Regulatory Sequences in the 5' Flanking Region of Goat β-Casein Gene

Mu-Chiou Huang* and Jiunn-Shiuan Chao[1]

Department of Animal Science, National Chung-Hsing University, 250 Kao-Kung Road, Taichung, Taiwan, ROC

ABSTRACT : A goat β-casein gene was cloned and sequenced. Our previous study had determined the nucleotide sequences of the 5' flanking region and the structural gene including all 9 exons. In the present study, investigations were done on the regulatory sequences in the 5' flanking region of the goat β-casein gene by aligning and comparing it with the same gene from other mammals. The results showed that -200/-1 bp of the 5' flanking sequences contained six conserved clusters, in which the sites of gene expression regulated by the transcription factor and hormone might exist. It showed that fourteen glucocorticoid receptor elements, two cAMP responsive elements, two SV40 virus enhancer core sequences, two OCT-1 binding elements and one CTF/NF-1 binding element were dispersed in the 5' flanking region of goat β-casein gene. Our findings are perhaps valuable for the elucidation of the molecular mechanisms that control the expression of the goat β-casein gene. *(Asian-Aust. J. Anim. Sci. 2001. Vol 14, No. 11 : 1628-1633)*

Key Words : Promoter, Regulatory Sequences, β-Casein Gene

INTRODUCTION

A variety of factors regulate the expression of milk protein genes, including hormones, cell-cell and cell-substratum interactions and the DNA sequences themselves (Doppler et al., 1989; Eisenstein and Rosen, 1988; Haslam, 1988; Kanai et al., 1993; Schmidhauser et al., 1990). In the eukarytic genome, the 5' flanking region of a gene is the primary sequence promoting its expression. In addition, regulatory elements are required for regulating the tissue- and stage-specific expression of the mammalian gene (Mitchell and Tjian, 1989). Casein genes display both tissue and stage specific patterns of expression. They are good models to study promoter function and the regulatory elements required for their expression, and also promote foreign genes to produce biologically proteins in the milk of transgenic animals. The 5' flanking region and the regulatory elements are essential for casein gene expression. It has been found that there are several transcription factors which could bind to some important regulatory elements of the casein gene (Groenen et al., 1992; Meier and Groner, 1994; Schmitt-Ney et al., 1991). Such phenomenon involves regulating signal transduction pathway or hormonal or cell-substratum interactions. But these transcription factors and genes are almost all obtained from the rodent model. Neither the regulatory elements of goat casein genes nor the transcription factors of goat mammary gland have been found. Even if the same regulatory sites in these casein genes and the same transcription factors occur in the mammary glands of all mammals; they could have some differences in expression patterns. The biological functions of mammary gland exhibit some differences between rodents and ruminants (Ann et al., 1995), and the casein ratio is different between them, i. e., αs1-casein is lacking in goat milk (Jenness, 1980).

In our previous work we cloned the goat β-casein gene from a genomic library (Huang and Lin, 1996). In this study, we aligned the sequences of the 5' flanking regions of the β-casein genes from goat, bovine, rabbit, rat and mouse and investigated the regulatory sequences of the goat β-casein gene.

MATERIALS AND METHODS

Positive clones of recombinant phages containing goat β-casein DNA fragments were screened from a genomic DNA library (Clontech Lab. Inc.) by plaque hybridization *in situ* according to the method of Maniatis et al. (1989). The DNA of positive clones were digested with SalI and the goat β-casein DNA fragments were isolated and constructed into pUC18 plasmid for restriction enzyme mapping. A restriction map was made from goat β-casein genomic DNA whose length was about 20 kb (Huang and Lin, 1996). The nucleotide sequence of goat β-casein genomic DNA was analyzed using the Sequenase® Kit (United States Biochemical Corp., Ohio, USA) by the dideoxy chain termination method. Our previous study had determined the nucleotide sequences of the 5' flanking region and the structural gene including all 9 exons (Chao and Huang, 1997). In the present study, the sequence of the goat β-casein gene and the casein gene sequences of bovine (Bonsing et al., 1988), rabbit (Thepot et al., 1991), mouse (Yoshimura and Oka, 1989) and rat (Jones et al., 1985) were keyed into a computer (Acer) for alignment studies using

* Address reprint request to Mu-Chiou Huang. Tel: +886-4-22850223, Fax : +886-4-22860265, E-mail: mchuang@mail.nchu.edu.tw

[1] Hsin-Chu Branch Station, Taiwan Livestock Research Institute, Council of Agriculture. No.1 Lane 51, Ta-Hu Road, Hsin-Chu, Taiwan, ROC.

Received March 20, 2001; Accepted June 15, 2001

the PC/GENE computer software (IntelliGenetics Co., CA, USA). Finally, the regulatory sequences of the goat β-casein gene were investigated by comparing the sequences of the goat with those of other species.

RESULTS AND DISCUSSION

Most of the DNA sequences mediating the β-casein gene expression reside in the 5' flanking region (Doppler et al., 1989; Schmidhauser et al., 1990; Yoshimura and Oka, 1990b). When the 5' sequence of the goat β-casein gene from -1 bp to -1700 bp (Chao and Huang, 1997) was compared with the first 1.7 Kb of the 5' flanking region of the bovine β-casein gene (-1 to -1,700) (Bonsing et al., 1988). It was found that there was a 94% homology between them. Aligning the first 200 bp of the 5' flanking region of five β-casein genes, revealed that their sequences were highly conserved. Six conserved clusters denoted by boxes A, B, C, D, E and F were also observed (figure 1). We could postulate these conserved sequences may affect the expression of the goat β-casein gene. Box A represents the TATA box, in which RNA polymerase II and TATA box-binding factors can correctly bind with the TATA box for initiating transcription. CAAACCAC, the first eight bases in box B, were identical with the virus enhancer core sequence (Weiher et al., 1983). The right half of box B was identical with the corresponding site in the 5' flanking

```
      -190       -180       -170       -160
CAATTTCATTTCCTTTATCAGATTGTGAGTTA-TTCCTGTTAA-AATGCT
 .....A..........T.........A...-....CT....-......
 ..----.............AG.CC.A...A...C.G...TG.CTC......
 .CTGACA.G......C.C...C..C...A..GC.G...TG.TT-....TC
  GGGCA.G......A.C...C..C...A..GC.G...TG.TT-....TA
         Box F
-150       -140       -130       -120       -110
CCCCAGAATTTCTGGGGACAGAAAAATAGGAAGAATTCATTTCCTAATCA
 ..........T.............................T.......
 .................A...T..-.GA.T....A......T.......
 .............T....A.......-...A.....AC.....T.......
 ...T.........T....A...C..-...A.....AC.....T.......
     Box E                                 Box D
-100        -90        -80        -70          -60
TGCAGATTTCTAGGAATTCAAATCCACTATTGGTTTTATTT---CAAACC
 ...........................................---......
 .ATG..G....T......G...CTC.AA..C.A....T...TTT......
 ..TGA.C....T.......-..GGA-...T..-.AA.A--.CTTA.G....
 C.T..C....T......G..GGG-...T..T.AG.A--.CTTA......
         Box C
    -50        -40        -30        -20        -10
ACAAAATTAGCATGCCATTAAATACTATATATAAACAGCCACTAAATCAG
 ...................................A....A.......
 .............T........GGAG.......G..T..CCA...G...
 .....-........T......G..TGG......T....T...AG.G..T.
 .............T......G.G.AG......T..G.T...AG.G..T.
Box B                          Box A
ATCATT     Goat
......     Bovine  (Bonsing et al.,1988)
...GAAC    Rabbit  (Thepot et al.,1991)
..AGACC    Rat     (Jones et al.,1985)
..AGGAC    Mouse   (Yoshimura and Oka,1989)
```

Figure 1. Comparison of the nucleotide sequences of the 5' flanking regions (-200/-1 bp) in goat, bovine, rabbit, rat and mouse β-casein genes. Dots represented nucleotides identical to goat β-casein gene. Dashes were introduced to optimally align homologous sequences. The conserved regions were underlined.

region of the bovine αs_2-casein gene. The OCT-1 transcription factor can bind to this site of the bovine αs_2-casein gene (Groenen et al., 1992), thereby implying that OCT-1 may bind to this region of the β-casein gene as well. A previous study noted a mammary gland-specific nuclear factor (MGF) which specifically binds in boxes C and E of the rat β-casein gene (Schmitt-Ney et al., 1991). But subsequently, MGF was shown to be a STAT protein (signal transducers and activators of transcription) and designated STAT5 (Wakao et al., 1994). The STAT5 binding event is indispensable for hormonal message and for mediating the β-casein gene expression. The STAT5 binding activity is more efficient and important in box C. As indicated by aligning the region of casein genes, this region is highly conserved among casein genes (figure 2). From previous evidence that STAT5 also binds to the same site in the bovine αs_2-casein gene, we can infer that STAT5 may bind to the goat β-casein gene (Groenen et al., 1992). Several nuclear DNA binding proteins which repress β-casein gene expression have already been identified: the pregnancy-specific mammary nuclear factor (PMF), single-stranded DNA-binding transcription regulator (STR) and YY1, whose binding sites in the β-casein gene were located between -9 to +4 and -364 to -349 for PMF binding, -194 to -164 for STR binding, and -120 to -110 for YY1 binding (Altiok and Groner, 1993; Lee and Oka, 1992; Meier and Groner, 1994) respectively. PMF is as a mediator for progesterone represses β-casein gene expression during pregnancy. The DNA-binding activity of STR is down regulated during lactation and is sequestered by β-casein mRNA. Lactogenic hormone can induce a trans-acting activity to prevent STR binding DNA sequence. Mutation of the DNA binding site for YY1 led to a stronger lactation-associated activation complex from lactating mammary gland (Meier and Groner, 1994). On the other hand, mutation of the STAT5 binding site caused an increase of YY1 DNA-binding activity. There is a counteraction relationship between YY1 and STAT5. Comparing the sites of five β-casein genes reveals that the right half region of the YY1 binding site and the left half region of the STR binding site were more conserved between them and resided in box D and box F, respectively. PMF can transport the message of progesterone to effectively repress β-casein gene expression during pregnancy (Altiok and Groner, 1993). In addition, lactogenic hormone and MGF can inhibit STR and YY1 DNA-binding activity to the β-casein gene, respectively (Lee and Oka, 1992; Meier and Groner, 1994). The aforementioned studies suggest that the β-casein promoter and the transcription factors are involved in goat β-casein gene expression.

The β-casein gene could have additional regulatory sequences other than those reported for transcription factor binding sites.

Glucocorticoid receptor binding element

The sequence of the glucocorticoid receptor binding element (GRE) is usually GT(A/T) CANNNTGTC(C/T)CT (Beato, 1987), and the right half region is more conserved. Glucocorticoid can stimulate the rat and mouse β-casein gene expression (Doppler et al., 1989; Eisenstein and Rosen, 1988; Yoshimura and Oka, 1990a). Therefore, in this study, we investigated whether GRE exists in the goat β-casein gene. Similarity to the overall sequences of the GRE has not been found; however, many sites were found to be identical with the TGTC(C/T)CT sequence. Sequences smilar to the TGTTCT sequence in the goat β-casein gene were located at nt -871 to -866, -359 to -354, 3,358 to 3,363, 7,160 to 7,165 and 8,272 to 8,277. Concordance to the reverse and complementary sequences of the TGT(C/T)CT sequence, AGGACA were located at nt -3,141 to -3,136, 3,840 to 3,845 and 6,478 to 6,483, and AGAACA sequence located at nt 120 to 125, 554 to 559, 1,026 to 1,031, 2,342 to 2,347,

Gene	Species	Sequence	Reference
β-casein	Goat	ATTTCTAGGAATTCA	
	Bovine	----------	(Bonsing et al., 1988)
	Rabbit	--G--------T----------G--	(Thepot et al.,1991)
	Rat	--C--------T----------A--.	(Jones et al.,1985)
	Mouse	--C--------T----------G--	(Yoshimura and Oka, 1989)
α-casein	Rat	--A--------TA----------T--	(Yu-Lee et al.,1986)
αs1-casein	Bovine	--A--------TA----------T--	(Yu-Lee et al., 1986)
	Rabbit	--A--------TA----------T--	(Jolivet et al., 1992)
αs2-casein	Bovine	--C--------TA----------	(Groenen et al.,1993)
γ-casein	Rat	C----------TA----------G	(Yu-Lee et al.,1986)

Figure 2. Comparison of STAT5 binding sequences among goat β-, bovine β-, rabbit β-, rat β-, mouse β-, bovine αs1-, bovine αs2-, rabbit αs1-, rat α- and rat γ-casein gene promoter sequences. Dashes indicated nucleotides identical to goat β-casein gene.

3,962 to 3,967 and 5,699 to 5,704 (figure 3).

CAMP responsive element

Cyclic AMP (cAMP) induces somatostatin transcription through the CRE binding protein (CREB) binding to the cAMP responsive element (CRE) within the somatostatin gene followed by the phosphorylation of the CREB by cAMP-dependent protein kinase (Montminy and Bilezikjian, 1987). In addition, cAMP can effectively repress the lactogenic hormone induction of milk protein (Perry and Oka, 1980). Moreover, the sequences at nt -2,714 to-2,707 and -3944 to -3937 in the goat β-casein gene (figure 3) were identical with CRE, TGACG(C/T)(C/A) (G/A) (Montminy et al., 1986).

SV40 virus enhancer core sequence

CAAT-box/enhancer binding protein (C/EBP) does not only bind to the enhancer core sequence of the SV40 virus, hepatovirus and polyomavirus, but also can stimulate the transcription of these viral genes. However, excessive C/EBP, which occupy other positive transcription factor binding sites, repress the transcription (Dikstein et al., 1990; Pei and Shih, 1990). In addition to the conserved element of the 5' flanking region of five β-casein genes located between -62 to -55, the sequences at nt 3022 to 3029 in the goat β-casein gene (figure 3) were the same as the enhancer core sequence of GTGG(A/T)(A/T)(A/T)G (Weiher et al., 1983).

OCT-1 binding element

The common sequence of the OCT-1 binding element has been shown to be ATGCAAAT. This sequence presented both at nt -51 to -44 in five β-casein genes reverse and complementary, and at nt 3661 to 3668 in the goat β-casein gene (figure 3). OCT-1 and OCT-2 can bind to the OCT-1 binding element which has been found in any cell type and only B cell, respectively (Pierani et al., 1990). Moreover, OCT-1 can stimulate human H2b transcription

Element	Goatβ-casein gene	nucleotide location
Glucocoticoid receptor responsive element GT(A/T)CANNNTGT(C/T)CT CA(T/A)GTNNNACA(G/A)GA	5'-TGTTCT-3'	-871~866
	5'-TGTTCT-3'	-359~-354
	5'-TGTTCT-3'	3358~3363 (intron Ⅳ)
	5'-TGTTCT-3'	7160~7165 (intron Ⅶ)
	5'-TGTTCT-3'	8272~8277 (intron Ⅷ)
	5'-AGGACA-3'	-3141~-3136
	5'-AGGACA-3'	3840~3845 (intron Ⅳ)
	5'-AGGACA-3'	6478~6483 (exon Ⅶ)
	5'-AGAACA-3'	120~125 (intron Ⅰ)
	5'-AGAACA-3'	554~559 (intron Ⅰ)
	5'-AGAACA-3'	1026~1031 (intron Ⅰ)
	5'-AGAACA-3'	2342~2347 (intron Ⅱ)
	5'-AGAACA-3'	3962~3967 (intron Ⅳ)
	5'-AGAACA-3'	5699~5704 (intron Ⅵ)
cAMP responsive element(CRE) TGACG(C/T)(C/A)(G/A)	5'-TGACGTCA-3'	-2714~-2707
	5'-TGACGTCG-3'	-3944~-3937
SV40 virus enhancer core sequence (C/EBP binding site) GTGG(A/T)(A/T)(A/T)G CACC(T/A)(T/A)(T/A)C	5'-CAAACCAC-3'	-62~-55
	5'-GTGGAAAG-3'	3022~3029 (exon Ⅳ)
Octamer(OCT-1) binding element ATGCAAAT TACGTTTA	5'-ATTAGCAT-3'	-51~-44
	5'-ATGCAAAT-3'	3661~3668 (intron Ⅳ)
Homologies with the CTF/NF-1 binding element (T/C)GG(A/C)N_{5-6}GCCAA	5'-TGGGGACTGGGCCAA-3'	-255~-241

Figure 3. Regulatory sequences of the goat β-casein gene

(Fletcher et al., 1987) and OCT-2 enhances immunoglobin light and heavy chain transcription (Pierani et al., 1990).

CTF/NF-1 binding element

The CTF/NF-1 proteins found in any cell type mediate among many genes, such as by inducing both MMTV (Bruggemeier et al., 1990) and α-globin transcription and adenovirus replication (Santoro et al., 1988). The same element was found at nt -255 to -241 in the goat β-casein gene (figure 3).

CONCLUSION

Defining the regulatory sequences of the β-casein gene is useful for understanding the mechanisms of gene expression. We searched for putative regulatory sequences of the goat β-casein gene by aligning its sequence with those of the same gene from other mammals and comparing with previously reported transcription factor binding sites. Our results are perhaps valuable for the elucidation of the molecular mechanisms that control the expression of the goat β-casein gene.

REFERENCES

Altiok, S. and B. Groner. 1993. Interaction of two sequence-specific single-stranded DNA-binding proteins with an essential region of the β-casein gene promoter is regulated by lactogenic hormones. Mol. Cell. Biol. 13:7303-7310.

Ann, J. K., N. Aoki, T. Adachi, Y. Mizuno, R. Nakamura and T. Matsuda. 1995. Isolation and culture of bovine mammary epithelial cells and establishment of gene transfection conditions in the cells. Biosci. Biotech. Biochem. 59:59-64.

Beato, M. 1987. Induction of transcription by steroid hormones. Cell 56:335-344.

Bonsing, J., J. M. Ring, A. F. Stewart and A. G. Mackinlay. 1988. Complete nucleotide sequence of the bovine β-casein gene. Aust. J. Biol. Sci. 41:527-537.

Bruggemeier, U., L. Rogge , E. L. Winnaker and M. Beato. 1990. Nuclear factor 1 acts as a transcription factor on the MMTV promoter but competes with steroid hormone receptor for DNA binding. EMBO J. 9:2233-2239.

Chao, J. S. and M. C. Huang. 1997. Sequencing of the goat β-Casein genomic DNA. J. Agri. Assoc. China 179:1-14.

Dikstein, R., O. Faktor and Y. Shaul. 1990. Hierarchic and cooperative binding of the rat liver nuclear protein C/EBP at the hepatitis B virus enhancer. Mol. Cell. Biol. 10:4427- 4430.

Doppler, W., B. Groner and R. K. Ball. 1989. Prolactin and glucocorticoid hormones synergistically induce expression of transfected rat β-casein gene promoter constructs in a mammary epithelial cell line. Proc. Natl. Acad. Sci. USA 86:104-108.

Eisenstein, R. S. and J. M. Rosen. 1988. Both cell substratum regulation and hormonal regulation of milk protein gene expression are exerted primarily at the posttranscriptional level. Mol. Cell. Biol. 8:3183-3190.

Fletcher, C., N. Heintz and R. G. Roeder. 1987. Purification and characterization of OTF-1,a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell 51:773-781.

Groenen, M. A. M., R. J. M. Dijkhof, J. J. van der Poel, R. van Diggelen and E. Verstege. 1992. Multiple octamer binding sites in the promoter region of the bovine αs2-casein gene. Nucleic Acids Res. 20:4311-4318.

Groenen, M. A. M., R. J. M. Dijkhof, A. J. M. Verstege and J. J. van der Poel. 1993. The complete sequence of the gene encoding bovine αs2-casein. Gene. 123:187-193.

Haslam, 1988. Cell to cell interactions and normal mammary gland function. J. Dairy Sci. 71:2843-2854.

Huang, M. C. and C. S. Lin. 1996. Cloning and restriction mapping of goatβ-casein gene. J. Agri. Forest. 45:83-94.

Jenness, R. 1980. Composition and characteristics of goat milk: Review 1969-1979. J. Dairy Sci. 63:1605-1630.

Jolivet, G., E. Devinoy, M. L. Fontaine and L. M. Houdebine. 1992. Structure of the gene encoding rabbit αs1-casein. Gene. 113:257-262.

Jones, W. K., L.-Y. Yu-Lee, S. M. Clift, T. L. Brown and J. M. Rosen. 1985. The rat casein mutligene family: Fine structure and evolution of the β-casein gene. J. Biol. Chem. 260:7042-7050.

Kanai, A., N. Nonomura, M. Yoshimura and T. Oka. 1993. DNA-binding proteins and their cis-acting sites controlling hormonal induction of a mouse β-casein:CAT fusion protein in mammary epithelial cells. Gene. 126:195-201.

Lee, C. S. and T. Oka. 1992. A pregnancy-specific mammary nuclear factor involved in the repression of the mouse β-casein gene transcription by progesterone. J. Biol. Chem. 267:5797-5801.

Maniatis, T., E. Fritsch and J. Sanbrook. 1989. Molecular cloning: A laboratory manual. Cold Spring Harbor Lab., NY.

Meier V. S. and B. Groner. 1994. The nuclear factor YY1 participates in expression of the β-casein gene promoter in mammary epithelial cells and is conteracted by mammary gland factor during lactogenic hormone induction. Mol. Cell. Biol. 14:128-137.

Mitchell, P. J. and R. Tjian. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science 245:371-378.

Montminy, M. R., K. A. Sevarino, J. A. Wagner, Gl. Mandel and R. H. Goodman. 1986. Identification of a cyclic-AMP-responsive element within the rat somatostatin gene. Proc. Natl. Acad. Sci. USA 83:6682-6686.

Montminy, M. R. and L. M. Bilezikjian. 1987. Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature. 328:175-178.

Pei, D. and C. Shih. 1990. Transcriptional activation and repression by cellular DNA-binding protein C/EBP. J. Virol. 64:1517-1522.

Perry, J. W. and T. Oka. 1980. Cyclic AMP as a negative regulator of hormonally induced lactogenesis in mouse mammary gland organ culture. Proc. Natl. Acad. Sci. USA 77:2093- 2097.

Pierani, A., A. Heguy, H. Fujii and R. G. Roeder. 1990. Activation of octamer-containing promoters by either octamer-binding transcription factor 1 (OTC-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. Mol. Cell. Biol. 10:6204-6215.

Santoro, C., N. Mermod, P. C. Andrews and R. Tjian. 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication : cloning and expression of multiple cDNAs. Nature. 334:218-224.

Schmidhauser, C., M. J. Bissell, C. A. Myers and G. F. Casperson. 1990. Extracellular matrix and hormones transcriptionally regulate bovine β-casein 5' sequence in stably transfected mouse mammary cells. Proc. Natl. Acad. Sci. USA 87:9118-9122.

Schmitt-Ney. M., W. Doppler, R. K. Ball and B. Groner. 1991.β-casein gene promoter activity is regulated by the hormone mediated relief of transcriptional repression and a mammary-gland specific nuclear factor. Mol. Cell. Biol. 11:3745-3755.

Thepot, D., E. Devinoy, M. L. Fontaine and L. M. Houdebine. 1991. Structure of the gene encoding rabbitβ-casein. Gene. 97:301-306.

Wakao, H., F. Gouilleux and B. Groner. 1994. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. EMBO. J. 13:2182-2191.

Weiher, H., M. Konig and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. Sci. 219:626-631.

Yoshimura, M. and T. Oka. 1989. Isolation and structural analysis of the mouse β-casein gene. Gene. 78:267-275.

Yoshimura, M. and T. Oka. 1990a. Hormonal induction of β-casein gene expression : requirement of ongoing protein synthesis for transcription. Endocrinology. 126:427-433.

Yoshimura, M. and T. Oka. 1990b. Transfection of β-casein chimeric gene and hormonal induction of its expression in primary murine mammary epithelial cells. Proc. Natl. Acad. Sci. USA 87:3670-3674.