$  CnD, Ann. Genet. Sel. Anim. 10 (1978) 313–327.
- [37] Hayes H., Petit E., Mapping of the  $\beta$ -lactoglobulin gene and of an immunoglobulin M heavy chain-like sequence to homoeologous cattle, sheep, and goat chromosomes, Mamm. Genome 4 (1993) 207–210.
- [38] Hayes H., Petit E., Bouniol C., Popescu P., Localisation of the alpha-S2-casein gene (CASAS2) to the homeologous cattle, sheep and goat chromosomes 4 by in situ hybridization, Cytogenet. Cell Genet. 64 (1993) 282–285.
- [39] Jackson R.J., Cytoplasmic regulation of mRNA function: the importance of the 3' untranslated region, Cell 74 (1993) 9–14.
- [40] Jackson R.J., Standart N., Do the poly(A) tail and 3' untranslated region control mRNA translation? Cell 62 (1990) 15–24.
- [41] Johnsen L.B., Rasmussen L.K., Petersen T.E., Berglung L., Characterization of three types of human  $\alpha_{s1}$ -casein mRNA transcripts, Biochem. J. 309 (1995) 237–242.
- [42] Jolivet G., Meusnier C., Chaumaz G., Houdebine L.M., Extracellular matrix regulates  $\alpha_{s1}$ -casein gene expression in rabbit primary mammary cells and CCAAT enhancer binding protein (C/EBP) binding activity, J. Cell. Biochem. 82 (2001) 371–386.
- [43] Kamiński S., DdeI RFLP at the 5' region of bovine  $\kappa$ -casein gene, J. Appl. Genet. 37 (1996) 173–178.
- [44] Kamiński S., Identification of *SduI* polymorphisms within the 5' flanking region of the bovine  $\alpha$ -lactalbumin gene, Anim. Sci. Pap. Rep. 17 (1999) 23–27.
- [45] Kamiński S., Associations between polymorphism within regulatory and coding fragments of bovine  $\kappa$ -casein gene and milk performance traits, J. Anim. Feed. S9 (2000) 435–446.
- [46] Kamiński S., Józefowicz T., SSCP polymorphism within the promoter of the bovine  $\alpha_{s1}$ -casein gene, J. Anim. Feed. 9 (2000) 73–79.
- [47] Kamiński S., Zabołewicz T., Associations between bovine  $\beta$ -lactoglobulin polymorphism within coding and regulatory sequences and milk performance traits, J. Appl. Genet. 41 (2000) 91–99.
- [48] Koczan D., Hobom G., Seyfert H.M., Genomic organization of the bovine  $\alpha_{s1}$ -casein gene, Nucleic Acids Res. 18 (1991) 5591–5596.
- [49] Leroux C., Analyse du polymorphisme du gène caprin codant la caséine  $\alpha_{s1}$  et des produits de sa transcription. Application au développement d'une procédure de typage précoce des animaux, Ph.D. thesis, Université d'Orsay-Paris XI, 1992.
- [50] Leroux C., Martin P., The caprine  $\alpha_{s1}$ - and  $\beta$ -casein genes are 12-kb apart and convergently transcribed, Anim. Genet. 27 (1996) 93.
- [51] Leroux C., Mazure N., Martin P., Mutation away from splice site recognition sequences might *cis*-modulate alternative splicing of goat  $\alpha_{s1}$ -casein transcript. Structural organization of the relevant gene, J. Biol. Chem. 267 (1992) 6147–6157.
- [52] Lum L.S., Dovc P., Medrano J.F., Polymorphisms of bovine  $\beta$ -lactoglobulin promoter and differences in the binding affinity of activator protein-2 transcription factor, J. Dairy Sci. 80 (1997) 1389–1397.
- [53] Lundén A., Nilsson M., Janson L., Marked effect of  $\beta$ -lactoglobulin polymorphism on the ratio of casein to total protein in milk, J. Dairy Sci. 80 (1997) 2996–3005.
- [54] Mahé M.F., Manfredi E., Ricordeau G., Piacère A., Grosclaude F., Effet des variants de la caséine  $\alpha_{s1}$  caprine sur les performances laitières : analyse intradescendance de boucs de race Alpine, Gen. Sel. Evol. 26 (1994) 151–157.
- [55] Malewski T., Computer analysis of distribution of putative *cis*- and *trans*-regulatory elements in milk protein gene promoters, Biosystems 45 (1998) 29–44.
- [56] Malewski T., Zwierzchowski L., Computer-aided analysis of potential transcription-factor binding sites in the rabbit  $\beta$ -casein gene promoter, Biosystems 36 (1995) 109–119.
- [57] Malewski T., Zwierzchowski L., Changes of tissue-specific transcription factors in the rabbit mammary gland during pregnancy and lactation, Tsitológia i Genetika 31 (4) (1997) 58–69.
- [58] Mao I.L., Buttazozoni L.G., Aleandri R., Effects of polymorphic milk protein genes on milk yield and composition traits in Holstein cattle, Acta Agriculturae Scand. Sect. A (Anim. Sci.) 42 (1992) 1–7.
- [59] Martin P., Polymorphisme génétique des lactoprotéines caprines, Lait 73 (1993) 511–532.
- [60] Martin P., Leroux C., Characterization of a further  $\alpha_{s1}$ -casein variant generated by exon skipping. Proc. 24th International Society of Animal Genetics Conference, Prague, Abstract E43, 1994, p. 88.
- [61] Martin P., Brignon G., Furet J.P., Leroux C., The gene encoding  $\alpha_{s1}$ -casein is expressed in human mammary epithelial cells during lactation, Lait 76 (1996) 523–535.

- [62] McLean D.M., Graham E.R.B., Ponzoni R.W., McKenzie H.A., Effects of milk protein genetic variants on milk yield and composition, *J. Dairy Res.* 51 (1984) 531–546.
- [63] Medrano J.F., Aquilar-Cordova E., Genotyping of bovine kappa-casein loci following DNA sequence amplification, *Biotechnology* 8 (1990) 144–146.
- [64] Medrano J.F., Aquilar-Cordova E., Polymerase chain reaction amplification of bovine  $\beta$ -lactoglobulin genomic sequences and identification of genetic variants by RFLP analysis, *Anim. Biotech.* 1 (1990) 73–77.
- [65] Nagy E., Maquat L.E., A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance, *Trends Biochem. Sci.* 23 (1998) 198–199.
- [66] Ng-Kwai-Hang K.F., Hayes J.F., Moxley J.E., Monardes H.G., Association of genetic variants of casein and milk serum proteins with milk, fat, and protein production by dairy cattle, *J. Dairy Sci.* 67 (1984) 835–840.
- [67] Ng-Kwai-Hang K.F., Hayes J.F., Moxley J.E., Monardes H.G., Relationships between milk protein polymorphisms and major milk constituents in Holstein-Friesian cows, *J. Dairy Sci.* 69 (1986) 22–26.
- [68] Pappalardo M., Rando A., Di Gregorio P., Masina P., Ramunno L., *MseI* RFLP in the 5' DNA region of the goat  $\beta$ -casein gene, *Anim. Genet.* 28 (1997) 238–246.
- [69] Passey R., Glenn W., Mackinlay A., Exon skipping in the ovine alpha S1-casein gene, *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 114 (1996) 389–394.
- [70] Perez M.J., Leroux C., Bonastre A.S., Martin P., Occurrence of a LINE sequence in the 3' UTR of the goat  $\alpha_{s1}$ -casein E-encoding allele associated with reduced protein synthesis level, *Gene* 147 (1994) 179–187.
- [71] Persuy M.A., Printz C., Medrano J.F., Mercier J.C., One mutation might be responsible for the absence of beta-casein in two breeds of goats, *Anim. Genet.* 27 (1996) 96.
- [72] Persuy M.A., Printz C., Medrano J.F., Mercier J.C., A single nucleotide deletion resulting in a premature stop codon is associated with marked reduction of transcripts from a goat beta-casein null allele, *Anim. Genet.* 30 (1999) 444–451.
- [73] Pinder S.J., Perry B.N., Skidmore C.J., Savva D., Analysis of polymorphism in the bovine casein genes by use of the polymerase chain reaction, *Anim. Genet.* 22 (1991) 11–20.
- [74] Rando A., Pappalardo M., Capuano M., Di Gregorio P., Ramunno L., Two mutations might be responsible for the absence of beta-casein in goat milk, *Anim. Genet.* 27 (1996) 31.
- [75] Rando A., Di Gregorio P., Ramunno L., Mariani P., Fiorella A., Senese C., Marletta D., Masina P., Characterization of the CSN1A<sup>G</sup> allele of the bovine  $\alpha_{s1}$ -casein locus by the insertion of a relict of a long interspersed element, *J. Dairy Sci.* 81 (1998) 1735–1742.
- [76] Remeuf F., Influence du polymorphisme génétique de la caséine  $\alpha_{s1}$  caprine sur les caractéristiques physico-chimiques et technologiques du lait, *Lait* 73 (1993) 549–557.
- [77] Rijnkels M., Kooiman P.M., Krimpenfort P.J.A., de Boer H.A., Pieper F.R., Expression analysis of the individual bovine  $\beta$ -,  $\alpha_{s2}$ - and  $\kappa$ -casein genes in transgenic mice, *Biochem. J.* 311 (1995) 929–937.
- [78] Rijnkels M., Kooiman P.M., de Boer H.A., Pieper F.R., Organisation of the bovine casein gene locus, *Mamm. Genome* 8 (1997) 148–152.
- [79] Rosen J.M., Li S., Raught B., Handsell D., The mammary gland as a bioreactor: factors regulating the efficient expression of milk-protein-based transgenes, *Am. J. Clin. Nutr.* 63 (1996) 627S–632S.
- [80] Rosen J.M., Wyszomierski S.L., Hadsell D., Regulation of milk protein gene expression, *Ann. Rev. Nutr.* 19 (1999) 407–436.
- [81] Schild T.A., Geldermann H., Variants within the 5'-flanking regions of bovine milk-protein-encoding genes. III. Genes encoding the Ca-sensitive casein  $\alpha_{s1}$ ,  $\alpha_{s2}$  and  $\beta$ , *Theor. Appl. Genet.* 93 (1996) 887–893.
- [82] Schild T.A., Wagner V., Geldermann H., Variants within the 5'-flanking regions of bovine milk protein genes. I.  $\kappa$ -Casein-encoding gene, *Theor. Appl. Genet.* 89 (1994) 116–120.
- [83] Simpson K.J., Bird P., Shaw D., Nicholas K., Molecular characterisation and hormone-dependent expression of the porcine whey acidic protein gene, *J. Mol. Endocrinol.* 20 (1998) 27–35.
- [84] Smith C.W., Chu T.T., Nadal-Ginard B., Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns, *Mol. Cell. Biol.* 13 (1993) 4939–4952.
- [85] Szymanowska M., Malewski T., Zwierzchowski L., Single-base polymorphism in the regulatory region of the bovine  $\alpha_{s2}$ -casein gene influences binding of transcription factors, 27th FEBS Meeting, 30 June–5 July 2001 Lisbon, Portugal, *FEBS J.*, Abstr. PS3–101, 2001, p. 86.
- [86] Szymanowska M., Malewski T., Zwierzchowski L., Association between sequence polymorphism in 5'-flanking region of bovine  $\alpha_{s2}$ -casein gene, transcription factor binding and milk composition. Symposium Franco-Polonais "Mécanismes de Régulation chez l'Animal et l'Homme en Croissance" 25–26 September, Paris, France, *Proceedings*, C14 p. 37/73, 2001.
- [87] Threadgill D.W., Womack J.E., Genomic analysis of the major bovine casein genes, *Nucleic Acids Res.* 18 (1990) 6935–6942.
- [88] Topper Y.J., Freeman C.S., Multiple hormone interactions in the developmental biology of the mammary gland, *Physiol. Rev.* 60 (1980) 1049–1106.

- [89] Valentine C.R., The association of the nonsense codons with exon skipping, *Mutation Res.* 411 (1998) 87–117.
- [90] Van Eenennaam A.L., Medrano J.F., Differences in allelic protein expression in the milk of heterozygous  $\kappa$ -casein cows, *J. Dairy Sci.* 74 (1991) 1491–1496.
- [91] Vassal L., Delacroix-Buchet A., Bouillon J., Influence des variants AA, EE et FF de la caséine  $\alpha_{s1}$  caprine sur le rendement fromager et les caractéristiques sensorielles de fromages traditionnels: premières observations, *Lait* 74 (1994) 89–103.
- [92] Voelker G.R., Bleck G.T., Wheeler M.B., Single-base polymorphisms within the 5' flanking region of the bovine  $\alpha$ -lactalbumin gene, *J. Dairy Sci.* 80 (1997) 194–197.
- [93] Vonderhaar B.K., Ziska S.E., Hormonal regulation of milk protein gene expression, *Ann. Rev. Physiol.* 51 (1989) 641–652.
- [94] Wagner V.A., Schild T.A., Geldermann H., DNA variants within the 5'-flanking region of milk-protein-encoding genes. II. The  $\beta$ -lactoglobulin-encoding gene, *Theor. Appl. Genet.* 89 (1994) 121–126.
- [95] Watson C.J., Gordon K.E., Robertson M., Clark A.J., Interaction of DNA-binding proteins with a milk protein gene promoter in vitro: Identification of a mammary gland-specific factor, *Nucleic Acids Res.* 19 (1991) 6603–6610.
- [96] Wilson T., Treisman R., Removal of poly(A) and consequent degradation of *c-fos* mRNA facilitated by 3' AU-rich sequences, *Nature* 336 (1988) 396–399.
- [97] Winklehner-Jennewein P., Geymayer S., Lechner J., Welte T., Hanson L., Geley S., Doppler W., A distal enhancer region in the human  $\beta$ -casein gene mediates the response to prolactin and glucocorticoid hormones, *Gene* 217 (1998) 127–139.
- [98] Yahyaoui M.H., Pena R.N., Sánchez A., Folch J.M., Rapid communication: polymorphism in the goat  $\beta$ -lactoglobulin proximal promoter region, *J. Anim. Sci.* 78 (2000) 1100–1101.
- [99] Zhang J., Sun X., Qian Y., LaDuca J.P., Maquat L.E., At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation, *Mol. Cell Biol.* 18 (1998) 5272–5283.
- [100] Zhang J., Sun X., Qian Y., LaDuca J.P., Maquat L.E., Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm, *RNA* 4 (1998) 801–815.
- [101] Zwierzchowski L., Structure, expression and engineering of milk protein genes, *Biotechnology (Poland)* 33 (1996) 81–94.