Structure, expression and engineering of milk protein genes

Lech Zwierzchowski
Institute of Genetics and Animal Breeding
Polish Academy of Sciences
Jastrzębiec

Milk protein genes are attractive models for studying tissue- and physiological stage-specific gene regulation and they also have potential applications in biotechnology. The expression of milk protein genes changes in different physiological stages of a mammal — in pregnancy and lactation, and is regulated by hormones, growth factors and extracellular matrix.

The main milk proteins are caseins. In ruminants milk, four casein species have been identified — caseins αS1, αS2, ß and κ. Milk of other mammals contains caseins which are analogous to those of cow's milk. However, concentrations of these caseins may differ substantially in different mammalian species. For example, only traces of α caseins have been found in human milk. Caseins exist in the milk in the form of colloidal structures — micelles. They may be precipitated from the milk either by lowering pH to 4.5 or by treatment with rennet - a proteolytic enzyme specificially digesting the κ-casein polypeptide. The soluble fraction of milk obtained after casein precipitation — so called whey — also contains milk proteins. Alpha-lactalbumin the ubiquitous milk protein, is a component of the lactose synthase complex. In ruminants milk, another major whey protein is β-lactoglobulin, while rodent and rabbit milk contains whey acidic protein (WAP) instead.

Casein genes of several mammalian species have been cloned and sequenced. The most offen studied β-casein gene has been cloned from seven species including man, ruminants and rodents (Tab. 1). The length of casein genes differs from almost 20 kb for goat αS1 to less than 7000 bp for mouse β-casein. The coding sequences of bovine αS1 and αS2 casein genes are distributed between 19 and 18 exons, respectively. Bovine κ-casein gene consists of 5 exons only, while all sequenced so far β-casein genes, except for the human gene, are composed of 9 exons (the human gene contains 8 exons). All known casein genes are single copy genes; no pseudogenes have been found.

TABLE 1
CLONED AND SEQUENCED CASEIN GENES

Gene	Mammal species	Sequence	Length of the gene	Reference
αS1-casein	rat	partial	app. 15 000 bp	(1)
αS1-casein	cattle	full	17 508 bp	(2)
αS1-casein	goat	partial	app. 19 500 bp	(3)
αS1-casein	rabbit	full	15 975 bp	(4)
αS2-casein	cattle	full	18 483 bp	(5)
γ-casein	rat	partial	app. 15 000 bp	(6)
β-casein	cattle	full	8498 bp	(7)
β-casein	rat	full	7234 bp	(8)
β-casein	sheep	almost full	8730 bp	(9)
β-casein	rabbit	full	9875 bp	(10)
β-casein	goat	partial	9000	(11)
β-casein	mouse	full	6794 bp	(12)
β-casein	man	full	10 466 bp	(13)
κ-casein	cattle	partial	13 070 bp	(14)

Genetic analysis has shown that the so called calcium-sensitive caseins αS1, αS2 and β (these caseins can be precipitated by calcium ions due to their high phosphate content) are linked together in the mammalian genome and they compose a small gene family. In the cattle genome, these casein genes form a gene cluster occupying approx. 200 kb on the chromosome no. 6 (15). It is believed that all calcium-sensitive casein genes arised from one ancestral gene during evolution, by processes of mutation, exon shuffling and duplication. Although different in length and exon/intron structure, all casein genes contain fragments of homologous sequences. This includes 5' upstream regulatory sequences as well as sequences coding for signal peptide and so called phosphorylation blocks.

Whey protein genes of several species have also been cloned and sequenced (Tab. 2). For example, the entire sequence of α-lactalbumin gene of six species has been published. Whey protein genes are shorter than casein genes and their length varies between 4723 bp for bovine β-lactoglobulin to approx. 2000 bp for most α-lactalbumin genes. Besides of functional whey protein gens also pseudogenes have been found. Five α-lactalbumin pseudogenes have been identified in bovine and ovine genomes (18).

TABLE 2
CLONED AND SEQUENCED WHEY PROTEIN GENES

Gene	Mammal species	Sequence	Length of the gene	Reference
α-lactalbumin	rat	full	2514 bp	(16)
α-lactalbumin	man	full	2363 bp	(17)
α-lactalbumin	cattle	full	2023 bp	(18)
α-lactalbumin	goat	full	2039 bp	(19)
α-lactalbumin	mouse	full	2300 bp	(20)
α-lactalbumin	guinea pig	full	2035 bp	(21)
β-lactoglobulin	sheep	full	4662 bp	(22)
β-lactoglobulin	cattle	full	4723 bp	GenBank Nr 14 710
WAP	rabbit	full	1724 bp	(23)
WAP	mouse	partial	3200 bp	(24)
WAP	rat	partial	2800 bp	(24)

During the last several years, at the Institute of Genetics and Animal Breeding, we have studied regulation of expression of milk protein genes in the rabbit mammary gland. Rabbit genes encoding caseins αS1 and β and WAP have been cloned and sequenced in INRA, Jouy-en-Josas by L-M. Houdebine and his coworkers (Tab. 1 and 2; Fig. 1). The longest of these genes, the αS1 casein gene, is almost 16 000 bp long and contains 19 exons, mostly very short. The β-casein gene contains 9875 bp and its coding sequences consist of 9 exons of the size ranging between 21 and 510 bp. WAP gene is only 1800-bp long and contains 4 exons of similar length.

Expression of milk protein genes in the rabbit mammary gland changes at various stages of pregnancy and in lactation. We showed that the expression of casein genes, as measured by accumulation of their mRNAs, begins at mid-pregnancy, increases approx. 10-times between days 15 and 25 of pregnancy and then increases again 3-fold at the beginning of lactation (day 5). The time frame of the expression of the WAP gene is slightly different than that of caseins. The expression of WAP begins later in pregnancy and its increase occurs mainly in lactation. This increase in the expression of milk protein genes is due, at least partially, to the increase in the gene transcription rate. The *in vitro* transcription of casein and WAP genes in the nuclei isolated from the mammary glands of rabbits is hardly detectable on day 15 of pregnancy but it increases manyfold during pregnancy and lactation (25). The transcription rates of the β-casein gene increase more than 10-fold between day 15 and 25 of pregnancy and then increase more than twice between days 25 of preganacy and day 5 of lactation. In pregnancy, transcription of the αS1 gene predominates, but in lactation transcription of β-casein gene exceedes that of αS1-casein.

Stabilization of mRNAs might be another factor promoting increased expression of appropriate genes. We compared the stability of β-casein mRNA in the explant cultures of the rabbit mammary gland. Explants derived from mammary glands of late pregnant rabbits (day 28) were cultured up to 40 hrs in the presence of actinomycin D (inhibitor of transcription) either with or without prolactin. It was shown that in the presence of prolactin the concentration of b-casein mRNA was still high after 40 hrs of culture, whereas in the absence of the hormone the ammount of β-casein mRNA decreased during the culture and the message completely disappeared after 32 hrs.

Early experiments carried out mostly on mammary explants of mice or rabbits established the minimal combination of hormones able to induce and support expression of milk protein genes (26). In the rabbit mammary gland, explants expression of casein genes is maximally induced in the presence of insulin, hydrocortisone and prolactin (27).

Prolactin is a member of the hormone family that also includes growth hormone (GH) and placental lactogens. These hormones regulate many physiological processes important for animal growth, differentiation and reproduction. It has been shown that some of placental lactogens demonstrate lactogenic activity similar to that prolactin and can be substitutes for prolactin in most *in vivo* and *in vitro* experiments with mammary gland. The *in vivo* lactogenic activity of growth hormone, another hormone of the family, has been recognized for a long time. It is well known that exogenous administration of bovine growth hormone, either natural or recombinant, increases milk production in lactating dairy cows. However, it is not clear whether this property of growth hormone results from direct stimulation of mammary gland to synthesize milk constituents or from a general anabolic effect. In most *in vitro* experiments, growth hormones (except human growth hormone) failed to induce mammary cell differentiation and casein synthesis.

We studied the effect of different growth hormones on the expression of milk protein genes — caseins αS1 and β and of WAP in the mammary gland explants derived from pregnant rabbits and found that growth hormones derived from bovine, ovine, porcine, rat and human pituitaries were able to induce expression of the milk protein genes in the mammary gland explants (28). At high hormone concentration the effect of some growth hormones was comparable to that of prolactin. However, the lowest effective concentration of bovine GH (bGH) was 0.1 μg/ml, and it exceeded 10-times that of prolactin which was effective at the concentration of 0.01 μg/ml. Preparations of recombinant bGH, obtained from three different sources — those produced by recombinant bacteria (gift of Monsanto or Lilly Research Company) and produced by transfected mammalian cells (gift of Dr. J. Kopchick of Ohio University) were also effective although their activity was lower than that of pituitary bGH. Different preparations of pituitary growth hormone were also active in inducing expression of casein polypeptides, thus indicating that both transcription and translation might be stimulated.

We assumed that the observed effect of GHs on the expression of milk protein genes is an intrinsic property of the GH molecule. This view was confirmed in several control experiments showing that GH preparation used in these experiments appear to be devoid of prolactin contamination as shown by SDS-polyacrylamide electrophoresis. Moreover, these GH preparations failed to stimulate proliferation of lymphoma Nb2 cells; the test detects prolactin at pg concentrations. The effect of bGH on the expression of milk protein genes was partially inhibited by bGH antagonist M8 — a product of mutated bGH gene (also a gift of Dr. J. Kopchick). And finally, small but reproducible effect of recombinant bGH was shown. Thus, we were able to show a direct effect of growth hormone on the expression of major milk protein genes in the rabbit mammary gland cultures. Similar results have also been reported recently at NIH by Barbara Vonderhaar and her colleagueas (29). They showed that bovine growth hormone induced both morphological differentiation and β-casein gene expression in cultures of the whole mammary glands of young virgin mice primed by estrogen and progesterone.

However, these experiments do not explain the possible mechanism of GH action. Most somatogenic actions of GH are believed to be mediated by insulin-like growth factors (IGFs). However, we were unable to show any effect of IGF-1 in this system. Human recombinant IGF-1, at concentrations ranging from 10 to 100 ng/ml, cannot substitute or modify the effect of GH on the expression of milk protein genes. Moreover, no specific GH receptors have been described in the mammary glands of mice or rabbits. We were also unable to detect binding sites for iodinated GH in the membranes isolated from mammary glands of pregnant or lactating rabbits. Receptors for prolactin are abundant in the rabbit mammary gland. However, the action of

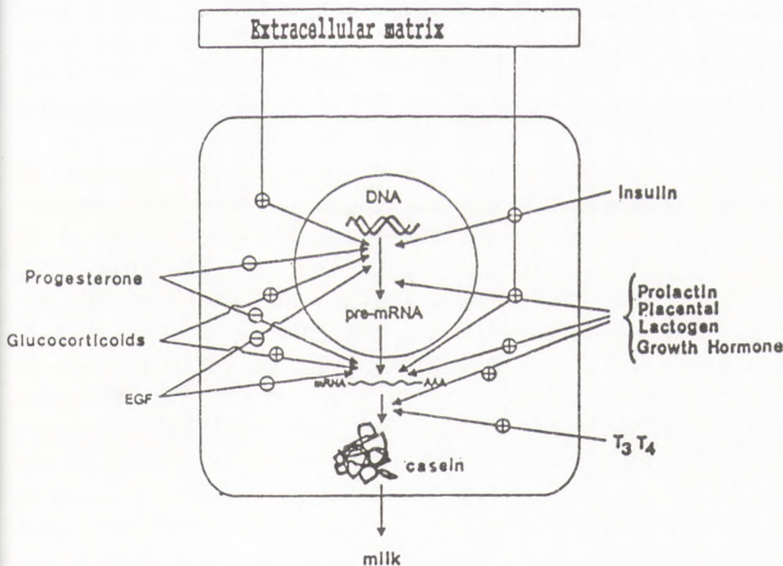

Fig.1. Schematic representation of the effects of hormones and growth factors on different stages of milk protein gene expression.

GH through prolactin receptors is unlikely since bovine, ovine or rat GH are unable to displace prolactin from its mammary receptors.

Taking together these and other results, the following scheme representing the influence of hormones and growth factors on the different stages of casein gene expression has been proposed (Fig. 2). Prolactin seems to be the most universal lactogenic hormone — it promotes all stages of casein gene expression — transcription, mRNA stabilization, translation, and post-translational modifications of casein polypeptides. In their actions on casein gene expression prolactin may be substituted by placental lactogens and, at least in some experiments, by growth hormone. Glucocorticoids (eg. cortisol) amplify while progesterone or EGF attenuates the lactogenic actions of prolactin. Insulin was shown to stimulate casein gene transcription only, while thyroid hormones were active in stimulating translation of some milk protein mRNAs, e.g. α-lactalbumin.

The molecular mechanism of action of any of the lactogenic hormones has not been fully elucidated. Steroid hormones, eg. glucocorticoids and progesterone, are believed to act through intracellular receptors that directly bind regulatory sequences in gene promoters. Specific regulatory sequences — GRE, PRE — have been identified in promoters of milk protein genes. Nevertheless, their functional significance has not been proved. Polypeptide hormones — insulin, prolactin, growth hormone — are known to bind cell surface receptors and their action on the gene expression assumes existence of intracellular mediator(s). Many experiments have been carried out in different laboratories, including my labotarory, to identify such mediators. Based on the results of these experiments, several intracellular transducers have been suggested to mediate hormone action on the expression of milk protein genes: inositol triphosphates, diacyloglycerol, calcium ions, protein kinases, polyamines and others. However, no definite conclusion can be drawn from these experiments. Indeed, tyrosine kinases, protein kinase C and Ca^{2+} seem to mediate or modulate the prolactin signal transmission from the membrane receptors to the genome. However, also other still unknown mediators are probably involved.

Transgenic animals may represent a convenient model for investigating the expression of milk protein genes in the mammary gland. The first transgenic mice expressing foreign protein in their mammary glands were produced by Simons end coworkers at AFRC Genetic Research Station, Edinburgh (30). These mice expressed ovine β-lactoglobulin, protein which is not normally expressed in rodents. After that, in many laboratories transgenic animals, mostly transgenic mice, were produced carrying heterologous milk protein genes or gene constructs (for references see 31). Transgenic mice carrying whole heterologous β-casein, WAP or α-lactalbumin genes have been generated. Moeover, transgenic pigs have been produced expressing mouse WAP, the protein which is not normally expressed in swine mammary gland. Several different promoters have been used to target heterologous gene to the mammary gland. The most used are mouse WAP promoter, bovine α-S1 casein promoter and ovine β-lac-

toglobulin promoter. In some of these animals expression of a transgene occured exclusively in the mammary gland, in others ectopic expression was also shown, eg. in salivary glands or in the brain. In spite of these difficulties, transgene technology enables to recognize both distal and proximal regulatory sequences in the milk protein genes. Moreover, such transgenic animals carrying structural genes of human proteins fused to mammary specific promoters are considered as living "bioreactors" designated to produce human protein of pharmaceutical use (32). A great progress has recently been done in this respect. Transgenic farm animals have been produced expressing human blood clotting factors, α-antitrypsin, plasminogen activator or lactoferrin. In some of these animals, expression of human proteins provides tens of mgs per ml of milk; enough to consider it a commercial use.

In order to investigate the targeted expression of a foreign gene in the mammary gland, we generated transgenic rabbits carrying the βCasLacZ gene construct (33). Some 2096-bp upstream sequences of the rabbit β-casein gene were fused to the 3.4-kb *E. coli* LacZ reporter gene and SV40 polyadenylation site. The integration of the transgene was detected by Southern-blot and PCR analysis. Transgenic rabbits transmitted the transgene to their progeny. However, our attempts to measure expression of the LacZ gene by northern-blot hybridization showed only some ectopic expression of the transgene in liver and spleen of transgenic rabbits; surprisingly, no expression of the LacZ gene was detected in the mammary glands of transgenic lactating or pregnant rabbits. The lack of expression was often reported in transgenic animals including those carring milk protein genes or gene constructs. The reason might be e.g. incorporation of the injected gene in the reverse orientation or its incorporation into inactive chromatin. Alternatively, it might have been due to transgene hypermethylation. Methylation plays an important role in the regulation of gene expression. In general, it is believed that high methylation attenuates gene expression while demethylation of variuos genes was shown to promote expression. We analysed methylation of LacZ gene in our transgenic rabbits using a pair of methylation-sensitive restriction nucleases — MspI and HpaII. This analysis showed heavy methylation of cytosine residues in the LacZ transgene in DNA isolated from different tissues including the mammary gland (33). Only cytosines at positions +1303 and +2638 seem to be unmethylated. These results suggest that methylation might inhibit the transgene expression in βCasLacZ transgenic rabbits. This view was further supported by *in vitro* experiments. Explants derived from pregnant transgenic rabbits were cultured in the presence of lactogenic hormones with or without 5'-azacytidine, a demethylating agent. The LacZ transgene expression, comparable to that of the endogenous β-casein gene, was detected in the mammary gland explants cultured for 2 days in the presence of insulin, hydrocortisone and prolactin with 5'-azacytidine at 10 or 25 μM concentration. Thus, our results show that the expression of a transgene bearing milk protein gene promoter might be regulated by its methylation and that in some experiments hypermethylation might block transgene expression.

Control of gene expression at the transcriptional level requires the sequence-specific interactions of *trans* activators or repressors with regulatory elements in gene promoters. One possible way to identify such regulatory sequences is a computer analysis. Milk protein gene promoters were analysed in search of sequences characteristic for this gene family and for consensus sequences potentially binding various transcription factors. Based on such computer analysis, different authors described 11 consensus sequences which can be found in all or nearly all milk protein genes. These sequences are listed in Tab. 3. They are named by names of the authors who described these sequences for the first time.

TABLE 3
HOMOLOGOUS SEQUENCES IDENTIFIED IN MILK PROTEIN GENES

Name	Sequence	Reference
milk box	RGAAGRAAANTGGACAGAAANTCAACGTTTCTA	(21)
Groenen structure	$G(N)_5TNT(N)_4TNNT(N)_6ANT(N)_{15}ANTNC(N)_4TTCCTGGACAC$ $ATTTCCTTT(N)_9TNANTNNNT(N)_5TNNNT(N)_4A(N)_4CNNNGA$ $ATNT(N)_4GGANANNANANT(N)_4AGAANA(N)_4TTTCNNAT(N)_7$ $ANTTCTTNGAATTNA(N)_{12}t(N)_{11}AAACCACANAATTAGCATNT$ $NA(N)_{13}TATAWAT$	(5)
Yu-Lee sequence	1. YYGTTKRAGA 2. MCYYAGAATYT 3. RGAASRAAWTVSAVAGAARHGAWTTTCYWAT 4. TTCTTAGAATT 5. RAAACCACARAATTAGCAT 6. RGTWTAWATAG	(1)
Oka box	box A CCCTAGAATTTCTGG box B TTCTTGAATTAA box C AAACCACAAAATTAGCATTTA	(12)

Symbols: A — adenine; C — cytosine; G — guanine; T — thymine; R — A or G (purine); Y — C or T (pyrimidine); S — G or C; W — A or T; K — G or T; M — A or C; B — C or T; D — A or C or T; V — A or C or G; N — A or C or G or T.

Using the HIBO DNASIS program and GenBank and EMBL databases we performed computer analysis of different milk protein gene promoters (34). The aims of this analysis were as follows:

1) to show which of the described milk-protein-gene-specific regulatory sequences are similar to others, and which are independent sequences,

2) to find out which of the known transcription factors can potentially bind the milk-protein-gene-specific regulatory sequences,

3) to identify the consensus sequences potentially binding the known transcription factors in the rabbit β-casein gene promoter.

Comparison of different motifs described for milk protein gene promoters

demonstrated that there exist 5 independent sequences which do not show significant homology to each other. These are: milk box, Groenen structure, Yu-Lee sequences 1 and 6, and Oka box A. Other motifs are either homologous to others or are included in longer sequences. Computer analysis of 2096 bp 5' upstream β-casein gene transcription initiation site showed that in the rabbit β-casein promoter only 4 out of 5 independent motifs typical of milk protein gene promoters occurred — Groenen structure, Yu-Lee 1 and 6 sequences and Oka box A. Surprisingly, no typical milk box was found. The absence of the typical milk-box motif may attenuate transcripion of the β-casein gene during pregnancy. In addition, our computer analysis revealed some one hundred and fifty putative binding motifs, representing 56 various consensus sequences. These sequences are located both inside and outside structures typical of milk protein gene promoters and include sequences homologous to mammary gland specific (MGF, MPBF, PMF and MAF), hormone specific (ARE, IRE, GRE/PRE) and ubiquitous transcription factors. Our analysis suggests that milk protein gene promoters can bind many transcription factors which assure their complex regulation during mammary gland differentiation.

Of course such computer analysis of milk protein gene promoters does not prejudge the involvement of any sequence in gene regulation. To do this, functional analysis must be performed including deletions or mutations of a putative regulatory sequence in a milk protein gene promoter fused to one of the reporter genes, e.g. LacZ, CAT or luciferaze gene, followed by measuring expression of the reporter gene in transfected cells or in transgenic animals. However, possibilities of this model are restricted by extreme difficulties in developing mammary gland epithelial cell lines that are consistently capable of expressing transfected milk protein genes. So far, two lines of mouse mammary cells have been used in transfection experiments — COMMAD and HC11, both expressing transfected rodent β-casein gene only. In addition, primary cultures of mammary epitelial cells have been used succesfully in some experimants. These experiments enabled identification of some mammary-specific transcription factors (Tab. 4) as well as some *cis*-regulatory sequences in milk protein gene promoters (for references see 31).

TABLE 4

TRANSCRIPTION FACTORS SPECIFIC FOR MAMMARY EPITHELIAL CELLS AND THEIR CONSENSUS SEQUENCES

Abbreviaton	Transcription factor	Consensus	Reference
MGF	Mammary Gland-specific Nuclear Factor	a) ACTTCTTGGAATT b) ANTTCTTGGNA	(35)
MAF	Mammary-Cell Activating Factor	GRRGSAAGK	(36)
MPBF	Milk Protein Binding Factor	GGTTCCNGGAACC	(37)
PMF	Pregnancy-specific Mammary Nuclear Factor	a) TGATATCA b) TGAAT$(N)_{47}$ATCA	(38)

The most intensively studied mammary gland specific factor is MGF, which was purified from mammary glands of different species. This factor was shown to be identical to Stat components of ISGF3 transcription factors which are signal transducers of interferons. As suggested by Groner and his colleagues (35), this factor is a mediator of prolactin action in mammary epithelial cells. Another mammary-specific transcripion factor — PMF — might be a mediator of the inhibitory action of progesterone on the expression of milk protein genes, as suggested by Lee and Oka (38).

The mammary-specific regulatory factors may also be characterized using DNA-protein binding experiments. Very few such experiments have been carried out with milk protein gene promoters. They were initiated in 1988 by Luboń and Hennighausen (39) who analysed binding of nuclear proteins to rat WAP and α-lactalbumin genes. However, in these early experiments neither proteins nor their binding sites were characterized in detail.

We also performed DNA-protein binding experiments to analyze regulation of β-casein gene in the rabbit mammary gland. The 2000-bp rabbit β-casein gene promoter and then its shorter fragments were cloned to analyze DNA-binding activity of nuclear proteins and to localize their binding sites using gel-shift technique. Our experiments showed binding of nuclear proteins to 204-bp fragment of β-casein gene promoter laying some 1200 bp upstream the transcription initiation site. According to our previous computer analysis, this promotor fragment contains putative MGF and PMF binding sites as well as several putative hormone-dependent binding sites. Gel-shift analysis showed binding of nuclear proteins to this fragment of β-casein promoter; extracts from cell nuclei isolated from mammary glands of late pregnant and lactating rabbits showed additional DNA-binding proteins as compared to day 15 of pregnancy. It may indicate that these proteins are functional in the regulation of β-casein gene expression in pregnancy and lactation. When a shorter DNA fragment (30 bp-long oligonucleotide), containing putative MGF binding site was used in gel-shift experiments, changes in binding activity during pregnancy and lactation were also shown. The MAF activity was only found in functionally differentiated mammary gland of late pregnant and early lactating rabbits. These results support the view of important regulatory function of MGF and MAF in milk protein gene regulation. On the other hand, only very small changes in PMF activity were shown in the rabbit mammary gland during pregnancy and lactation.

Engineering of milk protein genes is of great importance for biotechnology. Production of human therapeutic proteins in the mammary gland of transgenic animal has become a reality. Other great projects are in progress aiming to modify milk technological properties and the so called "humanization" of ruminants milk (Tab. 5). Despite the great progress, dificulties still exist, one of which being usually low expression of introduced genes in transgenic animals. That is why basic research on the regulation of milk protein genes may also have some further practical importance and may pay back in faster progress in animal biotechnology.

Fig. 2. Schematic representation of transcription factor binding sites in milk protein gene promoters.

Abbreviations: r, m, bβK — rat, mouse, bovine β-casein gene; bαS2K — bovine αS2-casein gene; m, rWAP — mouse or rat whey acidic protein gene; mαLA — mouse α-lactalbumin gene; oβLG — ovine β-lactoglobulin gene; GR — glucocorticoid receptor.

TABLE 5

POSSIBILITIES OF GENETIC ENGINEERING OF MILK PROTEIN GENES

<table>
<tr><th>Genetic manipulation</th><th>Expected effect</th><th>Expected benefits</th></tr>
<tr><td>Introduction of additional casein genes</td><td>Overproduction of different casein species in milk</td><td rowspan="4">— Increasing milk stability in high temperatures (eg. during milk sterilization).
— Improvement of quality of milk products (eg. cheeses, ice creams).
— Improvement of nutritious value of milk and milk products.
— Humanization of ruminant milk.
— Rendering ruminant milk consumable for individuals suffering from lactose intolerance or allergic to β-lactoglobulin</td></tr>
<tr><td>Targeted mutagenesis of casein genes</td><td>Alternation of amino-acid composition of caseins</td></tr>
<tr><td>Inactivation of α-lactalbumin gene. Introduction of bacterial β-galactosidase gene fused to milk protein gene promoter</td><td>Lowering of lactose content in milk</td></tr>
<tr><td>Inactivation of β-lactoglobulin gene</td><td>Production of β-lactoglobulin-free milk</td></tr>
<tr><td>Introduction of gene constructs containing milk protein gene promoters and human structural genes</td><td>Synthesis of heterologous proteins in the mammary gland of transgenic animals</td><td>Production of human therapeutic proteins in the milk of transgenic farm animals — goat, sheep or cow</td></tr>
</table>

Literature

1. Yu-Lee L. Y., Richter-Mann L., Couch C. H., Stewart A. F., Mackinlay A. G., Rosen J. M., (1986), *Evolution of the casein multigene family: conserved sequences in the 5' flanking and exon regions*, Nucl. Acid. Res., 14, 1883-1902.
2. Koczan D., Hobom G., Seyfert H. M., (1991), *Genomic organization of the bovine alpha-s1 casein gene*, Nucl. Acid. Res., 19, 5591-5596.
3. Leroux C., Mazure N., Martin P., (1992), *Mutations away from the splice site recognition sequences might cis-modulate alternative splicing of goat αS1-casein transcripts*, J. Biol. Chem., 267, 6147-6157.
4. Jolivet G., Devinoy E., Fontaine M. L., Houdebine L. M., (1992), *Structure of the gene encoding rabbit αs1-casein*, Gene, 113, 257-262.
5. Groenen M. A. M., Dijkhof R. J. M., Verstege A. J. M., van der Poel J. J., (1993), *The complete nucleotide sequence of the bovine αs2-casein gene*, Gene, 123, 187-193.
6. Yu-Lee L. Y., Rosen J. M., (1983), *The rat casein multigene family. I. Fine structure of the γ-casein gene*, J. Biol. Chem., 258, 10794-10804.
7. Bonsing J., Ring J. M., Stewart A. F., Mackinlay A. G., (1988), *Complete nucleotide sequence of the bovine β-casein gene*, Aust. J. Biol. Sci., 41, 527-537.
8. Jones W. K., Yu-Lee L. Y., Clift S. M., Brown T. L., Rosen J. M., (1985), *The rat casein multigene family. II. Fine structure and evolution of the beta-casein gene*, J. Biol. Chem., 260, 7042-7050.
9. Mercier J. C., Vilotte J. L., Provot C., (1990), *Structure and function of milk protein genes*, Ed. H. Gelderman, VCH, Weinheim, Germany.
10. Thepot D., Devinoy E., Fontaine M. L., Houdebine L. M., (1991), *Structure of the gene encoding rabbit β-casein*, Gene, 97, 301-306.

11. Roberts B., DiTullio P., Vitale J., Hehir K., Gordon K., (1992), *Cloning of the goat beta-casein encoding gene and expression in transgenic mice*, Gene, 121, 255-262.
12. Yoshimura M., Oka T., (1989), *Isolation and structural analysis of the mouse β-casein gene*, Gene, 78, 267-275.
13. Hansson L., Edlung M., Edlung A., Johansson T., Marklund S. L., Fromm S., Strömqvist M., Törnell J., (1994), *Expression and characterization of biologically active human extracellular superoxide dismutase in milk of transgenic mice*, J. Biol. Chem., 269, 5358-5363.
14. Alexander L. J., Stewart A. F., Mackinlay A. G., Kapelinskaya T. V., Tkach T. M., Gorodetsky S. I., (1988), *Isolation and characterization of the bovine κ-casein gene*, Eur. J. Biochem., 178, 395-401.
15. Threadgill D. W., Womack J. E., (1990), *Genomic analysis of the major bovine milk protein genes*, Nucl. Acid. Res., 18, 6935-6942.
16. Quasba P. K., Safaya S. K., (1984), *Similarity of the nucleotide sequences of rat α-lactalbumin and chicken lysozyme genes*, Nature, 608, 377-380.
17. Hall L., Emery D. C., Davies M. S., Parker D., Craig R. K., (1987), *Organization and sequence of the human α-lactalbumin gene*, Biochem. J., 242, 735-742.
18. Vilotte J. L., Soulier S., Mercier J. C., Gaye P., Hue-Delahaie D., Furet J. P., (1987), *Complete nucleotide sequence of bovine α-lactalbumin gene. Comparison with its rat counterpart*, Biochimie, 69, 609-620.
19. Vilotte J. L., Soulier S., Printz C., Mercier J. C., (1991), *Sequence of the goat α-lactalbumin-encoding gene: comparison with the bovine gene and evidence of related sequences in the goat genome*, Gene, 98, 271-276.
20. Vilotte J. L., Soulier S., (1992), *Isolation and characterization of the mouse α-lactalbumin encoding gene — interspecies comparision, tissue — specific and stage-specific expression*, Gene, 119, 287-292.
21. Laird J. E., Jack L., Hall L., Boulton A., Parker D., Craig R. K., (1988), *Structure and expression of the guinea-pig α-lactalbumin gene*, Biochem. J., 254, 85-94.
22. Ali S., Clark J., (1988), *Characterization of the gene encoding ovine β-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins*, J. Mol. Biol., 199, 415-426.
23. Thepot D., Devinoy E., Fontaine M. L., Hubert C., Houdebine L. M., (1990), *Complete sequence of the rabbit whey acidic protein gene*, Nucl. Acid. Res., 18, 3641.
24. Campbell S. M., Rosen J. M., Hennighausen L. G., Strech-Jurk U., Sippel A. E., (1984), *Comparison of the whey acidic protein genes of the rat and mouse*, Nucl. Acid. Res., 12, 8685-8697.
25. Zwierzchowski L., Sokół-Misiak W., Devinoy E., Houdebine L. M., Niedbalski W., Kleczkowska D., (1995), *DNA replication, β-casein gene transcription and DNase-I sensitivity of β-casein gene in the rabbit mammary gland nuclei*, in: *Intracellular Signalling in the Mammary Gland*, Eds. C. J. Wilde, M. Peaker, C. H. Knight, Plenum Press, 191-192.
26. Vonderhaar B. K., Ziska S. E., (1989), *Hormonal regulation of milk protein gene expression*, Annu. Rev. Physiol., 51, 641-652.
27. Devinoy E., Hubert C., Jolivet G., Thepot D., Clergue N., Desaleux M., Dion M., Servely J. L., Houdebine L. M., (1988), *Recent data on the structure of rabbit milk protein genes and on the mechanism of the hormonal control of their expression*, Reprod. Nutr. Develop., 28, 1145-1164.
28. Zwierzchowski L., Siadkowska E., Żebrowska T., (1993), *Effect of growth hormone on the expression of β-casein and whey acidic protein genes in rabbit mammary gland explants*, Anim. Sci. Pap. Rep., 11, 167-173.
29. Plaut K., Ikeda M., Vonderhaar B. K., (1993), *Role of growth hormone and insulin-like growth factor-I in mammary development*, Endocrinology, 133, 1843-1848.
30. Simons J. P., McClenaghan M., Clark A. J., (1987), *Alternation of the quality of milk by expression of sheep β-lactoglobulin in transgenic mice*, Nature, 328, 530-532.

31. Groenen M. A. M., van der Poel J., (1994), *Regulation of expression of milk protein genes*, Livestock Production Sci., 38, 61-78.
32. Clark A. J., Simons P., Wilmut I., Lathe R., (1987), *Pharmaceuticals from transgenic livestock*, TIBTECH, 5, 20-24.
33. Raczkowska M., Zwierzchowski L., Żebrowska T., Malewski T. and Went D. F., (1995), *Methylation of DNA and βCasLacZ gene expression in transgenic rabbits*, Beltsville Symp. XX. Biotechnology's Role in the Genetic Improvement of Farm Animals, abstr. P47.
34. Malewski T., Zwierzchowski L., (1995), *Computer-aided analysis of potential transcription-factor binding sites in the rabbit β-casein gene promoter*, BioSystems, 36, 109-119.
35. Wakao H., Goulieux F, Groner B., (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.
36. Mink S., Hartig E., Jennewein P., Doppler W., Cato A. C. B., (1992), *A mammary cell-specific enhancer in mouse mammary tumor virus DNA is composed of multiple regulatory elements including binding sites for CTF/NFI and a novel transcription factor, mammary cell-activating factor*, Mol. Cell. Biol., 12, 4906-4918.
37. Watson C. J., Gordon K. E., Robertson M., Clark A. J., (1991), *Interaction of DNA-binding proteins with a milk protein gene promoter in vitro: identification of a mammary gland-specific factor*, Nucl. Acid. Res., 19, 6603-6610.
38. Lee C. S., Oka T., (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.
39. Luboń H., Hennighausen L., (1988), *Conserved region of the rat α-lactalbumin promoter is a target site for protein binding in vitro*, Biochem. J., 256, 391-396.

Structure, expression and engineering of milk protein genes

Summary

Expression of milk protein genes is regulated by hormones, growth factors and extracellular matrix. Prolactin, the major lactogenic hormone, promotes all stages of casein gene expression. Besides prolactin, also growth hormone may directly induce expression of milk protein genes. Computer analysis, mutation experiments and DNA-protein binding experiments enabled identification of mammary-specific transcription factors and *cis*-regulatory sequences in milk protein gene promoters. Also, transgene technology enables recognition of regulatory sequences in milk protein genes. Moreover, transgenic animals carrying structural genes of human proteins fused to mammary specific promoters are considered as living "bioreactors" designated to produce human proteins for pharmaceutical use.

Key words:
milk, proteins, gene expression.

Address for correspondence:
Lech Zwierzchowski, Institute of Genetics and Animal Breeding, Polish Academy of Sciences, Jastrzębiec, 05-551 Mroków.