NOTE TO USERS

This reproduction is the best copy available.
DNA methylation of two milk protein genes in lactating and non-lactating bovine mammary gland tissues

by

XIAOLIANG WANG

Department of Animal Science
McGill University, Montreal, Quebec, Canada

May, 2008

A thesis submitted to the Graduate and Postdoctoral office in partial fulfillment of the requirements of the degree of

Master of Science

©Xiaoliang Wang, 2008
NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.
ABSTRACT

Master of Science Animal Science

Xiaoliang Wang

DNA methylation of two milk protein genes in lactating and non-lactating bovine mammary gland tissues

It is well known that DNA methylation in gene promoter regions inhibits gene transcription and that tissue-specific gene expression is partially under the control of this transcription regulatory mechanism. In this study, bovine mammary gland tissues were collected from individual animals in lactating and non-lactating stages to investigate the DNA methylation patterns in the κ-casein gene and α-lactalbumin gene core promoter regions using the bisulphite treatment in combination with polymerase chain reaction (PCR) sequencing. Different methylation status of each sample was classified into three categories, namely methylation at known transcription factor binding domains, methylation at core promoter non-binding domains and the absence of cytosine methylation. Real-time quantitative PCR was used to quantify the transcription levels of the κ-casein and α-lactalbumin genes from the collected samples. A comparative method was used and fold-change values were calculated based on the comparison of the normalized threshold values of samples from different physiological stages as well as on various methylation patterns observed in their core promoter regions. Statistical analyses showed that the expressions of the κ-casein and α-lactalbumin genes were significantly different in lactating and non-lactating mammary gland tissues. The methylation observed in the core promoter region of bovine α-lactalbumin gene was found to be associated with its gene expression. On the other hand, the methylation found in the core promoter region of bovine κ-casein gene did not have any effect on its gene transcript levels.
RESUME

M.Sc. Zootechnie

Xiaoliang Wang

Méthylation d'ADN de deux gènes de protéine du lait dans les tissus de la glande mammaires chez les bovins en lactation et non-produisant du lait

Il est reconnu que la méthylation d'ADN dans des régions d'instigateur des gènes empêche la transcription de ces gènes et que l'expression des gènes qui sont spécifique à certains tissus est partiellement sous contrôle de ce mécanisme de transcription. Dans cette étude, des tissus de la glande mammaires ont été rassemblés de différents bovins en lactation et non-produisant du lait pour étudier les modèles de méthylation d'ADN chez le gène de κ-caséine et dans les régions d'instigateur du gène de α-lactalbumine en utilisant un traitement de bisulfite en combinaison avec la réaction en chaîne de polymérase (PCR). Différent phases de méthylation de chaque échantillon a été identifié et classifié dans la catégorie méthylé ou non- méthylé. La méthode real-time quantitatif du PCR a été utilisée pour mesurer la transcription des gènes κ-caséine et de α-lactalbumine provenant des différents échantillons. Une méthode comparative a été utilisée et les différences numériques ont été calculées en basant sur la comparaison des valeurs normalisées. Les échantillons analysés provenaient des bovins en différents stages physiologiques aussi bien sur leurs divers cycles de méthylation observés dans les régions d'instigateur. Les analyses statistiques démontrent que l'expression des gènes de κ-caséine et de α-lactalbumine étaient différents dans les tissus de glande mammaires provenant des bovins en différentes étapes physiologiques. La méthylation observée dans la région d'instigateur...
du gène de α-lactalbumine a sérieusement affecté l’expression de ce gène. D’autre part, la
méthylation identifiée dans la région d'instigateur du gène de κ-caséine n'a eu aucun effet
sur la transcription de ce même gène.
ACKNOWLEDGMENTS

I would sincerely like to thank my supervisor, Dr. Xin Zhao, for giving me the opportunity to complete my Master degree at McGill. I am also very grateful for his scientific advice, useful guidance, patience, understanding, and support especially during stressful times of my study.

I would also like to thank Dr. Anthony V. Capuco from USDA for his collaboration on my research project in supplying bovine mammary gland tissues and always answering my questions concerning the samples. Also, I would like to acknowledge my committee members: Dr. Vilceu Bordignon and Dr. Arif F. Mustafa.

No doubt that Dr. Tarun Kumar Bhattacharjya is another important person whom I would like to acknowledge. I am extremely grateful to Tarun for his patience, technical advice and guidance, and data analyses. His perseverance, optimistic view of life and encouragements, were very often my driving energy for the difficult parts of my laboratory analysis.

Many thanks also go to my laboratory colleagues, namely Eveline, Aloy, Patrick, Shyam and Kamal, for their support and help. Especially I want to appreciate Shyam for his great help with the translation of my abstract and grammatical correction of my thesis. I would also like to thank Yaohua and Ellen, my best friends at McGill, who helped me a lot with my presentation. Without forgetting, I would also like to thank Bear, Jiaozi, Candy, Patty and Ivy, my best friends from China, who have always been very supportive.

Mr. Jianxiong Ye is another important person whom I owe my deepest thanks for his considerateness, patience and kindness.
Last not least, my deepest thanks and gratefulness go to my parents and families for their unconditional love and constant support. Those who have been herein acknowledged are the angels of my life.
TABLE OF CONTENTS

Abstract......................................................................................................................I
Resume........................................................................................................................II
Acknowledgments.....................................................................................................IV
Table of contents.......................................................................................................VI
List of tables...............................................................................................................X
List of figures.............................................................................................................XI
Abbreviations...........................................................................................................XIII

Chapter I

General Introduction

1.1 Introduction.......................................................................................................1
1.2 Hypotheses.........................................................................................................4
1.3 Objectives..........................................................................................................5

Chapter II

Review of literatures

2.1 Role of DNA methylation in the regulation of gene expression.....................6
2.1.1 Formation, distribution, and inheritance of DNA methylation..................6
2.1.2 Methods used to study DNA methylation.................................................10
2.1.3 The biological function of DNA methylation...........................................14
2.1.3.1 Relationship between DNA methylation and transcriptional activity..............................................................................14
2.1.3.2 Bilateral relation between DNA methylation and chromatin

2.1.4 Changes in DNA methylation during development

2.2 Mechanism of DNA demethylation

2.2.1 Passive and active DNA demethylation

2.2.2 Effects of demethylating reagents on the activation of silenced genes

2.3 Bovine milk protein genes

2.3.1 Bovine milk proteins

2.3.2 Milk protein synthesis

2.3.3 The role of lactogenic hormones in regulation of milk protein gene expression

2.3.4 DNA methylation and milk protein gene expression

2.3.5 Effect of 5' regulatory region methylation on gene expression

2.3.6 The core promoter of a gene

2.3.6.1 The target region in the core promoter of bovine κ-casein gene

2.3.6.2 The target region in the core promoter of bovine α-lactalalbumin gene

2.4 Quantitation of gene expression by real-time quantitative PCR (qPCR)

2.4.1 Real time quantitative reverse transcription PCR (qRT-PCR)

2.4.2 Data analysis in qPCR: absolute and relative quantification

2.4.3 Housekeeping genes in qRT-PCR

VII
Chapter III

Materials and Methods

3.1 Study of DNA methylation in bovine κ-casein and α-lactalbumin gene core promoter regions using bisulphite sequencing..........................39
  3.1.1 Genomic DNA preparation of bovine mammary gland tissues..............39
  3.1.2 Bisulphite conversion of the isolated DNA....................................40
  3.1.3 Design and modification of the primers in bisulphite PCR..................40
  3.1.4 PCR amplification of the bisulphite treated DNA.............................41
  3.1.5 Purification of the bisulphite PCR products...................................42
  3.1.6 Cloning and sequencing of the bisulphite PCR products.....................42

3.2 Quantification of κ-casein and α-lactalbumin gene expression in lactating and non-lactating bovine mammary gland tissue samples.........................45
  3.2.1 RNA isolation of bovine mammary gland tissue and complementary DNA (cDNA) synthesis.................................................................45
  3.2.2 Primer design and reverse transcription PCR (RT-PCR) of the target genes.........................................................................................46
  3.2.3 Quantification of target genes by qRT-PCR........................................47

3.3 Statistical analysis..................................................................................49

Chapter IV

Results

4.1 DNA methylation in the core promoter regions of bovine κ-casein and α-lactalbumin genes.................................................................50
  4.1.1 Verification of the bisulphite PCR products........................................50
4.1.2 Cloning ................................................................. 52
4.1.3 Methylated cytosines found in target bovine κ-casein gene core promoter region ................................................................. 55
4.1.4 Methylated cytosines found in target bovine α-lactalbumin gene core promoter region ................................................................. 57

4.2 Quantification of κ-casein and α-lactalbumin gene expressions in the lactating and non-lactating bovine mammary gland tissues ................................................................. 59
4.2.1 RT-PCR results of the target genes and the endogenous reference gene ................................................................. 59
4.2.2 Validation experiments in the qPCR ................................................................. 60
4.2.3 Quantification of the target genes’ expression ................................................................. 62
4.2.4 Melting curve analysis in the qPCR ................................................................. 68
4.2.5 Effect of physiological stage on the target genes’ expression ................................................................. 72
4.2.6 Effect of DNA methylation on the target genes’ expression ................................................................. 76

Chapter V
Discussion ................................................................. 80

Chapter VI
Concluding comments ................................................................. 87

References ................................................................. 88
LIST OF TABLES

Table 1. Primers used in the bisulphite PCR for amplification of target bovine κ-casein and α-lactalbumin gene core promoter regions.................................41

Table 2. Primers used in RT-PCR for amplification of bovine κ-casein and α-lactalbumin cDNA.................................................................47

Table 3. Positions of methylated cytosine found in target bovine κ-casein gene core promoter region of different samples...........................................56

Table 4. Positions of the methylated cytosine found in target bovine α-lactalbumin gene core promoter region of different samples........................................58

Table 5. Threshold cycle values of GAPDH gene and bovine κ-casein gene.............67

Table 6. Threshold cycle values of GAPDH gene and bovine α-lactalbumin gene.......67

Table 7. Fold-change value of bovine κ-casein gene expression in lactating cows relative to dry cows.................................................................73

Table 8. Fold-change value of bovine α-lactalbumin gene expression in lactating cows relative to dry cows.................................................................73

Table 9. Fold-change values of bovine κ-casein gene expression in different groups of animals with differential methylation status in the target core promoter region........................................................................77

Table 10. Fold-change values of bovine α-lactalbumin gene expression in different groups of animals with differential methylation status in the target core promoter region........................................................................77
LIST OF FIGURES

Figure 1. DNA methylation reactions.................................................................7

Figure 2. The classic model of maintenance of DNA methylation patterns in
somatic cells...........................................................................................................9

Figure 3. DNA methylation pattern is in a dynamic steady-state balance of DNA
methylation and demethylation.......................................................................9

Figure 4. Bisulphite conversion in combination with PCR amplification and
sequencing...........................................................................................................12

Figure 5. Possible modes of action of the demethylases.................................22

Figure 6. Known core promoter elements.......................................................32

Figure 7. Target region of bovine κ-casein gene core promoter with known transcription
factor binding domains ..................................................................................33

Figure 8. Target region of bovine α-lactalbumin gene core promoter with known
transcription factor binding domains ..............................................................34

Figure 9. PCR products of target bovine κ-casein and α-lactalbumin gene core promoter
regions after bisulphite conversion of the genomic DNA ..............................51

Figure 10. Colony PCR products of target bovine κ-casein and α-lactalbumin gene core
promoter regions .............................................................................................53

Figure 11. Plasmid PCR products of target bovine κ-casein and α-lactalbumin gene core
promoter regions .............................................................................................54

Figure 12. RT-PCR products of bovine κ-casein, α-lactalbumin and GAPDH genes......60

Figure 13. Validation experiment....................................................................61
Figure 14. qPCR amplification efficiencies of bovine κ-casein, α-lactalbumin and GAPDH genes.

Figure 15. Amplification plots of bovine κ-casein, α-lactalbumin and GAPDH genes.

Figure 16. Dissociation analysis of bovine κ-casein, α-lactalbumin and GAPDH genes.

Figure 17. Effect of physiological stage on the expressions of target genes.

Figure 18. Trends of expression of bovine κ-casein and α-lactalbumin genes with differential methylation status of the target core promoter regions.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>aprt</td>
<td>hamster adenine phosphoribosyltransferase</td>
</tr>
<tr>
<td>BRE</td>
<td>TFIIB recognition element</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>COBRA</td>
<td>combined bisulphite restriction analysis</td>
</tr>
<tr>
<td>Ct</td>
<td>threshold cycle number</td>
</tr>
<tr>
<td>DNMTs</td>
<td>DNA methyltransferases</td>
</tr>
<tr>
<td>DPE</td>
<td>downstream core promoter element</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehydes 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>HMTASE</td>
<td>histone methyltransferase</td>
</tr>
<tr>
<td>Inr</td>
<td>initiator</td>
</tr>
<tr>
<td>MBD2</td>
<td>methylation DNA binding protein 2</td>
</tr>
<tr>
<td>MeCP2</td>
<td>methylCpG-binding proteins</td>
</tr>
<tr>
<td>MTF</td>
<td>motif ten element</td>
</tr>
<tr>
<td>Ms-SNuPE</td>
<td>methylation-sensitive single nucleotide primer extension</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PGK1</td>
<td>phosphoglycerate kinase gene</td>
</tr>
<tr>
<td>PIC</td>
<td>pre-initiation transcription complex</td>
</tr>
<tr>
<td>QRT-PCR</td>
<td>quantitative real-time reverse transcription-polymerase chain reaction</td>
</tr>
</tbody>
</table>
RT-PCR reverses transcription-polymerase chain reaction
TBP TATA-box-binding protein
TF transcription factor
TFIIB transcription factor IIB
TFIID transcription factor IID
TK herpes thymidine kinase
TR trans-acting repressor
Chapter I

General Introduction

1.1 Introduction

The coding potential of the genome lies within the arrangement of four bases, namely adenine, cytosine, guanine and thymine. However, the so-called “minor” methylated bases can occur in DNA which provides additional information affecting phenotype but without changing the base sequences (Laird, 2003; Vanyushin, 2005). In mammalian cells, methylation plays a key role in mammalian development, X chromosome inactivation, gene imprinting and tumourgenesis (Jones and Takai, 2001; Laird, 2003). DNA methylation is an enzymatic reaction of which the main product is the 5-methylcytosine in CG dinucleotide sequence (Szyf, 2005; Vanyushin, 2005). CpG islands most frequently found in promoters are generally not methylated (Jones and Takai, 2001). Methylation observed in a promoter region usually results in the inactivation of gene expression. There are two mechanisms responsible for generation and maintenance of methylation, de novo methylation and maintenance methylation (Laird, 2003; Szyf, 2005). DNA methylation pattern is in a dynamic steady state of methylation and demethylation reactions catalyzed by DNA methyltransferases and demethylases, respectively (Szyf, 2005). In the past decade, DNA methylation analysis has been revolutionized by different technologies (Rein et al., 1998; Havliš and Trbušek, 2002).

Bovine milk is an important dietary source of nutrients, such as energy, high quality protein, and a variety of vitamins and minerals. Milk protein, one of most important milk components, is mainly composed of caseins, α-lactalbumin, and β-lactoglobulin. They are synthesized in the mammary secretory cells and make up about
90%-95% of the protein of bovine milk protein. Caseins are the major milk proteins found in milk only and constitute about 80% of total cow’s milk protein. The principal casein fractions are $\alpha_{s1}$ and $\alpha_{s2}$-caseins, $\beta$-caseins, and $\kappa$-caseins. Caseins are present as micelles in milk and provide amino acids as well as calcium and phosphorus to newborn calves for their proper development (Akers, 2002). The $\alpha_{s1}$-, $\alpha_{s2}$-caseins and $\beta$-caseins are referred to calcium-sensitive-caseins due to their precipitations in the presence of calcium ion. On the other hand, $\kappa$-casein is insensitive to the calcium while it plays a crucial role in stabilizing casein micelles. The protein $\alpha$-lactalbumin is found in milk of many species and is involved in the synthesis of lactose which is responsible for more than one-third of the osmotic pressure of normal milk and involved in monitoring the rate of milk secretion (Akers, 2002).

All bovine milk casein genes have been sequenced and different functional segments have been characterized [$\alpha_{s1}$ (Koczán et al., 1991); $\beta$ (Gorodetsky et al., 1988); $\alpha_{s2}$ (Groenen et al., 1993); $\kappa$ (Alexander et al., 1988)]. The casein genes are clustered on chromosome 6 with a locus order of $\alpha_{s1}$-$\beta$-$\alpha_{s2}$-$\kappa$ (Ferretti et al., 1990; Threadgill and Womack, 1990) and show similarity in the overall gene organization between species (Mercier and Vilotte, 1993). Specifically the three calcium-sensitive-caseins share common regulatory sequence in the promoter region (Mercier and Vilotte, 1993). The $\kappa$-casein has a different organization in the 5’ flanking region and is not evolutionarily related to other three casein genes (Alexander et al., 1988). On the other hand, $\alpha$-lactalbumin gene is segregated on another chromosome and has different regulatory region in the 5’ end from those in the casein genes (Threadgill and Womack, 1990). The expression of milk protein genes is mammary tissue- and stage-specific (Rijnkels et al., 1997).
Methylation study has been done on rat casein genes in different tissues including lactating mammary gland, mammary tumors and liver and the result showed that methylation occurred at CpG sites towards the 5' end of the casein genes originated from mammary tumors and liver (Johnson et al., 1981). The expression of rat κ-casein gene was also demonstrated to be inversely related to the methylation found in cDNA of this gene in non-expression tissues (Thompson and Nakhasi, 1985). Methylation patterns were also investigated on α-lactalbumin gene isolated from rat mammary tissues at different developmental stages and from rat mammary tumors. It was reported that the gene expression was inversely related to the occurrence of certain methylated bases found in the sequence (Qasba et al., 1982).

To our best knowledge, the methylation patterns in the promoter regions of bovine milk protein genes are not well studied. Therefore, the methylation status of milk protein gene promoter regions was examined in this study using bisulphite sequencing for bovine lactating and non-lactating mammary gland tissues. The expressions of two milk protein genes, the κ-casein and α-lactalbumin, were quantified using the real-time quantitative PCR and comparisons were made between lactating and non-lactating samples. The relationship between the milk protein gene promoter methylation and gene expression was analyzed.
1.2 Hypothesis

The hypotheses of the present study are as follows:

I. Methylation patterns in the core promoter regions of the \(\kappa\)-casein and \(\alpha\)-lactalbumin gene are different in bovine mammary gland tissues collected from lactating and non-lactating cows.

II. The expressions of \(\kappa\)-casein and \(\alpha\)-lactalbumin genes are different due to different methylation levels in bovine lactating and non-lactating mammary gland tissues.

III. The expressions of bovine \(\kappa\)-casein and \(\alpha\)-lactalbumin genes are inversely related to the DNA methylation in their core promoters.
1.3 Research objectives

I. To study the DNA methylation pattern in the core promoters of κ-casein and α-lactalbumin genes in lactating and non-lactating bovine mammary gland tissues.

II. To quantify the expressions of κ-casein and α-lactalbumin genes in lactating and non-lactating bovine mammary gland tissues.

III. To correlate the DNA methylation status with the expressions of κ-casein and α-lactalbumin genes in lactating and non-lactating bovine mammary gland tissues.
Chapter II

Literature Review

2.1 Role of DNA methylation in the regulation of gene expression

2.1.1 Distribution, formation, and inheritance of DNA methylation

In vertebrates, especially in mammals, almost all methylations are found at CpG dinucleotides (Jones and Takai, 2001). About 60 to 90% of CpG dinucleotides contain 5-methylcytosine in mammalian genome (Tweedie et al., 1997; Laird, 2003). Clusters of CpGs, called CpG islands, are first described as an unmethylated HpaII tiny fragment (HTF) by Bird (1986) and are often located near the 5' ends of genes (Laird, 2003). These islands are approximately of 500-base-pair window with a G: C content of minimum 55% and an observed CpG frequency of at least 0.65 (Laird, 2003). Usually they are unmethylated in normal cells, with the exception of those that are associated with imprinting genes and genes on the inactive X chromosome (Jones and Takai, 2001; Laird, 2003). Methylation of promoter CpG islands is associated with gene transcriptional silencing (Laird, 2003).

Methylation is an enzymatic reaction which occurs after DNA synthesis. DNA methyltransferases (DNMTs) catalyze the transfer of a methyl group from the methyl donor S-adenosylmethionine to the carbon-5 position of cytosine (Laird, 2003; Szyf, 2005). Two kinds of DNMTs have been proposed and they participate respectively in two differential mechanisms responsible for the generation and maintenance of DNA methylation pattern. First, DNMT1, the candidate maintenance DNMT, can copy a methylation pattern from a methylated parental strand template to an unmethylated
daughter strand (Razin and Riggs, 1980). It has a preference for hemimethylated DNA and discriminates between a hemimethylated site which is generated during replication of a methylcytosine and a site that is non-methylated on both strands of DNA. Second, de novo DNMTs (DNMT3A and DNMT3B) introduce methyl groups onto CG sites which were not previously methylated on the parental template strands of DNA. Methylation is a dynamic process. When DNMTs catalyze the methylation reaction to produce methylcytosines, demethylases can remove methyl groups to create unmethylated CG sites (Fig. 1).

**Fig. 1.** DNA methylation reactions (Syzf, 2005).
Two models have been proposed for maintenance of DNA methylation patterns in somatic cells. In a classical model, DNA methylation pattern is maintained in somatic cells post development (Fig. 2). Once a DNA methylation pattern is carved by de novo methyltransferases and demethylases, it is maintained by maintenance DNMT1 during DNA replication. In dividing cells, maintenance DNMT1 copies faithfully the methylation pattern from the paternal strand to the hemimethylated strands generated during replication (Araujo et al., 1998). Another model, in contrast to the classic model, proposes that DNA methylation pattern is in a dynamic steady state of methylation and demethylation catalyzed by DNA methyltransferases and demethylases. In this model, different transcriptional factors cooperate with each other to determine the state of DNA methylation. Firstly the trans-acting repressor (TR) will recognize and bind to the targeted sequences. This triggers the binding of histone modification enzymes such as histone deacetylase (HDAC) and histone methyltransferase (HMTASE) to particular genes. This event recruits DNMT and methylated DNA binding proteins such as methyl CpG-binding protein2 (MeCP2). This protein complex will tilt the balance towards DNA methylation. On the other hand, trans-acting activating factors (TAF) recruit histone acetyltransferases (HAT) resulting in histone acetylation. This will change the chromatin configuration and enhance the accessibility of DNA demethylases to the targeted genes. The balance will then be tilted towards demethylation (Fig. 3). Some data suggested that DNA methylation patterns are maintained by a combination of DNMT activities which target and recognize different sequence contexts rather than by copying of the methylation pattern of the template as predicted by the semi-conservation model of methylation (Szyf, 2001).
Fig. 2. The classic model of maintenance of DNA methylation patterns in somatic cells (Szyf, 2005).

Fig. 3. DNA methylation pattern is in a dynamic steady-state balance between DNA methylation and demethylation (Szyf, 2005).
2.1.2 Methods used to study DNA methylation

In the past few years, there has seen a tremendous advancement in the understanding of the functional consequences of DNA methylation and its interaction with chromatin structure as well as the transcriptional machinery. Many different approaches have been developed to determine the overall level of DNA methylation in a genome or to reveal the methylation state of particular nucleotide residues (Rein et al., 1998; Havliš and Trbušek, 2002). The choice of the proper method is influenced by the determination of global or regional methylation state. For the occurrence and abundance of any modified base found in an entire genome, they can be detected using total base composition and nearest neighbor analysis. These two methods provide widest range of modified base identification, but their ability to locate a position of a modified base is low. Some other methods, such as modification of base by bisulphite conversion, hydrazine or permanganate followed by PCR and sequencing, can merely target at methylcytosine. However, these methods can identify the position of a modified base precisely along a DNA strand. These methods can be chosen to analyze the genomes where the exclusive presence of methylcytosines has been established by analyses of total base composition.

DNA methylation analysis has been revolutionized by two technologies, namely bisulphite modification of DNA (Frommer et al., 1992) and methylation-specific PCR (MSP) (Herman et al., 1996). Frommer et al (1992) introduced PCR amplification and sequencing into the bisulphite conversion of genomic DNA and the conjunction of these technologies is widely applied to the quantitative or detailed analysis of 5-methylcytosine distribution over the whole genome (Laird, 2003). In this standard protocol, purified genomic DNA is denatured and modified with sodium bisulphite reagent. This chemical targets at the 5-cytosine and converts the unmethylated 5-cytosine into uracil without
affecting the intact status of the methylated cytosine. After treatment with bisulphate reagent, genomic DNA loses the original complementarity between two strands and PCR amplification of the region of interest requires appropriate PCR primers. Unmethylated cytosine is amplified as thymidine and methylated cytosine remains cytosine. The PCR product can be purified by gel electrophoresis and sequenced directly. Alternatively, the PCR product can be cloned into a plasmid vector and individual clones analyzed. In the sequencing procedure, unreacted cytosine is shown as cytosine, while converted cytosine is sequenced as thymidine (Fig. 4). The bisulphite method is easy to apply and it is very sensitive, requiring as little as 10 pg of genomic DNA. It is easy for quantification of methylated cytosine and can be used to scan any cytosine residual within the sequence of interest. The information yielded by sequencing the bisulphite PCR product can be important to determine the role of a 5-methylcytosine in a specific promoter in regulating the gene expression and its effects on replication region, transposable element or imprinting element (Rein et al., 1998).
As one of main methods used in the study of DNA methylation, bisulphite conversion has been improved with a number of technical innovations and variations. COBRA, combined bisulphite restriction analysis, is a combination of methylation-specific restriction enzyme (MSRE) analysis with the bisulphite protocol. New restriction enzyme sites are created by bisulphite modification, for example, bisulphite will convert the BstUl cleavage site CGCG in unmethylated DNA into TGTG, but not in methylated DNA (Sadri and Hornsby, 1996). The bisulphite-treated DNA is digested with MSRE BstUI which cuts methylated DNA and products are separated using gel electrophoresis and visualized by hybridization, using labeled oligos. The methylation levels in the original DNA sample are represented by relative amounts of digested and undigested PCR.
products in a linearly quantitative fashion. COBRA can be used to determine DNA methylation levels at specific loci with very high degree of quantitative accuracy (Xiong and Laird, 1997). Methylation-sensitive single nucleotide primer extension (Ms-SNuPE) is another alternative approach based on the bisulphite conversion (Gonzalo and Jones, 1997). In this approach, a single radioactive nucleotide positioned at 5' end to a specific cytosine is used to analyze the methylation at that particular site. Extension with dCTP reflects unconverted 5-methylcytosine and extension while dTTP reflects converted 5-cytosine. The ratio of methylation at a specific site can be measured by the ratio of the extension product using dCTP versus the extension product using dTTP.

Methylation-specific PCR is another option in investigation of DNA methylation status in CpG islands (Herman et al., 1996). Bisulphite-converted genomic DNA is amplified in the PCR reaction using primers designed to anneal specifically with either the methylated or unmethylated version of the treated sequence. Verification of the primer specificity is done by restriction enzyme digestion of the PCR amplification product. The sequences of the amplicons will be different due to differences in the treated DNA templates. So the DNA cleavage products will differ for PCR products amplified from methylated versus unmethylated DNA.
2.1.3 The biological function of DNA methylation

Cytosine methylation has a number of functions and a few are demonstrated. Between 60% and 90% of CpGs are methylated at the 5th position on the cytosine ring and these account for most of the methylcytosines in the vertebrate genome (Bird, 1986). Methylation of CpG-rich promoters is used by mammals to prevent transcriptional initiation and to ensure the silencing of genes on the inactive X chromosome, imprinted genes (Costello and Plass, 2001; Jones and Takai, 2001).

2.1.3.1 Relationship between DNA methylation and transcriptional activity

Repression of gene expression has been established by numerous transfection experiments using DNA methylated in vitro (Colot and Rossignol, 1999; Razin and Cedar, 1991). Mouse retrovirus experiments have strongly supported that DNA methylation can suppress transcription (Bird, 1984). When human γ-globulin gene DNA were methylated in the 5' flanking region in vitro and introduced into fibroblasts, the presence of this methyl group on exogenous gene sequence was sufficient to inhibit their expression (Busslinger et al., 1983). Transcriptional regulation is a complicated process involving the interaction of a large number of trans-acting factors with specific cis-acting DNA sequences and the local configuration of a gene may determine its accessibility to the various cellular factors required for transcription. Methylation provides one mechanism for stably altering the local structure of a gene and plays a role in the regulation of gene activity in animal cells (Cedar, 1988).

Most mammalian transcription factors have GC-rich binding sites and many have CpGs in their DNA recognition elements (Bird and Wolffe, 1999). Binding of some of the
transcription factors is impeded or inhibited by methylated CpG because methyl-CpG binding proteins, which incorporate efficiently with the methylated bases, share the same target sequences with the transcription factors (Nan et al., 1998; Bird 2002). Factors which interact with the liver-specific tyrosine aminotransferase gene in rat hepatoma cells have been demonstrated to be interfered by the binding of methyl moieties to DNA (Becker et al., 1987). Some factors such as Sp1, however, are indifferent to methylation status. Sp1 binds DNA and activates transcription regardless of its methylation or not (Holler et al., 1988). An alternative route by which CpG methylation can inactivate genes is to directly exclude the transcriptional machinery from methylated promoter region by altering protein-DNA interactions. For example the nucleosome stability or positioning could be influenced by the methyl-CpGs and this in turn denies access of transcription factors to a promoter (Bird and Wolffe, 1999). When DNA is transferred into cells by DNA-mediated gene transfer, it integrates into the genome and is always found in a DNasel-sensitive chromatin conformation. In contrast, methylated DNA goes into an insensitive structure that is DNasel-resistant and it is structurally similar to inactive genes of the cell (Keshet et al., 1986). This suggests that the methyl groups have a role in the placement of proteins required for building the correct conformation and this process probably does not involve sequence specificity (Razin and Cedar, 1991). A unique methyl-CpG binding protein (MeCP) which specifically binds to methylated residues has been identified and it presumably has a dramatic effect on overall chromatin structure (Meechan et al., 1989) which is required for methyl moieties to inhibit transcription (Buschhausen et al., 1987).

During development, one of the two X chromosomes in female cells is methylated (Goto and Monk, 1998). CpG sites in the promoters of the majority of genes on the
inactive X chromosome, including housekeeping genes (genes that are transcribed at a relatively constant level and whose products are typically needed for maintenance of the cell) are methylated and transcriptionally silent which presumably to ensure equivalent expression levels in male and female cells (Kass et al., 1997). Methylation is also involved in expression of imprinting genes which show differential activity depending on whether they are inherited from the father or the mother (Ceder, 1988; Costello and Plass, 2001). The activation of the normally silent paternal allele of the imprinted H19 genes in mutant mice which are deficient in DNA methyltransferase activity demonstrates that a normal level of DNA methylation is required for controlling differential expression of the paternal and maternal alleles of imprinted genes (Li et al., 1993).

Transfection experiments with an artificially methylated hamster adenine phosphoribosyltransferase (APRT) gene demonstrate that methylation of the 5’ region, which is a CpG island, reduces the level of transcription, but methylation of non-island regions shows no obvious correlation with inactivity (Keshet et al., 1985). There are two types of promoters in mammalian genes with respect to CpG: those that are constitutively nonmethylated and CpG rich, and those that are relatively CpG poor and, usually, methylated (Bird, 1992). Most of the promoters of housekeeping genes are CpG-rich promoters and they are methylation free (Antequera et al., 1990), while CpG-deficient promoters are invariably found at tissue-specific genes. DNA methylation occurring in CpG-deficient promoters can suppress activity of a gene in inappropriate cells without affecting high expression in appropriate cells (Bird, 1992). Extensive methylation of some CpG islands was observed in tissue culture cell lines, especially those which had been in culture for many years (Antequera et al., 1990).
2.1.3.2 Bilateral relation between DNA methylation and chromatin

DNA methylation pattern at any point in life could be considered as a balance of DNA methylation and demethylation and the direction of this balance is dictated by the chromatin structure. The factors which are responsible for an inactive chromatin structure enhance DNA methylation (Szyf et al., 2004). There are genetic data and biochemical evidence supporting the hypothesis that DNA methylation is dependent on chromatin structure from fungi to humans. For example, a histone methyltransferase is required for DNA methylation in Neurospora crassa (Lachner et al., 2001). Kryptonite, a histone methyltransferase gene specific to H3 Lys9, is required for methylation by the Arabidopsis DNA methyltransferase chromomethylase 3 (Jackson et al., 2002). In humans mutation of ATRX gene encoding a SW1/SNF chromatin remodeling protein causes diverse changes in the pattern of DNA methylation of several highly repeated sequences (Gibbons et al., 2000). All these evidences indicate that chromatin modification proteins are required for generating a DNA methylation pattern. However the relation between DNA methylation machinery and chromatin status is bilateral. DNA methylation enzymes are required as well as for the formation of chromatin structure. Knock out of DNMT1 in a human colorectal cancer cell line results in a global change in chromatin structure (Espada et al., 2004). The presence of DNA methylation brings about the deacetylation of histone H4 and methylation of H3 Lys9 indicating that DNA methylation is important in setting up the chromatin structure during development (Hashimshony et al., 2003). DNA methylation also takes part in the establishment of the histone acetylation profiles on tissue-specific and housekeeping genes during development by inducing decreased levels of chromatin acetylation (Eden et al., 1998).
The model that DNA methylation pattern is a steady state of methylation and demethylation implies that the DNA methylation pattern is dynamic and could be potentially responsive to any signaling pathway which might alter the chromatin structure in both dividing and nondividing cells. A simple mechanism through which inactivation of chromatin structure could possibly bring about de novo methylation is that the different chromatin modifying enzymes recruit DNMTs to genes. The chromatin modifying enzymes such as histone methyltransferases and histone deacetylases are recruited to specific genes by trans-acting repressors whose repertoire of interacting with a gene could be altered in response to different physiological and pathological signals resulting changes in the methylation pattern (Szyf, 2005). The continuous presence of such factors on a gene ensures that it is maintained in its methylated state (Szyf, 2005).

2.1.4. Changes in DNA methylation during development

Epigenetic mechanisms have been defined as a modification of the genome beyond the level of the genetic information and results in an alternation of gene expression (Laird, 2003; Smith and Murphy, 2004). As one of the components of epigenetic reprogramming machinery, DNA methylation is responsible for the regulation of imprinting gene expression in sperm and egg genomes. Its activity of modulating gene expression is dependent on DNMTs. Working together with several HDACs which control histone hypoacetylation-deacetylation regulating chromatin structures, they change the gene expression profiles during gametogenesis and early embryo implantation (Kierszenbaum, 2005).
Throughout gametogenesis, the paternal methylation is established beginning from spermatogonial stem cell whereas maternal-specific genomic imprints are established starting during follicular growth. Both processes are regulated by the activities of de novo methyltransferases DNMT3a, DNMT3b and DNMT3L. During spermatogenesis, HDACs and histone methyltransferases, including Suv39h, regulate chromatin structure causing the replacement of nucleosomal beaded chromatin pattern by smooth chromatin fibers and resulting in transcriptional silence of the genome (Kierszenbaum, 2005). After fertilization, demethylation occurs in both parental genomes at the blastocyst stage (Monk et al., 1987; Kierszenbaum, 2005). There is rapid demethylation of the paternal genomic DNA, while the methylation pattern of female genome persists and demethylation is observed after DNA replication (Smith and Murphy, 2004; Kierszenbaum, 2005). And this is followed by a wave of de novo methylation which takes place around the time of implantation in a lineage-specific manner (Monk et al., 1987; Kierszenbaum, 2005).

Tissue-specific genes, at some early stage in embryonic development, are methylated in almost all tissues. They only undergo demethylation and become actively transcribed in the tissue in which they are required for proper biological functions (Benvenisty et al., 1985). This new methylation pattern of a tissue-specific gene is faithfully maintained in this cell type (Pollack et al., 1980; Wigler et al., 1981). Some tissue-specific genes, exceptionally, are unmethylated at early oogonial stages in human embryos (Driscoll et al., 1990). In a gene transfection experiment, the 5’ CpG island of an artificially methylated APRT gene experienced demethylation in founder mice, whereas the non-island sequences of the gene remained methylated and even became de novo methylated at sites that are not originally in vitro (Frank et al., 1990). This result suggests
that the presence of an island demethylation activity may occur in the embryonic cells and these cells have the capability to recognize classes of gene sequences and reestablish their correct methylation pattern through a combination of demethylation and de novo methylation, which is then carried to somatic cells where it stably maintained (Frank et al., 1991).

2.2 Mechanism of DNA demethylation

2.2.1 Passive and active DNA demethylation

In contrast to the large amount of available information on DNA methylation, relatively little is known about DNA demethylation. Two types of demethylation events can be distinguished: global and local demethylation (Kress et al., 2001). In mammals, genome-wide demethylation occurs during gametogenesis, early development of embryo and in some differentiating cells (Brandeis et al., 1993). Site-specific demethylation, on the other hand, affects the expression of tissue-specific genes in the appropriate tissue and may probably involve the presence of an island demethylation activity in embryonic cells (Frank et al., 1991).

DNA demethylation can be either passive or active. Passive DNA demethylation occurs by inhibiting maintenance DNA methyltransferases throughout cycles of replication and demethylated cytosines could be simply resulted from an absence of maintenance methylation after replication. Such mechanism does not involve any demethylases, so it would be a passive pathway for removal of methyl groups from methylcytosines (Kress et al., 2001). In passive demethylation, several cell division and DNA replication are required to ensure that a significant part of the daughter molecules are effectively demethylated. Full double-stranded demethylation would occur in 50% of
the cells after at least two cell divisions (Razin and Riggs, 1980). Three rounds of replication demethylate only 87.5% of the CpG (Kress et al., 2001). Active demethylation is an enzymatic reaction which occurs in the absence of replication (Kress et al., 2001). It has been assumed that demethylation in vivo takes place through a passive mechanism whereby maintenance methylase is inhibited at specific sites. Indeed it has been proved in several instances that demethylation takes place through an active mechanism. In mice, a genome-wide demethylation of the zygotic paternal genome after fertilization appears to occur by an active mechanism, which is then followed by passive demethylation during cleavage stages (Mayer et al., 2000). Local demethylation of transfected DNA molecules that do not replicate has also been observed (Paroush et al., 1990; Frank et al., 1991). Local specific gene demethylation also occurs throughout embryonic development and in terminally differentiated cells. In human embryos, several tissue-specific genes are indeed unmethylated at early oogonial stages (Driscoll et al., 1990). In tissue culture cells, murine leukemia virus becomes unmethylated after the cells have undergone differentiation as a result of retinoic acid treatment (Razin and Ceder, 1991).

Three active demethylation mechanisms have been proposed but none has yet been proven to operate under relevant physiological conditions or gained wide acceptance (Fig. 5). The first mechanism is direct replacement of the methyl moiety by a hydrogen atom. The human methylation DNA binding protein 2 (MBD2) was reported to demethylate DNA by this mechanism (Bhattacharya et al., 1999; Ramchandani et al., 1999). The other two proposed mechanisms involve DNA repair processes. The second mechanism implicates a role for DNA glycosylases. These enzymes cleave the bond between the 5-methylcytosine base and the deoxyribose moiety in DNA. Resident repaired activity then repairs the cut site resulting in replacement of a 5-methylcytosine
with an unmethylated cytosine (Jost et al., 1995). The third mechanism proposed that the methylated nucleotide was removed by nucleotide excision and was replaced by an unmethylated cytosine (Weiss et al., 1996; 1997).

Fig. 5. Possible modes of action of the demethylases (Kress et al., 2001).
2.2.2 Effects of demethylating reagents on the activation of silenced genes

When 5-azacytidine, a cytosine analog, is incorporated into DNA, it cannot accept a methyl group like cytosine (Jones and Taylor, 1980; Creusot et al., 1982). Cells exposed to 5-azacytidine also inhibited maintenance-methylase activity (Tanaka et al., 1980; Creusot et al., 1982). Studies of specific genes have indicated that 5-azacytidine could activate or reactivate a range of previously silenced genes (Taylor and Jones, 1979; Mohandas et al., 1981; Venolia et al., 1982) strongly suggesting that a loss of methylation was the cause. The inactive endogenous virus gene in chicken AEV cells was switched on after the cells were treated with 5-azacytidine and the gene sequences were found to have undergone activation in different cell types (Croudine et al., 1981). For the X-linked genes, 5-azacytidine could activate one allele at high efficiency which was inactivated presumably by methylation (Mohandas et al., 1981; Harris, 1982). The activation potential of 5-azacytidine was best seen in 10T1/2 or NIH 3T3 cells in which treatment caused differentiation to three mesodermal cell types (Taylor and Jones, 1979). Other cytosine analogues which had no effect on differentiation were demonstrated not to affect DNA methylation (Jones and Taylor, 1980). These results suggested that alternations of methylation patterns had a major role in the switching of gene activity during development.

In many permanent cell lines, only those genes that are essential for life in culture conditions are free of methylation, while thousands of other genes have been stably inactivated presumably because of selection (Antequera et al., 1990). The effects of 5-azacytidine in culture could be explained by removal of the methylation from genes whose expression triggered differentiation (Gounari et al., 1987; Sneller and Gunter, 1987). The myogenesis-promoting gene MyoD1 has been found to possess a CpG island
that was methylated in 10T1/2 cells but demethylated upon differentiation of these cells into muscle under the influence of 5-azacytidine (Jones et al., 1990). The same result was observed in X chromosome-linked genes in human cell line. A specific zone within the PGK1 (phosphoglycerate kinase gene) promoter was required for transcription initiation and this DNA sequence underwent demethylation which was induced by the treatment of 5-azacytidine in expressing cells (Hansen and Gartler, 1990).

The 5-azacytidine not only affects the DNA methylation, but also influences other types of cell activities including protein, DNA and RNA synthesis and RNA processing (Jones and Taylor, 1980). The notion that these other effects might be important was suggested by observation that DNA methylation might sometimes be inhibited in vivo but not in vitro (Tanaka et al., 1980). Also, the level of 5% substitute with the analogue was not sufficient to explain 80% inhibition of cytosine methylation unless the incorporated base impeded the “walking” of a methylase along the DNA (Jones and Taylor, 1980). Competition experiments have shown that the addition of cytosine did not reduce the inhibitory effect of 5-azacytidine (Tanaka et al., 1980). A recent study has demonstrated that the effects of 5-aza-CdR, a DNA methylation inhibitor, on the transcriptome of a colorectal cancer cell line HCT116 were similar at one and five days’ exposure, presumably due to active demethylation (Gius et al., 2004). The fact that 5-aza-CdR causes active demethylation rather than passive demethylation is consistent with the working hypothesis that the DNA methylation pattern in mature tissues is steady state balance of constitutive methylation and demethylation reactions. According to this model, inhibition of DNMTs by 5-azacytidine leaves the demethylation reaction unopposed by DNA methylation and results in a new unmethylated state (Szyf, 2005).
2.3 Bovine milk protein genes

2.3.1 Bovine milk proteins

Milk is composed of many different components including fat, carbohydrate, and proteins. Proteins make up about 3.3% of bovine milk. True proteins excluding non-protein nitrogen are classified into three fractions: caseins in micelles, whey proteins in solution and fat globule membrane proteins on the surface of fat globules. In bovine milk, six major proteins have been characterized. These are the four caseins, $\alpha_s$, $\alpha_s$, $\beta$ and $\kappa$, and two whey proteins $\alpha$-lactalbumin and $\beta$-lactoglobulin. They are synthesized and excreted by the mammary epithelial cells during lactation and casein micelles are a main source of minerals and amino acids to the suckling calf (Akers, 2002).

Milk micelle structure is formed through the interaction among the three calcium-sensitive caseins, namely $\alpha_s$, $\alpha_s$, $\beta$ and stabilized by $\kappa$-casein. When it is cleaved by chymosin, the micelle aggregation is initiated resulting in curd formation and milk protein digestion. Similarly, the curd formed by casein aggregation is also important in making cheese. In this procedure $\kappa$-casein has a crucial role in micelle formation and stability (Alexander et al., 1992). The $\alpha$-lactalbumin protein occurs in all lactose-containing milks and influences lactose synthesis, which in turn affects the osmolarity of milk. It modifies the substrate specificity of a UDP-galactosyltransferase which usually involved in the synthesis of glycoproteins (Vilotte et al., 1987; Threadgill and Womack, 1990).
2.3.2 Milk protein synthesis

Protein synthesis in mammary epithelial cells follows the scheme including DNA replication, transcription of messenger RNA (mRNA) from DNA and translation of mRNA to form protein (Mepham, 1983; Akers, 2002). During replication, the two strands of DNA separate and duplicate by a base-pairing procedure. Replication takes place prior to cell division and has no direct role in protein synthesis. Transcription of mRNA from DNA depends on the binding of transcription factors to a site on the DNA adjacent to the start sequence called the promoter and produces a single strand molecule containing triplet of bases known as codon for a specific amino acid and. The mRNA molecules move into the cytoplasm whereby a specific sequence of three bases of transfer RNA (tRNA), known as anticodon, recognizes and binds to the corresponding codon on mRNA forming an amino acyl-tRNA complex (Akers, 2002). This step requires an amino-acid-activating enzyme that is specific to each amino acid and energy from ATP. The complex is brought together with mRNA by ribosome RNA (rRNA) for the synthesis of protein at ribosome. After being bound to ribosome by base pairing to rRNA, the codon on mRNA is recognized by the specific anticodon on tRNA which places the amino acid in the proper position to form the growing chain of amino acids by a peptide bond. This process is controlled by two different enzymes and the energy from guanosin triphosphate (Schmidt, 1976). Once the peptide bond is formed, tRNA is released into the cytoplasm and is free to pick up another amino acid. At a given time and a given strand of mRNA, multiple ribosomes can attach to only one single strand of mRNA. The ribosomes move along the mRNA to form several growing polypeptides at the same time while each ribosome contains one chain of new protein (Akers, 2002).
Before the milk proteins are secreted from the cell, the newly synthesized peptide sequences undergo several covalent modifications in the endoplasmic reticulum as well as in Golgi apparatus (Mepham, 1983; Akers, 2002). These post-translational modifications namely proteolytic processing, glycosylation and phosphorylation (Mepham, 1983) can significantly affect functioning of the protein. The proteolytic process in vivo occurs before the complete folding of the polypeptide chains to remove the signal peptides and serves to form the final three-dimensional structure of the complex protein structure (Mepham, 1983). Glycosylation is an enzymatic process which is important for the folding of proteins. Two types of glycosylation exist: N-linked glycosylation to the amide nitrogen of asparagine side chains and O-linked glycosylation to the hydroxy oxygen of serine and threonine side chains. It has been widely accepted that glycoproteins have some biological functions in stabilizing protein conformation, playing a role in membrane-membrane interaction and transporting proteins to definite sub-cellular compartments or organs (Mepham, 1983). Phosphorylation of proteins is another ubiquitous biological process which occurs after completion of polypeptide chains. It involves enzymatic recognition of specific sites corresponding to short amino acid sequences and conformational patterns. This biological process of phosphorylating caseins is essential in assembling stable micelles and the behavior of caseins as nutrients (Mepham, 1983). The synthesis of milk proteins is largely dependent on a supply of amino acids to mammary epithelial cells. Caseins, α-lactalbumin and β-lactoglobulin are synthesized in the udder and their sources of precursors from blood may be peptides, plasma proteins and free amino acids (Schmidt, 1976).
2.3.3 The role of lactogenic hormones in regulation of milk protein gene expression

The expression of genes controlling milk protein synthesis in lactating mammary gland cells is regulated by a complex of different lactogenic hormones. The relative amount of these hormones may inhibit or stimulate the gene expression. Prolactin and glucocorticoids play a major role in positive regulation (Gorodetsky et al., 1988). In a number of studies, glucocorticoids have been shown to enhance the induction of casein genes by prolactin (Choi et al., 1988; Doppler et al., 1989; Schmitt-Ney et al., 1991). In a study of the peptide hormone regulation of gene expression in rat mammary gland organ culture, Rosen et al. (1978) demonstrated that prolactin induced the accumulation of casein mRNA up to 48 hours. Moreover, hydrocortisone and progesterone participated in modulation of prolactin-induced accumulation of casein mRNA and the presence of hydrocortisone was necessary for maximal accumulation of casein mRNA. In contrast, casein gene and α-lactalbumin transcription which was induced by prolactin could significantly be inhibited by progesterone at pharmacologic doses in lactating tissue culture (Shamay et al., 1987). When mouse mammary epithelial cells were grown on floating collagen gels and exposed to insulin, cortisol and prolactin, the amount of casein secreted into the medium by floating collagen gels was 25 to 200 times greater in the presence of all three hormones than with insulin alone. In the presence of insulin, both prolactin and glucocorticoids are required for maximum induction of α-lactalbumin mRNA in the mouse. On the other hand, insulin and glucocorticoids together with epidermal growth factor can induce the synthesis of α-lactalbumin in the rat mammary gland even in the absence of prolactin. For most of the mammals, progesterone exhibits inhibition effect on α-lactalbumin synthesis. Exceptionally, it does not inhibit induction of α-lactalbumin mRNA accumulation in marsupials (Vonderhaar and Ziska, 1989).
2.3.4 DNA methylation and milk protein gene expression

Methylation of casein genes in rats has been studied using restriction enzyme analysis, such as isoschizomers MspI and HpaII (Johnson et al., 1981). Result showed that methylated CpG occurred towards the 5’ end of the γ-casein gene in the corresponding DNA isolated from mammary tumors as well as from liver which are the non-expressing tissues of these genes. Similarly, methylation was also observed at specific CpG sites in the α- and β-casein genes from hepatic and tumorous mammary tissues. An inverse relationship exists between methylation of the specific CpG sites within the casein genes and their expression. The same correlation between the degree of methylation and gene expression level was demonstrated between cell subpopulations within a single tissue for the rat γ- and β-casein genes (Johnson et al., 1983). Methylation of the rat κ-casein gene cDNA was studied in normal and neoplastic mammary gland and the expression of the rat κ-casein gene was shown to be inversely correlated with the degree of methylation of this gene (Thompson and Nakhasi, 1985). In bovine, the possible role of DNA methylation in the control of expression of the αs1-casein gene was investigated using restriction enzymes at different stages of lactation (Platenburg et al., 1996). The authors demonstrated that mammary gland-specific demethylation occurred at two sites in the 5’-flanking region of the gene and one in the 3’-flanking region in lactating cows. For the rat α-lactalbumin gene, differences in methylation patterns were observed in DNA fragments isolated from the mammary tissues at the different stages of development and from several mammary tumors. It was suggested by the authors that an inverse relationship existed between gene expression and occurrence of methylated bases (Qasba et al., 1982).
2.3.5 Effect of 5' regulatory region methylation on gene expression

The effect of DNA methylation in 5' regulatory region of a gene has been studied by in vitro methylated technology for the human γ-globulin gene and β-globulin gene (Busslinger et al., 1983; Yisrarli et al., 1988). The authors found that methylcytosine residues in the γ-globulin structural gene did not inhibit transcription of the gene, whereas methylation in the 5' region of the gene prevented its transcription (Busslinger et al., 1983). For the β-globulin gene, overall methylation at cytosines inhibited gene transcription. It was also shown that methylation located in the 5' and 3' ends of the gene had a negative regulatory effect on gene expression (Yisrarli et al., 1988). The effect of DNA methylation on the transcriptional activity of hamster APRT and herpes thymidine kinase (TK) genes was studied by using in vitro methylated construct containing these gene sequences (Keshet et al., 1985). It was found that the APRT gene was inhibited by CpG methylation in the 5' region but was unaffected by methylation at the 3' end. In contrast, DNA methylation at both the 5' promoter region and the 3' structural region of the TK gene inhibited gene transcription. The effect of promoter methylation was elucidated by studying the gene expression of chloramphenicol acetyltransferase in mammalian cells under the control of adenovirus type 12 promoters (Kruczek and Doerfler, 1983). These authors demonstrated that 5' methylation in the promoter region decreased or eliminated transcription, but methylation sites too far upstream or any sites downstream from the TATA site did not affect gene transcription (Kruczek and Doerfler, 1983).
2.3.6 The core promoter of a gene

In eukaryotes, gene expression is regulated by a wide variety of mechanisms. The RNA polymerase II core promoter is one of these regulatory components and the ultimate target of all the transcription factors (TF) involved in this transcription regulation machinery (Butler and Kadonaga, 2002). It is defined as "the minimal stretch of contiguous DNA sequence that is sufficient to direct accurate initiation of transcription by the RNA polymerase II machinery" (Butler and Kadonaga, 2002). The core promoter encompasses several sequence motifs including TATA box, initiator (Inr), TFIIB recognition element (BRE), and downstream core promoter element (DPE). There is a recently reported core promoter sequence, the MTF (motif ten element) (Deng and Roberts, 2006). The TATA box was the first element to be characterized and is recognized by the TBP (TATA-box-binding protein) subunit of TFIID (transcription factor IID) (Deng and Roberts, 2006). It's usually observed at the position of -31 to -26 related to the transcription initiation site (Butler and Kadonaga, 2002). The consensus sequence for the TATA box is TATAAA, but with one or two mismatches from this consensus may be functional in vivo (Butler and Kadonaga, 2002). The Inr encompasses the transcription start site +1 and is bound by the TFIID in a sequence-specific manner (Butler and Kadonaga, 2002; Deng and Roberts, 2006). The DPE is located precisely at +28 to +32 relative to the +1 position in the Inr and commonly found in TATA-less promoters. It is recognized and bound by the TFIID and mutation at this sequence results in a dramatic decrease in transcription (Butler and Kadonaga, 2002; Deng and Roberts, 2006). The BRE locates adjacent to the TATA box both upstream (BRE<sup>u</sup>) and downstream (BRE<sup>d</sup>). Both the sequences are bound by the TFIIB (transcription factor IIB) C-terminal domain and modulate the transcription potency of a promoter (Deng and Roberts, 2006). All these
motifs may not be found in all promoters. Some specific promoters may contain some or none of these sequence elements (Butler and Kadonaga, 2002). In gene transcription, different transcription factors can be assemble into a pre-initiation transcription complex (PIC) and TFIID and TFIIB are the first two factors that incorporate into the core promoter. They play a crucial role in the assembly of the PIC and the recognition of core promoter motifs (Butler and Kadonaga, 2002; Deng and Roberts, 2006).

![Diagram of transcription factors](image)

**Fig. 6.** Known core promoter elements. The consensus sequences of BRE\textsuperscript{u} and BRE\textsuperscript{d} are shown below (Deng and Roberts, 2006).

### 2.3.6.1 The target region in the core promoter of bovine κ-casein gene

The major glycoprotein synthesized by the mammary gland is κ-casein. In cows, it consists of a single polypeptide chain of 169 amino acid residues containing O-linked oligosaccharides in the C-terminal region (Mepham, 1983). Isolation and characterization of bovine κ-casein gene show that the gene consists of five exons distributed over a total length of approximately 13 kb and the promoter, together with exon 1, span 2250 base pairs. At the 5’end upstream of the gene transcription initiation site, the AT-rich sequence
TTTAATTA is found at positions -33 to -26. Downstream of this AT-rich sequence there is sequence GTCTCTGG, which is homologous with the BRE\textsuperscript{d} consensus sequence. CAAT element is observed at positions -78 to -75. Sequence TGACGCA found at -54 could be another sequence involved in regulation of gene expression. It has 6/7 homology with the sequence TGACTCTCA which has been shown to be the recognition site of transcription factor AP-1 (Alexander et al., 1988). In cattle, the calcium sensitive caseins show a clear evolutionary relationship to each other and there is complete conservation of the positions within the genes of the first two and the last two exons. However, this feature has not yet been seen in \(\kappa\)-casein gene. When comparing the first 200 base pairs of the 5' sequences of calcium-sensitive caseins, a variety of conserved motifs can be found, while none of them is contained in the corresponding portion of the \(\kappa\)-casein gene. These evidences support that \(\kappa\)-casein is evolutionarily divergent from the other three calcium-sensitive casein genes (Alexander et al., 1988).

Fig. 7. Target region of bovine \(\kappa\)-casein gene core promoter with known transcription factor binding domains. Bold-faced nucleotides indicate the identified sequences involved in the regulation of gene transcription. 1: CAAT element; 2: Homologous sequence of transcription factor AP-1 binding; 3: TATA box-like sequence (Alexander et al., 1988); 4: Homologous sequence of BRE\textsuperscript{d}.
2.3.6.2 The target region in the core promoter of bovine α-lactalbumin gene

The α-lactalbumin protein is glycosylated to various extents in different species. It is present in a small amount in bovine and mouse, but it is dominant in rabbit milk (Mepham, 1983). The complete sequence of bovine α-lactalbumin gene was studied and several regulation elements in 5′ regulatory region of the gene have been identified (Vilotte et al., 1987). Upstream from the transcription initiation site, the TATA box-like and the CAAT box were located at -24 bp and -117 respectively. These might be the recognition sites of polymerase II. At position -333 and -654, there are two sequences thought to be the possible binding sites of progesterone receptor. The sequence AGGCTTGATGCCA at position 672 is 85% homologous to the consensus sequence TGGCANNNTGCCA recognized by the TGGCA-binding protein which plays a role in the modulation of transcription of the lysozyme gene. Downstream of TATA box there is homologous sequence of BRE.

```plaintext
1 caacaagggac caagatact aagggacact tgttttctt catgccctggg ttgagtgggc
61 catgacatat gatgatgtac agtcccccttt ttc catatctgtgt atgtctctaa gaggagagag
121 gatgtgggcc tggacccttt tggactatttgc aacttgatatt cccctgagag
181 aaccttttgtt cctgaaatag gttggacaca tctgtgtctc tagaaccacat actaccagaa
241 acaacataaa taaagccaaa tgggaaacag gatcatgttt gtaacactct ttggcaggtt
301 aacaaataacct agatatgact gggattcgg gggaggaaga gaaaaagttgg ggtgaattac
361 ttgggaagag tcaatgtttg cttggtgttt ttactggcct cctctctctgt
421 gatgtgtaggc tttgtggcaag gcggcccttaag gcttttttcca caaataaga gagttgggag
481 cagttggttag
```
Fig. 8. Target region of bovine α-lactalbumin gene core promoter with known transcription factor binding domains. Bold-faced nucleotides indicate the sequences involved in the regulation of gene transcription. 1, 4: progesterone receptor binding site; 2: CAAT element; 3: Italics nucleotides indicate the TGGCA protein binding site; 5: TATA box-like sequence (Vilotte et al., 1987); 6: Homologous sequence of BRE
d.

2.4 Quantitation of gene expression by real-time quantitative PCR (qPCR)

2.4.1 Real-time quantitative reverse transcription PCR (qRT-PCR)

Real-time quantitative Polymerase Chain Reaction (qPCR) is a rapid, robust and highly sensitive method to quantify specific nucleic acid targets which may be DNA, cDNA or RNA. In real-time PCR, reactions are characterized by the point in time during the cycling when the amplification of a target is first detected rather than the amount of target accumulated after a fixed number of cycles. The synthesis of PCR products is measured by the intensity of fluorescence signal from the geometric phase of the amplification reaction during which all of the components required for the PCR are excess and exact doubling of products is occurring. The reaction is very specific and precise. The increase in reporter fluorescence signal is directly proportional to the number of amplicons generated.

To detect the real-time amplification of target sequences, fluorescent markers are utilized in the assay. The two most widely used fluorescent chemistries for qPCR are TaqMan chemistry and SYBR Green I. The TaqMan chemistry uses a fluorogenic probe to enable the detection of a specific PCR product as it accumulates during PCR cycles. It takes advantage of the 5' exonuclease activity of the Taq DNA polymerase to generate detectable distinguishable signal (Holland et al., 1991). SYBR Green is a small molecule
that binds to double-stranded DNA and can be categorized as intercalators or minor-groove binders (Applied Biosystems, 2002a). During the PCR, Taq DNA Polymerase amplifies the target sequence. The SYBR Green I dye then binds to amplicons and exhibits as much as 1000 fold higher fluorescence than its unbound form. As PCR progresses, more amplicons are generated and more binding sites are created for SYBR Green, which results in increased fluorescence signal. Compared to TaqMan chemistry, SYBR Green is easier to use and no probe is necessary. The primary disadvantage of the SYBR Green I dye chemistry is that it may generate false positive signals because the SYBR Green I dye binds to all double-stranded DNA sequences including nonspecific products. When SYBR Green is used, the primers used for amplification should be as specific as possible and a melting curve analysis should always be performed to identify the nonspecific amplification. This can be done by gradually increasing the temperature within a defined range programmed with a real-time PCR machine at the end of the reaction. The nonspecific products give distinguishable peaks with lower melting temperatures compared to those of the correct products.

2.4.2 Data analysis in qPCR: absolute and relative quantification

Two methods have been proposed in qPCR to analyze the data. Absolute quantification means that the absolute copy number of the target sequence is measured. Relative quantification means that a quantitative difference in the copy number between two samples, experimental and control is measured by normalizing both samples to an endogenous reference. In comparative quantification, a threshold cycle value (Ct) of each sample will be generated and ΔCt is calculated against the Ct of an endogenous gene to normalize the data. Then ΔΔCt is calculated for the difference in gene expression of a
specific transcript between two samples in comparison and the change in expression of the gene is described as an arbitrary unit, usually the fold-change $2^{-\Delta Ct}$ value (Livak and Schmittgen, 2001).

2.4.3 Housekeeping genes in qRT-PCR

An endogenous reference DNA is needed to normalize quantitation of the target sequence in order to minimize the variations between samples due to differences in the amount of total RNA added to each reaction, variable reverse transcription efficiencies and the pipetting errors. Traditionally, housekeeping genes (e.g. GAPDH, cyclophilin, β-actin and 18s rRNA) are selected as internal standards for data normalization. There are general criteria for internal control selection. Because the endogenous control is added to normalize the differences in cDNA loaded into real-time PCR reaction, the level of its expression must be similar in all sample in the study and it is critical to determine if the experimental treatment or invention is altering the expression level of the candidate endogenous control gene (Applied Biosystems, 2002a). A validation test must be conducted to evaluate the amplification efficiencies of the target gene and the endogenous reference gene in the same sample. The normalized ΔCt values are plotted vs. log input amounts of different templates to create a semi-log regression line. Both the target and endogenous control should be amplified at an approximately equal level and the reference gene can be considered to be used in the study when the absolute value of the slope of ΔCt vs. log input is less than 0.1 (Applied Biosystems, 2002a).

In this study, we were interested in the role of core promoter methylation in regulating the expression of bovine κ-casein and α-lactalbumin genes at different physiological stages. Bovine mammary gland samples were collected from lactating and
non-lactating animals. Bisulphite PCR and sequencing were applied to investigate the methylation profiles in the target core promoter regions of the two genes. The expression of the two genes was quantified using qRT-PCR. Finally, the DNA methylation status in the target regions were correlated to the gene expression profiles to explore the effect of DNA methylation in the core promoter regions on the genes’ expression.
Chapter III

Material and Methods

3.1 Study of DNA methylation in bovine $\kappa$-casein and $\alpha$-lactalbumin gene core promoter regions using bisulphite sequencing

3.1.1 Genomic DNA preparation of bovine mammary gland tissues

Mammary gland tissues from 12 lactating and 8 non-lactating Holstein cows were from two previous experiments and the tissue sampling has been described in detail (Capuco et al., 1997; Capuco et al., 2001). The lactating cows included 4 animals each on 14th day, 90th day and 240th day of lactation representing early, mid and late stage of lactation. All of these animals were non-pregnant. The non-lactating cows were comprised of 4 animals from the 7th day and 2 animals each from the 25th and 53rd day after cessation of milking (dry). All the dry animals were pregnant. Because milking was terminated at 60 days before expected parturition, these times equate to 53, 35 and 7 days prepartum. All cows were free of intramammary infection.

Approximately 100mg tissue was ground with a prechilled mortar and pestle to a fine powder, and then suspended in 1.2 ml of the digestion buffer containing 100mM NaCl, 10mM Tris-Cl at pH 8.0, 25mM EDTA at pH 8.0, 0.5% SDS and 0.12mg proteinase K (Invitrogen Life Technologies, Burlington, ON, Canada). Digestion reactions were incubated at 50°C overnight with gentle agitation. The digested samples were extracted with an equal volume of phenol/chloroform to remove its protein residues. DNA was then precipitated with the ice-cold 100% ethanol and rinsed with the room temperature 70% ethanol to remove any salt contamination. The DNA samples were dissolved with autoclaved ultra-pure water. The concentration and quality of isolated
DNA were determined by UV absorption measurement at 260nm using NanoDrop spectrophotometer (ND-1000, NanoDrop Technologies, Wilmington, USA) and integrity of the isolated DNA was verified by agarose gel electrophoresis method. Approximately 2 µg of the purified genomic DNA was used in the bisulphite conversion reaction.

3.1.2 Bisulphite conversion of the isolated DNA

Bisulphite conversion of isolated DNA samples was performed using the EpiTect Bisulphite kit (Qiagen Inc., Mississauga, ON, Canada) according to the manufacturer’s recommendations. The reactions were done in a final volume of 140 µl containing 85 µl of the Bisulphite Mix, 35 µl of the DNA Protect Buffer and 2 µl of the isolated DNA (approximately 2 µg). The conversions were incubated in a thermal cycler with the following conditions: 99°C for 5 min, 60°C for 25 min, 99°C for 5 min, 60°C for 85 min, 99°C 5 min and 60°C for 175 min. The bisulphite converted DNA was bound to EpiTect spin column and desulfonated and washed with a series of buffers supplied with the kit. The purified DNA was eluted with 20 µl of elution buffer provided.

3.1.3 Design and modification of the primers in bisulphite PCR

When we designed primers for the amplification of the target sequences, the cytosine residues were assumed to be modified as thymidines and primer composition was AT-rich. The primers were designed to flank the known transcription factor binding domains and hormone binding domains in the core promoter regions of bovine κ-casein and α-lactalbumin genes.
Table 1. Primers used in the bisulphite PCR for amplification of target bovine κ-casein and α-lactalbumin gene core promoter regions.

<table>
<thead>
<tr>
<th>Target sequences</th>
<th>Primer sequences (5’ to 3’)</th>
<th>Product (bp)</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bovine κ-casein</td>
<td>F: ATTGAGATTGATGTAAGATG</td>
<td>207</td>
<td>AF097400</td>
</tr>
<tr>
<td></td>
<td>R: TCCTTGTGACCATCAACTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bovine α-lactalbumin</td>
<td>F: TAATAAGGAATTAAGATAAAG</td>
<td>489</td>
<td>X06366</td>
</tr>
<tr>
<td></td>
<td>R: TACCAACTACTCACCTC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F: forward primer; R: reverse primer. Underlined bases refer to the modified nucleotides from C to T for amplifying the bisulphite treated DNA template.

3.1.4 PCR amplification of the bisulphite treated DNA

PCR reactions were conducted with the primers listed in Table 1 for the two target sequences using the Matercycler® ep PCR machine (Eppendorf, Mississauga, ON, Canada). Amplification was carried out in duplicates to ensure sufficient PCR product in the following ligation reaction. In a final volume of 15 µl, each reaction contained 1×PCR buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 0.1 µM of each primer, 1U Taq polymerase (Invitrogen Life Technologies, Burlington, ON, Canada) and 2 µl of modified DNA template. For the κ-casein gene target sequence, PCR condition was performed with an initial denaturation at 94°C for 6 min, 35 cycles of the following steps: 94°C for 45 s, 58°C for 45 s, and 72°C for 1 min, and a final extension at 72°C for 10 min to facilitate the addition of an A overhang to the PCR product. For the α-lactalbumin gene target sequence, PCR condition was performed with an initial denaturation at 94°C for 6 min, 35 cycles of the following steps: 94°C for 45 s, 56°C for 45 s, and 72°C for 1 min, and a
final extension at 72°C for 10 min. The PCR products were resolved on an agarose gel to verify the target sequences.

3.1.5 Purification of the bisulphite PCR products

The amplified products for both target sequences were purified using QIAquick Gel Extraction Kit (Qiagen Inc., Mississauga, ON, Canada) according to the manufacturer’s recommendation. The PCR products were firstly resolved and excised from the agarose gel. Gel slice was, then suspended in a specified buffer with a three volume of buffer to one volume of gel slice ratio. The reaction was incubated at 50°C until the gel slice was completely dissolved. One gel volume of isopropanol was added to increase the DNA yield. The sample was then applied to the QIAquick column in a supplied 2-ml collection tube and centrifuged for 1 min for DNA binding. The DNA pellets were washed by serial buffers supplied with the kit to remove the residual agarose gel and salt contamination. Purified DNA fragments were eluted with 20 µl of 10mM Tris-Cl at pH 8.5. The purified PCR products were quantified using the Nanodrop spectrophotometer and verified with gel electrophoresis.

3.1.6 Cloning and sequencing of the bisulphite PCR products

Cloning of the purified PCR products was accomplished using a QIAGEN PCR Cloning Kit (Qiagen Inc., Mississauga, ON, Canada) according to the manufacturer’s recommendation with some slight modifications. The PCR product was first ligated to a pDRIVE Cloning vector. The ligation reaction was done in a total volume of 10 µl containing 50 ng of the pDrive Cloning vector, 4 µl of the purified PCR product and 5 µl of 2× Ligation Master Mix. The ligation-reaction mixture was incubated for at least 1
hour at 4°C. For the transformation, the QIAGEN EZ competent cells were thawed on ice and the ligation-reaction mixture was added by carefully dispensing it into the cells. Then the cells were incubated on ice for 5 min and heated at 42°C for 30 s for the uptake of the recombinant DNA. The reaction was incubated on ice for another 2 min for the recovery of the transformed cells. After this 250 µl of room temperature SOC medium was added and one hour of 37°C incubation with shaking was performed to attain the maximum number of live cells for plating. One hundred µl and 20 µl of each transformation mixture were added onto separate Luria Bertani (LB)-agar plates containing 30 µg/ml of ampicillin to ensure good separation of colonies for subsequent single-colony isolation. The transformation mixture was plated using a sterile bent glass rod or a specialized spreader. The plate was incubated at room temperature until the transformation mixture was absorbed into the agar. Then the plate was inverted and incubated at 37°C overnight. The colony PCR, instead of the blue/white screening which may yield false positive result, was performed to identify the insertion of target sequence into the cloning vector. At least ten individual colonies were randomly selected as the template to scan the positive colonies of each sample. The colony PCR was done in a final volume of 15 µl containing 1×PCR buffer, 1.5 mM MgCl₂, 0.2 mM dNTP, 0.1 µM of each primer, and 1U Taq polymerase (Invitrogen Life Technologies, Burlington, ON, Canada). The PCR condition was an initial denaturation at 94°C for 6 min, 35 cycles of the following steps: 94°C for 45 s, 58°C for κ-casein gene and 56°C for α-lactalbumin gene for 45 s, and 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were resolved on an agarose gel with ethidium bromide staining. The positive colonies with amplified PCR products were inoculated for small overnight cultures in 5 ml of LB medium containing 30 µg/ml of ampicillin for the downstream plasmid
purification procedure. The recombinant plasmids were prepared using QIAGEN QIAprep Spin Miniprep Kit (Qiagen Inc., Mississauga, ON, Canada) according to the manufacturer's recommendation. After the overnight culture, bacterial cells were pelleted and resuspended in 250 µl of Buffer P1 containing RNase A and LyseBlue reagent. Two hundred and fifty µl of Buffer P2 was added and reaction was mixed thoroughly to achieve a homogeneously colored suspension. Then 350 µl of Buffer N3 was added and the suspension was mixed immediately and thoroughly until all trace of blue was gone and the suspension was colorless indicating that SDS had effectively been precipitated. The reaction was centrifuged for 10 min and the supernatant was applied to the QIAprep spin column to bind the DNA. The bound DNA was washed with 0.5 ml of Buffer PB and 0.75 of Buffer PE. Then 50 µl of 10mM Tris-Cl at pH 8.5 was added for DNA elution. The purified plasmids were quantified using the Nanodrop spectrophotometer and used as templates in the plasmid PCR to again confirm the insertion of the target sequence. The reaction was done in a final volume of 15 µl containing 1×PCR buffer, 1.5mM MgCl₂, 0.2mM dNTP, 0.1 µM of each primer, and 1U Taq polymerase (Invitrogen Life Technologies, Burlington, ON, Canada). The PCR condition was an initial denaturation at 94°C for 6 min, 35 cycles of the following steps: 94°C for 45 s, 58°C for κ-casein gene and 56°C for α-lactalbumin gene for 45 s, and 72°C for 1 min, and a final extension at 72°C for 5 min. The PCR products were resolved on an agarose gel with ethidium bromide staining. The plasmids were then sequenced to investigate the DNA methylation patterns within the target regions (McGill University and Genome Quebec Innovation Centre Sequencing).
3.2 Quantification of κ-casein and α-lactalbumin gene expression in lactating and non-lactating bovine mammary gland tissue samples

3.2.1 RNA isolation of bovine mammary gland tissue and complementary DNA (cDNA) synthesis

The total RNA was isolated from bovine mammary gland tissue using the TRIzol® Reagent with the PureLink™ Micro-to-Midi System (Invitrogen Life Technologies, Burlington, ON, Canada) according to the manufacturer’s recommendation. Approximately 100mg tissue was finely ground using mortar and pestle, and then homogenized with approximately 1 ml of TRIzol® Reagent. The lysate was incubated for 5 min at room temperature to allow complete dissociation of nucleoprotein complexes. Chloroform at the volume of 0.2 ml per 1 ml of TRIzol® Reagent was added to the lysate and the tube was vigorously mixed by shaking with hand for 15 s, followed by 2-3 min incubation at room temperature. The mixture was then centrifuged at 12,000×g for 15 min at 4°C for separation into different phases with an aqueous colorless upper layer containing the RNA. Approximately 400 µl of this upper phase was transferred to a fresh tube and an equal volume of 70% ethanol was added to obtain a final ethanol concentration of 35%. For binding of the RNA, up to 700 µl of the prepared sample was transferred to an RNA Spin Cartridge pre-inserted in a collection tube and centrifuged at 12,000×g for 15 s at room temperature. The bound RNA was washed using 700 µl of Wash Buffer I and then twice using 500 µl of Wash Buffer II with ethanol. For RNA elution, the cartridge was inserted into a supplied RNA Recovery Tube and 30 µl of RNase-free water was added to the center of the spin cartridge. After 1 min incubation at room temperature, the spin cartridge was centrifuged at 12,000×g for 2 min to collect the elute containing the purified RNA. RNA qualification and quantification were done by
UV absorption measurement at 260nm using NanoDrop spectrophotometer. Approximately 1 µg of the RNA sample was used in cDNA synthesis. This reaction was performed using SuperScript™III First-Strand Synthesis SuperMix for qRT-PCR Kit (Invitrogen Life Technologies, Burlington, ON, Canada) according to the manufacturer’s recommendation. The RT reaction was performed in a 20 µl volume containing 10 µl of 2×RT Reaction Mix, 2 µl of RT Enzyme Mix and variable volumes of RNA sample depending on the concentration of each sample (approximately 1 µg). The reaction was first incubated at 25°C for 10 min, and then at 50°C for 30 min. The reaction was terminated at 85°C for 5 min. 2 U of RNase H was added and the reaction was incubated at 37°C for 20 min to remove any unreacted residual RNA. The cDNA sample was used in real-time PCR to quantify the expression of the two target genes.

3.2.2 Primer design and reverse transcription PCR (RT-PCR) of the target genes

Primers were designed for the amplification of milk protein genes and the endogenous reference gene, Glyceraldehyde-3-phosphate-dehydrogenase (GAPDH). The GAPDH was used for data normalization. RT-PCR was conducted using the designed primers listed in Table 2. The amplification was carried out in a final volume of 25 µl containing 1×PCR buffer, 1.5mM MgCl₂, 0.2mM dNTP, 0.1 µM of each primer, 1U Taq polymerase (Invitrogen Life, Technologies, Burlington, ON, Canada) and 2 µl of cDNA synthesized from RT reaction. The PCR reactions were first incubated at 94°C for 2 min, and then cycled 30 times with denatured at 94°C for 45 s, annealing at 58°C for 45 s, and extension at 72°C for 1 min. A final extension step at 72°C for 5 min was added at the end. The RT-PCR products were verified on an agarose gel with ethidium bromide staining.
Table 2. Primers used in RT-PCR for amplification of bovine κ-casein and α-lactalbumin cDNA

<table>
<thead>
<tr>
<th>Target sequences</th>
<th>Primer sequences (5’ to 3’)</th>
<th>Product (bp)</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>κ-casein cDNA</td>
<td>F: ATTGCTAGTGCGAGCCTAC</td>
<td>118</td>
<td>NM174294</td>
</tr>
<tr>
<td></td>
<td>R: CTTGGACTGTGTGATCTCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-lactalbumin</td>
<td>F: AACATCTCCTGTGACAAGTTTC</td>
<td>141</td>
<td>NM174378</td>
</tr>
<tr>
<td>cDNA</td>
<td>R: CACAGAGCCACTGATCCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAPDH cDNA</td>
<td>F: CCCAGAAGACTGTGGATG</td>
<td>121</td>
<td>NM001034034</td>
</tr>
<tr>
<td></td>
<td>R: AGCTCAGGGATGACCTTG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F: forward primer; R: reverse primer.

3.2.3 Quantification of target genes by qRT-PCR

Real-time qRT-PCR was performed for both of the milk protein genes and for the endogenous reference gene with the cDNA templates synthesized from the total RNA using the Thermal Cycler Stratagene Mx3000P machine with Platinum® SYBR® Green qPCR SuperMix UDG Kit (Invitrogen Life Technologies, Burlington, ON, Canada). The reactions were prepared in duplicates with a final volume of 25 µl containing 12.5 µl of Platinum® SYBR® Green qPCR SuperMix, 0.5 µl of ROX Reference Dye, which is a passive reference dye used to normalize SYBR® Green fluorescent signal and provides a stable baseline in multiplex reactions (User instruction, Platinum® SYBR® Green qPCR SuperMix UDG, Invitrogen, 2006), 0.2 µM of each primer and 2 µl of cDNA. The qPCR condition was initiated at denaturation step at 95°C for 10 min, 40 cycles of denaturation...
at 95°C for 30 s, annealing at 58°C for 1 min, and extension at 72°C for 30 s. Following the amplification, a dissociation melting curve analysis was conducted with programming the PCR machine from 55°C to 95°C to detect the possible nonspecific products. Any distinguishable extra peaks present at about 75°C in the graph would be considered to be the nonspecific qPCR products.

Fluorescence threshold was determined by default method at 32.5% with Stratagene software for Mx3000P Real-time PCR machine. To validate the comparative quantification Ct (ΔCt) method, validation experiment was conducted by amplifying the serially diluted cDNA of the target genes and GAPDH gene from the same sample and plotting the log of different concentrations of cDNA from serial dilutions against the ΔCt (Ct target − Ct reference). The slope of the curve was used to determine the amplification efficiencies of the target genes and endogenous reference gene, whose absolute value should be <0.1 when a validation experiment is considered to be passed. The averages of Ct values of each sample generated in duplicate qPCR reactions were used in calculating the fold-changes of gene expression in the lactating samples compared to the non-lactating samples. The calculation was done as follows:

$$\text{Fold-change} = 2^{-\Delta\Delta C_t}$$

where

$$\Delta\Delta C_t = \Delta C_t_{\text{test sample}} - \Delta C_t_{\text{control}},$$

and

$$\Delta C_t = C_t_{\text{target in sample}} - C_t_{\text{reference in sample}}$$

$$\Delta C_t_{\text{control}} = C_t_{\text{target in control}} - C_t_{\text{reference in control}}$$
3.3 Statistical Analysis

A one way ANOVA was performed to analyze the effects of DNA methylation in milk protein gene promoter and the stage of lactation on the expression of bovine κ-casein and α-lactalbumin genes (SYSTAT Version 6.0.1). Due to the lack of mammary gland samples, the animals were classified into two categories, namely lactating and non-lactating. The methylation patterns of the target promoter were grouped into three categories, namely methylation at known transcription factor binding domains, methylation at known transcription factor non-binding domains and the absence of cytosine methylation.
Chapter VI

Results

4.1 DNA methylation in the core promoter regions of bovine κ-casein and α-lactalbumin genes

4.1.1 Verification of the bisulphite PCR products

After bisulphite modification and purification, the DNA samples were amplified with the modified primer sets using regular PCR protocol. The products were verified on agarose gel and the target fragments were found to be amplified with the expected sizes, specifically 207 bp for the κ-casein and 489 bp for the α-lactalbumin genes (Fig. 9). No nonspecific products were observed in electrophoretic analysis of the amplicons.
(a)

(b)
Fig. 9. PCR products of target bovine κ-casein and α-lactalbumin gene core promoter regions after bisulphite conversion of the genomic DNA. (a) Lane 1, DNA ladder (New England Biolab Inc., Pickering, ON, Canada); Lanes 2 to 4, amplification of bovine κ-casein gene target region from different samples. (b) Lanes 1 to 3, amplification of bovine α-lactalbumin gene target region from different samples; Lane 4, DNA ladder.

4.1.2 Cloning

Each purified bisulphate PCR product was cloned into a pDRIVE vector following UA ligation strategy and at least 10 colonies from each individual sample of each gene were selected for identification of the insert. Initially the colony PCR was performed to screen the positive clones, which was later confirmed by the plasmid PCR (Fig.10 and 11). The recombinant plasmid for each sample was sequenced from both directions of the insert with SP6 and T7 primer.
Fig. 10. Colony PCR products of target bovine κ-casein and α-lactalbumin gene core promoter regions. (a) Lanes 1, 5, 8 and 10 are positive colonies with the insertion of bovine κ-casein gene target region; Lane 6, DNA ladder (New England Biolab Inc., Pickering, ON, Canada). (b) Lanes 2, 4, 5, 9 and 10 are positive colonies with the insertion of bovine α-lactalbumin gene target region; Lane 6, DNA ladder.
Fig. 11. Plasmid PCR products of target bovine \(\kappa\)-casein and \(\alpha\)-lactalbumin gene core promoter regions. (a) Lane 1, DNA ladder (New England Biolab Inc., Pickering, ON, Canada); Lanes 2 to 7, the amplification of bovine \(\kappa\)-casein positive plasmids. (b) Lanes 1 to 6, the amplification of bovine \(\alpha\)-lactalbumin positive plasmids; Lane 7, DNA ladder.
4.1.3 Methylated cytosines found in target bovine κ-casein gene core promoter region

The target sequences of each gene for all samples were aligned and compared with the original target sequences available at NCBI. The methylation pattern of the animals was classified into three groups, namely methylation at known transcription factor binding domains, methylation at known transcription factor non-binding domains and the absence of cytosine methylation.

A number of methylcytosines were identified in the target core promoters of individual animals. In the target κ-casein core promoter (Table 3), cytosine methylation was observed at Ap1 site of three lactating and one dry animals, whereas two lactating animals showed methyl cytosine at the transcription factor IIB recognition element downstream relative to TATA box. Additionally, methylated cytosines were found outside those binding domains, but within the target region, in four lactating and one dry animal. In one lactating animal, methylated cytosine was detected both at core promoter binding and non-binding domains. Interestingly, in one cow of 240 days lactating, two methylated cytosines were found at Ap1 binding site. During non-lactating (dry period) period, two animals showed methylated bases at the core promoter of which one 53-day dry (7 day prepartum) animal revealed methylation at the Ap1 binding site. Nine animals were found to be unmethylated in the target region, only three of which were lactating (3/12 lactating cows) while the majority of these cows were non-lactating (6/8 non-lactating cows).
Table 3. Positions of methylated cytosine found in target bovine κ-casein core promoter region of different samples.

<table>
<thead>
<tr>
<th>Animal number</th>
<th>Physiological stages</th>
<th>Status of methylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1229</td>
<td>L14</td>
<td>175 &amp; 176</td>
</tr>
<tr>
<td>#9495</td>
<td>L14</td>
<td>99&lt;sup&gt;A&lt;/sup&gt;</td>
</tr>
<tr>
<td>#9668</td>
<td>L14</td>
<td>43</td>
</tr>
<tr>
<td>#9883</td>
<td>L14</td>
<td>Yes</td>
</tr>
<tr>
<td>#1204</td>
<td>L90</td>
<td>120&lt;sup&gt;T&lt;/sup&gt;</td>
</tr>
<tr>
<td>#1213</td>
<td>L90</td>
<td>99&lt;sup&gt;A&lt;/sup&gt;</td>
</tr>
<tr>
<td>#1231</td>
<td>L90</td>
<td>Yes</td>
</tr>
<tr>
<td>#9833</td>
<td>L90</td>
<td>104 &amp; 159</td>
</tr>
<tr>
<td>#1005</td>
<td>L240</td>
<td>Yes</td>
</tr>
<tr>
<td>#1026</td>
<td>L240</td>
<td>53</td>
</tr>
<tr>
<td>#1086</td>
<td>L240</td>
<td>97&lt;sup&gt;A&lt;/sup&gt; &amp; 110&lt;sup&gt;A&lt;/sup&gt; 32 &amp; 47</td>
</tr>
<tr>
<td>#1203</td>
<td>L240</td>
<td>114&lt;sup&gt;T&lt;/sup&gt;</td>
</tr>
<tr>
<td>#7878</td>
<td>D53</td>
<td>97&lt;sup&gt;A&lt;/sup&gt;</td>
</tr>
<tr>
<td>#7460</td>
<td>D53</td>
<td>Yes</td>
</tr>
<tr>
<td>#7674</td>
<td>D53</td>
<td>Yes</td>
</tr>
<tr>
<td>#7465</td>
<td>D53</td>
<td>Yes</td>
</tr>
<tr>
<td>#7479</td>
<td>D35</td>
<td>Yes</td>
</tr>
<tr>
<td>#7692</td>
<td>D35</td>
<td>Yes</td>
</tr>
<tr>
<td>#7863</td>
<td>D7</td>
<td>145</td>
</tr>
<tr>
<td>#7688</td>
<td>D7</td>
<td>Yes</td>
</tr>
</tbody>
</table>
L= Lactating stage; D= Dry stage; Numbers under the status of methylation indicate the positions of the methylcytosine found in the target core promoter region of different samples. $M^{b}$ represents cytosine methylation at the known transcription factor binding domains; $M^{b}$ represents cytosine methylation at the known transcription factor non-binding domains; $M^{u}$ represents the absence of cytosine methylation. Superscripts indicate the type of transcription factor binding site in the target core promoter region where A: AP1 binding site; T: cytosine sequence adjacent to TATA box.

4.1.4 Methylated cytosines found in target bovine $\alpha$-lactalbumin gene core promoter region

The sequence analysis of bisulphite converted region showed the presence of methylation of cytosine residues at various locations in target bovine $\alpha$-lactalbumin gene core promoter (Table 4). Methylated cytosines were found at the progesterone binding site of three lactating and non-lactating animals each, of which two or more cytosine residues were methylated in two lactating and two dry ones. The methylcytosine at other binding sites like CAAT element, sequence adjacent TATA box and TGGCA protein bite was observed in two lactating animal each while one dry animal of 53 days showed methylation at TGGCA protein binding site. All the non-lactating samples showed methylation at the target region in this study. On the other hand, the absence of cytosine methylation was observed in only three of the lactating animals.
Table 4. Positions of methylated cytosine found in target bovine α-lactalbumin gene core promoter region of different samples.

<table>
<thead>
<tr>
<th>Animal number</th>
<th>Physiological stages</th>
<th>Status of methylation</th>
<th>M^b</th>
<th>M^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1229</td>
<td>L14</td>
<td>413^P, 434^P</td>
<td>274 &amp; 449</td>
<td></td>
</tr>
<tr>
<td>#9495</td>
<td>L14</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>#9668</td>
<td>L14</td>
<td>85^P, 386^C &amp; 461^T</td>
<td>66, 225, 226 &amp; 313</td>
<td></td>
</tr>
<tr>
<td>#9883</td>
<td>L14</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>#1204</td>
<td>L90</td>
<td></td>
<td>228</td>
<td></td>
</tr>
<tr>
<td>#1213</td>
<td>L90</td>
<td></td>
<td>228</td>
<td></td>
</tr>
<tr>
<td>#1231</td>
<td>L90</td>
<td>408^P, 417^P &amp; 422^P</td>
<td>114, 268, 274 &amp; 285</td>
<td></td>
</tr>
<tr>
<td>#9833</td>
<td>L90</td>
<td></td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>#1005</td>
<td>L240</td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>#1026</td>
<td>L240</td>
<td></td>
<td>169 &amp; 215</td>
<td></td>
</tr>
<tr>
<td>#1086</td>
<td>L240</td>
<td></td>
<td>38 &amp; 41</td>
<td></td>
</tr>
<tr>
<td>#1203</td>
<td>L240</td>
<td>399^G</td>
<td>191 &amp; 455</td>
<td></td>
</tr>
<tr>
<td>#7878</td>
<td>D53</td>
<td></td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>#7460</td>
<td>D53</td>
<td>400^G</td>
<td>219</td>
<td></td>
</tr>
<tr>
<td>#7674</td>
<td>D53</td>
<td></td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>#7465</td>
<td>D53</td>
<td></td>
<td>66, 136 &amp; 215</td>
<td></td>
</tr>
<tr>
<td>#7479</td>
<td>D35</td>
<td>85^P</td>
<td>66, 144 &amp; 459</td>
<td></td>
</tr>
<tr>
<td>#7863</td>
<td>D7</td>
<td>91^P &amp; 447^P</td>
<td>41 &amp; 452</td>
<td></td>
</tr>
<tr>
<td>#7688</td>
<td>D7</td>
<td></td>
<td>144 &amp; 229</td>
<td></td>
</tr>
</tbody>
</table>
L= Lactating stage; D= Dry stage; Numbers under the status of methylation indicate the positions of the methylcytosine found in the target core promoter region of different samples. $M^b$ represents cytosine methylation at the known transcription factor binding domains; $M^b$ represents cytosine methylation at the known transcription factor non-binding domains; $M^u$ represents the absence of cytosine methylation. Superscripts indicate the type of transcription factor binding site in the target core promoter region where C: CAAT element; G: TGGCA binding site; P: progesterone binding site; T: cytosine sequence adjacent to TATA box.

4.2 Quantification of $\kappa$-casein and $\alpha$-lactalbumin gene expressions in the lactating and non-lactating bovine mammary gland tissues

4.2.1 RT-PCR results of the target genes and the endogenous reference gene

The cDNA synthesized from different samples of mRNA was amplified using conventional RT-PCR to identify the target fragments (Fig.12). PCR products were resolved on an agarose gel and finally, desired products were verified with the DNA marker.
Fig. 12. RT-PCR products of bovine $\kappa$-casein, $\alpha$-lactalbumin and GAPDH genes. Lanes 1 and 2, amplification of GAPDH partial cDNA (product size, 121bp); Lanes 3 and 4, amplification of bovine $\kappa$-casein partial cDNA (product size, 118bp); Lanes 5 and 6, amplification of bovine $\alpha$-lactalbumin partial cDNA (product size, 141bp); Lane 7, DNA ladder (New England Biolab Inc., Pickering, ON, Canada).

4.2.2 Validation experiments in the qPCR

The cDNA samples synthesized from total RNA isolated from bovine mammary gland tissues were diluted at serial concentrations and used as templates in qPCR validation experiments to test the amplification efficiencies of the target genes and the GAPDH gene (Fig.13). The Ct values for each sample for both target and GAPDH genes were generated from the qPCR output file and the Ct values for target genes were normalized by subtracting the Ct values of GAPDH gene for respective samples. The magnitude of slope with an absolute value is less than 0.1, indicating that amplification efficiencies of both the target and reference genes are approximately the same.
Fig. 13. Validation experiment. (a) Validation test of bovine κ-casein gene expression in the qPCR. (b) Validation test of bovine α-lactalbumin gene expression in the qPCR.
4.2.3 Quantification of the target genes' expression

As qPCR was performed in duplicates for each sample, the mean Ct values were calculated for each sample for both the target and GAPDH genes. Approximate linear prediction equations were obtained for bovine \( \kappa \)-casein, \( \alpha \)-lactalbumin and GAPDH genes, and the accuracy of prediction was estimated in terms of \( R^2 \) value. The PCR amplification efficiencies of the three genes were approximately 100% as the slope of the curve was approximate to -3.32, which indicated high quality PCR reactions (Fig. 14). Amplification plots were created for each sample in duplicates to determine the reproducibility of amplification of each gene (Fig. 15). The standard curves and amplification plots showed very high accuracy of estimating expression profiles for the target and reference genes and accordingly, mean Ct values were calculated for the animals at different physiological stages.

The Ct values for bovine \( \kappa \)-casein gene in lactating animals varied from 7.3 to 11.5 while the magnitudes in dry animals ranged from 9.5 to 20.1. In particular, the mean Ct values of 14, 90 and 240 days lactating animals were 8.9, 7.9 and 7.5 respectively, whereas in 7, 35 and 53 days dry animals the mean Ct values were 10.8, 11.5 and 16.3 respectively (Table 5). For bovine \( \alpha \)-lactalbumin gene, the Ct values in lactating and dry animals were found to be in the range of 8.0 to 12.6 and 12.8 to 22.2 respectively. Specifically, in 14, 90 and 240 days lactating animals the average Ct values were estimated as 9.5, 9.4 and 8.5 respectively, but in 7, 35 and 53 days dry animals, these values were 18.8, 14.7 and 18.8 respectively (Table 6).
(a) Amplification efficiency of GAPDH

\[ y = -3.4967x + 35.826 \]

\[ R^2 = 0.9973 \]

Log Input of cDNA

(b) Amplification efficiency of bovine \( \kappa \)-casein

\[ y = -3.3933x + 19.398 \]

\[ R^2 = 0.9927 \]

Log Input of cDNA
Fig. 14. qPCR amplification efficiencies of bovine κ-casein, α-lactalbumin and GAPDH genes. (a) Standard curves of GAPDH gene amplification. (b) Standard curves of bovine κ-casein gene amplification. (c) Standard curves of bovine α-lactalbumin gene amplification.
Fig. 15. Amplification plots of bovine $\kappa$-casein, $\alpha$-lactalbumin and GAPDH genes. Results from duplicate reactions were used to present the amplification plots. Y-axis represents the changes in fluorescence signal per unit of time; X-axis represents the Ct value. (a) Amplification plot of the GAPDH gene. (b) Amplification plot of bovine $\kappa$-casein gene. (c) Amplification plot of bovine $\alpha$-lactalbumin gene.
Table 5. Threshold cycle values of GAPDH gene and bovine $\kappa$-casein gene

<table>
<thead>
<tr>
<th>Physiology stages</th>
<th>Group Ct values of GAPDH gene</th>
<th>Group Ct values of bovine $\kappa$-casein gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means ± SD</td>
<td>Means ± SD</td>
</tr>
<tr>
<td>L14</td>
<td>22.25 ± 1.31</td>
<td>8.94 ± 1.80</td>
</tr>
<tr>
<td>L90</td>
<td>21.63 ± 0.49</td>
<td>7.94 ± 0.67</td>
</tr>
<tr>
<td>L240</td>
<td>21.42 ± 0.73</td>
<td>7.50 ± 0.12</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>21.77 ± 0.84</strong></td>
<td><strong>8.26 ± 0.87</strong></td>
</tr>
<tr>
<td>D53</td>
<td>23.30 ± 2.63</td>
<td>16.33 ± 3.40</td>
</tr>
<tr>
<td>D35</td>
<td>19.56 ± 0.77</td>
<td>11.54 ± 1.61</td>
</tr>
<tr>
<td>D7</td>
<td>21.17 ± 0.94</td>
<td>10.80 ± 0.26</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>21.34 ± 1.45</strong></td>
<td><strong>13.08 ± 1.76</strong></td>
</tr>
</tbody>
</table>

L: Lactation stage; D: Dry stage.

Table 6. Threshold cycle values of GAPDH gene and bovine $\alpha$-lactalbumin gene

<table>
<thead>
<tr>
<th>Physiology stages</th>
<th>Group Ct values of GAPDH gene</th>
<th>Group Ct values of bovine $\alpha$-lactalbumin gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means ± SD</td>
<td>Means ± SD</td>
</tr>
<tr>
<td>L14</td>
<td>22.64 ± 1.33</td>
<td>9.48 ± 2.10</td>
</tr>
<tr>
<td>L90</td>
<td>21.98 ± 1.05</td>
<td>9.37 ± 0.51</td>
</tr>
<tr>
<td>L240</td>
<td>21.41 ± 0.91</td>
<td>8.53 ± 0.54</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>22.01 ± 1.09</strong></td>
<td><strong>9.13 ± 1.07</strong></td>
</tr>
<tr>
<td>D53</td>
<td>23.17 ± 2.87</td>
<td>18.84 ± 2.97</td>
</tr>
<tr>
<td>D35</td>
<td>21.15 ± 0.93</td>
<td>14.73 ± 2.74</td>
</tr>
<tr>
<td>D7</td>
<td>19.24 ± 1.06</td>
<td>18.84 ± 0.73</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>21.19 ± 1.62</strong></td>
<td><strong>17.45 ± 2.15</strong></td>
</tr>
</tbody>
</table>

L: Lactation stage; D: Dry stage.
4.2.4 Melting curve analysis

The specificity of the amplified products was verified by melting curve analysis and gel electrophoresis (Fig.16). The amplification with the designated primer sets showed a single peak approximately at 82°C for bovine κ-casein gene, at 88°C for bovine α-lactalbumin gene and at 88°C for the GAPDH gene, respectively. At a lower melting temperature of approximately 75°C, there was no distinguishable extra peak for each gene amplification in our analysis. This indicated the high specificity of each primer set used in this study.
Fig. 16. Dissociation analysis of bovine κ-casein, α-lactalbumin and GAPDH genes. Results from duplicate reactions were used to present the melting curves. The Y-axis represents the changes in fluorescence signal per time unit; The X-axis represents the temperature. Amplicons of each gene were resolved on agarose gel to verify the target fragments and any possible nonspecific products. Melting curve analysis was presented in (a) for the GADPH gene, (c) for bovine κ-casein gene and (e) for bovine α-lactalbumin gene. (b) Gel verification of the GAPDH gene where Lane 1, DNA ladder (New England Biolab Inc., Pickering, ON, Canada); Lanes 2 and 3, amplifications of the gene from duplicates. (d) Gel verification of bovine κ-casein gene where Lane 1, DNA ladder; Lanes 2 and 3, amplifications of the gene from duplicates. (f) Gel verification of bovine α-lactalbumin gene where Lane 1, DNA ladder; Lanes 2 and 3, amplifications of the gene from duplicates.
4.2.5 Effect of physiological stage on the target genes’ expression

The calculated ΔCt values were used for estimating fold change of expression between two groups of animals. The physiological stage showed significant (P<0.05) effect on expression of the κ-casein and α-lactalbumin genes in bovine mammary gland tissues. For bovine κ-casein gene, lactating animals expressed 34-folds more than the dry ones (Table 7). Among the lactating animals, the 240th day lactating cows showed 1.52 fold more expression than the 14th day lactating ones and 1.17 times more than the 90th day lactating individuals. Among the dry animals, the 35th day dry cows showed 1.02 fold more expression than the 7th day and 3.86 fold more than the 53rd day dry animals. The 7th day dry cows also expressed 3.75 times more than the 53rd day dry animals. Among the lactating group, the lactating cows showed more or less similar expression levels of this gene, while a gradually decreasing trend of gene expression was observed in the animals from the dry stage (Fig. 17).

In the case of bovine α-lactalbumin gene, the lactating animals showed approximately 515-fold more expression than the dry animals (Table 8). Within the lactating groups, the 14th day lactating animals expressed the gene 1.45 times more than the 90th day lactating ones and 1.214 fold more than those from 240th day lactation. On the other hand, the 240th day lactating cows showed 1.19 fold more expression than the 90th day lactating cows. Among the dry groups, the 35th day dry cows showed 2.19 times more expression than the 7th day dry cows and 1.12 fold more than the 53rd day dry ones. The comparison of expression between animals of 7th day and 53rd day dry samples revealed that the 53rd day dry animals showed 4.05 fold more expression than the 7th day dry cows (Fig. 17).
Table 7. Fold-change value of bovine κ-casein gene expression in lactating cows relative to dry cows.

<table>
<thead>
<tr>
<th>Physiology stage</th>
<th>Average group ΔCt values</th>
<th>ΔΔCt</th>
<th>Fold-change (2^-ΔΔCt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactating stage</td>
<td>-13.64 ± 0.72\textsuperscript{a}</td>
<td>-5.09</td>
<td>34.06</td>
</tr>
<tr>
<td>Dry stage</td>
<td>-8.55 ± 2.13\textsuperscript{b}</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\textsuperscript{a,b} Different superscripts indicate significance at P<0.05.

Table 8. Fold-change value of bovine α-lactalbumin gene expression in lactating cows relative to dry cows.

<table>
<thead>
<tr>
<th>Physiology stage</th>
<th>Average group ΔCt values</th>
<th>ΔΔCt</th>
<th>Fold-change (2^-ΔΔCt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactating stage</td>
<td>-12.88 ± 0.90\textsuperscript{a}</td>
<td>-9.01</td>
<td>515.56</td>
</tr>
<tr>
<td>Dry stage</td>
<td>-3.87 ± 1.95\textsuperscript{b}</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\textsuperscript{a,b} Different superscripts indicate significance at P<0.05.
Fig. 17. Effect of physiological stage on the expressions of target genes. (a) Gene expression of bovine κ-casein and α-lactalbumin during lactation and dry stages. Y axis represents the ΔCt values. X axis represents the physiological stages. Different superscripts indicate significance at P<0.05. (b) Trend of expression of bovine κ-casein and α-lactalbumin genes in different physiological stages. X axis represents different physiological stages. Y axis represents the ΔCt values of different physiological groups of samples.
4.2.6 Effect of DNA methylation on the target genes’ expression

DNA methylation did not show any significant effect on the expression of bovine κ-casein gene. The average ΔCt values in animals with cytosine methylation at known transcription factor binding and non-binding domains of the gene were estimated as 13.07 ± 1.85 and 11.82 ± 3.05 respectively, while animals with absence of cytosine methylation showed an average ΔCt value of 11.04 ± 3.21 (Table 9). For bovine α-lactalbumin gene, cytosine methylation in the target region revealed a significant effect (P<0.05) on the gene expression. The animals with unmethylated promoters showed 21 folds more expression than those with methylated promoters. Specifically, the animals with absence of promoter methylation showed 11 folds greater expression than the animals with methylation at the binding domains and 32 folds more expression than those with methylation at the non-binding domains (Table 10). However, the expression profile of the animals with cytosine methylation at the binding domains did not significantly differ from the individuals with cytosine methylation at the non-binding domains within the target region (Fig. 18).
Table 9. Fold-change values of bovine κ-casein gene expression in different groups of animals with differential methylation status in the target core promoter region

<table>
<thead>
<tr>
<th>Status of methylation</th>
<th>Average ΔCt</th>
<th>Fold-change (2^-ΔΔCt)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M^b - M^b</td>
</tr>
<tr>
<td>M^b</td>
<td>-13.07 ± 1.85</td>
<td></td>
</tr>
<tr>
<td>M^b</td>
<td>-11.82 ± 3.05</td>
<td>2.38</td>
</tr>
<tr>
<td>M^b</td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>M^b</td>
<td></td>
<td>0.58</td>
</tr>
</tbody>
</table>

M^b represents cytosine methylation at known transcription factor binding domains; M^b represents cytosine methylation at known transcription factor non-binding domains; M^b represents the absence of cytosine methylation. Fold-change values were calculated for the difference in gene expression between two groups of animals with different methylation status in the target region.

Table 10. Fold-change values of bovine α-lactalbumin gene expression in different groups of animals with differential methylation status in the target core promoter region

<table>
<thead>
<tr>
<th>Status of methylation</th>
<th>Average ΔCt</th>
<th>Fold-change (2^-ΔΔCt)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M^b - M^b</td>
</tr>
<tr>
<td>M^b</td>
<td>-9.81 ± 4.60^a</td>
<td></td>
</tr>
<tr>
<td>M^b</td>
<td>-8.31 ± 4.80^a</td>
<td>2.83</td>
</tr>
<tr>
<td>M^b</td>
<td></td>
<td>11.39</td>
</tr>
<tr>
<td>M^b</td>
<td></td>
<td>32.22</td>
</tr>
<tr>
<td>M^b</td>
<td>-13.32 ± 0.64^b</td>
<td></td>
</tr>
</tbody>
</table>
$M^{\text{th}}$ represents cytosine methylation at the known transcription factor binding domains; $M^b$ represents cytosine methylation at the known transcription factor non-binding domains; $M^U$ represents the absence of cytosine methylation. Fold-change values were calculated for the difference in gene expression between two groups of animals with different methylation status in the target region. $a^b$ Different superscripts indicate significance at $P<0.05$. 
**Fig. 18.** Trends of expression of bovine κ-casein and α-lactalbumin genes with differential methylation status of the target core promoter regions. The Y axis represents ΔCt values of different samples. The X axis represents the observed cytosine methylation in the target sequences, where $M^{+b}$: methylation at the known transcription factor binding domains; $M^{-b}$: methylation at the known transcription factor non-binding domains; $M^{U}$: the absence of cytosine methylation. Different superscripts indicate significance at $P<0.05$. 
Chapter V

Discussion

Bovine mammary gland tissues were collected from the lactating and non-lactating animals for detection of DNA methylation in the core promoter regions of κ-casein and α-lactalbumin gene. In the lactating group, 75% cows showed methylation of cytosine in the core promoter of κ-casein gene whereas only 12.5% dried-off animals exhibited the methylation at that region. For the α-lactalbumin gene, more than 80% lactating animals and all the non-lactating animals showed methylated cytosines in its core promoter region. To date several workers described the presence of methylated cytosines in the promoter of different milk protein genes including rat γ- and β-caseins (Johnson et al., 1985), rat κ-casein (Thompson and Nakhasi, 1985), bovine αs1-casine gene (Platenburg et al., 1996) and human lactoferrin (Teng et al., 2004); as well as onco and proto-onco genes (Serman et al., 2006; Carraway et al., 2007; Cortese et al., 2008; Ateeq et al., 2008; Pike et al., 2008; ). In the present study, we found methylated cytosine residues both in different known transcription factor binding and non-binding domains of the core promoters of κ-casein and α-lactalbumin gene in bovine. On the other hand, in some of the animals, there were no methylated cytosines in the same regions of interest. Finally, we measured the quantity of gene transcripts synthesized in mammary gland during specific physiological stages and have made an attempt to correlate that information with the occurrences of methylation in the target core promoter sites of the specific genes.

The mammary gland from lactating cows showed more than 500-fold greater abundance of α-lactalbumin mRNA than did that of dry cows. The κ-casein transcript
abundance was 34-fold greater in lactating than in non-lactating cows which is in line with earlier observation (Choi et al., 1988). In dairy cows, milking is typically terminated when cows are approximately 60 days before parturition. During this period, cows are in the final months of an about 287 day gestation so that counterbalancing effects of milk stasis and pregnancy result in the significant cell replacement (Capuco et al., 1997). Physiologically, involution in the mammary gland is slow with less loss of alveolar structure. Regression of the gland during dry period was also not totally uniform (Capuco et al., 1997). During this late stage of gestation when lactogenesis begins, mammary gland is functionally differentiating and preparing for successive lactation under the effects of pregnancy. It can be speculated that genes encoding different milk components are actively transcribed for copious milk synthesis and secretion. Our result shows low but detectable level of the target gene transcripts in the non-lactating mammary tissue. We therefore suggest that this small amount of target genes mRNA is from the newly renewed epithelial cells whose secretory activity has been initiated in the pregnant animals. In dry period κ-casein expression showed a declining trend from initial towards late stage of drying. Secreotions collected from prepartum in cows are rich in protein concentration reflecting the accumulation of immunoglobulins. As pregnancy progresses, the alveolar cells become more differentiated and the composition of mammary secretion changes (Akers, 2002). This partially explains the gradually decreasing expression of κ-casein as gestation continues because most the alveolar cells may participate in synthesizing and secreting colostrum. During the dry period the expression level of α-lactalbumin mRNA was much lower in tissue from cows during early versus later stages of the dry period. Because α-lactalbumin mRNA is a key determinant of milk volume and the expression of α-lactalbumin transcript and protein typically increases immediately prior to the onset of
copious milk production. Lactogenic effects of pregnancy explain the increased expression level of α-lactalbumin mRNA as parturition approaches. It has been reported that both the gene expression are regulated by different lactogenic hormones (Vonderhaar and Ziska, 1989). Some other mechanisms such as RNA processing and nucleocytoplasmic transport efficiency, and the rate of casein mRNA degradation may be involved in the induction of the casein genes during mammary differentiation (Choi et al., 1988).

The methylation patterns in both of the gene promoters were different during various physiological stages. One recent study advocated that the DNA methylation assays are not always accurate predictors of gene silencing (Pike et al., 2008). Moreover, DNA methylation in the promoter has been reported having a partial effect on regulation of the gene expression (Cortese et al., 2008). Other reports concluded that DNA methylation alone is often not sufficient to block transcription (Weih et al., 1991; Li et al., 2002), instead it is likely the nature of chromatin formed on a methylated template that renders its transcriptional inactivity. In our study, DNA methylation at biologically active domains of α-lactalbumin gene down-regulates its expression. This inverse relationship between promoter methylation and gene expression was in agreement to the findings on rat casein genes (Johnson et al., 1983; Thompson and Nakhasi, 1985). Additionally, DNA methylation patterns become different during different physiological stages. It is known that levels of several proteins and factors including hormones vary among different physiological stages and hormonal regulation of gene expression via DNA demethylation has been demonstrated in chicken vitellogenin gene (Wilks et al., 1982) and rat κ-casein gene (Thompson and Nakhasi, 1985). Tissue specific genes were revealed to undergo demethylation in their tissue of expression at a particular physiological stage during
development (Razin and Cedar, 1991). Further, at a certain point of development a methylation pattern of a gene is carved by a dynamic balance between methylation and demethylation reactions catalyzed by several DNA methyltransferases and demethyltransferases, which is responsive to physiological and pathological changes (Syzf, 2005). Thus we suggest that some other factors might have played certain roles in alternation of the cytosine methylation attributing differential expression patterns of the genes during various physiological stages.

The absence of methylation in the α-lactalbumin core promoter was associated with increased abundance of α-lactalbumin mRNA. However, variant DNA methylation patterns were observed in the target core promoter region in different animals regardless of physiological status. The mechanism by which DNA methylation interferes with gene transcription can be direct or indirect. Methylation of cytosine residues in the major groove of the double helix impedes the binding of basal transcription machinery or ubiquitous transcription factors to its recognition site in DNA (Bird and Wolffe, 1999). Alternatively, CpG methylation in promoter DNA excludes transcriptional machinery from the target sequence by influencing nucleosome stability or positioning to prevent the access of transcription factors to a promoter (Bird and Wolffe, 1999). In this study, all the non-lactating cows were methylated at various cytosines in the core promoter region of the α-lactalbumin gene. The methylation at the target region may inhibit the binding of transcription factors or other regulatory proteins at known binding domains in the core promoter, thereby reducing the transcription rate of the gene. Additionally, the gene expression of α-lactalbumin is hormonally controlled. Prolactin, together with glucocorticoids and insulin, promotes the synthesis of α-lactalbumin while progesterone inhibits production of this protein in milk (Vonderhaar and Ziska, 1989). As the dry
animals studied here were pregnant, the level of progesterone remained high, which upon binding on the respective binding domains of the promoter ultimately inhibited the expression of α-lactalbumin gene. In dry animals, a few methylated cytosines were found at the progesterone binding sites of the core promoter, which might partially interfere with the binding of progesterone to these sites. Especially both two 35-day prepartum cows were detected methylated at the progesterone receptor binding domains. Our data of gene quantification also showed that the expression of α-lactalbumin gene was higher in these two cows than that in the other dried off animals. It has been found that only mammary tissue from cows approaching 30 days before parturition was responsive to prolactin induction suggesting that at late pregnancy milk protein gene expression changes in responsiveness to endocrine stimuli in preparation for lactation (Sheehy et al., 2004). Therefore it can be speculated that DNA methylation prevents the binding of progesterone so that other lactogenic hormones might have more accessibility to the regulatory sequences. Furthermore, we observed a number of methylated cytosines in the respective binding domains of the core promoter of lactating cows. This may be a key factor inhibiting the binding of progesterone at these specific sites, thus accounting for the reduced sensitivity of lactating tissue to progesterone inhibition of milk secretion. However, methylation in other factor binding domains such as transcription factor IIB recognition element was also identified in the core promoter region of the gene. It has been reported that gene transcription can be initiated while it is methylated, which suggests that gene may be activated by factors which override the effects of methylation (Razin and Cedar, 1991). Besides the methylation discovered at the known transcription factor binding domains, we also observed some other methylcytosines beyond these binding domains in the core promoter region of interest. It was found that there is no
significant difference in the gene transcription levels between animals methylated at these binding and non-binding domains at the target region. The exact role of the methylation occurs at these non-binding domains has to be investigated further.

Methylation of rat κ-casein gene cDNA was studied in normal and neoplastic mammary gland and the expression of rat κ-casein gene was shown to be inversely correlated with the degree of methylation of this gene (Thompson and Nakhasi, 1985). The result presented here shows most of non-lactating mammary tissues are unmethylated at the core promoter region of the gene. Inversely, only few lactating cows were observed to be unmethylated at the same region. Differential levels of κ-casein expression concurrent to core promoter methylation either at known transcription factor binding or non-binding domains may suggest that only methylation at the promoter is not sufficient to regulate the gene expression during different physiological stages. Analysis of methylation in specific DNA sequences from several DNase I hypersensitive sites located upstream of the tyrosine aminotransferase gene showed that even demethylation of CpG dinucleotides occurred at upstream regions of the gene promoter, there was no protein binding to these sites in the nuclei (Weih et al., 1991). This observation suggested that the nucleosomal organization in chromatin structure dominantly determines the maintenance of an inactive gene state. The unmethylation of cytosine found in the non-lactating cows in this study may be not sufficient to confirm an actively transcribed structure of the gene attributing to the low level of the gene transcript. As discussed previously the animals were pregnant when the mammary gland samples were collected. The mammary epithelial cells were rapidly differentiating at late pregnancy and preparing for lactogenesis. The hormonal regulation of milk protein gene expression may determine the DNA methylation pattern of target promoter region by demethylation of specific cytosine
residues. On the other hand, a majority of the lactating cows were methylated in the core promoter region of κ-casein gene in the present study. During lactation the mammary gland undergoes little differentiation, but the levels of different lactogenic hormones change at different stages. The expression of the gene may be regulated by other factors such as lactogenic hormones whose effects may be more dominant than that of methylation. It has illustrated that the density of methyl-CpG near the promoter is a crucial determinant of repression mediated by DNA methylation. If the density of methylation is weak, which might decrease the affinity of methylcytosine binding protein such as MeCPs, gene transcription can be activated by a strong promoter strengthened by an enhancer without methylation loss (Boyes and Bird, 1992). Thus, another possible explanation why lactating cows expressed κ-casein gene without the loss of methylation might be that the density of methyl-CpG is low in its core promoter region because the promoter region of the gene is CpG-deficient.
Chapter VI

Concluding comments

Our study demonstrated that both bovine κ-casein and α-lactalbumin regulatory regions contained methylated cytosines in various known transcription factor binding and non-binding domains, whose distribution was different from lactating to non-lactating cows and showed a significant effect on expression of bovine α-lactalbumin while non-significant effect on expression of the κ-casein gene. However, to specify the exact role of DNA methylation on expression of these two important genes, physical dynamic study through epigenetic analysis including different biochemical assays needs to be conducted to study the interactions between DNA and chromatin in a concerted and efficient way. We suggest that not only methylation but demethylation events conferred by several intrinsic factors may also regulate the expression of bovine milk protein genes at a specific physiological stage.
References


Applied Biosystems, Foster City, Calif. 2002.


Appiled Biosystems. Foster City, Calif. 2002a. Real-time PCR vs traditional PCR.


with specific demethylase activity for mCpG DNA. Nature. 397: 579-583.


Thompson, M.D., H.L. Nakhasi. 1985. Methylation and expression of rat κ-casein gene in


