# Nucleosomal Organisation over the Ovine β-Lactoglobulin Gene

Simon A. Boa

Ph.D. Thesis
University of Edinburgh
1999

There was an Oxford Cleric too, a student,
Long given to logic, longer than was prudent;
The horse he had was leaner than a rake;
And he was not too fat, I undertake,
But had a hollow look, a sober stare;
The thread upon his overcoat was bare.
He had found no preferment in the church
And he was too unworldly to make search.
He thought far more of having by his bed
His twenty books all bound in black and red,
Of Aristotle and philosophy
Than of gay music, fiddles or finery.
Though a philosopher, as I have told,
He had not found the stone for making gold.
Whatever money from this friends he took
He spent on learning or another book
And prayed for them most earnestly, returning
Thanks to them thus for paying for his learning.
His only care was study, and indeed
He never spoke a word more than was need,
Formal at that, respectful in the extreme,
Short, to the point, and lofty in its theme.
The thought of moral virtue filled his speech
And he would gladly learn, and gladly teach.

The Canterbury Tales by Geoffrey Chaucer

(translated by Nevill Coghill)

# Declaration

I declare that the work presented in this thesis is my own and that the contribution of others has been clearly indicated.

# Abstract

The genetic material of all higher organisms from yeast to mammals is organised in the cell nucleus as a nucleoprotein complex called chromatin. The fundamental repeating unit of chromatin, which covers nearly the entire DNA, is the nucleosome. Each one comprises eight highly conserved protein subunits that sequester approximately 146bp of DNA. Nucleosomes facilitate the highly condensed packaging of DNA, most obvious in metaphase chromosomes, and also permit non-histone protein factors access to the DNA in order to facilitate DNA replication, transcription and repair.

For temporally and spatially specific gene activation to occur, chromatin remodeling factors, transcription factors and RNA polymerase and its associated factors must act in concert with the underlying nucleosome environment to effect transcription. In some instances, this has shown to be a complex relationship. Nucleosomes are stably positioned over transcription factor binding sites in some genes. This can prevent access and therefore repress gene activation. In other genes, a positioned nucleosome is required to wrap up DNA between separate transcription factor binding sites. Bringing the sites together allows the binding factors to act cooperatively in initiating transcription. Therefore, nucleosomes that are positioned over a specific DNA sequence can have an instrumental role in gene regulation.

To date, there have only been limited studies on the nucleosomal organisation of genes in their natural environment. The majority of these studies have concentrated on short regions of positioned nucleosomes spanning either repetitive DNA or the promoter regions of specific genes. However, nucleosome positioning over an entire gene domain may have a significant impact on its regulation and compaction. I have mapped the nucleosomal organisation over 10kb of a tissue specific, temporally regulated gene using the enzymatic probe, micrococcal nuclease and the chemical probe, cuprous phenanthroline. The ovine β-lactoglobulin (BLG) gene studied has a well characterised

developmental profile, a minimal transcriptional domain and has been used extensively as an expression cassette in transgenic animals to drive heterologous gene transcription.

When the gene is inactive, in the liver, it displays a tightly defined array of positioned nucleosomes that modulate between two specific phases over the gene domain. A similar, less tightly defined array is present when the gene is active, in the mammary gland, except over the promoter and actively transcribing regions. The same arrays are present over the BLG promoter region in transgenic mice in both active and inactive states. A monomer extension reaction provides *in vitro* evidence of the positioning signals that are determined by DNA sequence alone. These show an interesting correlation with the *in vivo* results.

A number of other milk protein genes have a similar pattern of key transcription factor binding sites over their promoter regions. If the nucleosome positions were conserved in these genes, with respect to these binding sites, it might suggest a role for positioned nucleosomes in their regulation. A total of three genes, each in two different organisms, have been analysed to test for a correlation.

# Acknowledgements


## Supervisors

My deep thanks go to Bruce Whitelaw for his superb and constant supervision throughout my project. His enthusiasm and insight into the field of mammary gland biology was a constant source of inspiration. Not only was he a supervisor, but also a friend.

Jim Allan, my excellent co-supervisor, provided me with copious quantities of food for thought and a welcoming environment in which to carry out my monomer extension experiments.

## Colleagues

Thanks go to Jean-Luc Vilotte for the goat β-casein construct pI2B and the mouse α-lactalbumin construct; Romi Pena and Armand Sanchez for the goat BLG plasmid pSX(12.5); Satish Kumar for the mouse β-casein plasmid pPolyIIIDβ and PPL Therapeutics for the human α-lactalbumin gene vector, pHA1.

Thanks also to Alisdair Fleming and Colin Davey for providing helper phage, Jim Allan for providing core histone proteins and Marietta Genchova for testing pBlueBLG (+) in the monomer extension assay. Melanie Edgar and Colin Davey were instrumental in teaching me this technique and providing troubleshooting solutions.

Matt Blair, Tom Skinner, Nick Gilbert and Richard Meehan assisted with computing problems. Francis Thomson and Carrie Owen looked after the mice with upmost skill. Janie Fenty and Viv Thomson washed all my dirty dishes with a smile rather than a dirty look!

## Family

Where to start?! This would not have been possible without the continued moral and financial support from my parents Andrea and Hunter and my brother Kevin. They

nursed and guided me through my 2 years of illness and persuaded me to continue with this work when all seemed lost . There are no better parents than mine.

## Friends

Many thanks to Colin Burnett (my social advisor), Diane Dinnis (for ice-cream), Ellen Tasker (my relationship advisor), Graham Davies (for beer), Ian Brown (for dirty dishes), Kenneth Dobie (for my ego), Elena Farini (for sanity), Richard Knox (for Ouzo), Alison Bean (for clean dishes), Silvia Alverez i Ribes (for glamour); Bruce, Jim, Steve Pells, Doug Strathdee, John Webster and Lorraine Young for proof reading; Tracey Shipman and Holly Bridge (for their affections) and all the PhD posse. Thanks to John Webster for his incessant humour and spare tenner.

A special thanks to Lorraine Young, Sally Buchanan, Chris Denning, Martin Reynard and Lin Wylie for driving me up the walls with beers to follow.

## Funding

The first three years of this project were funded by a BBSRC award from the Roslin Institute and the latter one and a half years were courtesy of her majesty's government.

# Contents

## *Chapter 2:* Materials and Methods

## *Chapter 6:* A 16kb BLG Transgene Organises a Normal Nucleosomal Array over the Promoter



## *Chapter 7: In vivo* and *in vitro* Investigation of Additional Milk Protein Gene Promoters which are Stat5 Regulated

## *Chapter 8:* Conclusions

## *Appendices*

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 The Discovery of Chromatin

The seminal event in molecular biology took place over a hundred years ago with the isolation of sperm nuclein from salmon in the Rhine in Germany (Miescher, 1874). This led to the isolation of associated basic 'histon' proteins in the same laboratory and the birth of the term 'chromatin'. Today this term defines DNA and its associated proteins, the majority of which are histone proteins. By 1885 it was believed that chromatin was the "substance which not only fertilises but also transmits the hereditary characteristics" (Hertwig, 1885). However, it took until the 1950s for it to be generally accepted that the DNA component, rather than the associated protein, was, by itself, the hereditary material (Avery *et al.*, 1944; van Holde, 1988). This realisation led to dramatic drop in the level of research into histone proteins, which only reversed in the early 1970s. From then until the present day the field has expanded rapidly as it has been shown that histones work in concert with DNA to co-ordinate correct cellular development and maintenance.

## 1.2 Histones

Histones are the fundamental units of chromatin structure. The five classes of histone protein, H1, H2A, H2B, H3 and H4, were discovered only after development of chromatographic fractionation methods and gel electrophoresis techniques. Their discovery can be largely credited to E.W.Johns and co-workers (Johns *et al.*, 1960; Bonner & Ts'o, 1964; Johns, 1964; Phillips, 1965) In 1973, Kornberg proposed a model which is still largely correct, explaining how the histone proteins interacted with each other and with the underlying DNA (Kornberg, 1974a; Kornberg *et al.*, 1974b). Further refinements were made shortly afterwards to include H1 (van Holde *et al.*, 1975; Noll *et al.*, 1977 ).

These early discoveries are still central to our understanding of histone interactions today. Histones H3 and H4 associate strongly via their histone-fold domains to form a stable $(H3/H4)_2$ tetramer. On to this, two stable heterodimers of H2A and H2B also bind, via a weaker interaction to form a histone octamer. This complex sequesters 146bp of DNA which is wrapped in 1.75 turns around the octamer in a shallow helical path. This results in the octomeric *core particle* or *nucleosome* (Figure 1) (Oudet *et al.*, 1975). The DNA does not form a uniform spiral around the octamer and is particularly deformed over the central three helical turns. This makes DNA packaging more efficient. Unstacking of adjacent base pairs induces this kink and produces a helical turn of 10.7bp, quite different from the average 10.2bp/turn found in the complete core particle (Wolffe *et al.*, 1995). Unstable TA steps are preferentially found at these sites which facilitates energetically favourable nucleosome formation (Simpson, 1991).

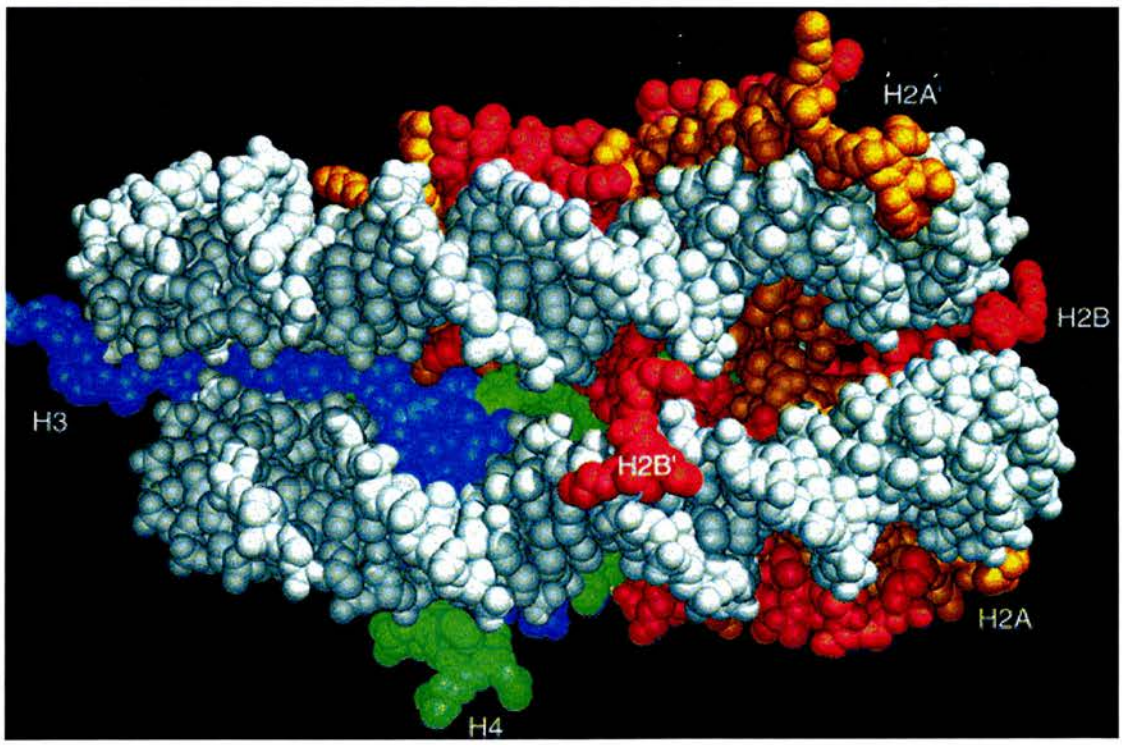



**Figure 1** The Nucleosome

(taken from Luger *et al.*, 1997: Crystal structure of the nucleosome core particle at 2.8 Å resolution)
The H2B (red) and H3 (blue) N-terminal tails pass through channels in the DNA superhelix (white) formed by aligned minor grooves. The other histones, H2A (gold) and H4 (green) are indicated.

Joining nucleosomes is a length of linker DNA which varies in size between species, tissue and gene activity. For example, in mouse liver and bone marrow it is about 55bp (Gaubatz *et al.*, 1979) and 64bp (Dean *et al.*, 1985), respectively. In frog liver and erythrocytes it is about 32bp and 40bp, respectively (Gottesfeld, 1980). Together, one length of this linker DNA and the core particle-associated DNA are known as the *nucleosome repeat length* which total approximately 200bp (Figure 2).

The 'linker' histone, H1, binds core particles and 20bp of the linker DNA exiting from them (Travers, 1999) to protect 168bp from digestion (Muyldermans & Travers, 1994). This association of a nucleosome with H1 and the protected DNA is called a *chromatosome*.

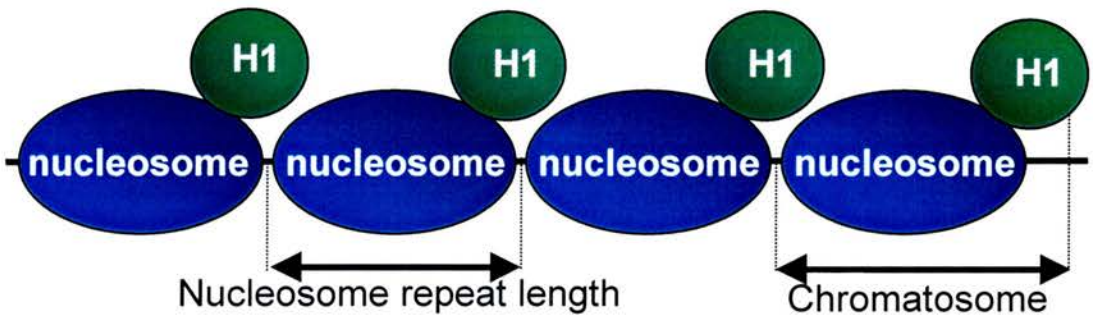



**Figure 2** One Nucleosome Repeat Length and a Chromatosome

The physical properties of nucleosomes depend on the monovalent and multivalent ion concentration in solution (Clark *et al.*, 1990), histone variants and histone modification states (**1.3**). The nucleosome has a disc-shaped structure of 5.6nm in height by 11nm in diameter when bound to DNA and has a mass of 206KDa. This structure has been refined to 2.8Å resolution using recombinant unmodified histone proteins from *Xenopus laevis* at physiologically relevant salt concentrations. The 3-dimensional structure can be viewed at:

http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?uid=8530&form=6&db=t&Dopt=s
using the Cn3D software at:
http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3dwin.html

### 1.2.1 Histone Domain Structure

The central C-terminal histone-fold domains of all four core histone proteins possess a high level of structural homology. The domain is formed through the interaction of three α-helixes connected by two loops, which forms a crescent-shaped heterodimer through the interaction of H3 and H4, and also H2A and H2B (Figure 3). Only the appropriate heterodimers are formed owing to internal packing constraints that are mainly caused by the different sizes of the first loop in the domain. Each histone-fold motif induces a 140° bend over 27 to 28bp of DNA as it arcs around their long axes. In total, this accounts for the folding of 121bp of DNA, including 4bp of linker between these units. The remaining DNA is bound by the H3 αN helix extension of the histone-fold to encapsulate the total 146bp. Binding is primarily to the phosphodiester backbone via hydrogen-bonding interactions (Luger *et al.*, 1997).

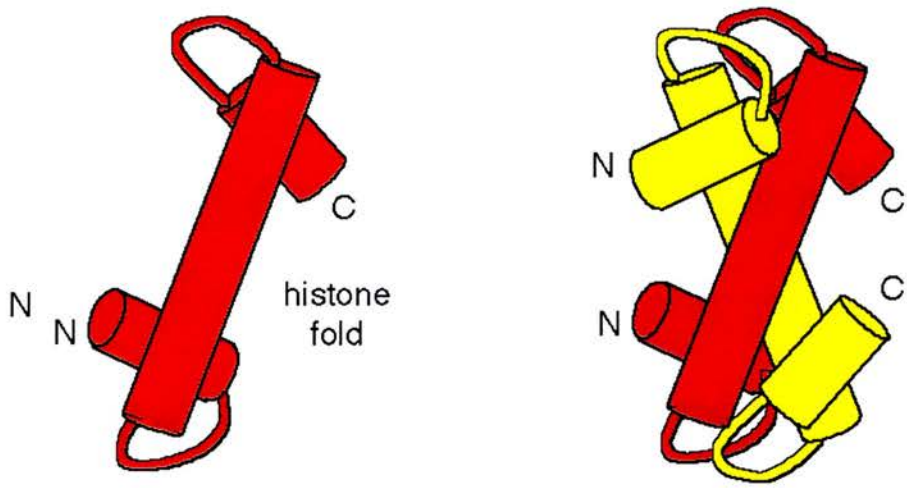



**Figure 3** The Histone Dimerisation Motif

Three α-helices linked by two loops forms a histone fold. The interaction of two histone folds generates a handshake motif between two different histone proteins. Courtesy of D.Pruss

The N-terminal tail regions of both H3 and H2B have random-coil segments that pass between the gyres of the DNA superhelix. The outer N-terminal tail regions of both H3 and H2B, which are located on the outside of the DNA superhelix, appear to be more flexible and probably interact with adjacent nucleosomes to assist in higher order

chromatin formation interactions (Luger *et al.*, 1997). They contain highly basic amino acid residues which are targets for acetylation (**1.3.1**).

The N-terminal domain of H2A passes over the DNA along a minor groove. The amino acid-base pair interaction within this groove may be important in modulating the nucleosomal association with particular DNA sequences. The H4 N-terminal tails appear to form divergent structures, although this may be because of a limitation in the crystallisation technique, resulting in diffuse electron density clouds. The only well-resolved region spanned from residues 20 to 26. It forms a strong association with a conserved domain of the H2A/H2B dimer on an adjacent nucleosome. This may be important in stabilising higher order structures.

### 1.2.2 Conservation Throughout Evolution

The length and structure of all core histone proteins have been generally conserved throughout evolution. This is most striking when comparing calf and pea H4, which differ only at two out of 102 residues (DeLange *et al.*, 1969). H3 shows a similarly high level of conservation throughout the protein sequence. Moreover, almost all H3 and H4 sequences examined to date have exactly defined lengths of 102 and 135 residues, respectively. Such outstanding consistency must reflect their critical role in nucleosome structure.

H2A and H2B are more divergent and even more-so is histone H1, which shows quite a low level of conservation, as compared with other proteins (Isenberg, 1978) (Figure 4). Interestingly, the sequence divergence is not evenly distributed along these protein sequences. Their lysine-arginine-rich N-terminal domains are far more divergent that the core of each molecule, with the exception of several reversibly modifiable residues. This may reflect the role of the core of these proteins in nucleosome structure. Deletion of the H2A and H2B N-terminal tails does not impair nucleosome formation and therefore these may be expected to be more divergent throughout evolution (Bohm *et al.*, 1984).

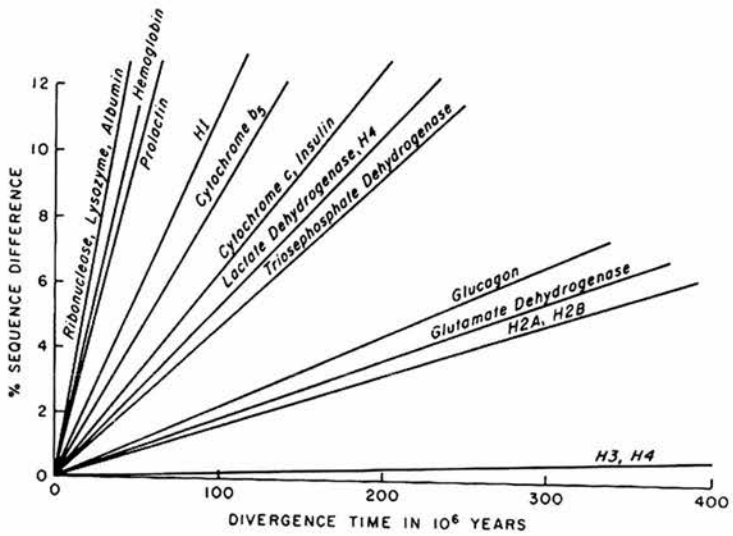



**Figure 4** Relative Conservation of Histone Proteins
Taken from Isenberg *et al.*, 1978. An approximate comparison for evolutionary rates in higher eukaryotes

## 1.3 Chromatin Modifications

Histones can be post-translationally modified with the covalent addition of acetyl-, phosphoryl-, methyl-, glycosyl-, ADP-ribosyl- and ubiquitin groups. Of these, acetylation is the most thoroughly studied and is strongly linked with transcriptional activation (Wolffe *et al.*, 1996; Grunstein, 1997). Moreover, transcription can be modulated by different histones types: histone *variants* are differentially expressed at different points in the cell cycle and the expression of *replacement* histones is upregulated during cellular differentiation. DNA methylation can also have a profound affect on transcription through the formation of an altered chromatin structure and chromatin remodelling machines can make chromatin structures more accessible to trans-acting factors by remodelling or sliding nucleosomal arrays. These points are discussed in more detail below.

### 1.3.1 Histone Acetylation

Acetylation has been observed in the four core histones in all animal and plant species examined (Candido & Dixon 1972). All the core histones can be reversibly

acetylated at the $NH_3^+$ group of lysine residues located exclusively at the N-terminal domain. Introduction of the acetyl group by histone acetyl-transferases (HATs) neutralises the residue's positive. This process is reversed by histone deacetylases (HDACs).

It has recently been discovered that some transcriptional regulators possess HAT activity. The first to be analysed in eukaryotes was the *Saccharomyces cerevisiae* transcriptional co-activator Gcn5 which is required for the expression of several genes *in vivo* (Chiang *et al.*, 1996; Georgakopoulos *et al.*, 1992; Martens *et al.*, 1996; Pollard *et al.*, 1997). It forms part of the Ada and SAGA complexes (Grant *et al.*, 1997). These may act as a bridge between trans-activators and general transcription factors since Ada2, which is present in both complexes, has been demonstrated to interact with certain DNA-binding activators and TATA-binding proteins (Barlev *et al.*, 1995; Silverman *et al.*, 1994).

The significance of this finding was reinforced by similar results obtained from work on mammalian transcription factors (Grunstein, 1997; Struhl, 1998), when p300/CBP was identified (Ogryzko *et al.*, 1996). It also possesses HAT activity and interacts with numerous DNA-binding regulatory proteins to mediate a broad range of effects on differentiation (Kawasaki *et al.*, 1998), transcription (Jenster *et al.*, 1997), the cell cycle (Ait *et al.*, 1998), viral activation (Benkirane *et al.*, 1998) and disease (Giles *et al.*, 1998).

Several aspects of nucleosome function can be modified by lysine acetylation. Most importantly, the affinity of histone N-terminal domains for DNA is significantly reduced (Hong *et al.*, 1993), demonstrated by increased DNaseI sensitivity. Acetylation may allow greater accessibility of transcription factors (Wolffe, 1995). Moreover, increased acetylation lessens the negative linking number of closed circular DNA *in vitro* (Norton *et al.*, 1989). Unacetylated histones in the same assay induce a more negative linking number on the closed circular DNA. This gyrase-like action may also effect transcription factor binding to the DNA. The association of nucleosomes with other non-histone proteins is also influenced by acetylation. For instance, the yeast silencing information regulators Sir3 and Sir4, and the general repressor Tup1, interact

preferentially with under-acetylated histone N-terminal domains (Edmondson *et al.*, 1996; Hecht *et al.*, 1995) .

*In vivo*, the correlation between lysine acetylation and transcriptional activation appears to be more elaborate. The acetylation of the N-terminal domain of histone H4 occurs at lysine residues K5, K8, K12 and K16. These residues are acetylated in a preferred order in mammalian cells, namely K16, followed by K12 and K8, followed by K5 (Turner *et al.*, 1989). In pol I and II genes, all sites are acetylated at equal levels with respect to bulk chromatin except K16, which is enriched. In non-coding, simple repeat heterochromatin, all sites are under-acetylated, but only K8 is strongly under-acetylated; K12 and K16 are acetylated only slightly below that of euchromatin (Johnson *et al.*, 1998). The acetylation patterns also differ in non-mammalian species such as *Drosophila melanogaster* (Turner *et al.*, 1992) and *Saccharomyces cerevisiae* (Braunstein *et al.*, 1996). In the latter case, the acetylation level of K5 on a gene promoter may be essential for regulating gene activity (Rundlett *et al.*, 1998). Therefore, the acetylation site on histone H4, rather than simply the level of acetylation, may also be a crucial factor.

### 1.3.2 Histone Phosphorylation

Phosphorylation of the N-terminal domains of the core histones also weakens their interactions with the DNA. When cells are stimulated to proliferate from a quiescent state with the addition of growth factors or phorbol esters, H3 is rapidly phosphorylated on the hydroxyl group of serine residues in its N-terminal domain (Mahadevan *et al.*, 1991). This would presumably allow easier access for transcription factors and more rapid gene activation. Although the role of histone phosphorylation is unclear, it is strongly correlated with the cell cycle (Gurley *et al.*, 1978; Hohmann, 1983) and in *Tetrahymena* it is required for proper chromosome condensation and segregation (Wei *et al.*, 1999).

### 1.3.3 Other Histone Modifications

The role of the methyl, ADP-ribosyl, glycosyl and ubiquitin group additions is rather unclear (Bradbury, 1992). Methylation occurs at the ε-group of lysine residues, predominantly at the N-terminal domains of H3 and H4. In one experiment, most sites were methylated in 12 day-old rats but no unmethylated sites could be detected in adult rats (Duerre *et al.*, 1977). It seems likely that this may simply be a part of nucleosome maturation. ADP-ribosylation, which is the primary means of glycosylation, may be important in DNA repair by nucleosome destabilisation (Althaus *et al.*, 1994). Attachment of a ubiquitin group can select a protein for degradation. Whether ubiquitinated histones follow this pathway or if ubiquitination is an important structural modification is unclear (Nickel *et al.*, 1989; Davies, 1994).

### 1.3.4 Histone Variants

In most organisms, histone genes are present in multiple copies. These are often, but not always, clustered in tandem arrays at one locus. Within these clusters are histone variants which differ often by only a few amino acids. These variants are classed as either replication-dependent, which form the majority of histones or minor (replacement) variants, which are present at low levels throughout the cell cycle and are replication-independent (Yu & Gorovsky, 1997; van Holde, 1988). For example, CENP-A is a vertebrate, replication-dependent H3 variant which associates with the inner kinetocore plate and is essential for correct chromosome segregation (Warburton *et al.*, 1997). Such histone variants can alter nucleosome structure (Simpson, 1981) and modulate gene activity (Gunjan & Brown, 1999; Takami *et al.*, 1997) but it is unclear which can alter nucleosome positioning (Yoda *et al.*, 1998)(**1.5**).

### 1.3.5 DNA Methylation

The addition of a methyl group to DNA at the cyclic carbon 5 of cytosines in CpG dinucleotides influences chromatin structure (Kass *et al.*, 1997a) and is stably maintained through replication by DNA methyl-transferases (Holliday, 1987). This epigenetic marker is highly associated with inactive chromatin although there is still

debate over whether its primary role is in gene regulation (Simmen *et al.*, 1999) or as a repressor of parasitic DNA (Bestor, 1998). These are not mutually exclusive and are reviewed by (Colot *et al.*, 1999).

Methylated DNA transfected into mammalian cells is assembled into a nuclease resistant structure containing unusual nucleosomal particles (Keshet *et al.*, 1986). Despite cleavage in the linker region, they migrate as large nucleoprotein complexes on an agarose gel held together by higher-order interactions. Moreover, individual nucleosomes appear to assemble more stably on methylated than unmethylated DNA. The replacement of histone H1 with a histone variant, MeCP2, may explain the distinct chromatin structure associated with methylated DNA (Keshet *et al.*, 1986; Nan *et al.*, 1997) (see below).

Methylated DNA can alter nucleosome positioning: by increasing the binding affinity at a particular site (Englander *et al.*, 1993), in higher order interactions (Godde *et al.*, 1996) or by nucleosome exclusion (Davey *et al.*, 1997). Methylation cannot occur within a strongly positioned nucleosome in *Saccharomyces cerevisiae in vivo* (Kladde *et al.*, 1994), however, this may not matter as methylation occurs in advance of nucleosome deposition after replication (Cusick *et al.*, 1983; Gruenbaum *et al.*, 1983). It is feasible that this methylation could be maintained by a SWI2/SNF2-like protein which might locally perturb the nucleosome to allow methyl-transferase access (Jeddeloh *et al.*, 1999). This nucleosome deposition time delay could also allow transcription factors access to methylated DNA, potentially overriding induced silencing (Kass *et al.*, 1997b).

MeCP2 is an abundant mammalian protein that binds to CpG methylated DNA *in vivo* and represses transcription (Lewis *et al.*, 1992; Cross *et al.*, 1997; Nan *et al.*, 1998). It possesses a transcriptional repression domain which associates with a corepressor complex containing SIN3A, a transcriptional repressor protein, and histone deacetylases (Jones *et al.*, 1998; Nan *et al.*, 1998). Transcriptional repression *in vivo* is relieved by the deacetylase inhibitor trichostatin A, indicating that deacetylation of histones (and/or of other proteins) is an essential component of this repression mechanism. A recent paper has shown that MeCP2 causally links DNA methylation and histone deacetylation (Wade *et al.*, 1999a).

### 1.3.6 Remodelling Machines

The perturbation of nucleosomal arrays can be induced by chromatin remodelling machines (Wade & Wolffe, 1999b). These can be classed into two distinct families, *SWI/SNF* and *ISWI*. SWI/SNF comprises a set of genes which were originally identified in yeast by their ability to oppose the inhibition of transcription by histones *in vivo* (reviewed in Bjorklund *et al.*, 1999). This complex was subsequently shown to make nucleosomes more accessible to transcription factors in an ATP-dependent fashion (Cairns *et al.*, 1994; Cote *et al.*, 1998). A number of other related complexes have been discovered, also in different species (Table 1).

By contrast, the ISWI family of proteins do not cause this perturbation of nucleosome structure, but do induce ATP-dependent nucleosome movement or *sliding* to enable trans-acting factors to bind the underlying DNA (Langst *et al.*, 1999; Hamiche *et al.*, 1999; Kornberg & Lorch, 1999). It has also been shown that both ACF and CHRAC have nucleosome spacing activities *in vitro* (Ito *et al.*, 1997; Varga-Weisz *et al.*, 1997). Both activities might reconcile if these complexes associate over a large array of nucleosomes in a regular manner, thereby spacing them evenly and causing them to shift *en mass*.

| Complex | Organism | Reference |
|---|---|---|
| **SWI/SNF family** | | |
| SWI/SNF | *S. cerevisae* | (Cairns *et al.*, 1994) |
| RSC | *S. cerevisae* | (Cairns *et al.*, 1996) |
| Brahma | *D. melanogaster* | (Dingwall *et al.*, 1995) |
| h SWI/SNF | *H. sapiens* | (Kwon *et al.*, 1994) |
| NRD | *H. sapiens* | (Tong *et al.*, 1998) |
| **ISWI family** | | |
| I SWI 1 & 2 | *S. cerevisae* | (Corona *et al.*, 1999) |
| NURF | *D. melanogaster* | (Tsukiyama *et al.*, 1994) |
| CHRAC | *D. melanogaster* | (Varga-Weisz *et al.*, 1997) |
| ACF | *D. melanogaster* | (Ito *et al.*, 1997) |
| RSF | *H. sapiens* | (LeRoy *et al.*, 1998) |

**Table 1** Chromatin Remodelling Machines

## 1.4 Nucleosomal Compaction of DNA

One of the central roles of nucleosomes is to package our DNA. Each cell contains 2 metres of DNA in its nucleus, which is between 2 and 10μm in diameter. In order to be accessible and yet fit inside this limited field, it must be compacted approximately 10,000 fold, leading to concentrations exceeding 50mg/ml and regions forming semi-crystalline states.

Chromatin structure above the level of the nucleosome can be critical in determining the transcriptional status of genes and genetic loci (Fletcher *et al.*, 1996). However, a great deal of controversy surrounds the interaction of nucleosomes to form higher order structures and the location of the linker DNA and linker histones within it (Felsenfeld *et al.*, 1986; van Holde *et al.*, 1995). Much of the accumulated data relates to the visualisation of chromatin fragments by electron microscopy in varying salt conditions.

The ionic environment is known to strongly influence compaction (Gerchman *et al.*, 1987; Leuba *et al.*, 1994a; Widom, 1986). Chromatin appears as a zigzag fibre of nucleosomes at low (1mM) ionic strength (Leuba *et al.*, 1994a) and gradually gets more compact, like a flat ribbon of 25nm diameter, as the concentration increases to 5mM. Approaching physiologically relevant concentrations of 100mM NaCl, the chromatin is condensed into irregular rod-like structures with a diameter of approximately 30nm (Figure 5). It is thought that much of the chromatin in the cell nucleus of higher eukaryotes exists in this form (Wolffe, 1995).

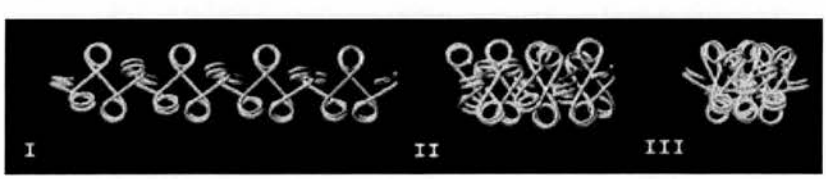



**Figure 5** Nucleosome-Mediated DNA Compaction

(Taken from Bednar *et al.*, 1998: Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin). Progressive compaction of the fibre through increasing ionic concentrations. I=5mM, II=15mM; III=80mM mono-valent cation concentration.

### 1.4.1 The Solenoid Model of Chromatin Compaction

Klug, Koller and colleagues developed the classical model of the first step in chromatin condensation (Thoma *et al.*, 1979). They proposed that a simple solenoid structure is formed by approximately six nucleosomes winding round per turn, with a total pitch of 11nm (Figure 6). This would be facilitated by the co-operative interaction of H1 histones down the central axis of the solenoid to form a 30nm fibre. Whether the co-operative polymerisation of H1 down the central axis mediates the compaction or if it is solely due to local linker DNA-histone interactions is still debatable (Leuba *et al.*, 1993; Leuba *et al.*, 1994b).

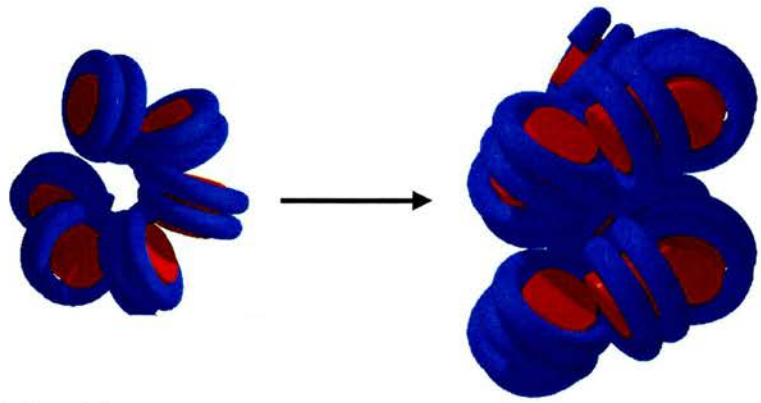


**Figure 6** The Solenoid

A simple 6-nucleosome solenoid structure and a 10-nucleosome solenoid structure forming the 30nm fibre. Courtesy of D. Pruss.

### 1.4.2 The Woodcock Model of Chromatin Compaction

An alternative model, developed by Woodcock and colleagues, proposes that the zigzag array forms a condensed ribbon containing two parallel rows of nucleosomes (Woodcock *et al.*, 1984) (Figure 7). Coiling of the ribbon, again facilitated by the linker histones, generates a 30nm fibre. Indeed, linker histones and the N-terminal tails of the core histones are essential for compaction to occur (Allan *et al.*, 1982). Recently, the use of electron cryo-microscopy (EC-M) has allowed them to take a snapshot of reconstituted nucleosome arrays in an unfixed and unstained solution conformation (Bednar *et al.*, 1998). This technique allows the compaction process to be followed much further (Figure 8) than conventional EM studies, where resolution of linker DNA

and nucleosomes becomes impossible at very early stages (Zlatanova *et al.*, 1998). These high-resolution images confirm their original hypothesis and are in accordance with alternative techniques employed by other groups (Bordas *et al.*, 1986; Kubista *et al.*, 1990; Staynov, 1983).

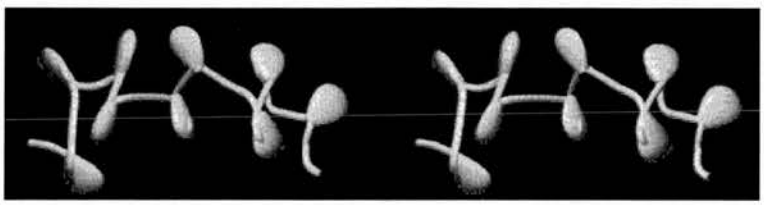


**Figure 7** The Woodcock Chromatin Fibre Model

(Taken from Bednar *et al.*, 1998) 3D stereo pair of a 9-nucleosome segment model of chromatin. Nucleosomes and their associated 'stems' are represented by pear-shaped solids, all of which point toward the fibre interior.

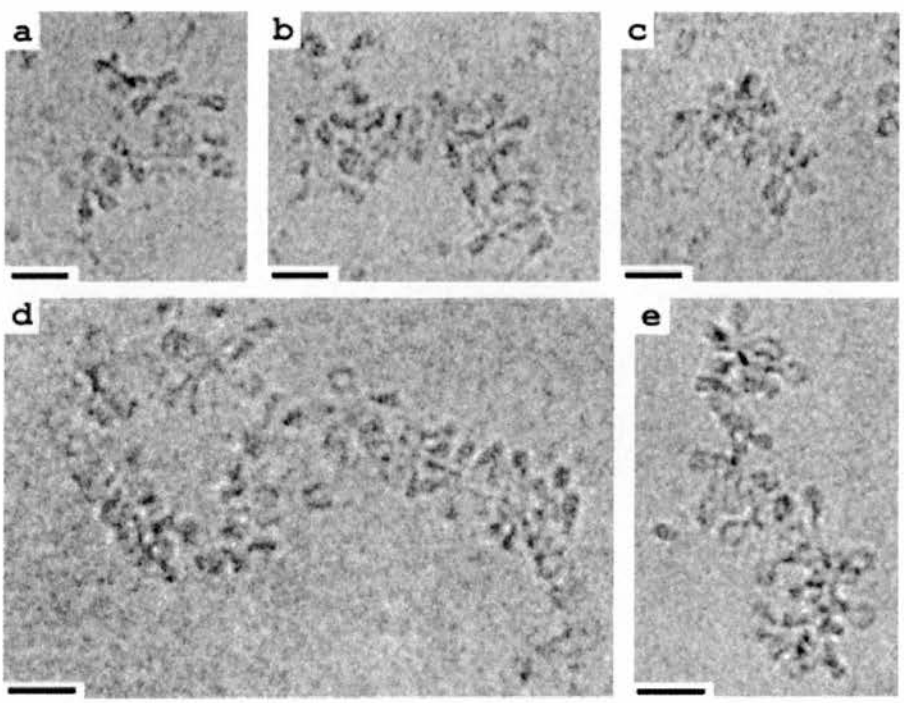



**Figure 8** EC-M-Visualised Nucleosome Compaction

(Taken from Bednar *et al.*, 1998) EC-M compaction process of nucleosomes. (a-c) chromatin in 40mM ions and (d-e) in approx. 15mM ions. Bar=30nm.

Although only limited advances have been made in studying higher order structures *in vivo* (Simpson, 1999), E-CM might prove effective in conjunction with a

novel and straightforward method of isolating chromatin fragments in yeast. For example, the two sets of recombinase and recombination sequences in *S. cerevisiae* (*R, RS* and *FLP1, FRT*) have been employed by two groups to study particular chromatin regions. The strategy used was similar in both cases: two recombinase sequences were inserted on either side of a test genomic sequence and the corresponding recombinase enzyme was activated at a specific time via its inducible promoter. The recombinase then proceeded to excise the intervening genomic sequence as a covalently closed circle that is presumed to maintain its topological properties of the chromosomal locus (Bi *et al.*, 1997; Cheng *et al.*, 1998). These groups used the chromatin fragments to assess topological linking number with respect to nucleosome number or silencing elements. It might also be possible to examine the higher order structure of these chromatin circles using E-CM.

Above the level of the 30nm fibre, two principle models account for the large degree of packing required for compaction into metaphase chromosomes. The first suggests a helical folding of the 30nm fibre into a 250nm fibre, which also folds helically into the observed metaphase chromosome arms (Sedat *et al.*, 1978). The second suggests an organisation of the 30nm fibre into loops that protrude radially from the chromosome axis (Gasser *et al.*, 1986) (Figure 9). The first model does not require additional non-histone scaffolding proteins in order to generate the metaphase chromosome structure and therefore it postulates that no discrete higher order structure exists above the 30nm fibre (Belmont *et al.*, 1987; Belmont *et al.*, 1989).

On the contrary, the second model requires anchoring proteins to fix the radial loops of the 30nm fibre to an underlying nuclear matrix structure. The loop sizes appear to be approximately 100kb in size (Cook *et al.*, 1975; Marsden & Laemmli, 1979; Jackson *et al.*, 1990) and are anchored at matrix attachment sites (MARs or SARs). The assay used to determine these sites is still controversial as it relies on the association of DNA fragments with the nuclear matrix after many highly disruptive steps, such as high salt concentrations and DNA digestion. These DNA sequences have no defined consensus sequence, but are A/T rich. It is thought that H1 or topoisomerase II might bind particularly strongly to these regions (Adachi *et al.*, 1989; Izaurralde *et al.*, 1989)

and generate a distinct gene domain which could influence its regulation (Hart *et al.*, 1998).

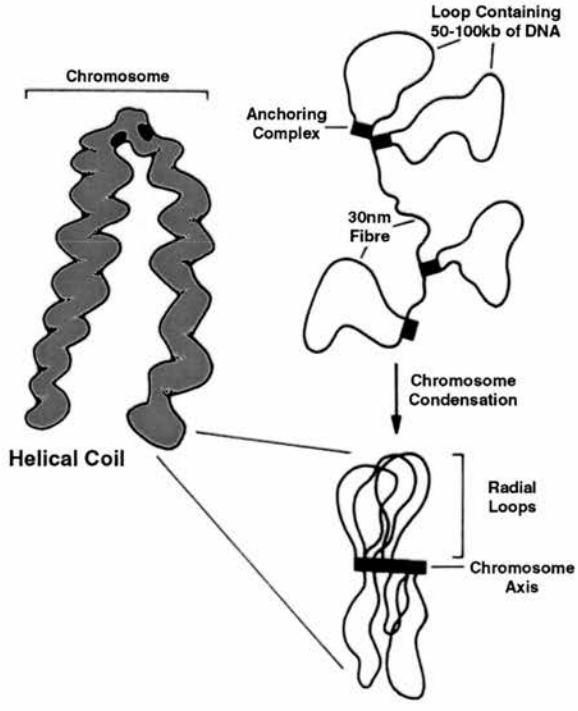



**Figure 9** The Laemmli Model of Higher Order Compaction

The 30nm fibre is anchored to the nuclear scaffold at 'matrix attachment regions'. During chromosome condensation at mitosis, these regions join to form the chromosome axis and metaphase structure. Modified from A.P.Wolffe: Chromatin Structure and Function, 2nd ed.

### 1.4.3 Nucleosomes Within a Higher Order Structure

Chromatin in a higher order structure is stabilised by its interaction with histone H1 (Ramakrishnan, 1997) which appears to be located within the central axis of the 30nm fibre (Dimitrov *et al.*, 1987; Cattini & Allan, 1988). This stabilisation is more pronounced in inactive regions of chromatin (Weintraub, 1984). It is still a matter of some debate as to the exact location of histone H1 relative to the nucleosome dyad (Vignali & Workman, 1998) and its role in regulating gene regulation is even more unclear (Widom, 1998)! However, it is depleted in actively transcribing MMTV (Bresnick *et al.*, 1992) and different variants of H1 can modulate MMTV expression (Talasz *et al.*,1998; Gunjan & Brown, 1999)

### 1.4.4 High Mobility Group Proteins

HMG proteins are among the most abundant and ubiquitous non-histone chromosomal proteins (Wolffe, 1995). Their association with the chromatin fibre may lead to an alteration in its structure which may mediate their effects on replication, nucleosome assembly and transcriptional activity. They are classed into three major families: HMG 1/2, HMG 14/17 and HMG I/Y.

The HMG 1/2 proteins can bind DNA and stimulate the initiation of transcription (Tremethick & Molloy, 1988; Cotmore & Tattersall, 1998), but do not facilitate transcriptional elongation. This appears to be the role of HMG 14/17 proteins.

Transcriptional elongation is prone to pausing and is combated by factors within the transcriptosome (Bengal *et al.*, 1991) and by accessory factors (Brown *et al.*, 1996; Hartzog *et al.*, 1998; LeRoy *et al.*, 1998; Orphanides *et al.*, 1998; Orphanides *et al.*, 1999). HMG14 and HMG17 preferentially bind to nucleosomes and also have a stimulatory role in transcriptional elongation (Crippa *et al.*, 1993; Ding *et al.*, 1994; Ding *et al.*, 1997; Trieschmann *et al.*, 1998).

The HMG I/Y proteins are involved in attenuating the transcriptional activities of a number of factors (Du *et al.*, 1993) and since they can induce positive supercoiling in the DNA strand, may be involved in nucleosome spacing (Reeves & Nissen, 1993; Nissen & Reeves, 1995; Reeves & Wolffe, 1996).

The above descriptions of chromatin structure and composition encompass both single genes and large domain regions. A number of factors interact within this framework to mediate nucleosome spacing, remodelling and sliding (**1.3.1**). These factors must frequently negotiate positioned nucleosomal arrays.

## 1.5 Nucleosome Positioning

Perhaps owing to the relative difficulty of examining the effect of higher order structures on gene activity, much more progress has been made at the nucleosomal level of gene organisation. It was initially believed that most nucleosomes could move freely up and

down the DNA strand, passively interacting with the DNA (Prunell *et al.*, 1978). This mobility is also seen *in vitro* over a range of temperatures and at high and low ionic concentrations (Meersseman *et al.*, 1992).

Since then, the majority of studies on the nucleosomal architecture of chromatin *in vivo* have shown that most nucleosomes are *positioned.* This occurs when nucleosomes associate strongly with the DNA wrapped around them, such that they appear not to move from that location in biochemical assays. Nucleosome substitutions may take place, but for most of the time, that particular 146bp of DNA is associated with a nucleosome.

In addition to *translational* positioning described above, a nucleosome can also be *rotationally* positioned. This refers to the angular orientation of individual base pairs on the nucleosome surface. This can be important in allowing or excluding factor access. For instance, if a transcription factor binding site is obscured by facing towards the surface of a nucleosome, its binding ability may be limited (Figure 10).

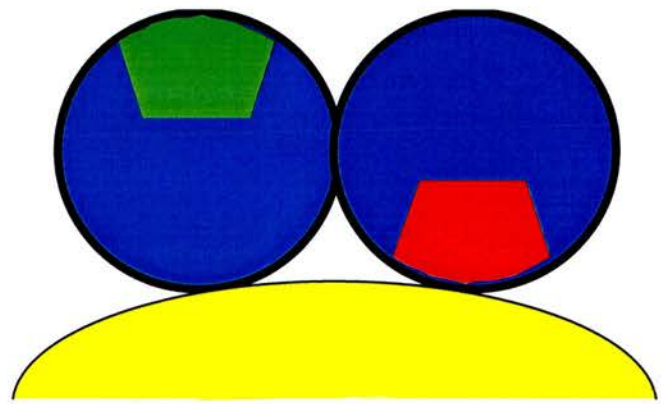


**Figure 10** Rotational Positioning of Transcription Factor Binding Sites
A cross-section through part of the nucleosome (yellow) and the DNA strands (blue). Transcription factor sites that are positioned away from the nucleosome are accessible (green) unlike those facing towards the nucleosome (red).

Kornberg argued that since all cells in an organism contain the same DNA sequences and since the nucleosomal repeat length can vary between different tissues in an organism, it is impossible for precise positioning to occur for a given nucleosome in every cell type (Kornberg *et al.*, 1988). This is quite correct when comparing nucleosome positions between tissues in an organism. Nonetheless, this does not exclude

the possibility that in cells of a given type, most nucleosomes adopt the same positions. Therefore, it is possible that all nucleosomes *are* positioned, but that these positions are cell type specific. This in itself could produce an epigenetic form of gene regulation.

## 1.6 Positioning Mechanisms

There are several possible mechanisms which might lead to nucleosome positioning:

### 1.6.1 Nucleosome Constraint or Tethering by DNA-Bound Factors

Binding of non-histone proteins to specific recognition sequences on the DNA might create some degree of regular chromatin structure. Two non-histone proteins, bound a limited distance apart, might limit the extent of nucleosome movement between them, thereby acting to constrain a finite array. Alternatively, non-histone proteins might interact with adjacent nucleosomes, tethering an array which extends away from them.

A vast number of transcription factors are now known to bind to cognate sites on DNA. Some of these can bind without necessitating nucleosome removal when their sites are rotationally positioned away from the nucleosome surface (Figure 10) (Beato *et al.*, 1995). The glucocorticoid receptor (GR) binds to the mouse mammary tumour virus promoter in this way, leading to gene activation. It is the only factor discovered to date which binds nucleosomal and naked DNA with equal affinity. Its binding does not lead to nucleosome positioning though, as the array is already set before GR binds. Other factors such as the winged helix transcription factor HNF3 are capable of binding within a positioned nucleosome. It binds the mouse serum albumin gene enhancer prior to gene activation (Shim *et al.*, 1998) and is necessary for cell type specific expression (Liu *et al.*, 1991). It may be involved in actually setting this position (Shim *et al.*, 1998).

Other factors can interact with remodelled or pre-set chromatin structures which were generated prior to the inducing stimulus. These can be divided into two groups: those which are able to bind to altered nucleosomes and those which can only bind nucleosome-free regions (Beato *et al.*, 1997). Nuclear factor 1 (NF1) is one such transcription factor which cannot interact with its cognate site in a nucleosome, although

it can bind in the linker region (Spangenberg *et al.*, 1998). Although it has a high affinity for naked DNA, it can only bind nucleosomal DNA once chromatin remodelling has occurred (Eisfeld *et al.*, 1997). This is most likely because NF1 almost completely surrounds the circumference of the DNA helix and makes numerous contacts with the bases and the backbone (Eisfeld *et al.*, 1997). This class of transcription factor probably does not position nucleosomes either, since they only gain access to most cognate sites via additional factors or chromatin remodelling machines to assist in transcriptional activation.

Factors which are found constitutively bound to nucleosome-free DNA are much more likely candidates to set a nucleosome array structure. These are thought to bind immediately after DNA replication at mitosis, when they compete for access with histones. These are found in the promoter regions of genes which require a rapid response. For instance, the GAL1/GAL10 regulatory region of the galactosidase gene contains a pre-formed, inactive complex of GAL4 and GAL80 (Lohr and Lopez, 1995). In the presence of galactose, the GAL80 repressor is removed, leading to rapid induction. As these factors are bound when the nucleosomal architecture is forming, they could easily restrict or curtail nucleosome movement and form an array.

### 1.6.2 Nucleosome Positioning from a Replication Origin

A potential, but untested, mechanism for positioning is location with respect to a replication origin. As the replication complex moves along the DNA, histone proteins are deposited on the newly synthesised strands to form nucleosomes. It may be that they remain in the location where they are deposited; this location is therefore related to where replication was initiated. Active replication origins can vary depending on cell type, and the time of replication initiation also varies between euchromatin and heterochromatin. This would alter the position of nucleosome arrays and could possibly be used to test this model. Additional mechanisms can impose a regular nucleosomal array after deposition (Cairns, 1998). However, unless this replication-generated conformation is locked into position immediately, this model seems rather improbable.

The DNA sequence alone is probably sufficient to shift large nucleosomal arrays into the most thermodynamically stable position.

### 1.6.3 DNA Sequence-Directed Positioning

The DNA sequence is generally believed to be the primary determinant of nucleosome positioning. However, there is no unique positioning consensus sequence; rather it appears that the structure formed by the DNA may either attract or repel nucleosomes. DNA base pairs are not the static structures normally seen in the double helix model. They can twist, roll and slide in varying amounts depending on the base composition (Calladine & Drew, 1992). For instance AA/TT sequences are most stable in a low roll configuration, as they prefer to stack by forming an additional hydrogen bond in the major groove. This leads to a more rigid structure. GG/GC sequences are most stable in a high roll configuration, as the strong imbalance of electrical charge in the guanine ring discourages them stacking directly over one-another.

This stability is reflected in the average position of these sequences within nucleosomes. When DNA wraps round the core particle, it must bend approximately 45° for every complete turn of the helix (approximately 10bp). This severe contortion favours alternating structures to create a bend. AA/TT sequences are found in a 10bp periodicity within nucleosomes where the minor groove faces inwards at a low roll angle. Likewise, GC sequences are found where the minor groove faces outwards at a high roll angle. By creating an intrinsic bend, nucleosome packing is much more energetically favourable. In contrast, long $(A)_n$ or $(T)_n$ sequences can exclude nucleosomes because of the relatively high energy required to bend them (Calladine & Drew, 1992). These sequences are often found in the linker region that is relatively straight.

## 1.7 Nucleosome Positioning can Modulate Gene Function

Nucleosomes are frequently positioned around DNA sequences with important functional roles (Becker, 1994). These specific chromatin structures have been found in

the vast majority of genes in which nucleosome positioning has been analysed (Wolffe, 1995). Two extensively studied examples are reviewed here. A regulatory role for a positioned nucleosome is identified in the first example, the mouse mammary tumour virus. This nucleosome prevents access of ubiquitous transcription factors to the promoter until the correct inducing signal appears. The second example of the *Drosophila* hsp26 gene examines how the folding of intervening DNA can keep transcription factor sites accessible and possibly assist in gene activation.

### 1.7.1 The Mouse Mammary Tumour Virus

The mouse mammary tumour virus (MMTV) is one system in which nucleosome positioning plays a critical role in gene regulation. Its 3′ long terminal repeat (LTR) is incorporated into a positioned nucleosomal array when inactive, in both episomes and when integrated into a mouse chromosome (Bresnick *et al.*, 1990; Richard-Foy *et al.*, 1987). Binding of glucocorticoid hormone receptor (GR) remodels one of these nucleosomes and GR's co-operative interaction with NF1 activates transcription (Figure 11).

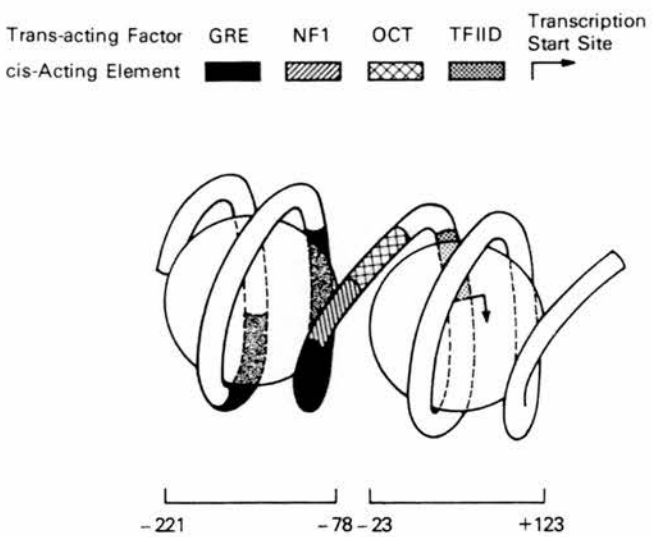



**Figure 11** Chromatin Organisation of the MMTV LTR

Positioned nucleosomes 'A' (+123 to -23) and 'B' (-78 to -221) are indicated. The transcription factor binding sites on the associated DNA are shown with respect to these nucleosomes and the start site. Courtesy of A.P.Wolffe: Chromatin Structure and Function, 2nd ed.

Six nucleosomes are positioned over the 3′LTR and are designated 'A' to 'F' (Richard-Foy *et al.*, 1987). Nucleosome 'A' is closest to the 5′Cap site. Nucleosome 'B', next to it, contains the GR cognate binding site and is remodelled at induction. The exact positions of these nucleosomes have been widely debated as they differ depending on: temperature, construct size, *in vivo* versus *in vitro* results (Archer *et al.*, 1992), the model system (Mymryk *et al.*, 1995), integration site and type of construct (Flaus *et al.*, 1998). Nucleosome 'A' appears to be assembled in a tightly defined translational position by A/T-rich regions at its borders. This short sequence can exclude nucleosomes by forming a rigid structure that requires an excess of free energy in order to be packaged. Nucleosome 'B', on the other hand, can assemble at over 13 different locations that vary by over 60bp (Fragoso *et al.*, 1995). All of these positions still encapsulate the GR binding site and have a consistent rotational position, which is essential for GR to bind effectively (Eisfeld *et al.*, 1997).

As the DNA is intrinsically distorted over nucleosome 'B', nucleosome 'A' assembles first, and is thought to lock 'B' into a fixed translational position. Which of the 13 translational positions is chosen may depend on the DNA sequence in a particular MMTV copy and on other proteins or cell-specific epigenetic mechanisms (Ostrowski *et al.*, 1983).

Glucocorticoid hormone receptor, by itself, does not lead to significant levels of activation. Octameric transcription factor (OTF) and NF1 are essential for full activation. As the GR's cognate site faces outward from the nucleosome, it can easily gain access. OTF and NF1 only bind naked DNA strongly. In the case of NF1, it makes DNA contacts at almost all points around the DNA duplex, necessitating nucleosome disruption (Blomquist *et al.*, 1996). Therefore, it is only when GR has disrupted nucleosome 'B' that OTF and NF1 can gain access and act co-operatively to mediate full activation (Chavez *et al.*, 1997).

How GR remodels the nucleosome is still unclear. One possible method is by recruitment of a HAT. A moderate increase in histone acetylation levels in cell culture by the histone deacetylase inhibitor, Trichostatin A, remodels the nucleosome without the need for GR. Indeed, it displays the same DNaseI digestion pattern as GR-mediated

activation (Bartsch *et al.*, 1996). Alternatively, the H2A/H2B dimer(s) could be removed from nucleosome 'B'. The H3/H4 tetramer occupies the same position on the DNA as the octamer and the rotational orientation is essentially the same. The tetramer displays the same DNase I digestion pattern as the induced octamer (Spangenberg *et al.*, 1998). The same pattern is seen in SWI/SNF treated octamers, indicating that the SWI/SNF complex may modify nucleosomes by H2A/H2B removal. GR may mediate nucleosome removal by recruiting SWI/SNF activity (Ostlund *et al.*, 1997).

### 1.7.2 The Drosophila hsp26 Gene

Raising the temperature of a *Drosophila melanogaster* fly to 34°C leads to rapid transcriptional activation of the hsp26 heat-shock gene. The promoter region of the gene contains a positioned nucleosome which is located between two DNase I hypersensitive sites encompassing the essential heat-shock elements (HSEs) (Figure12). The nucleosome position is determined not only by the underlying DNA sequence, but also by the neighbouring sequences. These contain multiple (CT)n.(GA)n repeats which bind GAGA factor (GAF) which may in turn assist in exact nucleosome positioning (Fedor *et al.*, 1988; Simpson, 1991) and in forming the local chromatin architecture (Katsani *et al.*, 1999; Wilkins *et al.*, 1999).

Heat-shock transcription factor (HSF) also binds within these two HSEs and is required to initiate transcription. The induction process is very rapid as RNA polymerase II is already bound at the start site and is in a paused conformation about 20bp downstream of the start site. An additional role of GAF may be to initiate RNA polymerase II movement via phosphorylation (Lis *et al.*, 1993; O'Brien *et al.*, 1994) once HSF has bound. However, it is not clear how all of these factors work in unison (Wilkins *et al.*, 1997).

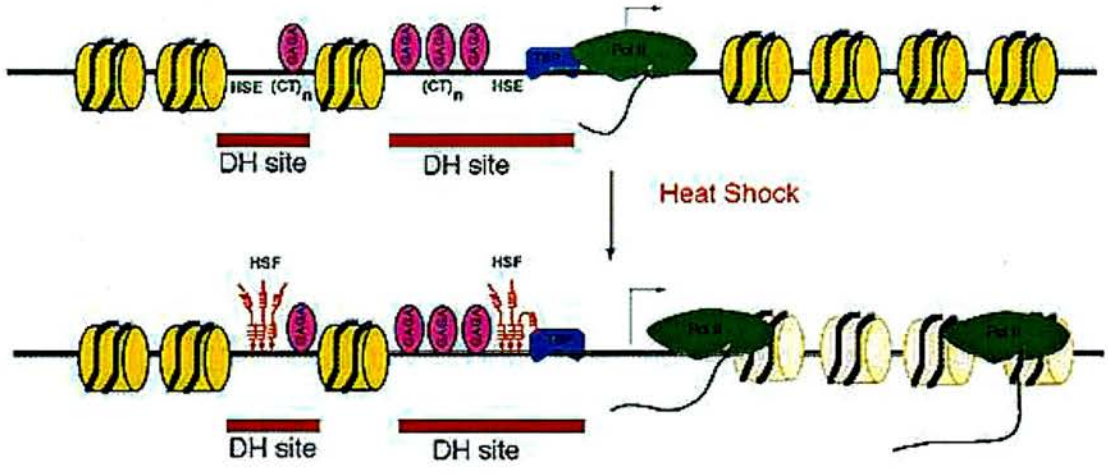



**Figure 12** Chromatin Structure of the Drosophila hsp26 Gene

The gene is assembled into a precise chromatin structure prior to heat shock. A specifically positioned nucleosome (yellow disk) lies between the DNase I hypersensitive sites (DH sites; red bars). GAGA is bound to high affinity sites (magenta) and the TATA binding protein (TBP; blue) is present with RNA polymerase II (green) prior to initiation. After heat shock, heat shock transcription factor (red) binds to the HSEs and triggers transcription; there is a perturbation of the downstream nucleosome array during transcription (white). Courtesy of S.Elgin.

The positioned nucleosome situated between the DH sites wraps up the DNA to bring the sites into contact (Figure 13). However, this nucleosome is not essential for gene activation (Lu *et al.*, 1995). The underlying DNA sequence can be deleted, inverted, or replaced by 'random sequence' bacteriophage lambda DNA without impairing transcriptional activity. Transcriptional activity is impaired if a nucleosome is artificially positioned over one of the DH sites. Therefore, the role of this positioned nucleosome may be to prevent the DH sites being otherwise occluded by nucleosomes. Other studies have shown that bringing distant sites into close contact by nucleosome folding can lead to higher expression levels (Laybourn *et al.*, 1992).

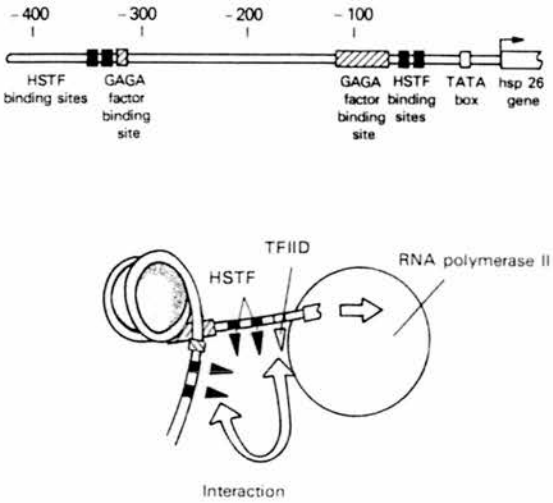



**Figure 13** Chromatin Organisation over the *Drosophila* hsp26 Promoter

A nucleosome is positioned between GAGA/HSTF binding sites. By wrapping up the intervening DNA it brings them into contact next to the RNA polymerase. Courtesy of A.P.Wolffe: Chromatin Structure and Function, 2nd ed.

## 1.8 Studying Chromatin: Finding a Sympathetic System

Chromatin structure has been studied in some detail in a variety of genes, most notably: (1) MMTV in mouse (**1.7.1**), (2) *PHO5* in yeast (Svaren & Horz, 1997), (3) the silent mating type locus in yeast (Weiss & Simpson, 1998), (4) *hsp*26 in *Drosophila* (**1.7.2**), (5) lysozyme in chicken (Huber *et al.*, 1996; Kruger *et al.*, 1999), (6) the globin locus in chicken (Davey *et al.*, 1995), (7) 5S RNA genes in several species (Roberson *et al.*, 1989), (8) repetitive sequence DNA in species such as mouse (Widlund *et al.*, 1997) and (9) the serum albumin gene enhancer in mouse (McPherson *et al.*, 1996).

All of these systems have their advantages and drawbacks. For example, although MMTV is particularly amenable to study, the exact nucleosome positions it adopts are dependent on its integration site *in vivo*, which may not be reflected *in vitro* (Flaus & Richmond, 1998). The yeast *PHO5* gene is inducible, which is useful when assessing nucleosome alignment before and during gene activation. However, like MMTV, *hsp*26 and serum albumin, only six or fewer positioned nucleosomes in the

array have been studied, whereas those positioned out-with this region may also be influential. Lastly, the repetitive and 5S RNA genes are rather specialised cases and may not relate to inducible pol II-transcribed genes. Perhaps the most interesting studies have been on the avian lysozyme and globin genes. Both of these are eukaryotic, developmentally regulated, tissue specific and the array structure has been studied over a large region of each gene. Unfortunately, very few studies have focused on long-range nucleosomal structures in mammalian genes.

## 1.9 The Mammary Gland

The mammary gland is a particularly amenable system in which to study the developmental regulation (Groenen *et al.*, 1994) and chromatin structure of a single copy gene. Mammary-specific genes are regulated by hormonal (Forsyth, 1986) and extra-cellular matrix (ECM) interactions (Streuli *et al.*, 1991) which have been well characterised. These also control the rounds of cellular differentiation, growth and apoptosis which occur during successive lactations. Unlike most tissues which develop in the embryo, the mammary gland develops mostly in the adult. This makes the study of gene activation easier because of the abundance of tissue and because it can be analysed at various points during activation. Tissue culture systems are available, which can help elucidate the influences of hormones and transcription factors (Burdon *et al.*, 1994).

As the gland is not essential for viability, it can be modified at will by generation gene knock-outs (Kumar *et al.*, 1994; Stinnakre *et al.*, 1994), gene knock-ins (Stacey *et al.*, 1995) and gene modifications (Whitelaw *et al.*, 1991; Whitelaw *et al.*, 1992). In addition, transgenic mouse models are available which have been instrumental in defining the essential regulatory regions of various mammary-specific genes (Mercier & Vilotte, 1993; Wilmut & Whitelaw, 1994).

## 1.10 Milk Protein Genes

Milk comprises two major classes of protein, namely the caseins and the wheys. These two classes of proteins are defined by their ability to acid precipitate and their relative

abundance varies considerably between species (Jenness, 1986). All of the caseins and four of the whey proteins (α-lactalbumin, β-lactoglobulin, whey acidic protein and lactoferrin) are synthesised in the mammary gland, whereas serum albumin, lysozyme and the immunoglobulins are imported from the blood. There follows a brief summary of the milk protein-encoding genes which are relevant to this thesis (reviewed in Mercier & Vilotte, 1993).

### 1.10.1 The Caseins

The casein genes are clustered together in a locus in sheep (Leveziel *et al.*, 1991), goat (Hayes *et al.*, 1992; Hayes *et al.*, 1993) , mouse (Geissler *et al.*, 1988) and human (McConkey *et al.*, 1996) species. The most extensively studied milk protein gene is β-casein which is under the control of lactogenic hormones (Doppler *et al.*, 1990; Goodman & Rosen, 1990) and the extra-cellular matrix (Schmidhauser *et al.*, 1990; Schmidhauser *et al.*, 1992). Both prolactin and glucocorticoid activate expression whereas progesterone inhibits it. This results in the up-regulation of β-casein expression during early lactation and its subsequent inactivation at weaning and involution. The gene comprises nine exons spanning approximately 8.5kb. A 3.8kb promoter region appears to be sufficient to drive tissue-specific and developmentally regulated expression of β-casein or a reporter construct (Cerdan *et al.*, 1998), although only a 338bp promoter region is sufficient to confer hormone responsiveness to mouse cells in tissue culture (Schmitt-Ney *et al.*, 1991).

### 1.10.2 Whey Acidic Protein

Like the casein genes, the induction and maintenance of WAP expression depends on the synergistic action of glucocorticoid, prolactin, insulin and the ECM. mRNA is found at low concentrations in virgin and early pregnant animals (Pittius *et al.*, 1988) and increases by 1000-fold at mid-lactation (Mercier & Vilotte, 1993). The gene comprises a 2kb transcription unit which contains four exons. A minimal promoter region of 175bp is sufficient to stimulate transcription of mouse cells *in vitro* (Lubon *et al.*, 1989).

### 1.10.3 α-Lactalbumin

The presence of prolactin, insulin and glucocorticoid (Funder, 1989) are required for gene induction *in vivo*. Unlike β-casein, α-lactalbumin levels remain low during pregnancy and rise rapidly to a maximum after parturition (Nardacci *et al.*, 1978). This may partly result from their responses to differing concentrations of glucocorticoid hormone (Ono & Oka, 1980).

The transcription unit spans approximately 2.5kb and includes four exons (Hall *et al.*, 1987). It appears that only 400bp of 5' sequence and 340bp of 3' sequence may be required for mammary-specific expression and correct developmental regulation, but not for high level or copy number-related expression (Soulier *et al.*, 1992).

### 1.10.4 β-Lactoglobulin

BLG constitutes the majority of milk whey protein in ruminants. It is also present in a wide variety of other animals, including horses, pigs, dogs, cats, dolphins and kangaroos (Hambling *et al.*, 1992). However, it is not present in rodents, rabbits or humans. In ruminants it consists of a mature polypeptide chain of 162 amino acid residues, containing five cysteine residues, four of which form intra-chain di-sulphide bridges (Kold *et al.*, 1983). The three-dimensional structure of bovine BLG has been determined at 2.8A resolution (Papiz *et al.*, 1986) and possesses an eight-stranded, anti-parallel β-barrel core motif. Although its function is unknown, it displays high structural homology to the human serum retinol binding protein (Newcomer *et al.*, 1984), suggesting a role in vitamin A transport. In milk, BLG exists predominantly as a stable dimer, and like other members of the divergent lipocalin superfamily, it binds, and may transport, a variety of hydrophobic molecules (Futterman *et al.*, 1972).

BLG is expressed in a temporally regulated manner, specifically in mammary gland secretory epithelial cells. Expression is barely detectable in the virgin sheep (Whitelaw *et al.*, 1998) but after conception increases gradually to constitute up to 0.25% total poly(A) mRNA at mid-pregnancy, which rises to 5% at parturition (Mercier *et al.*, 1993). After weaning, expression levels fall dramatically in accordance with apoptosis and involution.

## 1.11 Cloning and Manipulation of the β-Lactoglobulin Gene

Interest in ovine mammary-specific gene regulation was stimulated by the prospect of generating transgenic sheep, which could produce valuable human pharmaceutical proteins in their milk (Clark *et al.*, 1989; McClenaghan *et al.*, 1991; Simons *et al.*, 1988; Wilmut *et al.*, 1991; Wilmut *et al.*, 1994). Mice were used initially to test the performance of transgene constructs in an *in vivo* environment (Archibald *et al.*, 1990; Simons *et al.*, 1987). The sheep β-lactoglobulin (BLG) gene was chosen as it has a high expression level in sheep and although not endogenous to mice, displays a similar expression profile in transgenic animals (Simons *et al.*, 1987) (Harris *et al.*, 1991) (Wilde *et al.*, 1992) but also (Whitelaw *et al.*, 1998). This has the additional advantage that analysis would not be complicated by interference from a similar endogenous gene.

The cloning and sequencing of the sheep BLG cDNA (Gaye *et al.*, 1986) lead to the isolation of the gene from genomic DNA two years later (Ali *et al.*, 1988a). Four clones were isolated, two of which, BLG-A and BLG-B, were characterised by restriction enzyme and sequencing analysis. A comparison of the two alleles demonstrated that they differ by only 1bp in the coding region, resulting in a Tyr/His change (Harris *et al.*, 1988). The BLG-A clone, used in this study, is a 16.2kb SalI-SalI fragment which comprises 4.2kb of 5' sequence, a 4.7kb transcription unit and 7.3kb of 3'flanking sequence (Figure 14). The transcription unit comprises seven exons; exons I and II encode an 18 amino acid signal and exon VII is non-coding.

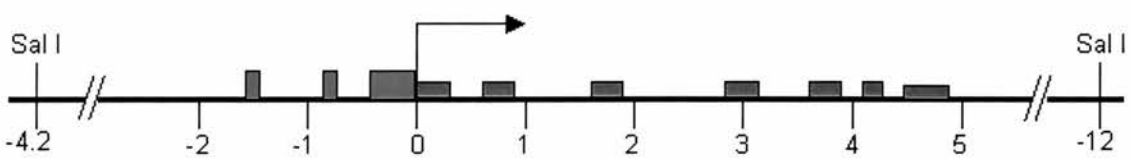



**Figure 14** The Ovine BLG Gene
The bent arrow indicates the start site. Green boxes represent exons; red are DNaseI hypersensitive sites. Numbering is in kilo-base pairs.

In order to identify whether the clone contained the entire gene domain necessary for correct regulation *in vivo*, it was introduced into mice by pronuclear microinjection

(Simons *et al.*, 1987) to create lines 7 and 14. Transgenic mouse line 45 was also created which contained only 1.9kb of 3' flanking sequence at the 3' end resulting in a smaller 10.5kb transgene. All lines expressed high levels of BLG mRNA in the mammary tissue with levels roughly proportional to copy number. To further refine the gene domain, a series of BLG gene fragments were made containing 1.9kb of 3' flanking sequence and either 4.3kb, 3.3kb, 2.0kb, 800bp, 406bp, 146bp or 79bp of 5' flanking sequence (Whitelaw *et al.*, 1992). All lines containing 406bp or more of 5' flanking sequence displayed high level, copy-number related mRNA expression levels. The constructs that were further resected to leave 146bp or less of 5' sequence resulted in a dramatic decrease in the frequency of transgene expression. This lead to the conclusion that the BLGΔDp construct, which contains 406bp of 5' sequence and 1.9kb of 3' sequence of the ovine BLG gene, is sufficient to direct efficient expression to the mammary gland of transgenic mice. Further analysis revealed that the 406bp BLG 5' flanking region alone was sufficient to direct expression of a reporter gene to the mammary gland (Webster *et al.*, 1995).

As with most transgenes, the intronic regions are also required for efficient expression (Brinster *et al.*, 1988; Palmiter *et al.*, 1991). Deletion of all six introns (Whitelaw *et al.*, 1991) or either the first two, middle two, or last two introns (Webster *et al.*, 1997) abolished copy number dependent expression. Moreover, expression levels were much lower and some lines did not express BLG. This variability in expression was not evident *in vitro*.

### 1.11.1 Position Independent Expression

It is intriguing that the transgenic mouse lines generated which contained the BLGΔDp construct (or larger flanking regions), all expressed BLG in a copy number related manner. This is known as *position independent expression* and does not occur in the majority of other transgenes. Normally transgene expression levels are highly variable and substantially lower than predicted (Bonifer *et al.*, 1996). This may be caused by insertion into or near heterochromatin or by heterochromatinisation (Dorer *et al.*, 1994; Henikoff, 1998) and is exacerbated by elevated copy-numbers (Garrick *et al.*,

1998). The position independent expression effect may be caused by DNA sequences which inhibit heterochromatinisation or induce an active domain. This undoubtedly involves the creation of a permissive chromatin environment. Alternatively, the transgene itself may direct its insertion into a euchromaic 'open' chromatin domain (Bishop, 1996)

### 1.11.2 Transgene Rescue

Not only does BLG appear to be immune to many of these silencing effects, it can rescue other transgenes from inactivation. BLG-driven cDNA transgenes, which on their own suffered these silencing affects, displayed a significant improvement in the frequency and level of expression when co-injected with a BLG transgene (Clark *et al.*, 1992). This was dependent on BLG transcription but not its translation (Yull *et al.*, 1997).

It appears that the tissue-specific expression of both transgenes is influenced by their joint promoter regions. A poorly expressed MMTV-driven transgene was rescued by co-injection with a BLG transgene. Since the MMTV-LCR normally drives expression in numerous tissues, it was expected to do the same once rescued. Surprisingly, parallel tissue specificity was observed between both transgenes in the double-transgenic mouse lines generated (Langley *et al.*, 1998). Therefore, although BLG is required to rescue expression, its expression profile is also influenced by the other transgene. This was particularly evident when it was co-injected with prokaryotic sequences which appear to be a foci for silencing (Clark *et al.*, 1997).

## 1.12 BLG Regulation

### 1.12.1 Regulatory Elements

Like most milk protein genes, transcription factors and steroid hormones act together to initiate BLG transcription (Topper *et al.*, 1980; Osborne *et al.*, 1995). There are five known DNase I hypersensitive sites over the gene which vary depending on the gene activation state (Whitelaw & Webster, 1998) (Figure 15). HSIII comprises the

proximal promoter region and is essential and sufficient for temporally regulated, tissue-specific expression of a heterologous gene (Webster *et al.*, 1995). However, this does not confer a copy number-related, stable expression profile.

Interestingly, HSIV and HSV are only present until just before parturition, after which point BLG transcription is upregulated. Factors interacting at these sites may be required to set a permissive domain structure. Once this is completed, they may no longer be required, which correlates with the subsequent lack of hypersensitivity.

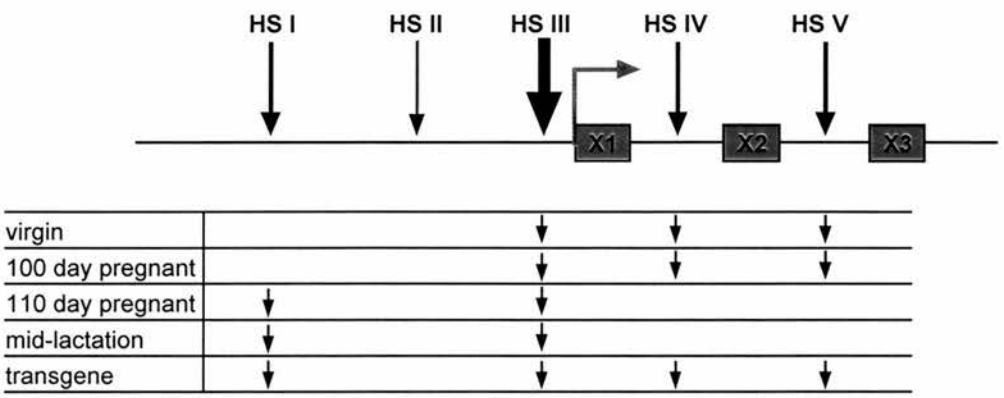



**Figure 15** DNase I Hypersensitive Sites over the BLG Gene

(Modified from Whitelaw & Webster, 1998) The ovine BLG gene is depicted with its transcriptional start site (green arrow) and exons (green boxes). Hypersensitive sites (vertical arrows) vary in intensity (arrow thickness; HSII is erratic) relative to each other. The pattern in lactating mammary tissue of transgenic mouse line BLG14 is shown for a comparison.

Another element which influences transcription is the matrix attachment region 3' of the transcription unit (Whitelaw *et al.*, 1999). It up-regulates basal expression levels in HC11 cells and transgenic mice, although it does not confer copy number-dependent expression.

### 1.12.2 Transcriptional Control

Factors which associate with the proximal promoter region, HSIII, have been investigated. Stat5 and two different types of NF1 associate with this region (Watson *et al.*, 1991). NF1 is a ubiquitously expressed family of transcription factors which appear, in most cases, to be general activators of transcription. Signal transducers and activators of transcription (Stats), on the other hand, are activated in the response to extracellular signals, such as cytokines, growth factors and steroid hormones.

## 1.13 Stats Matter

Seven different Stat genes have been identified so far in mammals, named Stat1 to Stat6 with two different forms of Stat5 (a and b) (Hoey & Schindler, 1998). They all have highly conserved structural domains (Becker *et al.*, 1998) (Figure 16) and in general bind as dimers to a palendromic recognition sequence. This recognition sequence contains a central spacer region which ranges from 4 to 6bp and partly determines their sequence specificity (Seidel *et al.*, 1995). Both forms of Stat5 bind recognition sequences with a 3bp internal spacer.

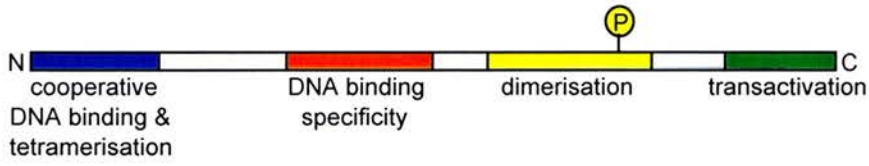



**Figure 16** Domains of a Stat Protein

Although both 'a' and 'b' isoforms possess 95% sequence identity, they have different targets for gene regulation (Teglund *et al.*, 1998). Stat5a appears to be the major activator of milk protein genes (Han *et al.*, 1997; Liu *et al.*, 1997) and predominates, relative to Stat5b, in the mammary tissue. Stat5b predominates in muscle tissue (Liu *et al.*, 1995), can be activated through an independent pathway (Kazansky *et*

*al.*, 1999) and is involved in sexual dimorphism (Udy *et al.*, 1997). However, Stat5b can partly rescue milk gene expression in an absence of Stat5a activity (Liu *et al.*, 1998).

### 1.13.1 Stat5 and Chromatin

Although all Stats must function within a chromatin environment, it is unknown whether they can bind efficiently within a nucleosome. Most transcription factors bind naked DNA much more easily than nucleosomal and some, such as NF1, can only bind remodelled chromatin or within the linker region (Eisfeld *et al.*, 1997). This appears to be related to the numerous contacts it makes on all sides of the DNA helix, resulting in steric exclusion.

The dimerisation structure of Stat5 is unknown, but that of Stat3 (Chen *et al.*, 1998) (Figure 17) and Stat1 (Becker *et al.*, 1998) (Figure 18) are. Since all Stats are highly conserved, it is reasonable to presume that the dimerisation structure of Stat5 is also similar.

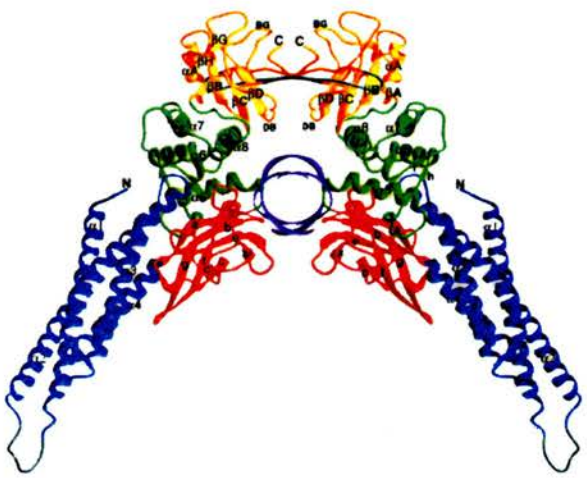


**Figure 17** A Stat3 Homodimer Interacting with its Cognate Site
Modified from Chen *et al.*, 1998. DNA helix, purple; DNA binding domain, red; coiled-coil domains which mediate co-operative DNA binding and Stat tetramerisation, dark blue; linker domain, green; SH2 domain, gold.

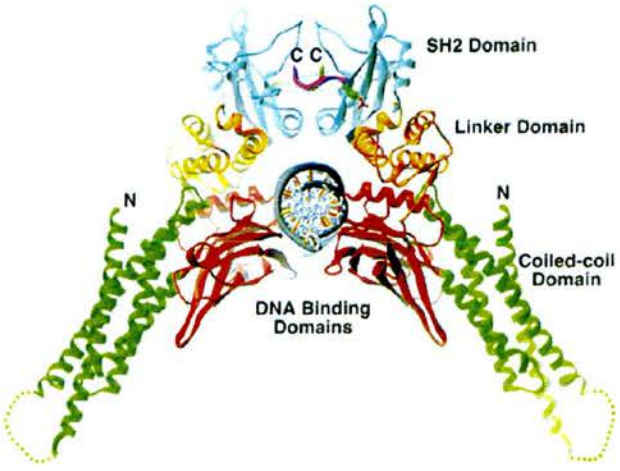



**Figure 18** A Stat1 Homodimer Interacting with its Cognate Site
Modified from Becker *et al.*, 1998. Domains are as indicated.

As can be seen from both figures, Stat dimers form extensive interactions right round the DNA helix in a clamp-like structure. Like NF1, such extensive interactions with all sides of the helix probably preclude binding within a nucleosome. Stat binding domains are similar in structure to those of NFκB and p53 (Chen *et al.*, 1998). Indirect evidence suggests that the NFκB binding domain is incapable of binding efficiently within an unstructured chromatin environment (Cook *et al.*, 1995; Cook *et al.*, 1999).

### 1.13.2 Stat-Mediated BLG Activation

Stat5 is up-regulated during mammary development (Philp *et al.*, 1996) and binds DNA sites in many milk protein genes. In BLG there are three in the 408bp minimal promoter region at positions –93, -210 and –278bp with respect to the start site (Watson *et al.*, 1991). This transcription factor is part of the prolactin signal transduction pathway which initiates with the binding of prolactin to its receptor on the cell surface (Watson & Burdon, 1996) (Figure 19). This induces dimerisation of one or more of the receptor chains, bringing together two Jak kinases (Wilks *et al.*, 1991), owing to their association with the membrane proximal intracellular domain of the receptor. The proximity of the receptor chains and their associated Jaks results in tyrosine cross-phosphorylation of both Jaks and specific residues on the receptor. It is this

phosphorylation which is thought to recruit Stat5 monomers from the cytosol. Transient association with the receptor results in tyrosine phosphorylation of Stat5 protein, oligomerisation with at least one other Stat protein (Barahmand-Pour *et al.*, 1998) and subsequent translocation into the nucleus. Once in the nucleus, it binds DNA sites to activate transcription.

An *in vitro* binding assay showed that the three BLG sites had varying affinities for Stat. The three sites at –93bp, -210bp and –278bp have relative affinities of 44:8:1, respectively. Mutation of these sites abrogates the hormonal response *in vitro* (Burdon *et al.*, 1994; Demmer *et al.*, 1995) and *in vivo*. These authors found that Stat binding was not essential for mammary expression, but was required for maximal activity. However, they only showed an absence of Stat binding *in vitro*, not *in vivo,* therefore it is still feasible that Stat5 is essential for BLG activation. Alternatively, there may be that additional factor(s) as yet uncharacterised which are required for mammary-specific expression.

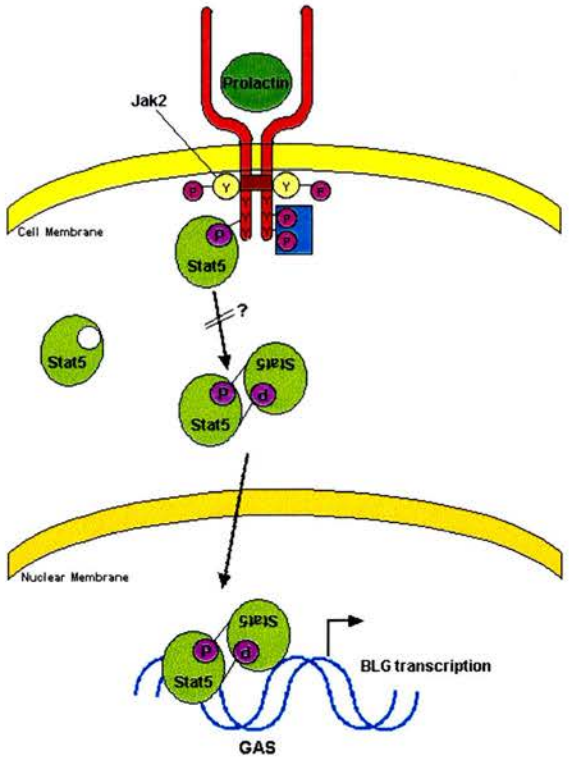



**Figure 19** The Jak/Stat Endocrine Pathway
See text for details

## 1.14 Project Overview

Both the hormonal signalling pathways and the DNase I hypersensitive sites have been well characterised with respect to the BLG gene during its induction. However, nothing is known regarding its nucleosome structure. I have completed a preliminary analysis of the nucleosome structure throughout the ovine BLG gene as an initial step to understanding the potential role of nucleosomes in the gene's regulation. Micrococcal nuclease was used as a probe to test for the presence of nucleosomes and a regular nucleosome repeat length. Further analysis using the chemical probe cuprous phenanthroline tested for discrete nucleosome positions on the gene. This *in vivo* work was complemented by an *in vitro* monomer extension technique which quantitatively and qualitatively tests for DNA-directed nucleosome positioning of reconstituted histone proteins. Both *in vivo* and *in vitro* approaches were used to analyse the promoter regions of a number of other genes which display a similar pattern of Stat5 site consensus sequences.

# Chapter 2

# Materials and Methods

## 2.1 General DNA Manipulation and Analysis

### 2.1.1 Digesting DNA with site-specific restriction enzymes

All genomic and plasmid DNA was digested with various restriction enzymes as per the manufacturers' guidelines (New England Biolabs; Roche Diagnostics) with their buffers, except for the following:

- 2-3 fold excess of enzyme was used to digest plasmid DNA in one hour
- 6-12 fold excess of enzyme was used to digest genomic DNA overnight

### 2.1.2 Isolation of plasmid DNA from agarose gels

10% loading buffer (15% Ficoll, 0.25% Orange G) was added to the plasmid digests before loading onto an agarose gel of between 1.5 and 3% (w/v) in Tris-acetate/EDTA buffer (40mM Tris(hydroxymethyl) methylamine, adjusted to pH7.7 with glacial acetic acid; 1mM diaminoethanetetra-acetic acid disodium salt) or Tris-borate/EDTA buffer (50mM Tris(hydroxymethyl) methylamine, adjusted to pH8.8 with 16mM boric acid; 1mM diaminoethanetetra-acetic acid disodium salt). After electrophoresis DNA fragments were visualised using long wavelength UV light (304nm) which causes intercalated ethidium bromide (250μg/litre buffer) to fluoresce. The appropriate fragments were cut out using a fresh scalpel blade and finely sliced on a petri-dish.

Two methods were used to extract the DNA from the agarose gel matrix:

**Method 1:**

A hole was made in the bottom of a 0.5ml microcentrifuge tube and a small amount of glass wool pressed down towards the hole. The finely sliced agarose was pushed down into the tube which then was placed in a 1.5ml microcentrifuge tube (Figure 20). The column was frozen at -20°C for at least 30 minutes to break up the

agarose gel matrix and release the DNA. The extraction column was then micro-centrifuged at 14K rpm for 5 minutes to elute the DNA.

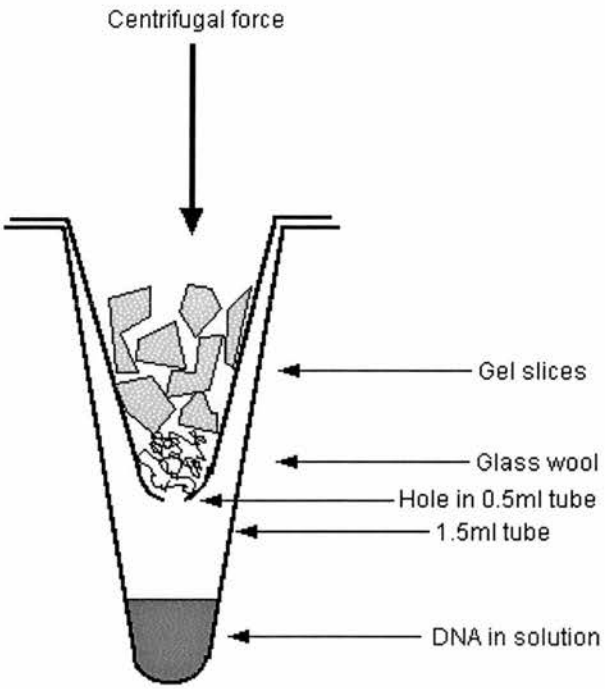



**Figure 20** DNA Extraction Column from Agarose

**Method 2:**

Qiagen QIAquick Gel Extraction Kit was used latterly for higher DNA yields and limited the need for further purification. Briefly, the gel slice was solubilised using QG1 solution and this solution was centrifuged through a column which bound the DNA. After a wash step to remove salt, the DNA was eluted by adding 30µl TE (10mM Tris, adjusted to pH8 with glacial acetic acid; 0.1mM EDTA) to the column and centrifuging for 1 minute.

### 2.1.3 Phenol:Chloroform Extraction

Phenol:chloroform extraction was used to extract and purify nucleic acids from a variety of sources (Current Protocols in Molecular Biology (2.1.1)), including the eluate from the DNA gel extraction spin column (Figure 20). Phenol solubilises hydrophobic protein groups while the aqueous solution solubilises hydrophilic groups. This strips protein from the DNA and aqueous solution, and leaves it in the

phenol:aqueous interface. Addition of chloroform assists in protein denaturation therefore aiding the dissociation from the DNA. The high density of chloroform also enhances the separation of phases, facilitating the removal of the aqueous phase with little cross-contamination from organic material.

An equal volume of 1:1 phenol (equilibrated to pH7.6 by repeated extraction with Tris buffer; Fisher Scientific UK) and chloroform was added to each sample in a microcentrifuge tube, mixed by inverting the tube and micro-centrifuged at 14K rpm for 5 minutes to separate the phases (see Appendix I for rpm to rcf conversion table). The top aqueous phase was transferred into a clean tube taking care not to disturb the proteinacious interface.

An equal volume of chloroform was added to this aqueous solution to remove traces of phenol which would otherwise inhibit subsequent enzymatic reactions. The two phases were mixed by inverting the tube and micro-centrifuged for 1 minute. The upper aqueous phase was transferred to a clean tube and the DNA precipitated using ethanol.

### 2.1.4 DNA Precipitation with Ethanol

DNA molecules in solution are surrounded by a 'cloak' of polarised water molecules. The negatively charged phosphate backbone of DNA attracts the hydrogen atoms in water molecules, polarising the water molecule and delocalising electrons towards the oxygen atom (Figure 21). This, in turn, attracts more water molecules and forms a shield keeping the DNA in solution. The association of the phosphate groups with ions such as $Na^+$ and $K^+$ in solution decreases this effect.

Addition of 100mM sodium acetate and two volumes of ethanol competed off these water molecules. Incubation at -20°C overnight or -80°C for 20 minutes forced the DNA out of solution. Micro-centrifugation at 14K rpm for 20 minutes pelleted the DNA. The supernatant was removed.

The sodium acetate was removed by adding 500µl 70% ethanol diluted with water. In subsequent enzymatic reactions it might otherwise generate star activity or cause inhibition. The salt can dissolve in 70% ethanol but the DNA cannot. After a 2 minute incubation the DNA was pelleted as before. The supernatant was removed and the pellet left to dry at room temperature.

This is a particularly critical stage. If the pellet has not dried enough, ethanol contamination will inhibit ligation reactions and endonucleases. If the pellet is too dry, plasmid DNA can be difficult, and genomic DNA impossible, to solubilise. I found that adding TE (10mM Tris HCl, 0.1mM EDTA) immediately once the rim of the pellet turned clear overcame this problem.

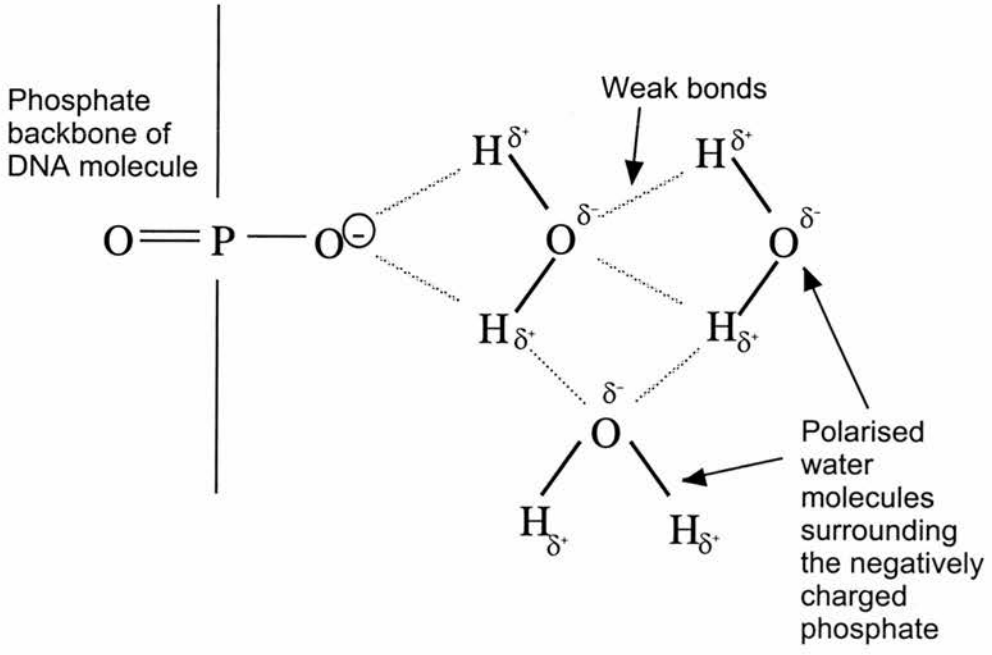
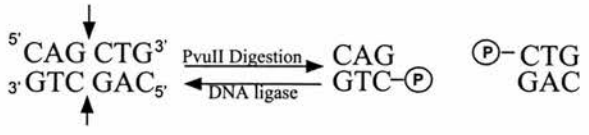


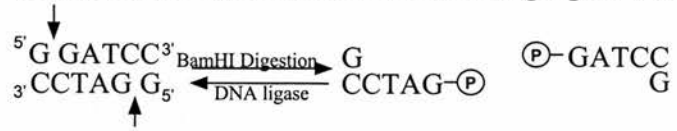

**Figure 21** Solubilisation of a DNA Molecule

### 2.1.5 DNA Ligation

Restriction enzyme generated DNA fragments possess either 'blunt' or 'sticky' ends which can be rejoined by DNA ligase. An enzyme such as ***Pvu*II** recognises the following sequence to produce a non-overhanging or 'blunt' end:




An alternative enzyme such as ***Bam*HI** produces overhanging or 'sticky' ends:




'Sticky-ended' fragments contain a few bases of overlap between strands and so can anneal prior to ligation. With the sites to be ligated in intimate contact, DNA ligase can quickly covalently link them together. Owing to this lack of homology, 'blunt-

ended' ligations require fifty times more enzyme and a warmer temperature in order for DNA ligase to work effectively.

'Blunt' end ligations were performed with a 1:3 parts plasmid : insert molar ratio, whereas, 'sticky' end ligations only required a 1:2 parts plasmid : insert molar ratio. Vector sequences were always phosphatased prior to ligation to prevent self-ligation by adding Shrimp Alkaline Phosphatase (SAP) (manufacturer: USB). Once the restriction digest was complete, 1µl (1Unit) SAP was added and incubated at 37°C for an additional 20 minutes.

A typical ligation reaction contained:

- 100ng vector and insert DNA (for 'sticky' end ligations)
- 500ng vector and insert DNA (for 'blunt' end ligations)
  -in 8µl $H_2O$
- 1µl 10x ligation buffer (Roche Diagnostics)
  (500mM Tris-HCl pH7.8; 100mM $MgCl_2$; 100mM DTT; 10mM ATP; 500µg/ml BSA)
- 1µl DNA ligase at 1U/µl (Roche Diagnostics)
  - total 10µl

'Sticky' end ligations were performed at 16°C overnight while 'blunt' end ligations were left at room temperature overnight.

### 2.1.6 Transformation of competent cells with plasmid DNA

Two primary cell types were employed during this work. (1) *E. Coli* DH5α competent cells (genotype: F-, F80δ*lac*ZΔM15, Δ(*lac*ZYA-*arg*F), U169, *deo*R, *rec*A1, *end*A1, *hsd*R17(rk-, mk+), *pho*A, *sup*E44, λ-*thi*-1, *gyr*A96, *rel*A1) were purchased from Gibco BRL and were used for all plasmid transformation procedures. (2) For *in vitro* reconstitution work *E. Coli* DH11S cells were used (genotype: F′, *mrc*A, Δ(*mrr hsd*RMS*mcr*BC), Δ(*lac-pro*AB), Δ(*rec*A1398), *deo*R, *rps*L, *srl*, *thi*, F′*pro*AB+, *lac*IqZΔM15). These provided the F′ pillus required for helper phage infection. Both cell types are competent when purchased.

The transformation procedure was the same for both strains. Cells were thawed on ice and gently mixed with a yellow pipette tip. About 10ng of ligated plasmid DNA was added to 50µl of cells and placed on ice for 30 minutes. After a

heat shock of 60 seconds at 42°C they were returned to ice for 2 minutes. 450µl of LB medium (In 1 litre: 10g Tryptone, 5g yeast extract, 5g NaCl, 2g $MgCl_2$) was added (without antibiotic) and incubated at 37°C for 1 hour with vigorous shaking.

Agar plates (15g agar per 500ml LB medium) were prepared containing 100µg/ml ampicillin to select for plasmid bearing bacteria. They were overlaid with 50µl 50mg/ml 5-bromo-4-chloro-3-indolyl-β–D-galactoside (X-gal) and 50µl 0.2M isopropyl-β–D-thiogalactopyranoside (IPTG) and left at 37°C to dry. 50-100µl of cells were spread on each plate using a sterile glass spreader. Plates were incubated overnight at 37°C.

Blue/white colony screening was used to identify bacterial colonies containing an insertion in the multiple cloning region (MCR) of pBluescript vectors. The *E.coli lacZ* gene codes for β–galactosidase which is induced with IPTG. It forms dark blue plaques when plated on *lac-* hosts such as DH5α in the presence of the chromogenic substrate X-gal. An insertion into the MCR disrupts the gene, forming white rather than blue colonies which can then be selected.

### 2.1.7 Preparation of Plasmid DNA

*Small-scale preparation*

Two different methods were used. Firstly, alkali lysis as described by Current Protocols in Molecular Biology (1.6.1):

5ml LB medium, containing 100µg/ml ampicillin in 15ml tubes was inoculated with a single white bacterial colony and incubated overnight at 37°C with vigorous shaking. 1.5ml of cells was pelleted at 14K rpm in a microcentrifuge for 30 seconds and the supernatant removed. They were resuspended in 100µl GTE buffer (50mM glucose, 25mM Tris-HCl pH8, 10mM EDTA pH8) leaving no cell clumps. After 5 minutes, 200µl of freshly prepared lysis solution was added (1% w/v SDS, 0.2M NaOH) and cells were left at 4°C for 5 minutes. Next, 150µl 5M potassium acetate solution was added and the mixture vortexed for 2 seconds before placing on ice for 5 minutes. The cell debris and chromosomal DNA were removed by micro-centrifuging at 14K rpm and transferring the supernatant into a fresh tube containing 800µl ethanol. After 2 minutes the plasmid DNA and RNA were pelleted by

centrifuging as before for 1 minute. A 70% ethanol wash removed excess salt and SDS. The dry pellet was resuspended in 30µl TE.

RNA was removed from the preparation by adding RNase A. The protocol recommends adding 1µl of 10mg/ml RNase A per digestion, however, this causes aberrant band shifting (Benore-Parsons *et al.*, 1997). Adding 0.25µl RNase A digested all the RNA in the majority of mini-preps without causing any shifting.

Secondly, the Qiagen QIAprep Spin Miniprep Kit was used, following the manufacturer's instructions:

Bacteria were grown and pelleted as in the alkali lysis protocol. 250µl of buffer P1 was added to resuspend the cells and remove RNA. 250µl of buffer P2 lysed the cells and the mixture was centrifuged for 10 minutes after addition of 350µl of buffer N3. The supernatant was applied to the QIAprep column and centrifuged for 30-60 seconds. The bound DNA was washed with 750µl of buffer PE and, once dry, eluted in 50µl of 10mM Tris-HCl pH8.5.

*Large-scale preparation*

The Qiagen QIAfilter Plasmid Maxi Kit was used for isolating large amounts of pure plasmid DNA. The protocol is described briefly as follows:

A single colony was grown overnight in 5ml LB medium plus 100µg/ml ampicillin, minipreped the next day and checked by digestion with an appropriate restriction enzyme. Approximately 500µl of this culture was used to inoculate LB medium for the bulk preparation. For low copy number plasmids such as pPolyIII, 500ml LB medium was inoculated. High copy number plasmids such as pUC and pBluescript were grown in 100ml LB medium for a similar yield. Ampicillin was added at 100µg/ml ampicillin as above for selection purposes.

The next day the bacteria were pelleted at 6K rpm for 10 minutes at 4°C in Sorvall Dry-Spin centrifuge tubes in an F-16 rotor. The supernatant was removed and the cells resuspended in 10ml of buffer P1, leaving no cell clumps. 10ml of buffer P2 lysed the cells over a period of 5 minutes after which 10ml buffer P3 was added to precipitate cell debris and chromosomal DNA. This mixture was chilled on ice for 15-20 minutes before being centrifuged as above at 13K rpm for 30 minutes.

Meanwhile, a Qiagen-tip 500 DNA binding column was prepared. 10ml buffer QBT was applied to the column and allowed to drain. Folded, pre-wetted filter paper was placed on top of the column to remove any suspended or particulate matter which might otherwise clog the column.

Once the bacterial lysate finished centrifuging, the supernatant was promptly loaded onto the filter paper and allowed to enter the column via gravity flow. Traces of RNA and protein were removed by two 30ml washes of buffer QC. Finally the DNA was salt-eluted from the column by adding 15ml of buffer QF. Addition of 10.5ml isopropanol precipitated the DNA while leaving most of the salt in solution.

The DNA was pelleted by centrifugation at 11K rpm for 30 minutes at 4°C in a 50ml Corex tube using a Sorvall SS-34 rotor. The remaining salt was removed by adding 5ml of 70% ethanol and centrifuging as above for 10 minutes. Once the pellet was dry, the DNA was dissolved in 300µl TE and its concentration determined by UV spectrophotometry.

## 2.2 Isolation of Mouse and Sheep Nuclei

### 2.2.1 Tissue Dissection

A Merino sheep was sacrificed in mid-lactation to ensure that the β-lactoglobulin gene was fully active in the mammary gland. Euthatal (Supplier: Vetdrug) was used to over-anaesthetise the animal while causing no tissue damage. Guidelines advise 1ml euthatal per 1.4kg sheep weight, but a 40-50ml injection is routinely given to ensure no complications. An average Merino sheep weighs between 50-60kg.

The entire mammary gland and a lobe of the liver were removed separately with clean scalpel blades. 50ml tubes were filled with each tissue which had been diced into approximately $1cm^3$ pieces. These were quickly frozen in liquid nitrogen to prevent cellular endonucleases degrading the genomic DNA.

For analysis of mouse β-casein and mouse α–lactalbumin gene promoter regions, a (CBA x C57Bl/6)F1 female mouse was culled at 10 days post-partum when lactation is at a maximum and milk proteins are most highly expressed. The mammary and liver tissues were immediately frozen in liquid nitrogen.

Goat liver and lactating mammary tissue was kindly provided by R. Pena, Unitat de Genètica, Fac. Veterinària, Bellaterra - UAB, Barcelona, Espagne. This was used to analyse goat BLG and β–casein genes in expressing and non-expressing tissues.

All tissues were stored in liquid nitrogen until required and an entire tube of frozen tissue was prepared at a time.

### 2.2.2 Nuclei Preparation

Tissue from each tube was placed in a mortar containing liquid nitrogen. At no time was it allowed to cool down from -192°C. Using a pestle, the tissue was ground down into a fine powder while constantly refilling the mortar with liquid nitrogen. This mix was poured into several 50ml tubes containing 10ml of ice-cold Tissue Resuspension Buffer (TRB: 15mM Tris-HCl pH7.4, 2mM EDTA, 1mM EGTA, 15mM NaCl, 60mM KCl, 0.15mM spermine, 0.5mM spermidine and 0.2% v/v IGEPAL-CA630)(modified from Whitelaw *et al.*, 1992). An additional 10ml of TRB was poured on top. The upper and lower layers of buffer assist in resuspending the tissue.

A plastic pastette was used to break up the frozen tissue into a homogenous solution which was kept chilled on ice. A Dounce Homogeniser was used to free nuclei from the cells by shearing the cells through a 150-250μm gap between the plunger and holder. Approximately 8ml of tissue was homogenised at a time. Liver tissue only required one stroke with the plunger owing to the low amount of connective tissue. The mammary tissue required 2 strokes.

The homogenised tissue was filtered through two layers of pre-wetted Miracloth paper to remove the majority of the cell debris. This filtered solution was centrifuged at 1K rpm for 10 minutes at 4°C to pellet the nuclei in a Jouan CR3000. Lower molecular weight cell debris remains in the supernatant which was removed to leave the pellet.

The nuclei pellet was resuspended in Nuclei Resuspension Buffer (NRB: 15mM Tris-HCl pH7.4, 2mM EDTA, 1mM EGTA, 15mM NaCl, 60mM KCl, 0.15mM spermine, 0.5mM spermidine) to a volume of 1ml and centrifuged at 6K rpm at 4°C for 5 minutes in a microcentrifuge. The supernatant was then removed

and the nuclei were resuspended in ice-cold storage buffer (50% glycerol, 75mM NaCl, 0.5mM EDTA, 20mM Tris-HCl pH7.9, 0.85mM DTT and 0.125mM PMSF) and frozen at -80°C until required.

### 2.2.3 Estimating Nuclei Concentration

Before use, the nuclei were removed from the -80°C freezer and thawed on ice. They were washed three times in NRB in preparation for micrococcal nuclease (MNase) digestion or three times in NBII (60mM KCl; 15mM NaCl; 15mM Tris-HCl pH7.4; 0.3M sucrose; 0.5mM spermidine; 0.15mM spermine & 0.1mM PMSF) in preparation for cuprous phenanthroline digestion (see below). A 10µl aliquot was diluted 100 fold in the same buffer, NRB or NBII, and 8µl of this diluted sample was loaded onto a haemocytometer to estimate the nuclei concentration.

A haemocytometer contains a grid pattern in which each large square contains 100nl of liquid between it and the cover slip once loaded with sample. The nuclei and grid pattern were visualised using a Nikon Phase Contrast microscope. Nuclei in four of these squares were counted and an average taken for a more accurate result. Multiplying the nuclei average by the volume of one square divided by 1ml (1ml/100nl or $1 \times 10^{-3} / 1 \times 10^{-7} = 10{,}000$) and the dilution factor of 100 gives the nuclei concentration per millilitre:

e.g. If the nuclei average from 4 squares is 300,

10,000 (vol. conversion) x 100 (dilution factor) x 300 (nuclei in that volume) =

$3 \times 10^{8}$ nuclei/ml

### 2.2.4 Protein-Free DNA Extraction

Sheep and mouse high molecular weight genomic DNA for positive and negative controls for Southern blotting were prepared from the isolated nuclei. They were washed three times in NBII buffer as above and then lysed by adding SDS to 1% volume and 10µl of 2mg/ml proteinase K.

After incubating at 37°C for at least four hours the DNA was phenol:chloroform extracted. This was made difficult because the DNA was of extremely high molecular weight. To ease extraction, it was sheared by syringing the DNA through a fine bore needle (0.2mm) several times prior to extraction. This

reduced the fragment sizes of DNA to no smaller than 50kb which did not interfere with subsequent manipulations.

After extraction, the DNA was ethanol precipitated, washed in 70% ethanol and resuspended, once dry, in 300µl TE. The concentration was calculated by spectrophotometry at 260nm at which wavelength its extinction co-efficient is 1mg/ml per 20OD Units.

### 2.2.5 Nuclear DNA Digestion using Micrococcal Nuclease

The washed and counted nuclei were digested by adding 120µl MNase (1800U) per $2x10^8$ nuclei in a volume of 1ml. The reaction was initiated by adding $CaCl_2$ to a final concentration of 9mM at room temperature. 150µl aliquots were removed at: 0, 2, 5, 7, 10 & 15 minutes and the MNase was inhibited by adding EDTA to 100mM. The nuclei proteins were digested by adding 10µl of 10mg/ml proteinase K and lipids were dissociated with SDS to 1% volume. After at least 4 hours, samples were phenol:chloroform extracted, ethanol precipitated, washed with 70% ethanol and resuspended in 300µl TE. The DNA concentration was calculated using spectrophotometry.

*Sheep Naked DNA Control with MNase*

In a total volume of 200µl NRB (without spermine and spermidine), 100µg naked sheep genomic DNA was digested with 30µl (450U) MNase in 9mM $CaCl_2$. The reaction proceeded at 4°C for 20 minutes when it was terminated with EDTA to 100mM. After phenol:chloroform extraction it was ethanol precipitated and washed in 70% ethanol. Finally it was resuspended in 75µl TE.

### 2.2.6 Nuclear DNA Digestion using Cuprous Phenanthroline

The following procedure was based on those developed by others (Cartwright & Elgin, 1982; Quivy & Becker, 1996).

Nuclei in storage buffer were thawed from -80°C on ice and micro-centrifuged at 3K rpm for 5 minutes. The pellet was resuspended in NBII buffer (60mM KCl; 15mM NaCl; 15mM Tris-HCl pH7.4; 0.3M sucrose; 0.5mM spermidine; 0.15mM spermine & 0.1mM PMSF)(Caplan *et al.*, 1987). The nuclei

were centrifuged again, and resuspended in NBII buffer. This wash step was repeated once more. The nuclei concentration was estimated using a haemocytometer (**2.2.3**) and 4.2 x$10^7$ nuclei were made up to 0.5ml with NBII buffer. This amount of nuclei is equivalent to 250µg of genomic DNA. They were incubated in a water bath at 22°C until they reached temperature.

The cuprous phenanthroline (OP-Cu) complex was prepared by mixing 50µl of 40mM 1,10-phenanthroline with 50µl of 9mM copper II sulphate (Cartwright & Elgin, 1982). After incubation at room temperature for 1 minute the complex was diluted to 2ml with water. 50µl of the complex was added to the nuclei and after 2 minutes the reaction was initiated by addition of 50µl of 58mM 3-mercaptopropionic acid. The reaction proceeded at 22°C and 95µl aliquots were taken at the following time points: 0, 10, 20, 30, 40 & 60 minutes.

Each aliquot was immediately quenched with 9µl of 28mM neocuproine. Once all time points had been completed, the nuclei were centrifuged at 3K for 5 minutes and the supernatant removed. The pellet was resuspended in 100µl of STOP MIX I (20mM Tris-HCl pH8; 20mM NaCl; 20mM EDTA; 1% SDS(w/v) & 600µg/ml proteinase K) and 50µl of STOP MIX II (150mM NaCl & 5mM EDTA), and incubated at 37°C for at least 3 hours. The DNA was deproteinated by phenol:chloroform extraction, followed by ethanol precipitation and a 70% ethanol wash. The DNA was resuspended in 300µl TE and its concentration calculated using UV spectroscopy.

20µg of each time point was digested with an appropriate restriction enzyme such as BamHI. A restriction site on the gene acted as a reference point from which nucleosome positions could be calculated.

*Sheep Naked DNA Control with OP-Cu*

The following were added to a 1.5ml microcentrifuge tube:

| | |
|---|---|
| sheep DNA | x µl (100µg) |
| 3M KCl | 6µl |
| 3M NaCl | 1.5µl |
| 1M Tris-HCl (pH 7.4) | 4.5µl |
| $H_2O$ | -to 300µl |

The conditions are the same as those for nuclear DNA digestion except for the omission of a number of chemicals: Sucrose was only required as a buffer to keep the nuclei intact and was not required for naked DNA stability. Spermine and spermidine are important in stabilising the nucleosomal structure of DNA in nuclei. These may have imposed unnatural structural constraints on naked DNA and so were also omitted (Pelta *et al.*, 1996).

The reaction was incubated at 22°C for approximately 5 minutes when 33µl $CuIISO_4$ - 1,10-phenanthroline complex was added. Two minutes later, once it had associated with the DNA, the reaction was initiated with 33µl 58mM 3-mercaptopropionic acid. After 75 seconds the reaction was quenched with 33µl neocuproine. After phenol:chloroform extraction it was ethanol precipitated and washed in 70% ethanol. Finally it was resuspended in 75µl TE.

20µg of this DNA was digested with an appropriate restriction enzyme. This digested the DNA at a site abutting the probe used in subsequent indirect end-labelling reactions.

## 2.3 Southern Blotting

### 2.3.1 Transfer onto Zetaprobe GT Membrane

By transferring the DNA from the agarose gel on to a membrane, it can be examined for the presence of specific DNA fragments. This is achieved by the flow of liquid through the gel and membrane. The DNA is carried with the liquid but cannot pass through the membrane and is trapped there by cross-linking (Current Protocols in Molecular Biology (2.9A)).

The gel did not require any treatment before blotting. Acid depurination is commonly used to fragment the DNA so that higher molecular weight fragments are transferred more easily. This was not needed as fragments up to 4.4kb could be visualised quite clearly. Pre-treatment with NaOH in order to hydrolyse the phosphodiester backbone at the site of depurination and to denature the DNA duplex was also not required. The DNA is transferred through the gel with 0.4M NaOH which denatures it and fixes it to the membrane.

The Southern Blotting apparatus (Figure 22) shows how the alkali solution can travel from the tank, via the 3MM paper wick, through the gel by the capillary action of the paper towels. The DNA in the gel is carried up to the membrane with the alkali solution and is fixed there by cross linking. The transfer process is fastest with small fragments of DNA, although all required sizes are transferred after 16 hours or overnight.

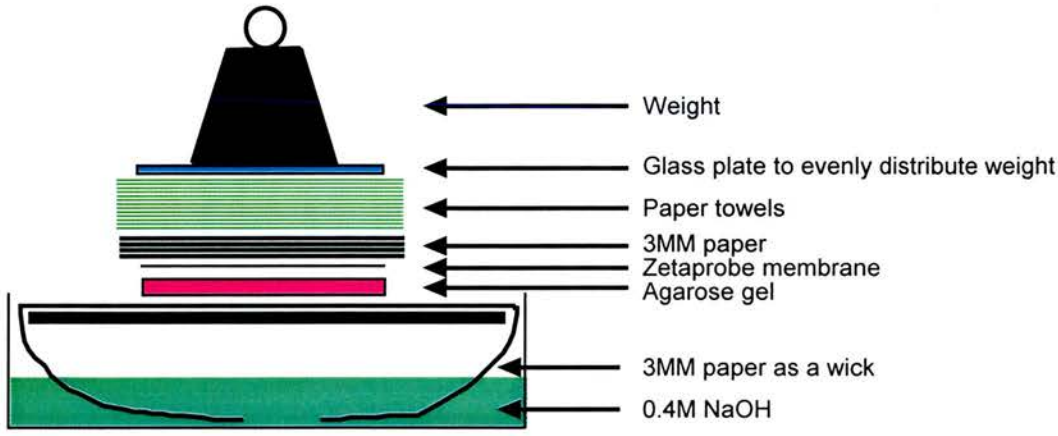



**Figure 22** Southern Blotting Apparatus

The three sheets of 3MM paper above the membrane prevent it being contaminated with any debris from the paper towels. The weight at the top of the stack was never more than 500g in order to prevent transfer being impeded

After transfer, the membrane was washed in 2x SSC (30mM tri-sodium citrate; 300mM NaCl; adjusted to pH7) to remove traces of agarose and fibres from the 3MM paper. It was then sandwiched between fresh 3MM paper and baked at 80°C for 1 hour. The membrane was stored between fresh 3MM paper, sealed in a plastic bag, until it was ready to be hybridised.

### 2.3.2 Probe Labelling for Hybridisation to Membrane

The 'random primed' labelling method was originally developed by Feinberg and Vogelstein (Feinberg & Vogelstein, 1983; Feinberg & Vogelstein, 1984). It is based on the hybridisation of oligonucleotides of all possible hexamer sequences to the denatured DNA to be labeled. $\alpha$-$^{32}$P labeled dCTP is incorporated into the complementary strand as it is synthesised by Klenow polymerase.

100ng of probe was made up to a volume of 11µl in distilled water. It was boiled for 5 minutes and placed on ice to cool for 2 minutes. 4µl of High Prime (Roche Diagnostics) was added which contains all constituents required for a 'random primed' labelling reaction except for radioactive dCTP. 5µl of $\alpha$-$^{32}$P labelled dCTP was added in the 'hot lab' and the mixture was incubated at 37°C for 10 to 20 minutes.

While the labelling reaction proceeded, a G-50 'Nick Column' (Pharmacia Biotech) was prepared as per the manufacturer's instructions. The sephadex column removes unincorporated nucleotides and thus reduces background on the final image. The storage buffer was removed and the column rinsed with TE (10mM Tris-HCl pH8; 0.1mM EDTA pH8). Fresh TE was then added and allowed to pass through the column by gravity flow.

Once the column was ready and the labelling reaction complete, the probe mixture was added to the top of the column with 400µl TE. This eluate was discarded. A second 400µl TE was added to the column and this eluate contained the labelled probe.

The probe was denatured by adding 50µl 4M NaOH just before addition to the hybridisation solution.

### 2.3.3 Testing Radiolabel Incorporation

2µl of the 400µl probe mixture was spotted on to a small piece of filter paper and placed in a plastic vial. A Wallac 1410 β–counter (Pharmacia Biotech) was used to analyse β–emission levels. All probes had to have a total activity of at least 1 x$10^8$ cpm.

### 2.3.4 Hybridisation

The baked Zetaprobe membrane was placed in a Techne FHB11 roller bottle and wetted for at least 5 minutes with 30ml hybridisation solution (0.5M $Na_2HPO_4$ pH7.2; 7% SDS). It was evenly coated as it rotated in a Techne Hybridiser HB1D oven at 65°C. Once the probe was ready, the old hybridisation solution was removed and 20ml fresh hybridisation solution was added to the roller bottle with the denatured probe. The hybridisation continued overnight for at least 16 hours.

### 2.3.5 Hybridisation Washes and Exposure on Film

Once the hybridisation solution was removed, the membrane was washed twice in 70ml wash buffer 1 (5% SDS; 40mM $Na_2HPO_4$ pH7.2) for 15-30 minutes per wash at 65°C. Two further, more stringent, washes removed the remainder of non-specifically hybridised probe. Likewise, 70ml wash buffer 2 (1% SDS; 40mM $Na_2HPO_4$ pH7.2) was added for 15-30 minutes per wash. Next, the membrane was removed from the roller bottle and immediately sealed in 2 layers of Saran wrap. The membrane was scanned with a Geiger counter to assess β–emission levels and was laid down on Afga film. The length of incubation at -80°C was proportional to the β–emission levels and was typically 1-3 weeks.

## 2.4 An *In Vitro* Assay for Nucleosome Positioning: Monomer Extension

### 2.4.1 Cloning Gene Promoter Regions

The following promoter region fragments were cloned into phagemid pBluescript II KS (-) in both orientations so that ssDNA could be made from both the upper and lower strands. Monomer extension could then be tested in both directions and examined to make sure the results correlate.

**Ovine β-lactoglobulin Gene**

(Genebank Accession No.X12817) (Harris *et al.*, 1988)

pSStg1XS contains the entire BLG locus from –4.2kb (***Sal*I**) to –6.6kb (***Xba*I**) inserted into pUC19 using the same sites in the vector. ***Rsa*I** digestion of this vector yields a 1.7kb fragment of ovine BLG from -864 (***Rsa*I**) to +882 (***Rsa*I**). This was inserted into pBluescript II KS (-) at the ***Eco*RV** site by blunt-ended ligation.

The plasmids produced were named pBlueBLG(+) and pBlueBLG(-). The +/- refers to the orientation of the insert: in pBlueBLG(+), the numbering of the insert increases in the same direction as those of the vector sequence. The numbering of the pBlueBLG(-) insert decreases as those of the vector increase.

**Caprine β -lactoglobulin**

Genebank Accession No. Z33881 (Folch *et al.*, 1994)

pSX(12.5) contains the goat β–lactoglobulin gene from approximately -6kb (***Sal*I**) to +6.9kb (***Xba*I**) with respect to the start site (kind gift from R. Pena, M. Josep and A. Sanchez , Unitat de Genètica, Fac.Veterinària, Bellaterra-UAB, Barcelona, Espagne). Using the same ***Sal*I** and ***Xba*I** sites, it was inserted into pBluescript II SK (+). Digestion with ***Bsr*BI** and ***Pvu*II** generated a 0.75kb fragment from -727 (***Bsr*BI**) to +31 (***Pvu*II**). Using blunt ended ligation it was inserted into the ***Eco*RV** site of pBluescript II KS (-).

Unfortunately, the fragment preferentially inserted in only one direction. This was the case in over 48 minipreps for no apparent reason. To overcome the problem a different approach was taken: the pBluescript II KS (-) vector contains ***Bss*HII** sites on either side of the multiple cloning region. Digestion with this enzyme and subsequent re-ligation should generate plasmids in which 50% contain the insert in the opposite direction. This was successful and produced pK'BLG(R) where 'R' stands for Reverse orientation, i.e. the numbering of the insert decreases as those of the vector increase. This was the complement of pK'BLG(D) generated earlier, where 'D' stands for Direct orientation. " K' " is an abbreviation for caprine.

**Caprine β-casein**

Genebank Accession No. M90559 (Roberts *et al.*, 1992)

pI2B contains 5kb of the caprine β–casein gene from approximately -3kb (***Sal*I**) to +2kb (***Not*I**). It was cloned into the pUC19 vector using the same restriction sites (kind gift from J-L. Villote, Jouy-en-Josas, France). Partial digestion with ***Bgl*II** generated a 0.45kb fragment from -309 (***Bgl*II**) to +139 (***Bgl*II**). This was inserted into pBluescript II KS (-) using a ligation-compatible ***Bam*HI** site to generate pK′bcasR and pK′bcasD.

**Mouse β-casein**

Genebank Accession Nos. X13484, X15991 & X15992 (Yoshimura *et al.*, 1989)

A 6.6kb ***Eco*RI** fragment of the β–casein gene from C57Bl/6 strain of mouse (Goodman and Rosen 1990) was cloned into the ***Eco*RI** site of pPolyIII-D (Lathe *et al.* 1987). The resultant plasmid which contains -1.4kb to +5.2kb the mouse β–casein gene was named of pPolyIIIDβ (kind gift from S. Kumar).

***Nsi*I** digestion of pPolyIIIDβ resulted in a 0.64kb fragment from -545 (***Nsi*I**) to +95 (***Nsi*I**). This was inserted into pBluescript II KS (-) using a ligation-compatible ***Pst*I** site to generate pmbcasR and pmbcasD.

**Mouse α-lactalbumin**

Genebank Accession No. M87863 (Vilotte and Soulier, 1992)

A ~600bp region of the mouse α–lactalbumin gene promoter was linked to the chloramphenicol acetyl transferase (CAT) gene coding sequence in a pB9 vector (kind gift from J-L. Villote, Jouy-en-Josas, France). This was excised with ***Bam*HI** which digests the plasmid on either side of the promoter region at -561 (***Bam*HI**) and +50 (***Bam*HI**) in the polylinker region of the plasmid.

This mouse α–lactalbumin gene fragment was ligated into pBluescript II KS (-) at the ***Bam*HI** site. Unfortunately, the fragment only inserted in one orientation to produce pmusalacR. The insert was rotated in the opposite direction by the same means as the caprine BLG fragment, by digestion with ***Bss*HII** and re-ligation, to produce pmusalacD.

**Human α-lactalbumin**

Genebank Accession No. X05153 (Hall *et al.*, 1987)

pHA1 (kind gift from PPL Therapeutics) contains a 7kb fragment of the human α–lactalbumin gene from -2kb (***Sau*3AI**) to +5kb (***Sau*3AI**) in the ligation-compatible BamHI site of the pUC18 vector.

Using ***Acc*I** and ***Pvu*II** restriction enzymes, the promoter region was excised from -570 (***Acc*I**) to +113 (***Pvu*II**) and inserted in one orientation into pBluescript II KS (-) between the ***Acc*I** and ***Eco*RV** sites to produce phumalacR. The fragment was

cloned in the reverse orientation, as before, by digestion with ***Bss*HII** and re-ligation to produce phumalacD.

### 2.4.2 Preparation of ssDNA

All ten clones were initially transformed into *E.coli* strain DH5α, tested by miniprep, grown in bulk culture and isolated by Qiagen QIAfilter Plasmid Maxi Kit. This DNA was used for nucleosome reconstitution. The same DNA was transformed into *E.coli* strain DH11S for subsequent preparation of ssDNA.

**Buffers**

All solutions were autoclaved unless indicated:

**TBG** (in 500ml $H_2O$): 6g bactotryptone, 12g yeast extract, 2ml glycerol, 1.16g $KH_2PO_4$, 4.79g $K_2HPO_4$ -add 5ml 2M glucose immediately before use

**SOB** (in 250ml $H_2O$): 5g bactotryptone, 1.25g yeast extract, 625µl 1M KCl, 833µl 3M NaCl, 5ml 500mM $MgCl_2$, 5ml 500mM $MgSO_4$

**PEG/NaCl** (in 500ml $H_2O$): 40% polyethylene glycol 8000, 2.5M NaCl

**5x M9 Salts** (in 200ml $H_2O$): 12.8g $Na_2HPO_4.7H_2O$, 3g $KH_2PO_4$, 0.5g NaCl, 1.0g $NH_4Cl$

**M9 Minimal Medium** (un-autoclaved):

3g agar in 150ml $H_2O$ (autoclaved) -keep at 50˚C while adding: 40ml 5x M9 Salts (also at 50˚C), 0.8ml 0.5M $MgSO_4$, 2.22ml 2M glucose, 20µl $CaCl_2$, 19.3µl of 10µg/µl thiamine, 193µl of 100µg/µl ampicillin

**TYP** (per litre $H_2O$): 16g bactotryptone, 16g yeast extract, 5g NaCl, 2.5g $K_2HPO_4$

**B Broth** (per litre $H_2O$): 10g bactotryptone, 8g NaCl

-add 5.56ml 2M glucose immediately before use

**B Top**: B Broth & 0.7% agar

**B Bottom**: B Broth &1.5% agar

### 2.4.3 Helper Phage Titre

M13KO7 helper phage stock was kindly provided by A. Fleming of the Pennings laboratory, Dept. of Biochemistry, University of Edinburgh.

A DHIIS plating culture was set up by inoculating 5ml of TYP with 100µl DHIIS M9 glycerol. It was incubated at 37°C for about 6 hours at 300rpm with shaking to reach mid to late log phase. Meanwhile, a 10-fold serial dilution of the helper phage stock was prepared. Assuming an initial concentration of $1x10^{8}$ to $1x10^{11}$pfu/ml, it was diluted seven times (10µl phage and 90µl B Broth at each stage) so that the last dilution contained between 10 and $1x10^{4}$pfu/ml.

10µl neat phage, 10µl of each phage dilution and 10µl B Broth, as a negative control, were transferred to the bottom of 4.5ml polypropylene tubes. 190µl of the DHIIS plating culture was added to each tube and left at room temperature for 5 minutes. 4ml of B Top (at 55°C) was added to each tube and then the mix was poured on to B Bottom plates, left to dry and incubated overnight at 37°C.

The number of plaques formed was counted and the concentration was estimated at $7.5x10^{9}$pfu/ml.

### 2.4.4 Isolation of Phagemid Containing DH11S F' Colonies

Individual colonies of transformed DHIIS bacteria were picked from LB agar plates and streaked onto M9 minimal medium plates. DHIIS bacteria are deficient for proline biosynthesis owing to a deletion in the *pro* operon. The F' episome contains the complete *pro* operon and as the minimal medium lacks a source of proline, only F' containing bacteria should grow. The F' episome is essential for ssDNA synthesis.

A small amount of LB agar carry-over is inevitable when re-streaking colonies. This will contain proline which is lacking in the minimal medium plates and will therefore allow F' deficient bacteria to grow. To be confident that each selected colony contains the F', two screens were performed. Four colonies were initially streaked per plate and positive colonies isolated. These were then re-streaked on fresh minimal medium plates (one colony per plate). From these, individual colonies were grown up in SOB plus 100µg/µl ampicillin, mini-prepped to test for the plasmid and then stored in 40% glycerol.

### 2.4.5 Phage Isolation

For each construct, 100ml TBG was dispensed into a sterile, plugged 500ml conical flask and warmed to 37°C. It was inoculated with 300µl DHIIS plasmid glycerol and then 1.3ml M13KO7 helper phage (~$1x10^{10}$pfu/ml). It was swirled after both additions and incubated at 37°C for 90 minutes, rotating at 300rpm.

100µl of 100µg/µl ampicillin and 140µl of 50µg/µl kanamycin were added before letting the incubation continue for another 14 hours. This was sufficient time to produce a maximum amount of ssDNA without any dsDNA contamination.

Once the incubation was complete, the culture was poured into a 100ml bottle, cooled on ice for 15 minutes and centrifuged at 7K rpm for 20 minutes in a JA14 Beckman rotor. The supernatant was then transferred to a fresh bottle and the centrifugation step repeated to pellet all remaining bacteria.

80ml of this supernatant was added to 20ml PEG/NaCl in a fresh, iced 250ml bottle. It was swirled to mix and incubated on ice for at least one hour. The cloudiness of the solution was a good indicator of the yield of phage particles and therefore ssDNA.

The phage particle precipitate was pelleted by centrifugation at 7K rpm for 30 minutes. The supernatant was removed and the sample centrifuged again for 5 minutes at 3K rpm and the supernatant removed. To remove the remaining supernatant the last centrifugation step was repeated, supernatant removed using a p200 pipetteman and the pellet resuspended in 1 to 2ml TE.

### 2.4.6 ssDNA Isolation

The QIAprep Spin M13 Kit is suitable for retrieving small amounts of ssDNA from phage supernatant. 3 columns, with 100µl phage supernatant per column, were sufficient to collect over 5µg of ssDNA. 2µg is sufficient for one monomer extension reaction.

After at least 2 minutes, 1µl MP buffer was added to 100µl phage supernatant to precipitate the phage particles from the solution. This was loaded on to the spin-column and centrifuged for 60 seconds at maximum. Addition of 700µl MLB buffer creates appropriate conditions for binding of the M13 ssDNA to the QIAprep silica-gel membrane and bacteriophage lysis begins. After centrifugation as above, 700µl

more of MLB buffer was added to complete the lysis and binding. After 60 seconds it was centrifuged as above.

Addition and centrifugation with 700µl PE wash buffer removed residual salt. The empty column was then centrifuged again to remove traces of ethanol which would perturb subsequent reactions. Lastly, the column was placed in a fresh 1.5ml tube. 100µl 10mM Tris-HCl (pH8.5) was added on to the column membrane to elute the ssDNA and after 10 minutes it was centrifuged as above to remove the solution.

To test for contaminating dsDNA, 2µl of each sample was diluted to 20µl with TE. 10µl of this was heated to 95°C for 5 minutes to denature any dsDNA. This was loaded on a 1% agarose/TAE gel alongside the remaining 10µl sample. A smear would appear in the heated sample lane if the sample were contaminated.

Each sample was quantified using DE81 chromatography paper which binds DNA. 1µl of each sample was spotted onto the paper next to a series of standards. This was immersed in 1xTBE containing 2.5ng/ml ethidium bromide for 10 minutes. Excess ethidium bromide was removed by five 10 minute washes in 1xTBE. The concentration was estimated visually under 304nm UV light. The concentration was also measured by electrophoresis through a 1.5% agarose/TAE gel using a Mass Ladder standard (Roche Diagnostics).

### 2.4.7 Core Particle Isolation Buffers

**TEP 2M** (300ml): 6ml 50xTE, 120ml 5M NaCl

**TEP 0.4M** (500ml): 10ml 50xTE, 40ml 5M NaCl

**TEP 16mM** (1200ml): 24ml 50xTE, 3.8ml 5M NaCl

### 2.4.8 Nucleosome Reconstitution by Salt Gradient Dialysis

Just before using the TEP buffers, 40µl of 250mM PMSF was added per 100ml buffer to inhibit proteases (Moss *et al.*, 1978). The dialysis sheet (6000-8000Da exclusion limit) was prepared in the same manner by soaking in 300ml water with 120µl 250mM PMSF.

Only one orientation of each construct was used for the reconstitution reaction. 30µg of each construct was linearised with ***Sca*I** restriction enzyme, which cuts the plasmid only once, far from the insertion site. After phenol: chloroform

extraction, ethanol precipitation and a 70% ethanol wash, it was raised to a concentration of 1µg/µl. 1µl was electrophoresed through an agarose gel to check for complete linearisation.

Each 25µg reaction contains 1xTE, 2M NaCl, 1µl PMSF, 16.8µl of 0.742µg/µl chicken core histones (a kind gift from the Allan laboratory, Dept. of Biochemistry, University of Edinburgh) and 25µg linearised plasmid DNA. The core histone: DNA ratio is 1:2 (w/w). Each reaction was incubated at room temperature for 10 to 20 minutes prior to loading onto the dialysis sheet in the Gibco BRL micro-dialysis system.

Once loaded, it was dialysed against a continual flow of the TEP 2M solution which was gradually diluted over a period of 5 hours with the TEP 0.4M solution. This process assembles nucleosomes on to the DNA. It was left overnight dialysing against the TEP 16mM solution which induces nucleosomes to adopt energetically favourable positions on the DNA.

### 2.4.9 Micrococcal Nuclease Digestion of Reconstitutes

Each reconstitution reaction was kept on ice and made up to 310µl with TEP 16mM solution. The exonuclease activity of micrococcal nuclease is reduced at 0°C (Hörz *et al.*, 1981). This reaction was initiated with 1.15mM $CaCl_2$ and 1.5 Weiss Units of MNase. This proceeded for 35 minutes on ice. A brief period of 2 minutes at 37°C stimulated MNase exonuclease activity which digested the DNA to the end of the core particle boundaries. The reaction was quenched with 5mM EDTA.

After phenol:chloroform extraction, ethanol precipitation and a 70% ethanol wash, the resuspended core particle DNA was electrophoresed through a 2% agarose/TAE gel and the 146bp monomer fragments isolated. Concentration was estimated using the Mass Ladder (Roche Diagnostics) on an agarose gel.

### 2.4.10 Core Particle Labelling with $^{32}P$

Polynucleotide kinase catalyses the transfer of a phosphate group from the γ position of ATP to the 5' hydroxyl terminus of polynucleotides (Midgley *et al.*, 1985; Current Protocols in Molecular Biology (3.10.2)) . By using $\gamma$-$^{32}P$ ATP, the core particle DNA can be radioactively end-labelled.

30-50ng of core particle (monomer) DNA is required per monomer extension reaction. In order to ensure reasonable recovery, at least 200ng was labelled in each case. A typical 15µl labelling reaction contained:

9µl core particle DNA (≥200ng)
1.5µl 10x buffer (500mM Tris-HCl (pH7.6), 100mM $MgCl_2$,
50mM DTT, 1mM spermidine & 1mM EDTA)
3.5µl $\gamma$-$^{32}$P ATP
1µl T4 Polynucleotide kinase

The reaction proceeded for 45-60 minutes at 37°C when it was terminated with 30mM EDTA. The protein and unincorporated nucleotides were removed via centrifugation through a Sephadex G50 column. The eluted core particle DNA was denatured with 0.2M fresh NaOH for 10 minutes.

Addition of 2M ammonium acetate (pH7.5) neutralised the solution and aided precipitation with 2.5x (v/v) ethanol overnight at -20°C. The precipitate was washed twice with 70% ethanol and resuspended in TE to 10ng/µl.

### 2.4.11 Monomer Extension Reaction

This technique was developed in the Allan laboratory in order to map the precise translational positions adopted by core histone octamers reconstituted onto long DNA sequences (Yenidunya *et al.*, 1994). This technique maps the boundaries of core particle DNA fragments protected by a histone octamer from MNase digestion and has been applied to the entire region of chicken β-globin gene (Davey *et al.*, 1995).

In the initial step, annealing is undertaken at a ≥2.5 fold excess of single stranded template over monomer DNA. This typically means 30-50ng monomer DNA and 1.5-2µg template ssDNA, depending on the size of the insert. As pBlueBLG +/- contain a large insert, 2µg ssDNA was required per reaction. All other constructs contain inserts of 450-750bp which require about 1.5µg ssDNA.

Each extension reaction requires a control in which water is substituted for the restriction enzyme. This allows the background to be eliminated from subsequent

data analysis. Therefore, at least two reactions are annealed each time and split later for the extension reaction. More than two reactions are required if more than one restriction enzyme is used during the extension step.

**Annealing Reaction**

For a double reaction, the following reagents were added to a PCR tube: 15µl ssDNA (1.5-2µg construct dependent), 10µl of 10ng/µl labelled, denatured monomer DNA and 25µl premix (100mM NaCl, 2mM DTT, 20mM Tris-HCl (pH7.5), 20mM $MgCl_2$). The reaction was mixed well and then denatured for 3 minutes at 95°C. The monomers were annealed at 80°C, dropping to 55°C, over a period of about 45 minutes. The reaction was placed on ice after reaching 55°C.

**Extension Reaction**

Two fresh, ice-cold PCR tubes were prepared containing 5 units Klenow DNA polymerase and either 20 units of the appropriate restriction enzyme or water as the control. 50µl ice-cold premix (0.1mg/ml BSA & 10µM dNTPs) was added to the annealing reaction and mixed well. Then, two 50µl aliquots of the reaction mix were transferred into the two enzyme-containing PCR tubes. After a thorough mix, the tubes were incubated at 37°C for 60 minutes in order for the extension reaction to proceed.

After phenol:chloroform extraction, ethanol precipitation and a 70% ethanol wash, the samples were resuspended in 8µl denaturing gel loading buffer (90% deionised formamide, 1xTBE, 9mM EDTA, 0.04% bromophenol blue & 0.04% xylene cyanol). Before loading, the counts were normalised with respect to each other by dilution in gel loading buffer and denatured at 95°C for 3 minutes.

### 2.4.12 Generation of Markers

**Lambda markers**

A 20µg sample of phage lambda DNA was digested with ***Dde*I** and another with ***Hinf*I**. They were phenol:chloroform extracted, ethanol precipitated, washed in 70% ethanol and resuspended in TE to 100ng/µl. They were end-labelled with $\gamma$-$^{32}$P ATP using polynucleotide kinase (as core particle labelling above),

phenol:chloroform extracted, ethanol precipitated, washed in 70% ethanol again and resuspended in 50µl denaturing gel loading buffer.

Before loading on the polyacrylamide gel, both markers were diluted with denaturing gel loading buffer to give a similar intensity with respect to the monomer extension samples.

**100bp Ladder**

The 100bp ladder (Fermentas) was labelled in the same way as the lambda markers. It was also resuspended in 50µl denaturing gel loading buffer.

Before loading on the polyacrylamide gel, the 100bp ladder was diluted with denaturing gel loading buffer to about 20% cpm with respect to the lambda markers to give a similar intensity.

**M13 Sequence markers**

2µg of M13 ssDNA was added to 1pmol universal forward primer (a 17mer adjacent to the MCR), 4µl of 10x reaction buffer (200mM Tris-HCl pH7.5, 100mM $MgCl_2$ , 250mM NaCl) and made up to 20µl with water. The annealling reaction proceeded at 65°C for 3 minutes, decreasing to 30°C over a period of 30 minutes. It was removed to room temperature and 2µl of the labelling mix was added (7.5µM dATP, 7.5µM dCTP, 7.5µM dGTP, 7.5µM dTTP) along with 2µl 100mM DTT, 3µl $\alpha$-$^{35}$S dATP and 4µl 10mM Tris HCl pH7.5. After mixing, 1µl (10units) T7 DNA polymerase was added and it was mixed once more.

It was incubated at room temperature for 10 minutes and then split in to two 15.5µl aliquots. C-termination mix was added to one (80µM dATP, 80µM dCTP, 80µM dGTP, 80µM dTTP, 8µM ddCTP and 50mM NaCl) and T-termination mix to the one (80µM dATP, 80µM dCTP, 80µM dGTP, 80µM dTTP, 8µM ddTTP and 50mM NaCl).

Five minutes at 37°C and then fifteen minutes at 50°C polymerised the DNA. Finally, they were phenol:chloroform extracted, ethanol precipitated, washed in 70% ethanol and resuspended in 56µl denaturing gel loading buffer. 2µl was loaded per lane.

### 2.4.13 Polyacrylamide Gel Electrophoresis

All samples were electrophoresed in a denaturing 6% polyacrylamide gel at 50W for about 5 hours (Yenidunya *et al.*, 1994). The gel was disassembled, fixed (10% acetic acid & 12% methanol solution) and washed in distilled water. After it was transferred on to 3MM Whatman paper, it was dried for 1 hour under vacuum in a Bio-Rad 583 gel drier. The gel was laid down on Kodak photographic film for typically 4 hours before developing. It was exposed for several hours on a Phosphorimaging plate for quantitative analysis.

## 2.5 Transgenic Mouse Analysis

In order to analyse the nucleosomal array over the promoter region in transgenic mouse line 14, a lactating female mouse had to be generated. This required testing the initial progeny for the transgene before they were of sufficient age to be bred.

A (CBA x C57Bl/6)F1 female mouse was mated with a (CBA x C57Bl/6) BLG14 transgenic male mouse. It contained two copies of a 16.2kb ***Sal*I** fragment encompassing the ovine BLG gene from -4.2kb to +12kb with respect to the transcriptional start site (Simons *et al.*, 1987; Dobie *et al.*, 1996). The resulting progeny were tail-tipped at 4-5 weeks old and this tissue was analysed by PCR for the BLG transgene.

The BLG primers amplify a 248bp segment of the 5' end of the BLG gene:

5' primer: 5' GCTTCTGGGGTCTACCAGGAA 3' (-222bp wrt start site)
3' primer: 5' TCGTGCTTCTGAGCTCTGCAG 3' (+26bp wrt start site)

HPRT primers serve as an internal positive control by amplifying a 332bp segment of that gene:

5' primer: 5' GAGTTCCGGAACTGCCTTTGGTG 3'
3' primer: 5' CTGTGCCACCGGGCGCATGG 3'

In both cases, equimolar amounts of both primers are added together to make a 20μM working solution.

The ~1cm of tail tissue was digested in 700µl tail-tip digestion buffer (100mM Tris-HCl (pH8.5), 5mM EDTA, 200mM NaCl, 0.2% SDS and 100µg/ml proteinase K) (Laird *et al.*, 1991). This was incubated overnight at 55°C with occasional mixing. An equal volume of isopropanol was added to the lysate and the sample mixed until precipitation was complete. The DNA was recovered by lifting the aggregated precipitate from the solution with a yellow tip. Excess liquid was dabbed off and the DNA resuspended in 150µl TE with agitation over a period of 30 minutes.

Meanwhile, the PCR mix was prepared (25µl / reaction):

2.5µl 10x PCR buffer (100mM Tris-HCl (pH8.3), 15mM $MgCl_2$, 500mM KCl)
0.25µl 20µM BLG primer set
0.25µl 20µM HPRT primer set
4µl 1.25mM dNTP mix
2µl DMSO
0.08µl Taq polymerase
- to 25µl with $H_2O$

1µl of the mouse DNA was added to 24µl of PCR mix in a 200µl PCR tube. After an initial denaturation step at 95°C for 2 minutes, it underwent 35 cycles of:

| | |
|---|---|
| 40secs | 95°C |
| 1 min | 65°C |
| 1 min | 72°C |

The final extension step was carried out for an additional 4 minutes. 10x gel loading buffer was added to the sample and the fragments were separated on a 2% agarose/ TAE gel.

A female mouse which was positive for the transgene was mated and mammary and liver tissues were isolated from it at mid-lactation.

# Chapter 3

# Initial Analysis of the Chromatin Environment of the Ovine BLG Gene

## 3.1 Introduction

Previous analysis of the BLG gene using DNase I has elucidated a number of regions particularly sensitive to digestion. There are three of these hypersensitive sites over the active promoter region (Figure 15): HS I (at -1.8kb), HS II (at -800bp) and HS III (at -300 to +100bp) (Whitelaw et al., 1992). It appears that the first two sites are not required for tissue-specific, temporally regulated expression in transgenes, but the last site is essential. This latter site covers a region which contains the majority of transcription factor binding sites, e.g. Stat5 and NF1, and the TATA box. DNase I hypersensitivity is thought to reflect a more open chromatin environment in which transcription factors have access to an otherwise repressive, compact domain.

In addition to these primary sites, HS IV (at +800bp) may be present in the first intron and HS V (at +1500bp) in the second intron. However, these are only present in virgin mammary chromatin until days 100-110 after conception. These sites may prepare the locus for activation (Struhl, 1999). It is unknown what factors bind to these hypersensitive sites. All five of these DNase I hypersensitive sites are mammary specific, none being detected in the liver.

Although DNase I hypersensitivity is a useful marker for a pre-active (Elgin, 1984) or active gene domains (Elgin, 1981; Elgin 1988), it tells us very little about the local nucleosomal chromatin structure which could have a direct relevance to transcription. The most commonly used probe to investigate nucleosomal DNA is micrococcal nuclease (MNase)(Thoma, 1992). It was discovered in 1957 by Cunningham and colleagues (Privat de Garilhe *et al.*, 1957) and was shown to: be activated by calcium, be very heat stable, display maximal activity at pH8.6 and have a molecular weight of 16.8kD (Cunningham, 1958; Alexander et al., 1961; Taniuchi *et al.*,

1967). The enzyme also possesses ribonuclease activity and both activities are inactivated with the addition of EDTA.

MNase digests DNA initially in an endonucleic manner (Horz & Altenburger 1981). To initiate cleavage, a groove in the enzyme binds three phosphate groups at the 3' side of the hydrolytic site on a single DNA strand (Cuatrecasas *et al.*, 1967). After cleavage, the exonucleic activity of the enzyme proceeds to chew away the DNA in a 5' to 3' direction. In chromatin, this process is blocked by nucleosomes (Noll *et al.*, 1977).

MNase preferentially digests chromatin in the linker region between nucleosomes which is more accessible (Thomas *et al.*, 1988; McPherson *et al.*, 1993; Truss *et al.*, 1995). Partial digestion reveals a 'ladder' of DNA fragments in multiples of approximately 200bp (Figure 23). After sufficient time, only 200bp monomer fragments are left, containing one nucleosome repeat length. These are eventually digested, when the exonucleic activity trims the DNA up to and then in to the core particle region. Lowering the reaction temperature decreases this exonucleic activity. However, this also increases the sequence preference of the enzyme. By cutting DNA at specific sites, often within core particles, it becomes less reliable as a probe for nucleosome positioning. Therefore, a balance must be reached between limited sequence preference, which generates unwanted fragments, and some exonucleic activity, which makes sizing the required fragments more difficult.

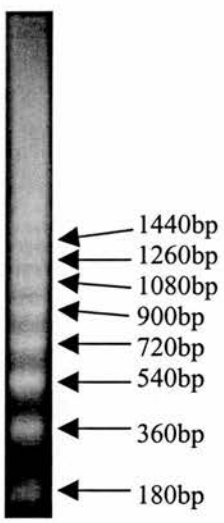



**Figure 23**

MNase digest of sheep chromatin. Ethidium bromide staining of the resulting DNA fragments separated by electrophoresis through a 1.5% agarose gel. Approximate fragment sizes are indicated in base pairs.

## 3.2 Results

In the first series of experiments I assessed the presence and repeat length of nucleosomes throughout the BLG gene in both its active state in mammary nuclei (Figures 28 - 31) and its inactive state in liver nuclei (Figures 32 - 35). Subsequently, I examined specific regions of the gene for positioned nucleosomes, again in the active state (Figures 41 & 42) and in the inactive state (Figures 43 & 44). In light of these results, I examine the limitations of MNase as a probe for nucleosome positioning and investigate other possibilities.

### 3.2.1 Digestion of Sheep Liver and Mammary Nuclei with MNase

Nuclei were prepared as in the materials and methods and were digested with MNase at a ratio of $1.1x10^5$ nuclei: 1 Unit of enzyme. Aliquots were taken at 0, 1, 2, 5, 10 and 15 minutes and the reaction was terminated with EDTA. The resulting DNA fragments were separated on a 1.5% agarose gel by electrophoresis, transferred to a nylon membrane by Southern blotting and end-labelled with a probe specific to the BLG gene region in question (Figure 24). In the initial series of experiments, nucleosome footprints span both directions along the DNA strand since an additional restriction enzyme was not used to delimit them to one direction. Therefore, where two probes exist, abutting either side of a restriction enzyme site, only one was used in this assay.

### 3.2.2 Controls

Three controls are required for each Southern blot probing. The negative control was mouse genomic DNA digested with a restriction enzyme whose site abuts the probe. This tests for non-specific binding of the probe. DNA was extracted from non-transgenic (CBA x C57Bl/6)F1 mouse stock lines. The positive control was sheep genomic DNA, again digested with the same restriction enzyme. Only a genomic fragment of the correct size should be seen when the probe is specific to the target sequence. The last control tests for the sequence specificity of MNase. Protein-free, 'naked' genomic DNA was partially digested with MNase to reveal a smear of differently sized fragments on an agarose gel. Any preferentially cleaved sites will appear as bands after probing the

Southern blot membrane. These bands may limit the interpretation of equivalent bands in the test lanes.

Naked sheep DNA was incubated with MNase and the reaction was stopped at various time points. The resulting fragments were Southern blotted and indirectly end-labelled with the 8900UP probe to examine which gave the most representative fragment sizes (Figure 25). The 4-minute time-point was chosen.

### 3.2.3 Probe Generation

Nine probes were designed to span the gene for contiguous nucleosome mapping (Figure 26). In this initial assay, testing for the presence and repeat length of nucleosomes, not all of these probes were used (see above). They were named according to their position and mapping direction as follows:

*Probe -2700UD*

Derived from plasmid pBJ1 which contains 4.2kb of ovine BLG 5' sequence in the pPolyIII-I vector. Digestion with ***Pst*I** generates 10 fragments; the 1420bp fragment was isolated and further digested with ***Eco*RI**. Out of three fragments generated, the **286bp** fragment was isolated, which spans from -3065bp to -2270bp with respect to the transcriptional start site.

*Probe -1700UP*

Derived from plasmid pBJ21 which contains 2kb of ovine BLG 5' sequence linked to the chloramphenicol acetyltransferase (CAT) coding region in the pPolyIII-I vector. Digestion with ***Hind*III** and ***Bam*HI** generates three fragments from which the **440bp** fragment was isolated. It spans from -1699bp to -2139bp.

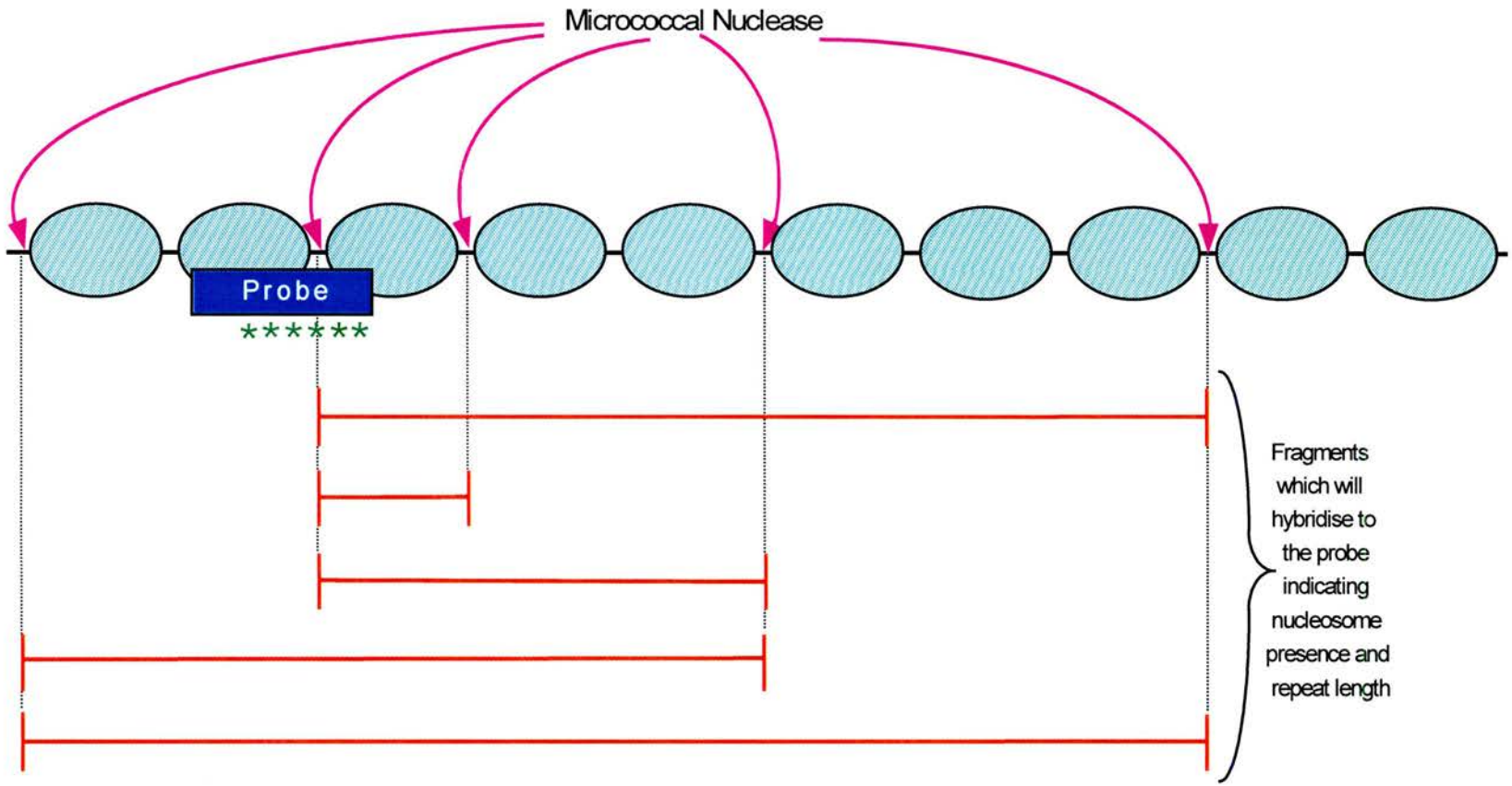



**Figure 24** A Schematic of the Technique to Test for the Presence of Nucleosomes

MNase partially digests chromatin (pink arrows), leaving fragments of DNA in multiples of 180bp (red lines). The radioactive probe (green & blue) hybridises to fragments spanning that region of BLG. Any distinct bands revealed after exposure to film indicate a DNA footprint. Footprints that are multiples of 180bp are most likely to represent nucleosomes associated with DNA (light blue).

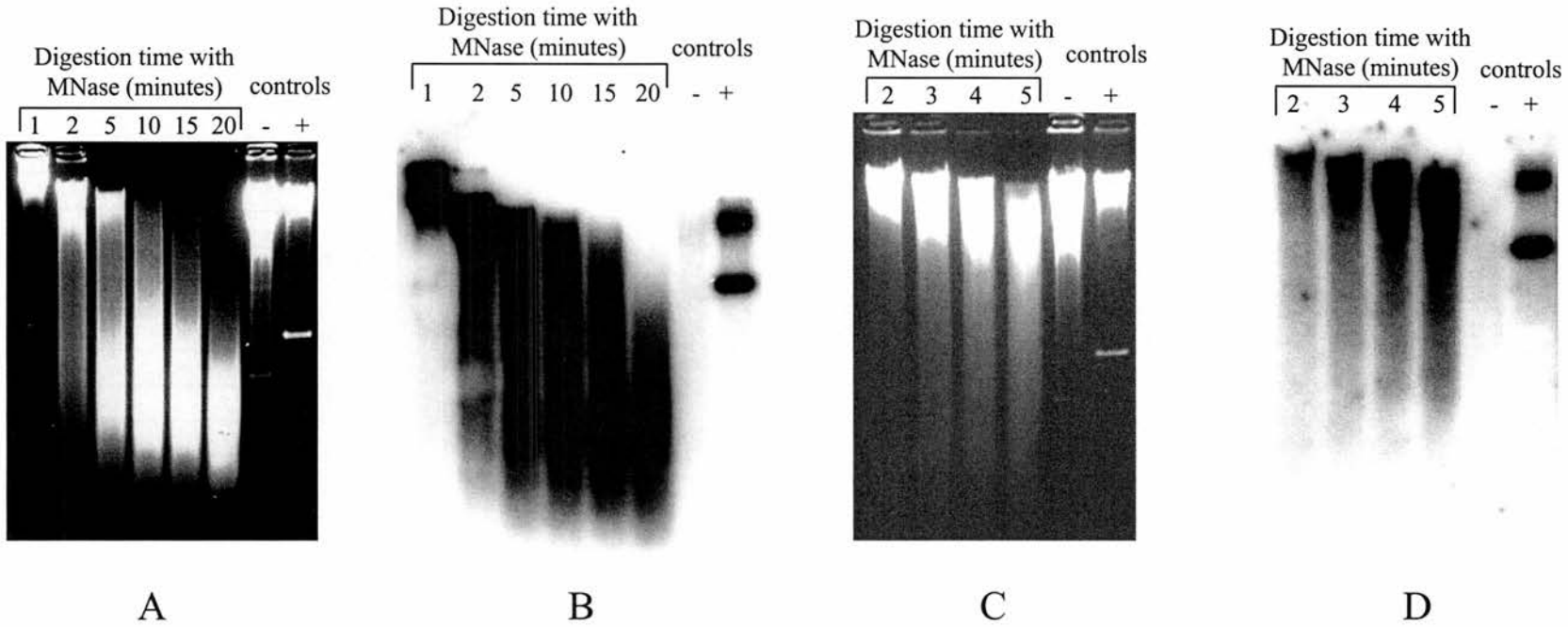



**Figure 25** The MNase Naked DNA Positive Control.

Sheep naked genomic DNA was digested with MNase over a period of 20 minutes. Aliquots were taken at several time points during the reaction and the fragments generated in each were separated by agarose electrophoresis (A). This gel was Southern blotted and hybridised with the 4700UP probe (Figure 26)(B). (C) and (D) are the same as (A) and (B) respectively, except that a narrower window of digestion times was used to refine the search. The positive control was sheep genomic DNA digested with ***Bam*HI** and the negative control was mouse genomic DNA, also digested with ***Bam*HI**. The probe abuts a ***Bam*HI** site, but the test lanes are only digested with MNase. Unfortunately, the probe was contaminated with a fragment adjacent to the other side of the ***Bam*HI** site, evident by the presence of two genomic bands in the positive control lane. However, this is inconsequential, as the aim of this experiment was to determine which time point produces the most representative smear of fragments over the gene region.

*Probe -1700DOWN*

Derived from plasmid pBJ58 which contains 408bp of ovine BLG 5' sequence linked to the chloramphenicol acetyltransferase coding region in the pPolyIII-I vector. Digestion with ***Taq*I** and ***Bam*HI** generates three fragments from which the **358bp** fragment was isolated. It spans from -1699bp to -1341bp.

*Probe -900DOWN*

Derived from plasmid pBJ2 which contains 3.1kb of ovine BLG 5' sequence in the ptgPolyI vector. Digestion with ***Taq*I** and ***Rsa*I** generates fourteen fragments from which the **367bp** fragment was isolated. It spans the region -864bp to -497bp.

*Probe +1700UP*

Derived from plasmid pBJ14 which contains 1.1kb of ovine BLG exon 3 and flanking introns in the pUC18 vector. Digestion with ***Eco*RI** and ***Pvu*II** generates five fragments from which the **411bp** fragment was isolated. It spans from +1252bp to +1663bp.

*Probe +2700UP*

Derived from plasmid pBLGΔSpΔ3' which contains 4.4kb of ovine BLG sequence from the promoter to the last intron in the pUC19 vector. Digestion with ***Bam*HI** and ***Hind*III** generates three fragments from which the **445bp** fragment was isolated. It spans from +2742bp to +2297bp.

*Probe +2700DOWN*

Derived from plasmid pBLGΔSpΔ3' and digested as above. The 2.3kb ***Bam*HI** fragment was isolated and further digested with ***Ban*I** to produce the **345bp** probe. It spans from +2742bp to +3087bp.

*Probe +4700UP*

Derived from plasmid pBJ34 which contains 2.8kb of ovine BLG including exon 1 fused to exon 5 and downstream sequence of 2.75kb, in the ptgPolyIII-I vector. Digestion with

***Bam*HI** and ***Nco*I** generates six fragments from which the **338bp** fragment was isolated. It spans the region +4755bp to +4417bp.

*Probe +4700DOWN*

Derived from plasmid pBJ34 as described above. Digestion with ***Bam*HI** and ***Rsa*I** generates seven fragments from which the **272bp** fragment was isolated. It spans from +4755bp to +5025bp.

### 3.2.4 Probe Testing

All probes were Southern blot-tested with the positive and negative controls prior to use to check for DNA contamination and ability to radioactively label. Only the -2700UD probe was not used in subsequent experiments. In the positive / negative control test it seemed to function perfectly. However, it appeared that the positive control had not been digested to completion with ***Pst*I** as there was a high degree of background. Therefore, sheep liver nuclei were partially digested with MNase and completely digested with ***Pst*I** and probed with -2700UD (Figure 27) to test for nucleosome positioning. The resulting image revealed that the probe was non-specific. As the negative control lane was blank and the positive control lane was identical to the test lanes, it appears that the probe was annealling to repetitive sequence DNA. For example, only a single genomic band is evident in the positive control lane in Figure 29 which shows that the probe is gene-specific. The high background and repetitive banding pattern seen in the positive control and test lanes in Figure 27 suggests that not only is it non-specific, but that it is annealing to a large range of repeating sequences.

There is weak homology to reverse transcriptase from -3170bp to -3660bp indicating the presence of an ancient long interspersed nuclear element. As the probe is located very close to this region, it is probably hybridising to the terminal end of the element and will therefore anneal throughout the genome at similar sites.

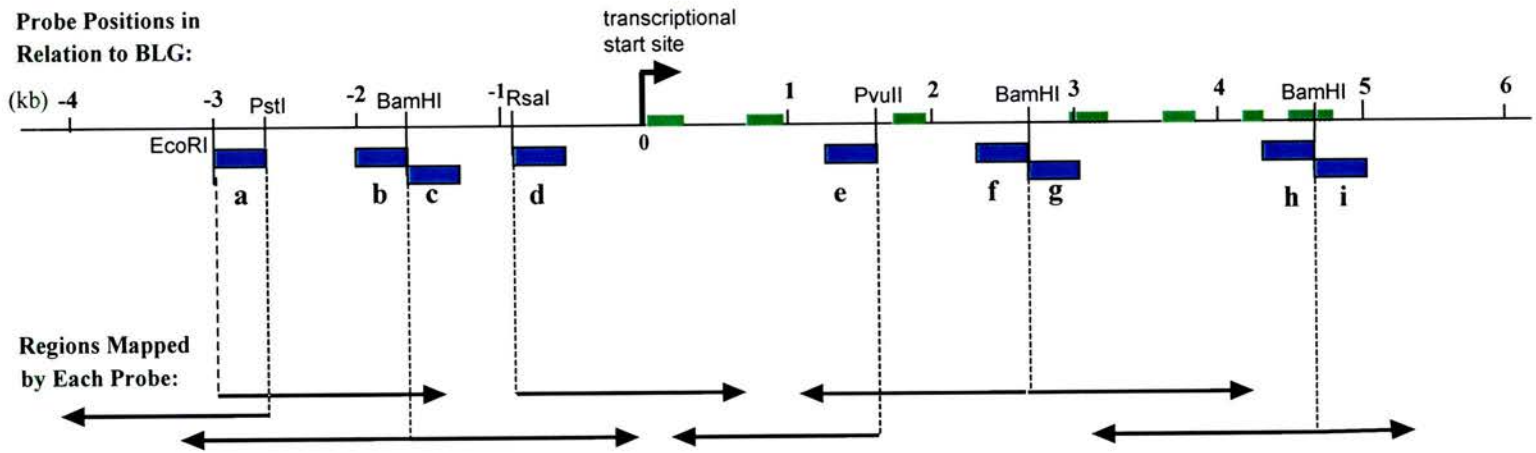



**Figure 26** Mapping Strategy over the Sheep BLG Gene

Probes (blue) are named relative to their location and mapping direction as follows (from left to right): -2700UD (a), -1700UP (b), -1700DOWN (c), -900DOWN (d), +1700UP (e), +2700UP (f), +2700DOWN (g), +4700UP (h) and +4700DOWN (i). Various restriction sites abutting the probes are indicated. The arrows below show the minimal area mapped by each probe in the indirect end-labelling assays in chapter 4. Exons are green.

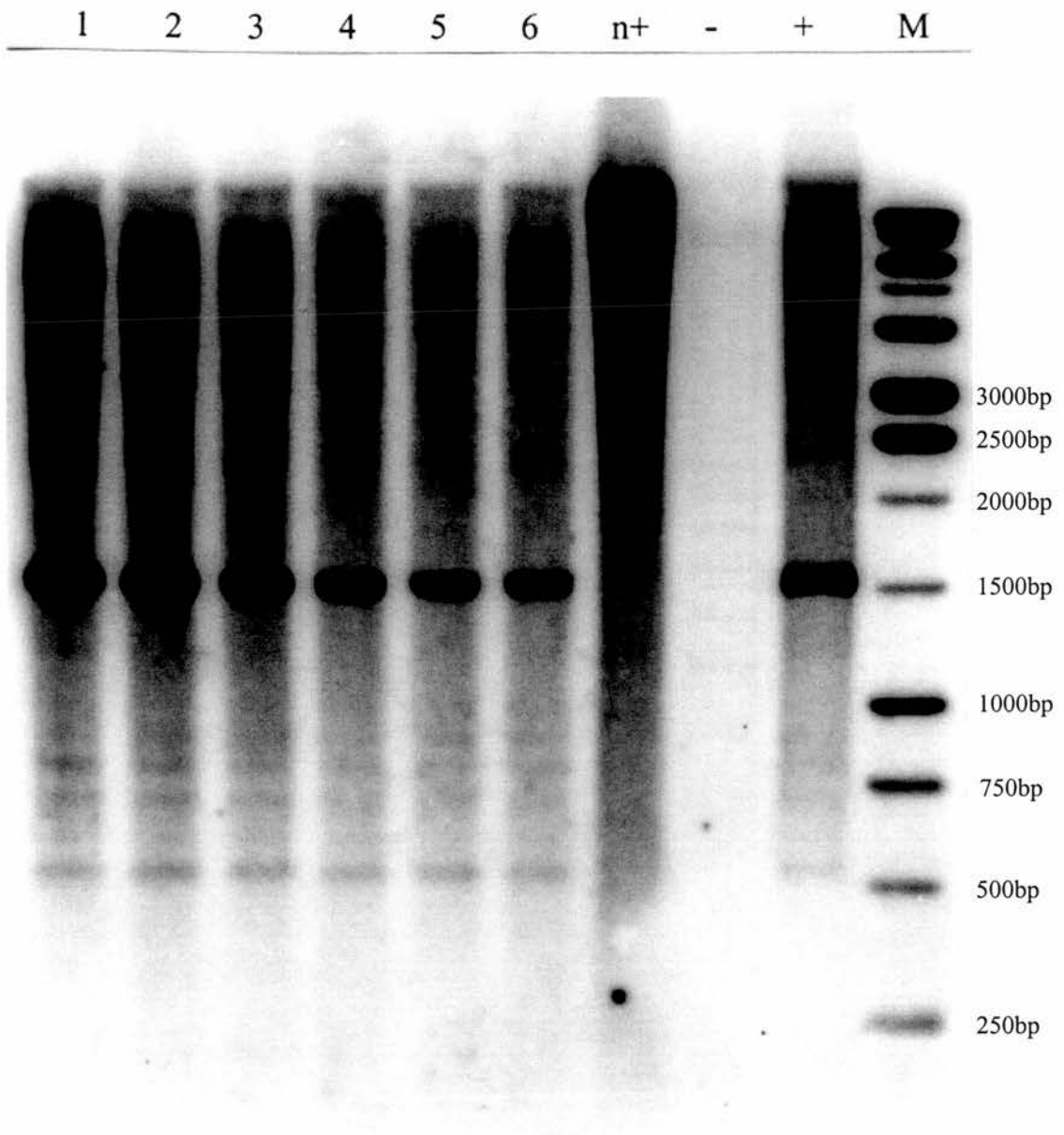



**Figure 27** Nucleosomes Over the -2.7kb Region of Ovine BLG in Liver

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Pst*I**. The naked DNA control (n+) was digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was -2700UD (Figure 26)

### 3.2.5 Micrococcal Nuclease Digestion of Ovine Mammary and Liver Chromatin

In order to examine the gene in its active and inactive states, both mammary and liver nuclei were analysed. The BLG gene locus was tested initially for the presence and repeat length of nucleosomes by partial digestion with MNase, without an additional restriction enzyme. Titrations consisting of six time points were performed with mammary nuclei which were hybridised to the following probes:

*-1700DOWN*.....(Figure 28)
*+1700UP*.....(Figure 29)
*+2700UP*.....(Figure 30)
*+4700DOWN*.....(Figure 31)

These included positive, negative and naked DNA controls and the band sizes were calculated in relation to the marker lane.

Liver chromatin was digested and electrophoresed in the same way and contained equivalent controls. The resulting Southern blots were labelled with the following probes:

*-1700DOWN*.....(Figure 32)
*-900DOWN*.....(Figure 33)
*+2700DOWN*.....(Figure 34)
*+4700DOWN*.....(Figure 35)

The size of bands generated from the partial digestion with MNase were calculated using either time point 2 (2 minutes digestion) or 3 (5 minutes digestion). These bands were related to those in the marker lane to calculate their size in base pairs. The results are tabulated for mammary chromatin in Table 2 and liver chromatin in Table 3. Photomicrographs of the ethidium bromide-stained agarose gels are displayed in Figure 36, prior to Southern blotting.

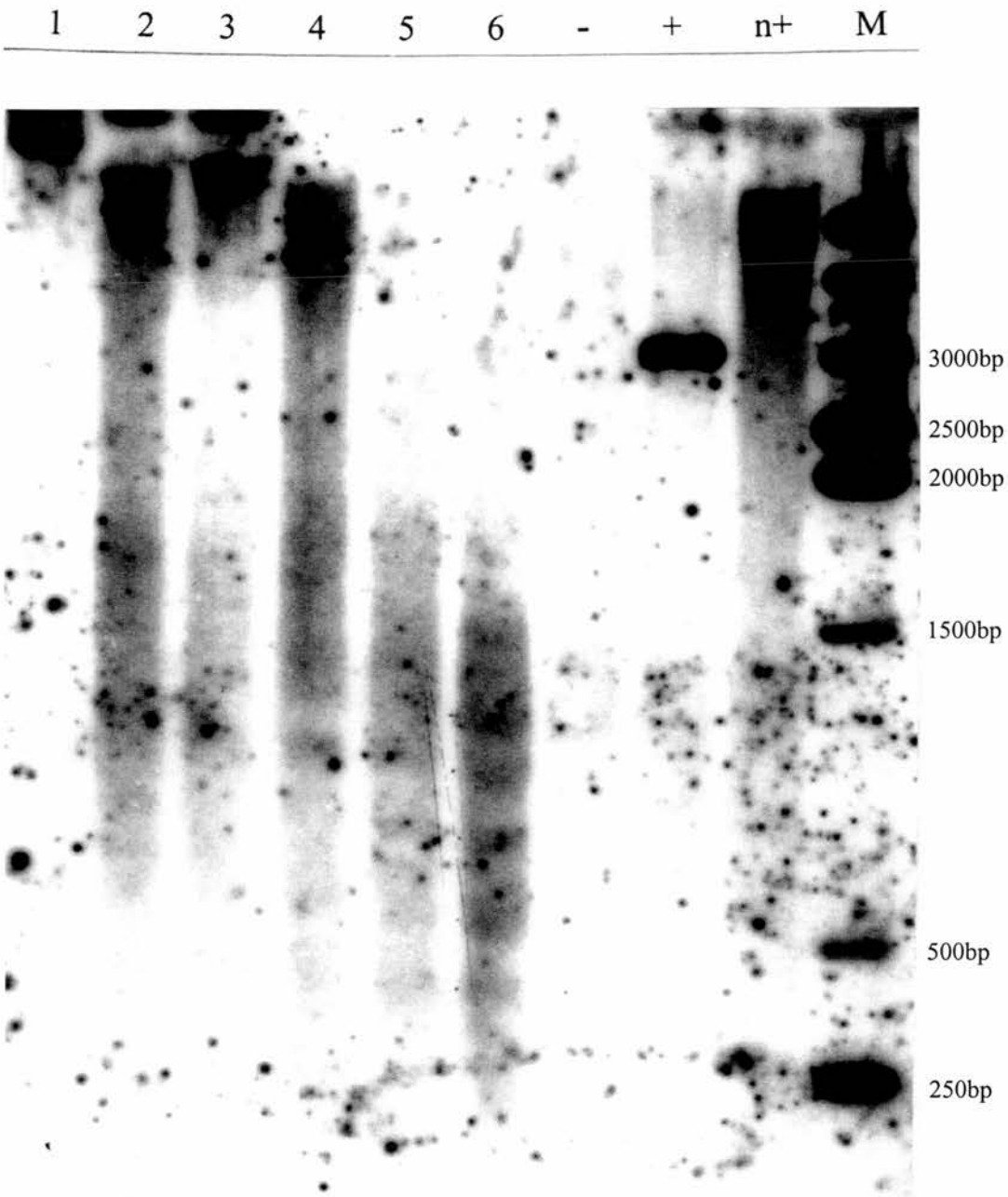



**Figure 28** Nucleosomes Over the -1.7kb Region of Ovine BLG in Mammary

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was -1700DOWN (Figure 26).

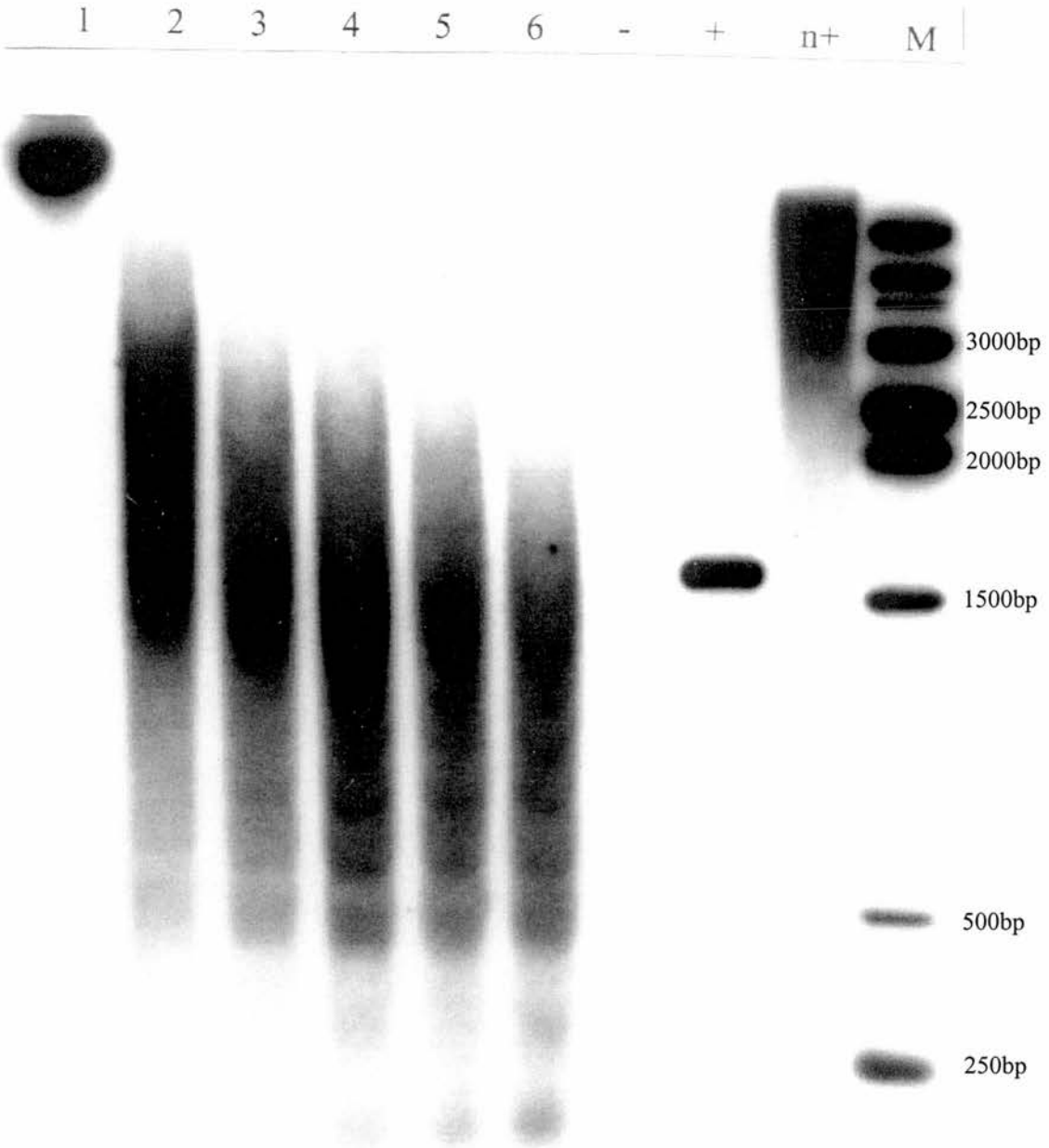



**Figure 29** Nucleosomes Over the +1.7kb Region of Ovine BLG in Mammary

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was +1700UP (Figure 26).

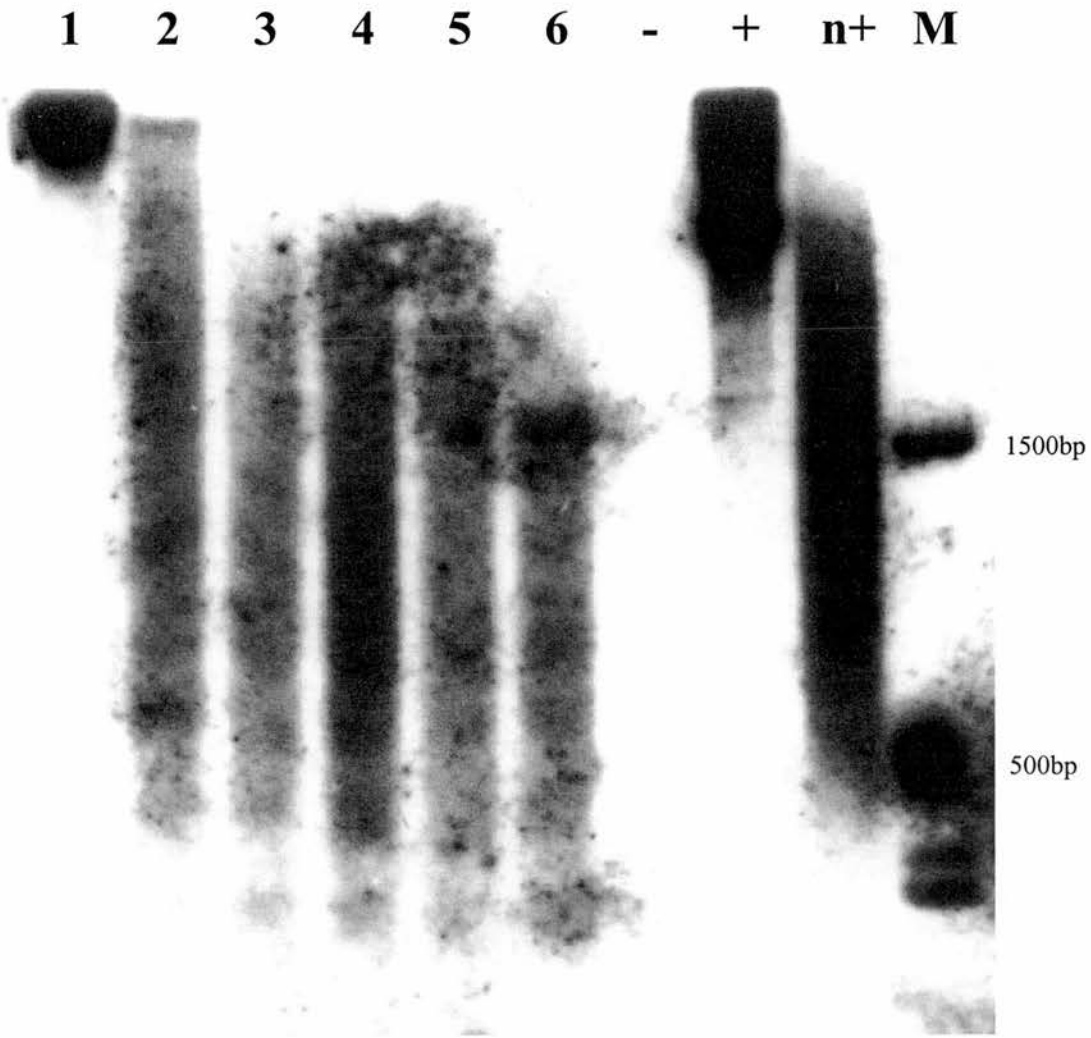



**Figure 30** Nucleosomes Over the +2.7kb Region of Ovine BLG in Mammary

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was +2700UP (Figure 26).

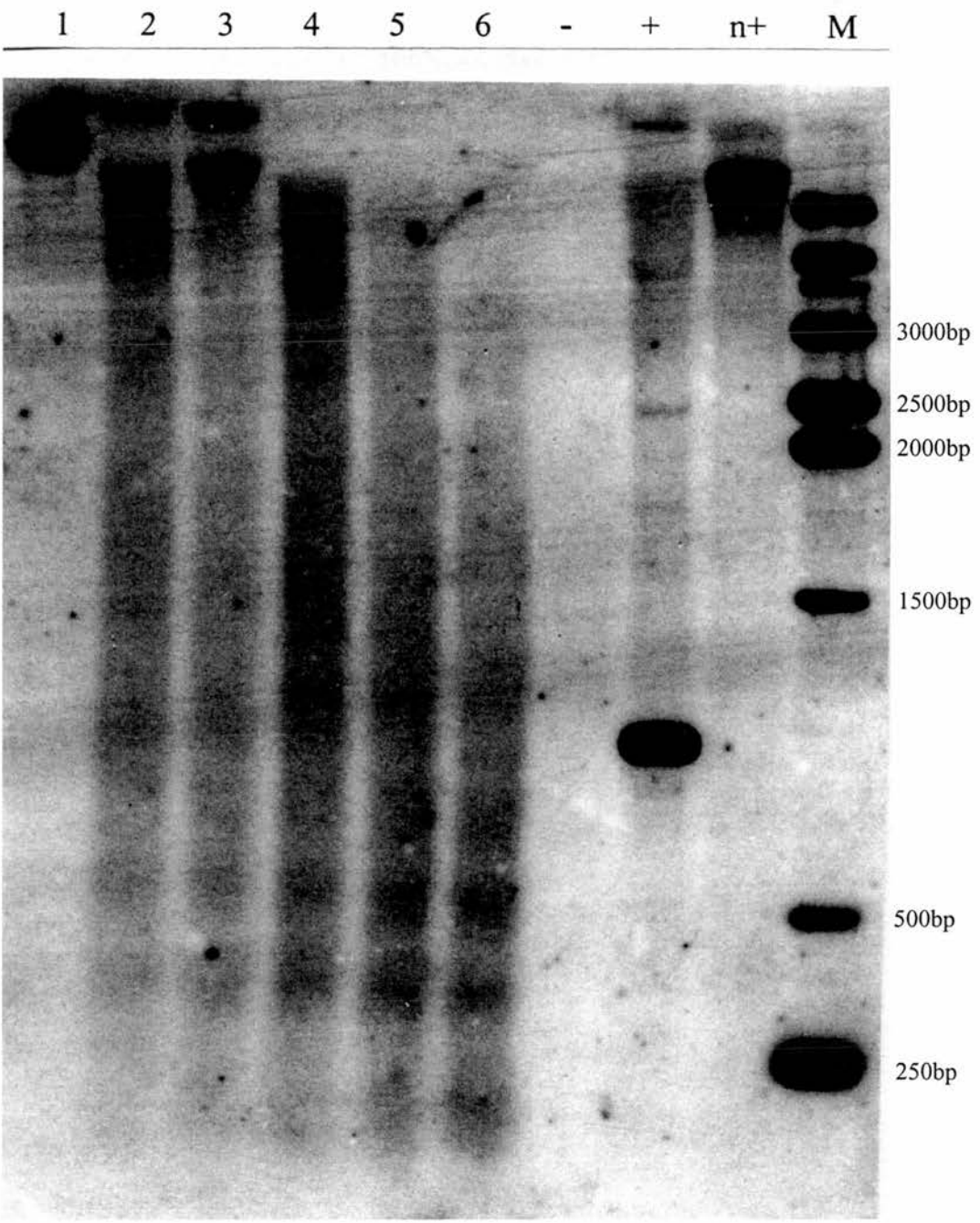



**Figure 31** Nucleosomes Over the +4.7kb Region of Ovine BLG in Mammary

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was +4700DOWN (Figure 26).

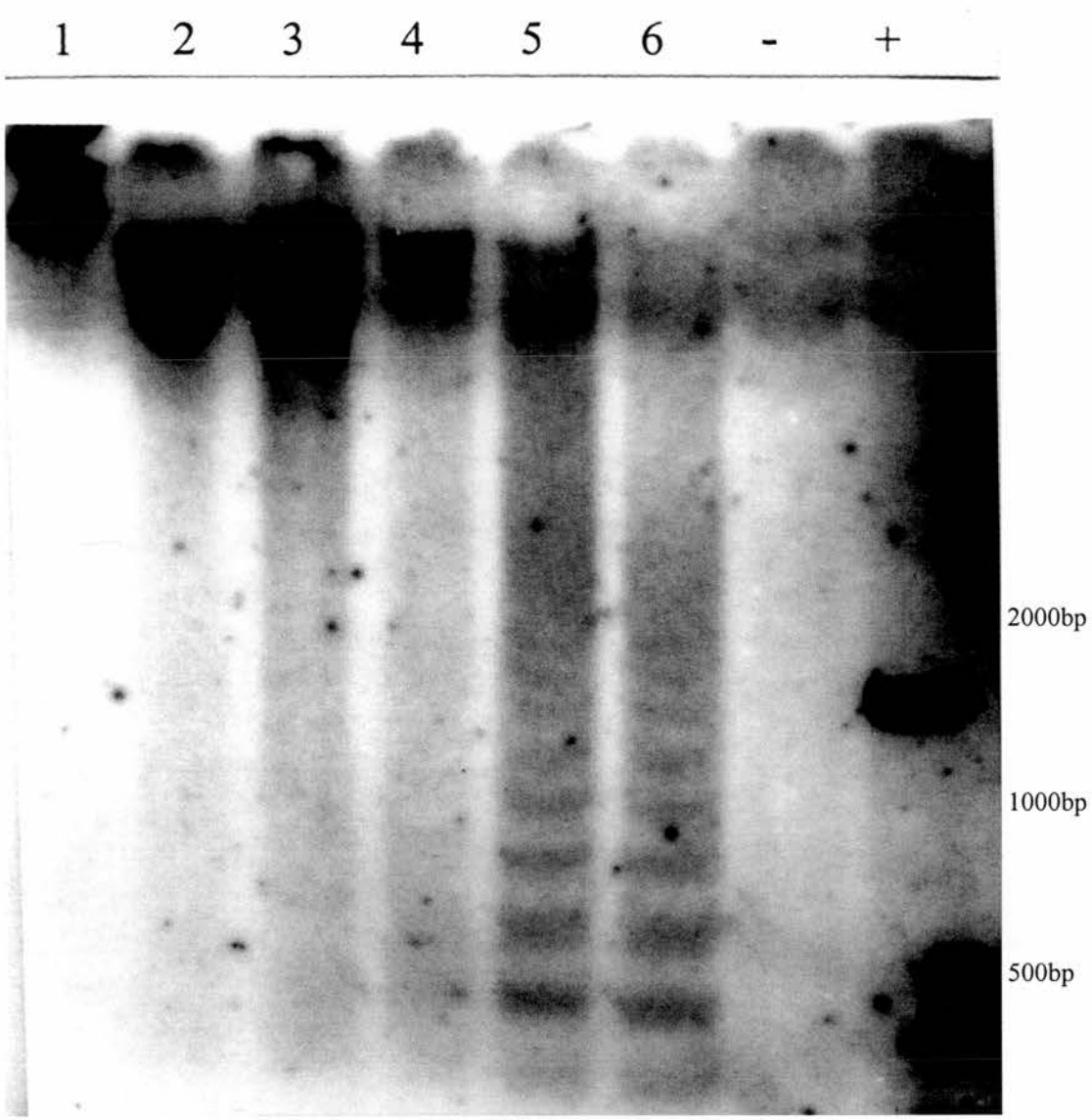



**Figure 32** Nucleosomes Over the -1.7kb Region of Ovine BLG in Liver

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Unfortunately, it was obscured by the marker lane. Sizes are shown in relation to a weaker exposure of the Promega 1kb marker lane. The probe used was -1700DOWN (Figure 26).

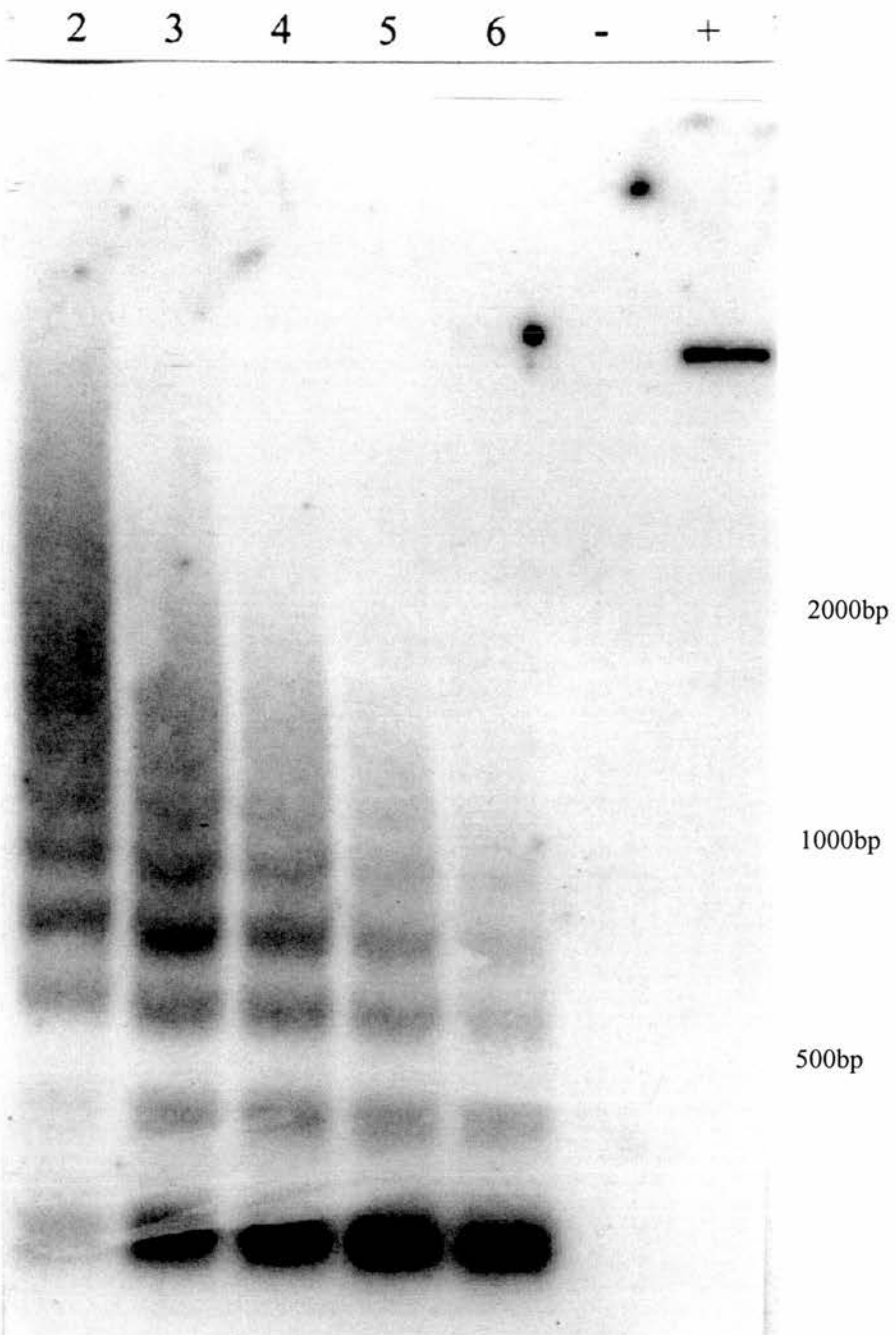



**Figure 33** Nucleosomes Over the -0.9kb Region of Ovine BLG in Liver

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Rsa*I**. The naked DNA control (n+) was only digested with MNase, but was too faint to be seen on this exposure. Sizes are shown in relation to a stronger exposure of the Promega 1kb marker lane (M). The probe used was -900DOWN (Figure 26).

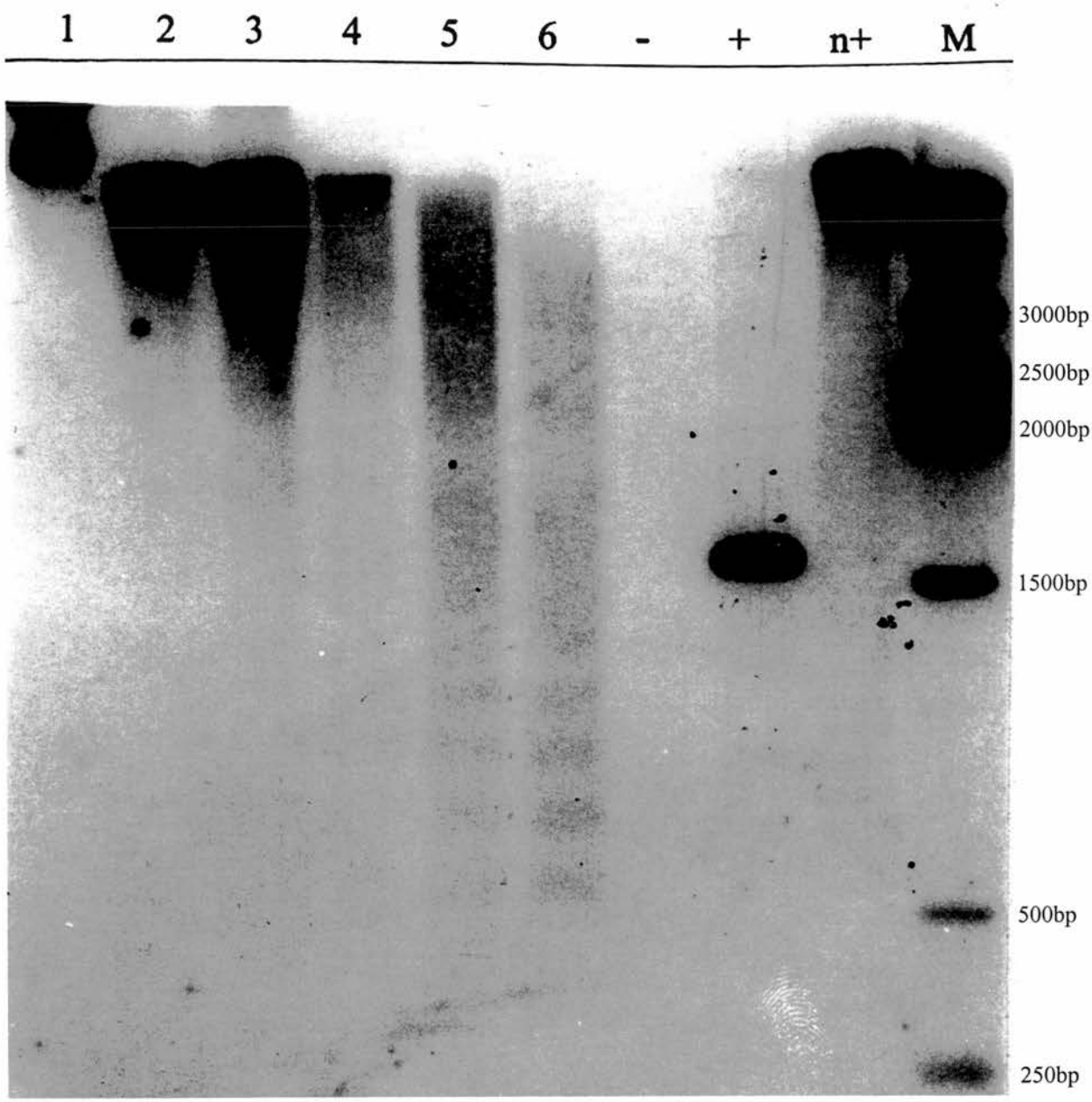



**Figure 34** Nucleosomes Over the +2.7kb Region of Ovine BLG in Liver

Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was +2700DOWN (Figure 26).

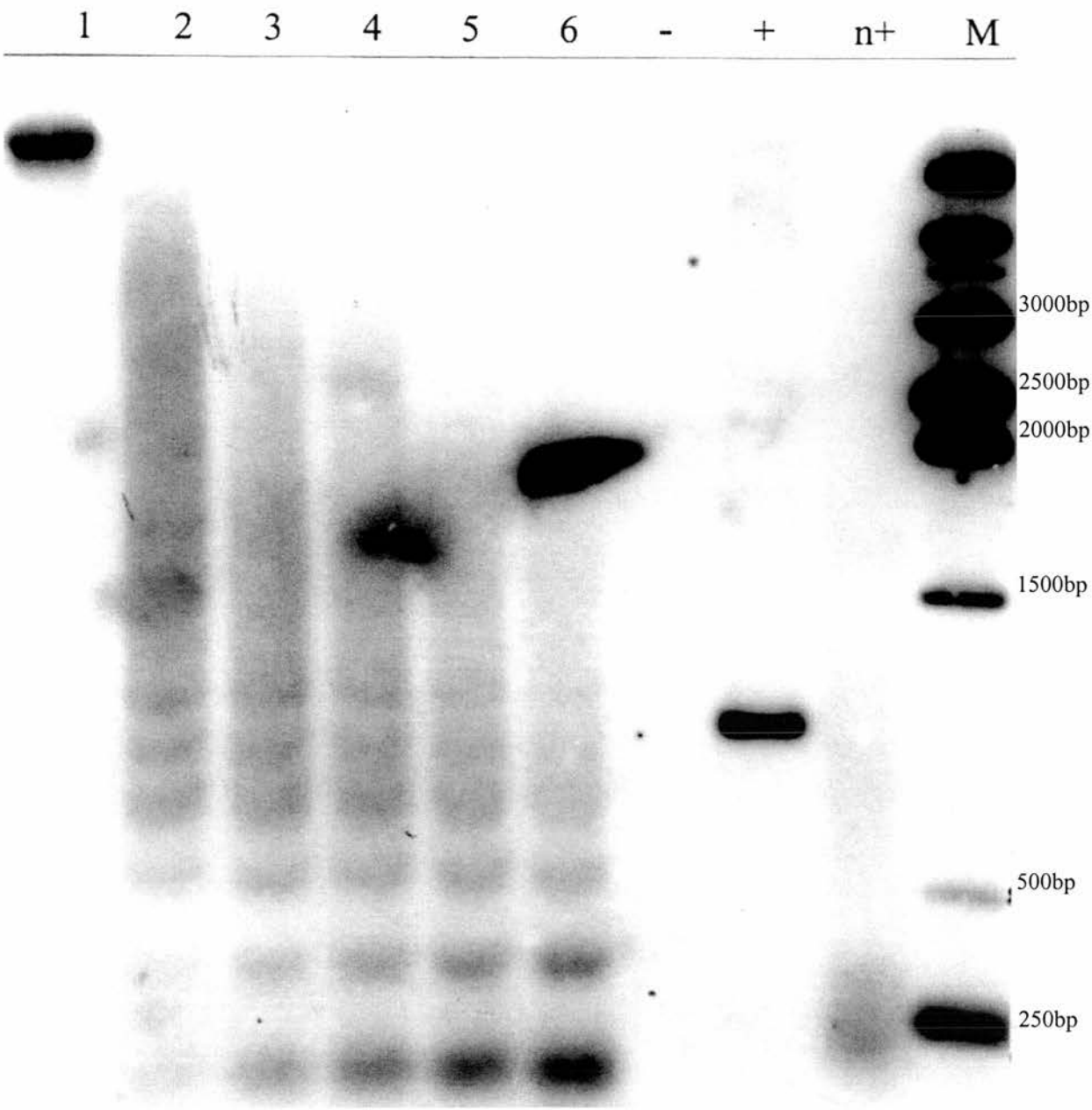



**Figure 35** Nucleosomes Over the +4.7kb Region of Ovine BLG in Liver

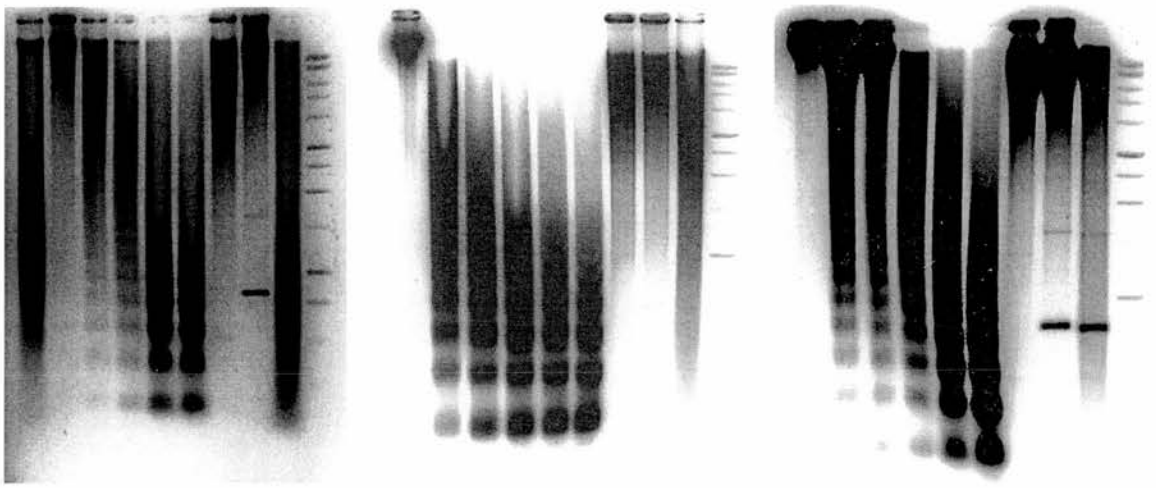
Lanes 1 to 6 show increasing digestion times with MNase (see **3.2.1**). The negative (-) and positive (+) controls were only digested with ***Bam*HI**. The naked DNA control (n+) was only digested with MNase. Sizes are shown in relation to the Promega 1kb marker lane (M). The probe used was +4700DOWN (Figure 26).

| Mammary Chromatin | | | | |
|---|---|---|---|---|
| genome (average) | -1700*DOWN* probe | +1700*UP* probe | +2700*UP* probe | +4700*DOWN* probe |
| 163 | 200 | 210 | 180 | 180 |
| 360 | 380 | 370 | 360 | 350 |
| 540 | 540 | 560 | 540 | 520 |
| 713 | 720 | 720 | 720 | 730 |
| 897 | 900 | 890 | 900 | 900 |
| 1093 | 1080 | 1060 | 1080 | 1060 |
| 1267 | 1260 | 1260 | 1260 | |
| 1454 | 1440 | 1440 | 1440 | |
| 1644 | 1620 | | | |

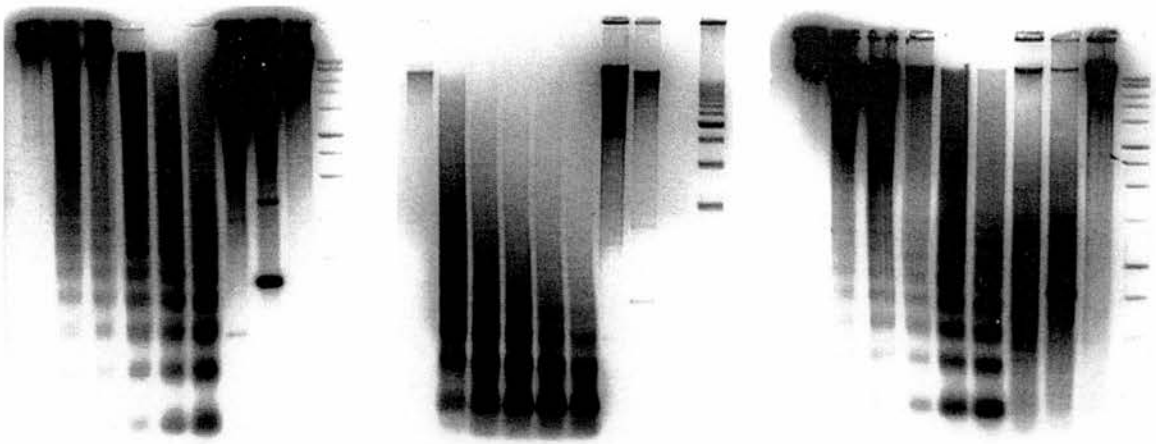
**Table 2** Fragment Sizes Generated from MNase Digestion of Mammary Chromatin with an Ovine BLG-Specific Probe

| Liver Chromatin | | | | |
|---|---|---|---|---|
| genome (average) | -1700*DOWN* probe | -900*UP* probe | +2700*DOWN* probe | +4700*DOWN* probe |
| 178 | 180 | 180 | ? | 170 |
| 363 | 360 | 350 | ? | 350 |
| 540 | 560 | 540 | 540 | 540 |
| 733 | 720 | 720 | 700 | 700 |
| 910 | 900 | 900 | 880 | 860 |
| 1090 | 1060 | 1080 | 1060 | 1050 |
| 1260 | 1260 | 1280 | 1260 | 1210 |
| 1440 | 1440 | 1440 | 1440 | |
| 1625 | 1620 | 1620 | 1620 | |
| 1800 | 1800 | | | |
| | 2000 | | | |

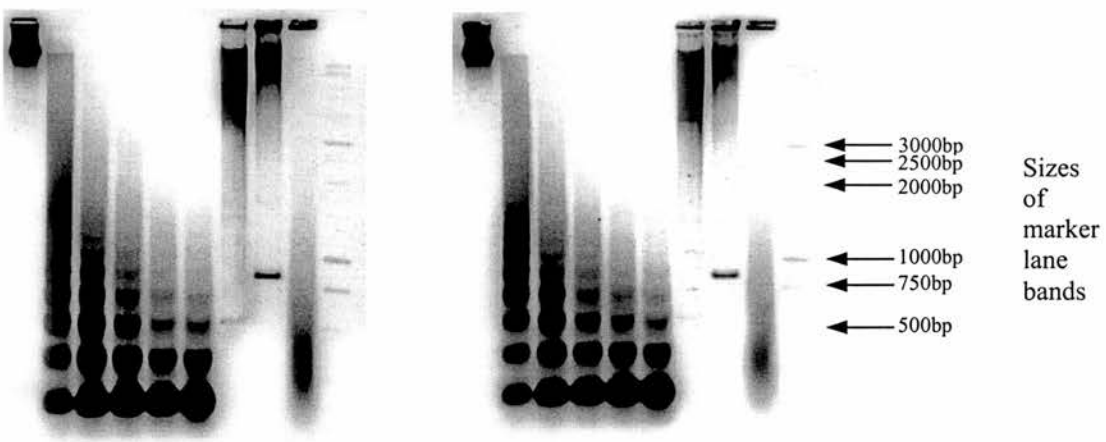
**Table 3** Fragment Sizes Generated from MNase Digestion of Liver Chromatin with an Ovine BLG-Specific Probe ('?' indicates that bands in this range cannot be seen)



-1700DOWN (mammary; figure 28) +1700UP (mammary; figure 29) +2700UP (mammary; figure 30)



+4700DOWN (mammary; figure 31) -1700DOWN (liver; figure 32) -900DOWN (liver; figure 33)




+2700DOWN (liver; figure 34) +4700DOWN (liver; figure 35)

**Figure 36** Ethidium Bromide-Stained Agarose Gels which Correlate with the Above Southern Blots See corresponding figure for details of each gel; the marker lane is the same in all

### 3.2.6 Analysis of Nucleosome Repeat Lengths

The size of each nucleosome multimer band from figure 36 was calculated for both mammary and liver chromatin. These multimer sizes were averaged from all gels and represent the mean size from the entire genome in that tissue. In addition, an average of all multimer sizes was calculated over the BLG gene for both mammary and liver chromatin (Table 4). These values were plotted relative to the band number in order to calculate the average repeat length for bulk chromatin and BLG chromatin in both mammary and liver tissues (Figure 37).

| | Nucleosome Band | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Mammary (genome) | 163 | 360 | 540 | 713 | 897 | 1093 | 1267 | 1454 | 1644 |
| Mammary (gene) | 193 | 365 | 540 | 723 | 898 | 1070 | 1260 | 1440 | 1620 |
| Liver (genome) | 178 | 363 | 540 | 733 | 910 | 1090 | 1260 | 1440 | 1625 |
| Liver (gene) | 177 | 353 | 535 | 710 | 885 | 1063 | 1253 | 1440 | 1620 |

**Table 4** Average Band Sizes for Gene-Specific and Bulk Chromatin Digests

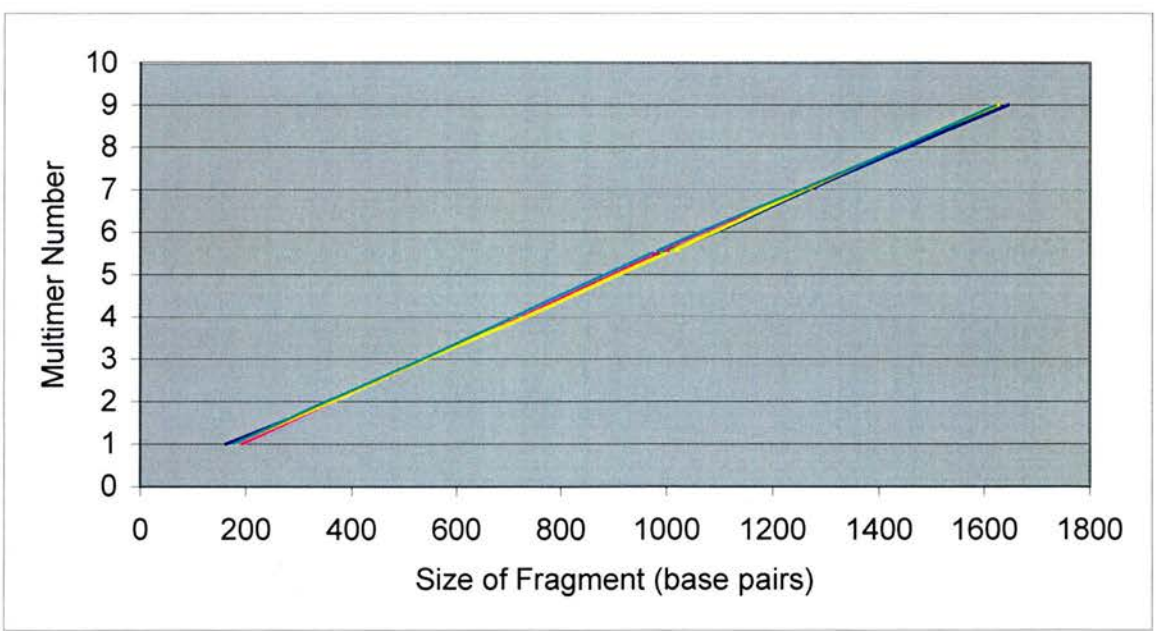



**Figure 37** Graphical Representation of the Nucleosome Repeat Lengths

Mammary genomic, dark blue; mammary gene, magenta; liver genomic, yellow; liver gene, cyan.

### 3.2.7 Discussion: Digestion of BLG with MNase

The average repeat lengths were calculated for the gene in its inactive and active states and for bulk liver and mammary chromatin (Table 5). There is a distinct difference between the nucleosome repeat lengths in bulk mammary and liver chromatin of 4.2bp. This must be due to tissue specific factors. The repeat length in liver chromatin is similar to that in the BLG gene where it only differs by 0.5bp. However, in mammary chromatin it differs by 6.7bp. This large difference may be caused by the association of the active gene with HMG14/17 rather that H1. These HMG proteins sequester 15bp of linker DNA whereas H1 sequesters 20bp (Goodwin *et al.*, 1979). Thus, this association might reduce the nucleosome repeat length over the active gene.

| | Mammary (genome) | Mammary (gene) | Liver (genome) | Liver (gene) |
|---|---|---|---|---|
| Repeat Length | 185.1bp | 178.4bp | 180.9bp | 180.4bp |

**Table 5** Nucleosome Repeat Lengths

Sizing these MNase digested fragments proved to be slightly problematic, because as the MNase titration proceeded, the repeat length decreased. This is caused by the exonucleic activity of the enzyme as it chews at the ends of the endonucleically cleaved fragments. This was particularly evident in monomer, dimer and trimer fragments as they lose more of their mass relative to the larger multimer fragments (Figure 38) and (Figure 33).

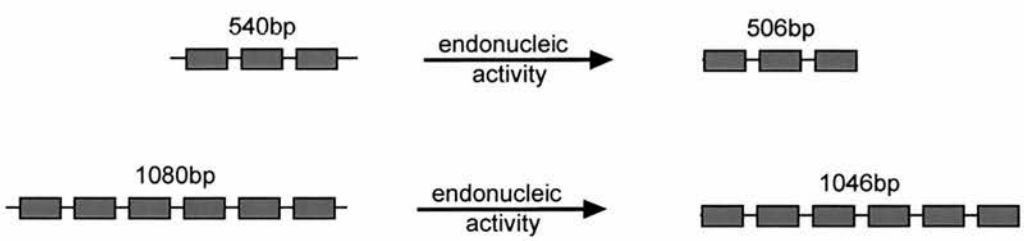



**Figure 38** Micrococcal Exonuclease-Induced Size Reduction of DNA Fragments

A trimer fragment of 540bp is reduced approximately by 6% to 506bp when the terminal linker regions are chewed away. However, a six-mer is only reduced by approximately by 3%. A red box indicates nucleosome-protected DNA.

This method is sufficient to map the BLG locus from -4kb to +6kb for the presence and repeat length of nucleosomes in both tissues, so long as nucleosomes extend out in both directions from the probe. For instance, when probed with *-900DOWN*, over ten multimers are evident in liver chromatin (Figure 33), a distance of 1800bp. If they extend in both directions, this maps from approximately -1800bp to the transcriptional start site. A large difference in the repeat length in either direction would be evident as two bands instead of one. A small difference would only be noticeable as a slightly broader band in large multimer fragments since this assay is not sensitive enough to distinguish small changes. Indeed, I estimate that it has an accuracy of approximately +/- 20bp.

Of course, this can only be true when nucleosomes extend outwards in both directions. If a nuclease resistant domain were present on one side of the probe (Figure 39), the only difference would be a reduction in the banding intensity after the domain had been reached. This could be very hard to determine in such an assay. Alternatively, if transcription factors or a protein complex existed on one side of the probe there could be diverse effects.

Bound transcription factors, such as NF1, protect much less DNA than a nucleosome. This could result in an altered array on either side, or a smear of fragments over that region owing to nucleosome sliding or a DNase I hypersensitive site. However, in order for NF1 to bind, chromatin remodelling may have to take place (Di Croce *et al*, 1999). These effects are more likely to result from a combination of interactions, such as the gene being pre-set for activation, nucleosome remodelling and then hierarchical transcription factor interactions.

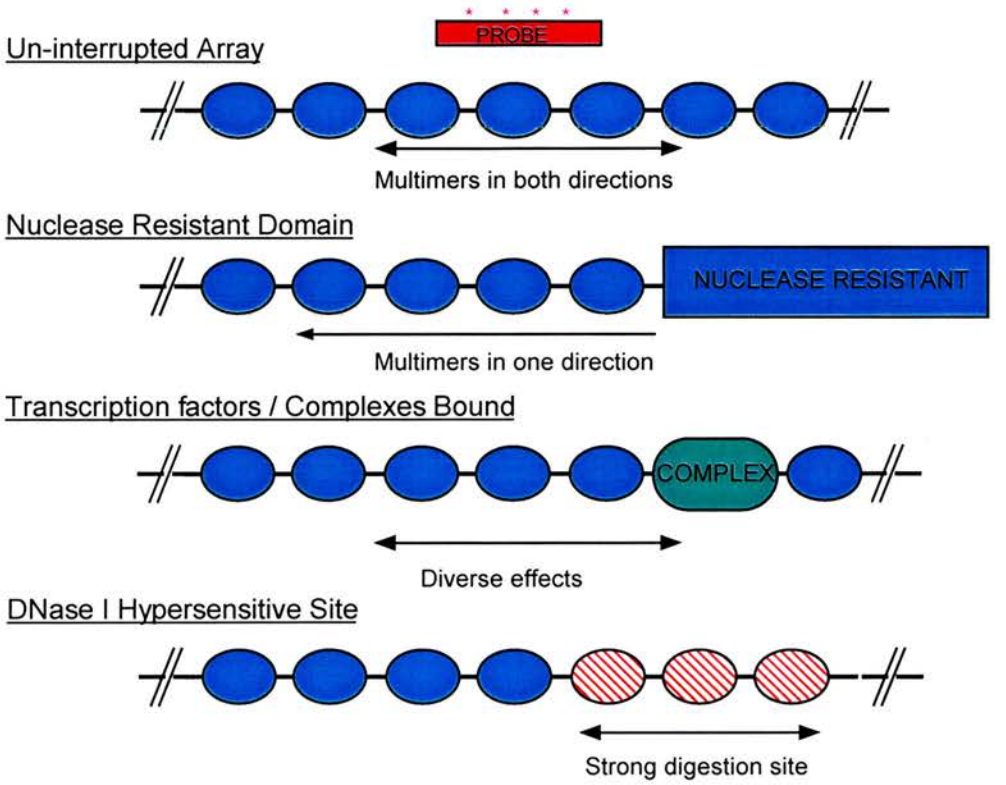



**Figure 39** Chromatin Structure can Alter Nucleosome Arrays Detected by MNase

None of these effects are obvious in these experiments. In order to examine the gene more closely and possibly visualise these alterations in the array, we must look in only one direction at a specific location.

### 3.2.8 Micrococcal Nuclease Analysis of Positioned Nucleosomes on the BLG Gene

In order to examine the nucleosomes in only one direction from the probe, a restriction enzyme must be used to delimit the probing direction (Figure 40). This abuts the probe and acts as a reference point to determine the MNase cut sites on the gene. In the following four examples for mammary (Figures 41 & 42) and liver chromatin (Figures 43 & 44), the digestion sites have been tabulated in table 6. The region of BLG mapped in each case is shown in figure 26.

The following results are displayed in a standard format: a photomicrograph of the ethidium-stained DNA is shown before transfer onto nitro-cellulose membrane. Once this had been probed, the radioactive bands were visualised on photographic film. A diagram shows how these bands relate to their position on the gene. This was accomplished by plotting their position on the image in relation to those of the maker lane bands. A second diagram indicates how these digestion sites might correlate to nucleosomes positioned on the gene region. See the legend below for the symbols relating to these diagrams.

### 3.2.9 Examination of the BLG Gene in Mammary Chromatin

*+2700DOWN Probing of Mammary Chromatin*

The MNase digestion sites show a regular periodicity over this gene region in a downstream direction from +2744bp (Figure 41). There is a phased array which begins at approximately +3020bp and continues for seven nucleosomes downstream until +4290bp. However, there is a digestion site missing in this array at position +4110bp. This might be caused by MNase sequence specificity, as the enzyme may find the DNA conformation difficult to cleave at this site. There are, however, no obvious stretches of sequence in this region which it might find intolerable. Alternatively, the partial digestion time points taken may simply not be best suited to reveal this digestion site.

In addition to this array, there is a single nucleosome positioned at +3070 to +3260. This could be the end of an alternative nucleosome array which extends up until this point. A map extending upstream from this point would answer this question.

*+4700DOWN Probing of Mammary Chromatin*

Owing to the small size of the genomic fragment generated from +4755bp to the next ***Bam*HI** site, only three nucleosomes can be visualised (Figure 42). These are in a different phase from the previous experiment and extend from +4925 to +5465. Apart from this array, there are two additional digestion sites at +5195 and +5445. Again, it is unclear what significance these have. As the MNase lane was not digested with ***Bam*HI**, it is impossible to say whether these are due to sequence specific digestion or are related to protein interactions with the DNA.

### 3.2.10 Examination of the BLG Gene in Liver Chromatin

*+2700DOWN Probing of Liver Chromatin*

This array (Figure 43) is identical to that seen in mammary chromatin (Figure 41). Again it starts at +3020 (+/- 20bp) and continues downstream for seven nucleosomes. Interestingly, the additional bands seen in the mammary sample do not appear in this one. There is an alternative site at +3120 which could mean that some other protein is binding this region or that it is again sequence-specific cutting. In this case, the nucleosome rotational positioning must differ such that these different sites are open to cleavage on the nucleosome's surface.

Like mammary chromatin, there is a digestion site missing, but in this case it is on the other side of the nucleosome at +1190bp! It is unclear why this might be, but perhaps a protein is binding the linker region and inhibiting MNase digestion on one of the two sites. As they are very close, it is feasible that it could be exchanged between sites depending on the tissue. Alternatively, as the liver chromatin is in a different conformation from the active gene, in higher order structure, certain sites may be less accessible to MNase than those in the active gene. This could lead to different cutting efficiencies.

An additional problem with MNase was the naked DNA control. In both this blot and the subsequent one, the control DNA has been almost completely degraded after ***Bam*HI** digestion. Further experiments revealed that MNase is responsible for this degradation, even although the sample was phenol :chloroform extracted, with particular

care to avoid contamination from the proteinacious interface. Only this lane suffered from this problem and it was prominent in the majority of blots. Digestion with ***Bam*HI** before MNase digestion would have alleviated this problem, but may have given rise to erroneous sequence-specific digestion sites, as the fragment sizes would be much smaller.

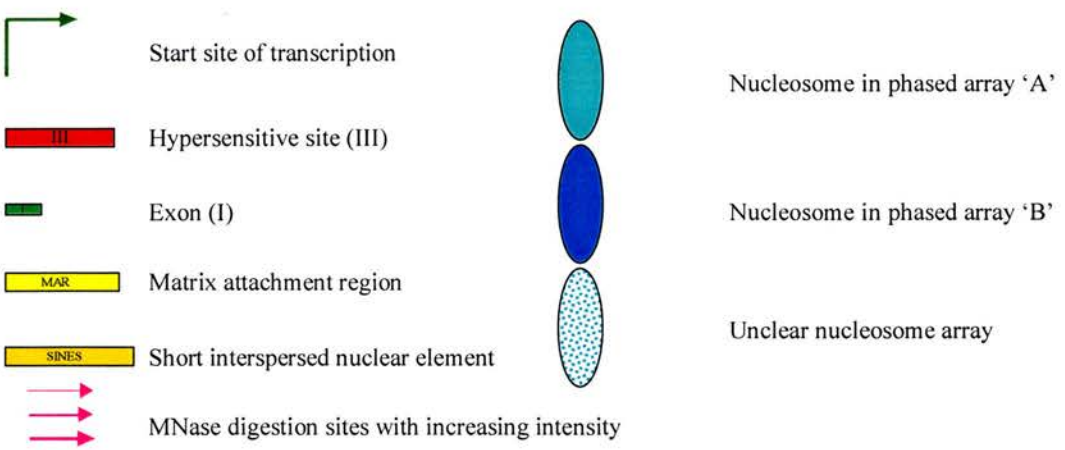
*+4700UP Probing of Liver Chromatin*

These results suggest that two alternative arrays exist from +4075 until exon 6 at +4560, at which point only one array persists (Figure 44). A number of digestion sites are present around this junction point, not all of which align with a nucleosome array. Chromatin remodelling could account for these extra bands. However, it is impossible to say without a full naked DNA control whether these extra sites exist because of MNase sequence-specific digestion.

This experiment displays several common features with nucleosomes positioned on mammary chromatin over the downstream region from +4.7kb (Figure 42). Notably, both the dark blue coloured arrays and the light blue coloured arrays align in a 180bp periodicity. It was initially unclear what the band at +5195bp in the mammary chromatin represented, but as it aligns with the light blue array in liver chromatin at +4475bp, it probably indicates the end of this alternative array.

Like the previous blots, a digestion site is absent at +3435bp. In addition to the previous possibilities, this may be due to chromatin remodelling over the exon 6 region.

LEGEND:



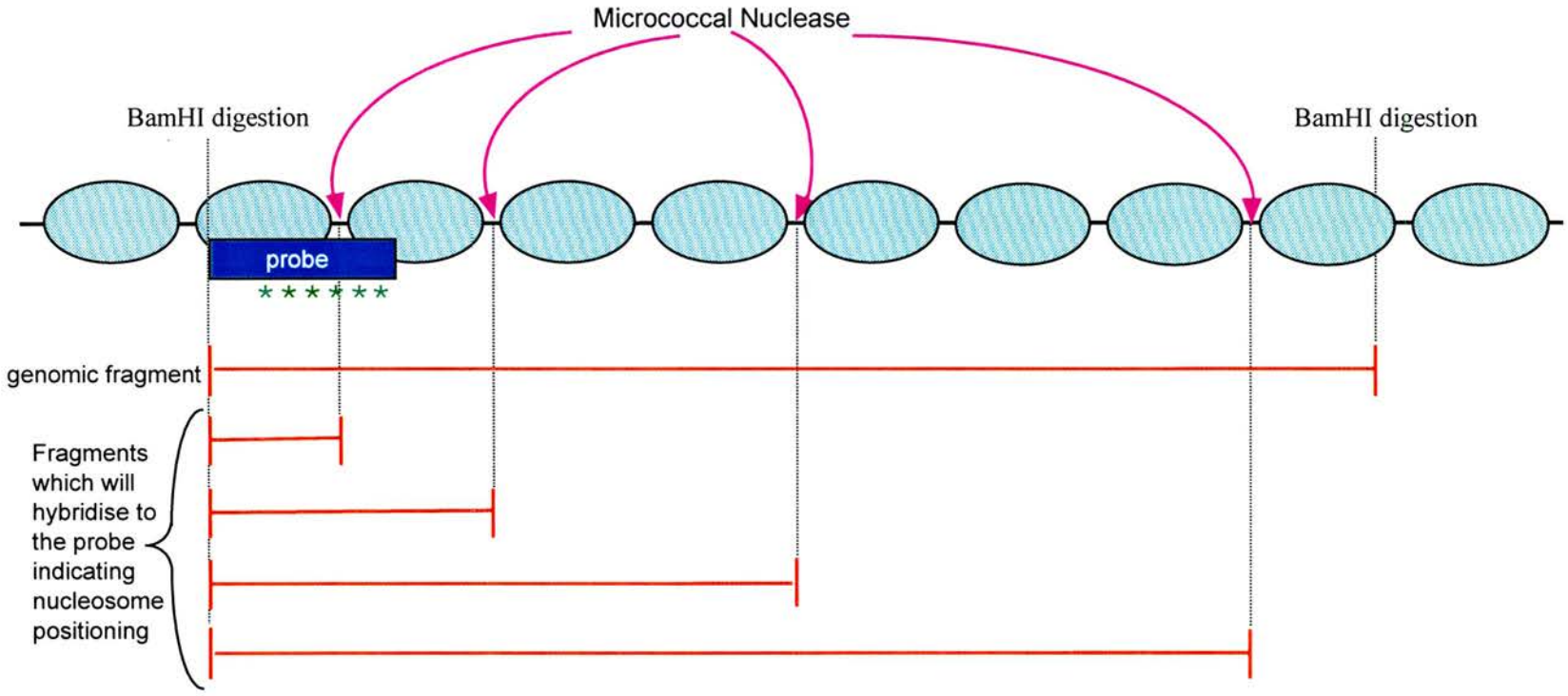
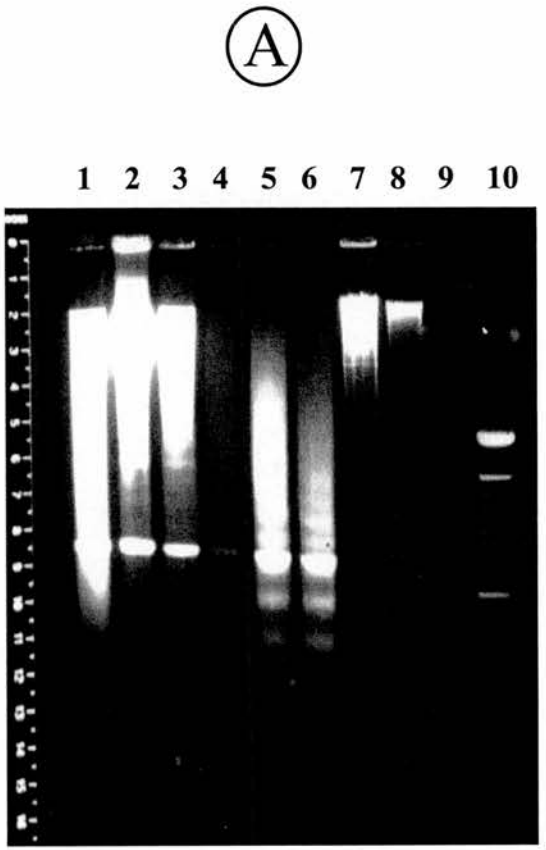



**Figure 40** A Schematic of the Indirect End-Labelling Technique to Test for Nucleosome Positioning

MNase partially digests chromatin leaving fragments of DNA in multiples of 180bp. Further digestion of the purified DNA with a restriction enzyme, such as ***Bam*HI**, acts as a reference point. After agarose electrophoresis and Southern blotting, the membrane is hybridised with a radioactively labelled probe which abutts the BamHI site on the BLG gene. Any distinct bands revealed after exposure to film indicates a DNA footprint. Footprints which are in multiples of 180bp are most likely to represent nucleosomes covering the DNA. No distinct bands suggests that nucleosomes are either not present or are motile along the DNA duplex. Fragments hybridising to the probe are red; the probe is dark blue, nucleosomes are light blue; the radioactive label is green; MNase cut sites are pink.

**Figure 41 Nucleosome Mapping Downstream from +2.7kb in the Ovine BLG Gene in Mammary Chromatin**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) MNase digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 0 (0 minutes)
2. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 1 (1 minute)
3. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 2 (2 minutes)
4. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 3 (5 minutes)
5. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 4 (10 minutes)
6. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 5 (15 minutes)
7. Negative control: mouse genomic DNA digested with ***Bam*HI**
8. Positive control: sheep genomic DNA digested with ***Bam*HI**
9. Naked DNA control: sheep genomic DNA digested with MNase and ***Bam*HI**
10. Marker lane: Promega 1kb ladder

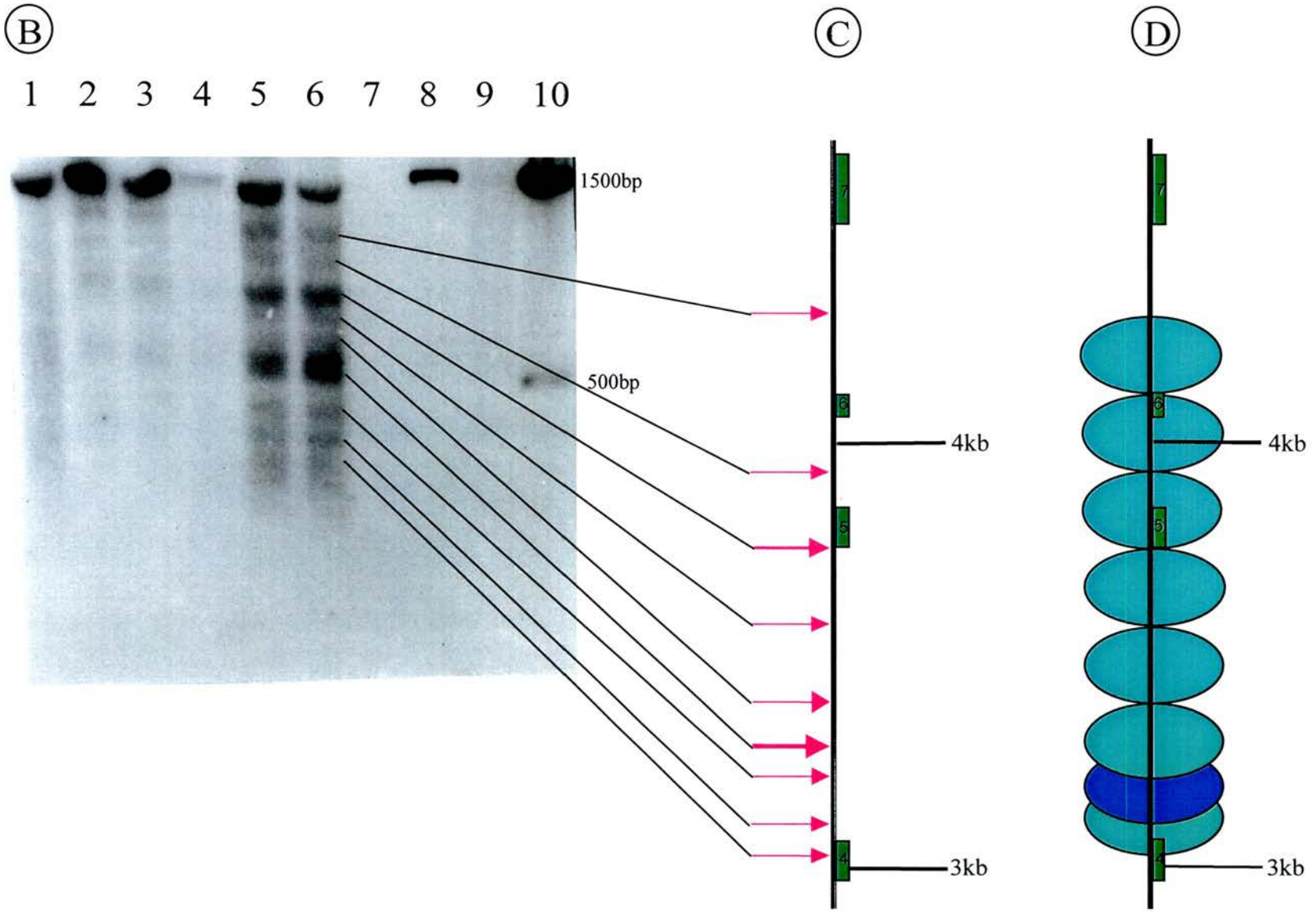
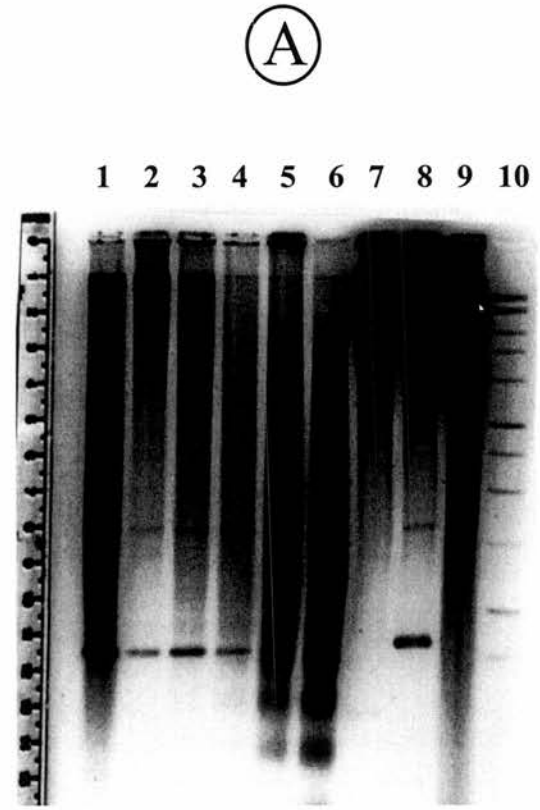

B
C
D
1 2 3 4 5 6 7 8 9 10
1500bp
500bp
4kb
3kb
4kb
3kb

**Figure 42 Nucleosome Mapping Downstream from +4.7kb in the Ovine BLG Gene in Mammary Chromatin**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) MNase digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 0 (0 minutes)
2. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 1 (1 minute)
3. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 2 (2 minutes)
4. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 3 (5 minutes)
5. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 4 (10 minutes)
6. Sheep mammary nuclei digested with ***Bam*HI** plus MNase; time point 5 (15 minutes)
7. Negative control: mouse genomic DNA digested with ***Bam*HI**
8. Positive control: sheep genomic DNA digested with ***Bam*HI**
9. Naked DNA control: sheep genomic DNA digested with MNase
10. Marker lane: Promega 1kb ladder

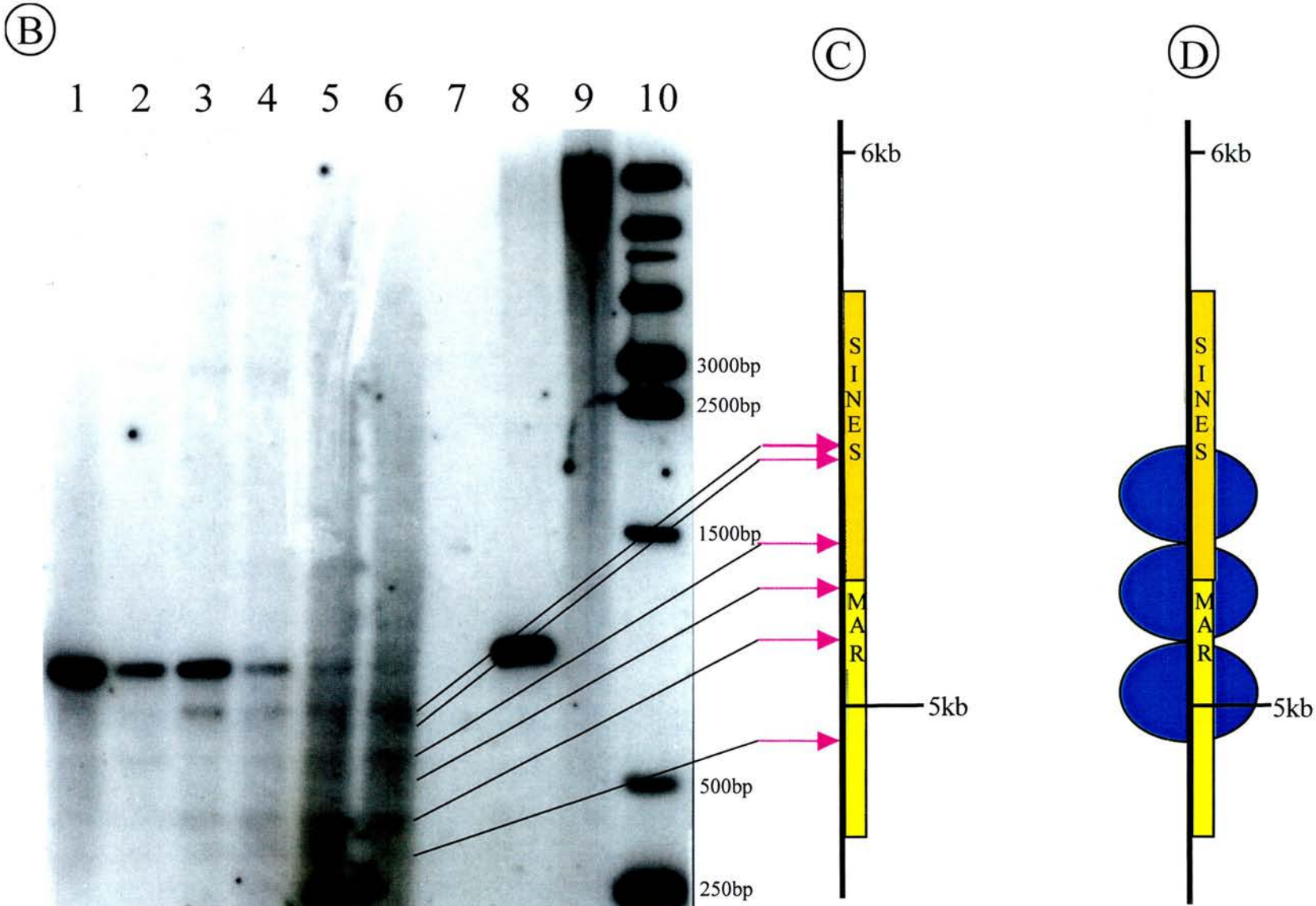
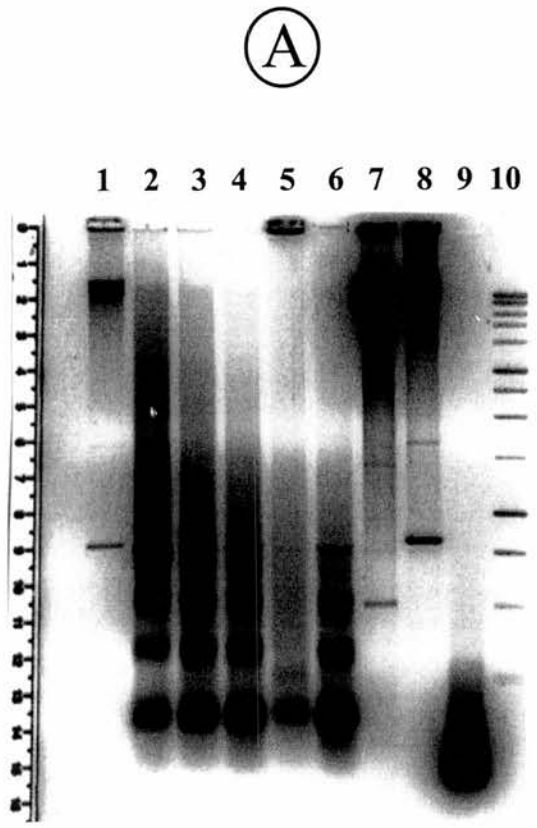

B
1 2 3 4 5 6 7 8 9 10
3000bp
2500bp
1500bp
500bp
250bp
C
6kb
SINES
MAR
5kb
D
6kb
SINES
MAR
5kb

**Figure 43 Nucleosome Mapping Downstream from +2.7kb in the Ovine BLG Gene in Liver Chromatin**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) MNase digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 0 (0 minutes)
2. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 1 (1 minute)
3. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 2 (2 minutes)
4. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 3 (5 minutes)
5. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 4 (10 minutes)
6. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 5 (15 minutes)
7. Negative control: mouse genomic DNA digested with ***Bam*HI**
8. Positive control: sheep genomic DNA digested with ***Bam*HI**
9. Naked DNA control: sheep genomic DNA digested with MNase and ***Bam*HI**
10. Marker lane: Promega 1kb ladder

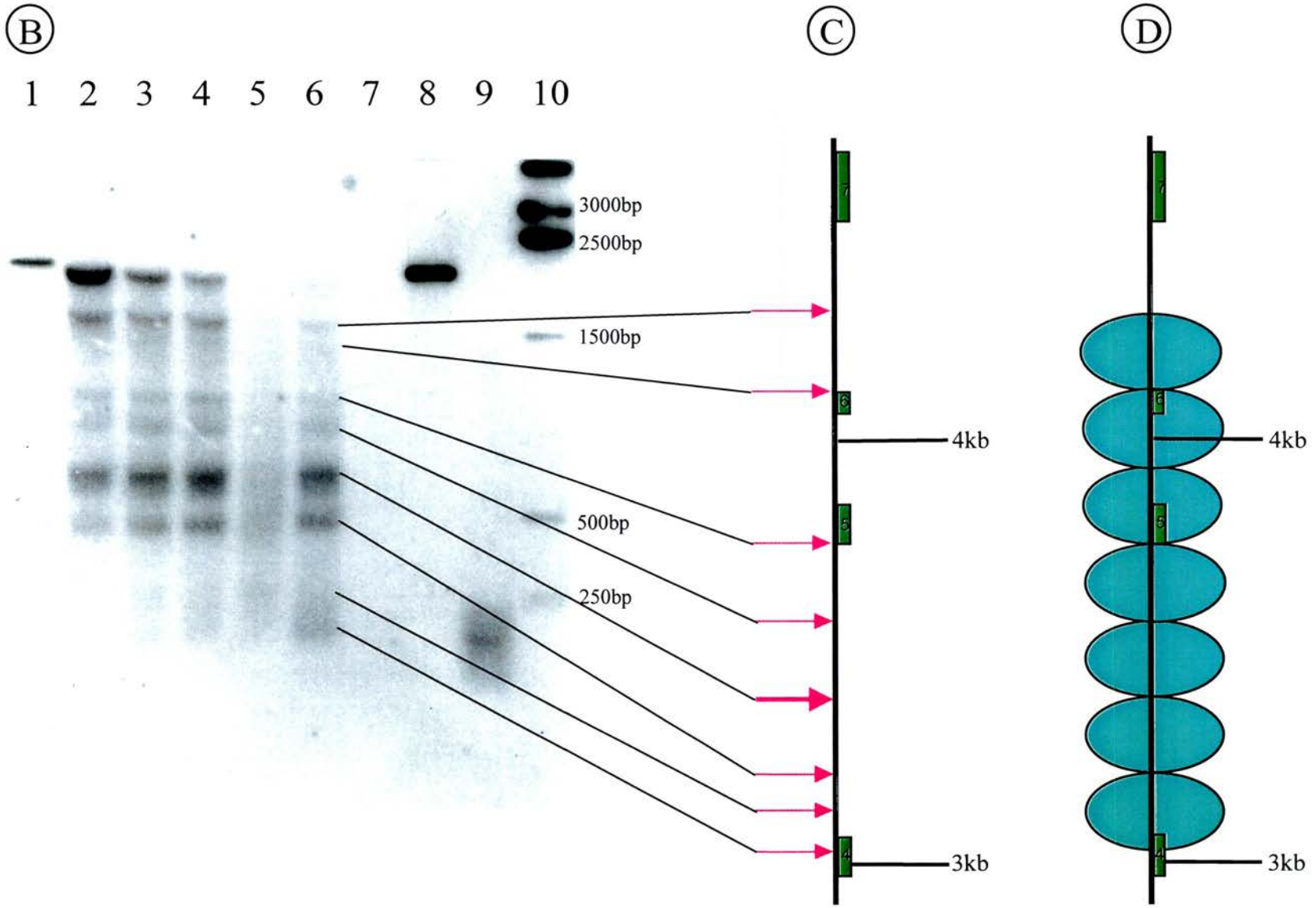
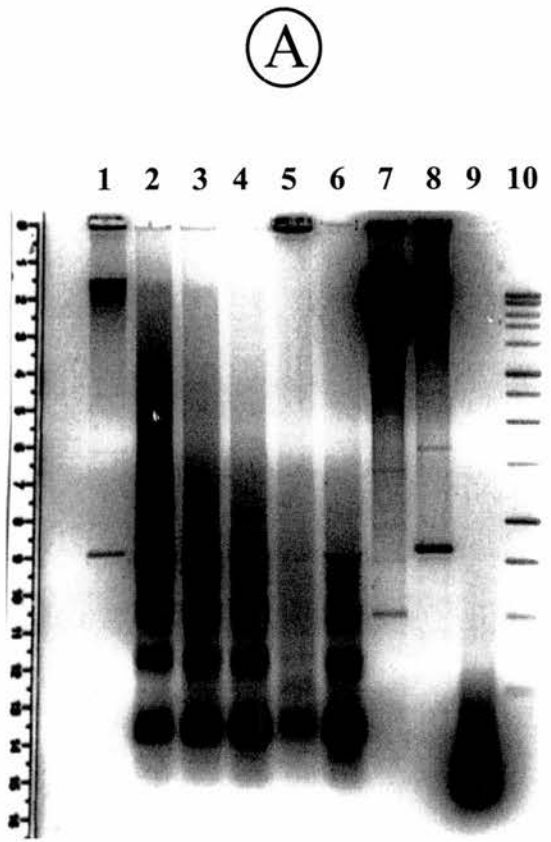

B
C
D
1 2 3 4 5 6 7 8 9 10
3000bp
2500bp
1500bp
500bp
250bp
4kb
3kb
4kb
3kb

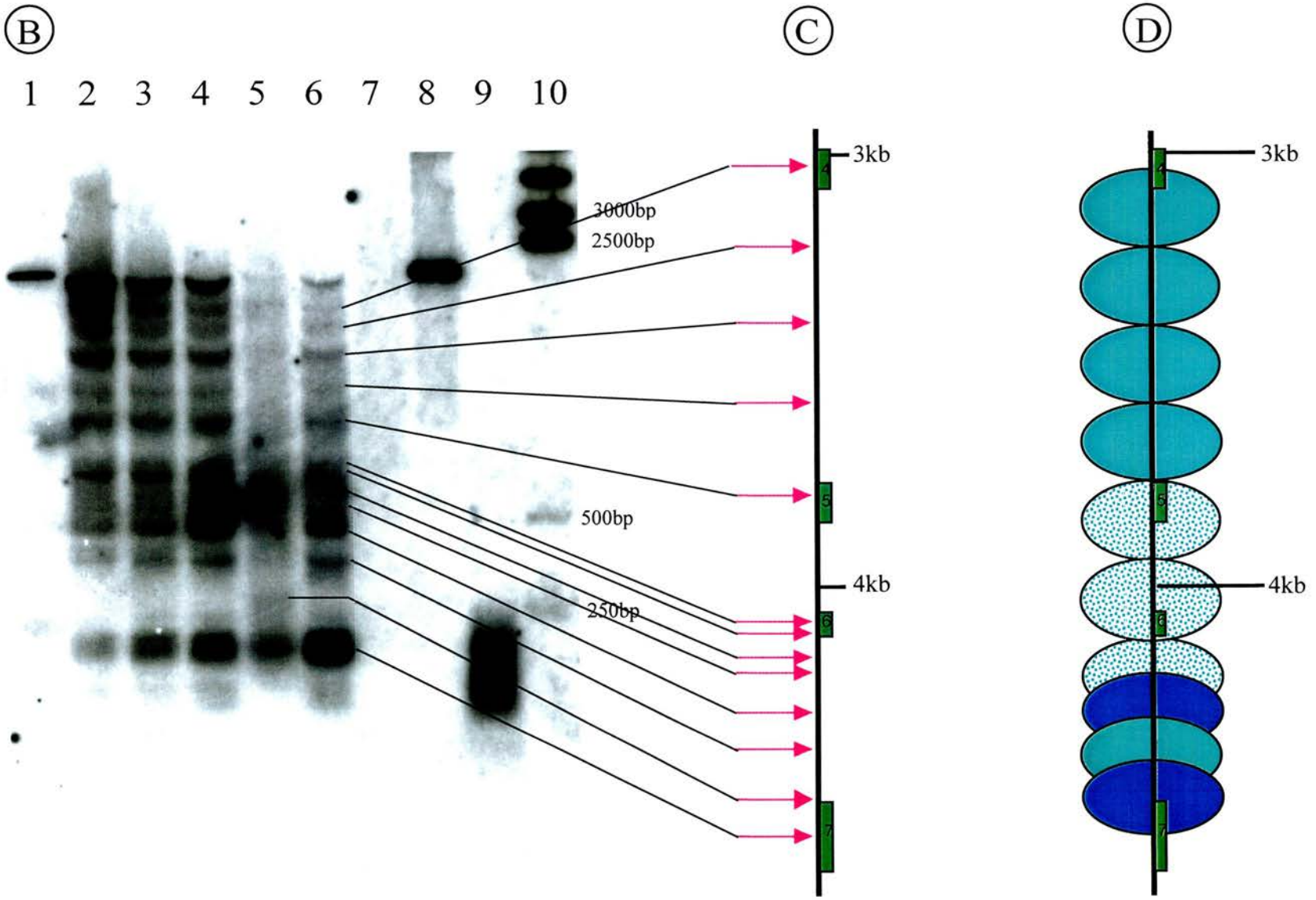
**Figure 44 Nucleosome Mapping Upstream from +4.7kb in the Ovine BLG Gene in Liver Chromatin**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) MNase digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 0 (0 minutes)
2. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 1 (1 minute)
3. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 2 (2 minutes)
4. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 3 (5 minutes)
5. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 4 (10 minutes)
6. Sheep liver nuclei digested with ***Bam*HI** plus MNase; time point 5 (15 minutes)
7. Negative control: mouse genomic DNA digested with ***Bam*HI**
8. Positive control: sheep genomic DNA digested with ***Bam*HI**
9. Naked DNA control: sheep genomic DNA digested with MNase and ***Bam*HI**
10. Marker lane: Promega 1kb ladder


B
C
D
1 2 3 4 5 6 7 8 9 10
3000bp
2500bp
500bp
250bp
3kb
4kb
3kb
4kb

| Mammary Chromatin | | | | Liver Chromatin | | | |
|---|---|---|---|---|---|---|---|
| +2700DOWN | | +4700DOWN | | +2700DOWN | | +4700UP | |
| Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene |
| 270 | +3010 | 170 | +4925 | 280 | +3020 | 195 | +4560 |
| 330 | +3070 | 350 | +5105 | 380 | +3120 | 280 | +4475 |
| 430 | +3170 | 440 | +5195 | 460 | +3200 | 380 | +4375 |
| 520 | +3260 | 530 | +5285 | 640 | +3380 | 460 | +4295 |
| 640 | +3380 | 690 | +5445 | 850 | +3590 | 550 | +4205 |
| 830 | +3570 | 710 | +5465 | 1010 | +3750 | 580 | +4175 |
| 1010 | +3750 | | | 1370 | +4110 | 640 | +4115 |
| 1190 | +3930 | | | 1550 | +4290 | 680 | +4075 |
| 1550 | +4290 | | | | | 960 | +3795 |
| | | | | | | 1160 | +3595 |
| | | | | | | 1370 | +3385 |
| | | | | | | 1540 | +3215 |

**Table 6** MNase digestion sites over selected areas of the ovine BLG gene

## 3.3 Overview and Discussion

The first series of experiments (Figures 28-35) show that BLG contains a regularly spaced array of nucleosomes in both the active gene, in mammary nuclei, and the inactive gene, in liver nuclei. Moreover, the repeat length of approximately 180bp shows an interesting correlation with gene activity.

The second series of experiments (Figures 41-44) indicate that nucleosomes are positioned on the BLG gene. These positions correspond to two different phases (nucleosomes coloured dark blue and light blue) which are not mutually exclusive. Presumably, one phase may exist in one cell while another exists in a neighbouring cell. Indeed, they could be in a constant state of flux, shifting between either conformation.

### 3.3.1 MNase as a Probe for Chromatin Structure

I have found that MNase is limited as a probe for chromatin structure. Firstly, there are a substantial number of additional digestion sites which do not correspond to the nucleosome array. Indeed, there are some bands missing which are expected to exist to complete the array. These discrepancies may exist because of the heterogeneity in the nuclei populations taken from both tissues. It is feasible that alternative positions are

adopted in different cell types within a given tissue. This would give rise to additional bands and possibly an absence of bands, where the nucleosome position is not defined. However, an equally likely explanation is that they are caused by the enzyme's sequence specificity.

This is further complicated because of problems with the naked DNA control. This control is necessary to highlight sequence-specific digestion sites which may be present in the test lanes. Even after careful phenol: chloroform extraction, MNase was consistently carried over into the restriction enzyme digestion where it digested the DNA almost completely. Multiple extractions failed to circumvent this problem. Even this control may not be sufficient to overcome the problems of MNase sequence-specific digestion as it has been shown to digest naked DNA in multiples of 200bp (Keene & Elgin, 1981).

As the DNA is undoubtedly in a different conformation in chromatin than in that of naked DNA, this control may not identify all the genuine sequence-specific digestion sites which are present in the nuclei. Unfortunately, this is a rather circular problem which cannot be easily rectified. However, in a histone reconstitution reaction, one group did show that the sequence specificity of the enzyme did not differ between nucleosomal and naked DNA (Laskey 1977) .

Interestingly, these sequence-specific digestion sites may have evolved through evolution and reflect a functional relationship between DNA and the nucleosomes which organise it (Keene & Elgin 1981).

### 3.3.2 MNase Sequence Specificity

Other groups have investigated this sequence specificity in detail. An analysis of mouse satellite sequence DNA and pBR322 plasmid DNA *in vitro* (Horz & Altenburger 1981) showed that:

1. The majority of cleavage sites consist of alternating T/A sequences, specifically TA rather than AT.
2. A cytosine residue at the 5' end of an A/T-region, equivalent to a guanine residue at the 3' end of this region, is required for preferential cleavage.

3. In addition, a G/C rich environment on either side of an A/T-region enhances cleavage.

At the same time another group (Dingwall & Lomonossoff, 1981) proved that in addition:

4. Nearly all tri-nucleotides released after digestion contain either an A or T residue at their 5' terminus (96%).
5. TT dinucleotides are hydrolysed 100 times more rapidly than GG dinucleotides.
6. It preferentially digests inverted repeats.
7. MNase digests only a limited number of sites and the frequency of cutting varies greatly between sites.

These results reflect those conducted on monkey α-satellite DNA in cell culture (Horz, Fittler 1983; Smith & Lieberman 1984).

Within the transcribed region, BLG is approximately 60% G/C-rich. In the flanking regions, which include 4kb 5' sequence and 1.9kb 3' sequence, this situation is reversed. This may reflect the problems experienced with MNase. Since it has such a strong sequence preference for TA sites (Figure 45), digestion within the coding region may be limited and may be reflected in the above experiments (Figure 43). In the flanking regions, it may be digesting heavily within the nucleosomes giving a high background (Figure 45). Clearly MNase is not an ideal probe for chromatin structure.

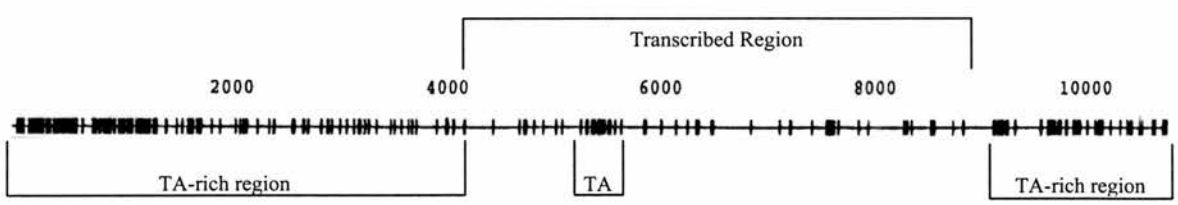



**Figure 45** TA di-nucleotides over the BLG Gene

Each vertical line represents a TA dinucleotide of which there are 297 in total over the gene domain

### 3.3.3 Alternative Probes for Chromatin Structure

Chemical nucleases are an alternative to enzymatic nucleases (Sigman, 1990a). These can be defined as complexes which are redox-active and can nick nuclei acids under physiological conditions by oxidative attack on the ribose or deoxyribose sugar. It is this ribose-directed reactivity which distinguishes them from other DNA-modifying chemical agents such as dimethyl sulphate. Unlike these which react preferentially with various bases, chemical nucleases are essentially sequence-non-specific as they react with the sugar backbone.

I considered a number of different types which are reviewed here.

*TMTA-Fe(III)*

(1,4,7-trimethyl-1,4,7-triazacyclononane)iron (III) chloride is a relatively new addition to the list of chemical nucleases (Ehmann *et al.*, 1998). It cleaves DNA in a sequence independent manner under physiological conditions. However, it is relatively unaffected by the non-sequence-specific histone-DNA interactions within a nucleosome complex and is therefore not limited to linker DNA digestion. Although unsuitable as a tool for nucleosome positioning, it intriguingly creates a DNA footprint over regions where sequence-specific trans-acting factors are bound to the DNA and therefore could be useful as a probe for transcription factors bound within a chromatin environment.

*Hydroxyl Radicals*

Iron (II)-mediated reduction of hydrogen peroxide releases a hydroxyl radical which cleaves the DNA sugar backbone with virtually no sequence specificity (Tullius *et al.*, 1986). It is ideal for generating very high resolution footprints of protein-DNA interactions, but this resolution is too high to make it suitable for nucleosome mapping. For instance, it can clearly show which side of the DNA molecule interacts with the bound protein.

*MPE-Fe*

Although hydroxyl radicals are, by themselves, unsuitable for nucleosome mapping, they can be useful when generated by a co-ordinated complex, such as methydiumpropyl-EDTA-Fe(II) (Hertzberg *et al.*, 1984). MPE-Fe intercalates into the DNA strand and produces a free radical in the presence of $O_2$. This can be supplied in the form of hydrogen peroxide. Although the optimum reaction conditions occur at 5M NaCl, it cleaves efficiently and almost sequence-independently at physiological conditions (Cartwright *et al.*, 1983). It has been used effectively as a probe for nucleosome positioning in cell culture to produce a very clear map (Benezra *et al.*, 1986) and is therefore an ideal candidate.

*Cuprous Phenanthroline*

Like MPE-Fe, the 1,10-phenanthroline-Cu(II) (OP-Cu) complex also interacts with the DNA within the minor groove and cleaves it by producing hydroxyl radicals. It possesses low sequence specificity and cleaves efficiently under physiological conditions. Initial experimenters pre-incubated nuclei with hydrogen peroxide as the oxygen donor and initiated the reaction by adding OP-Cu (Cartwright *et al.*, 1982). Although this method cleaved between nucleosomes with almost no sequence specificity, it also generated free radicals in solution and lead to cleavage within the nucleosomes. A subsequent method pre-incubated OP-Cu first at a moderately low concentration with nuclei, which lead to its association with the DNA. The later addition of mercaptopropionic acid as the oxygen donor meant that hydroxyl radicals were generated specifically at the sites of OP-Cu-DNA contact (Quivy *et al.*, 1996). Therefore, this method significantly decreases inter-nucleosomal cleavage events.

### 3.3.4 A New Probe for Chromatin Structure

In a review of several reagents for protein footprinting, Tullis and co-workers showed that MPE-Fe and OP-Cu provided the best resolution maps of the λ repressor-$O_R1$ complex (Tullius *et al.*, 1987). The comparison in this paper is shown in figure 46. Both are fairly comparable, except that MPE-Fe has fractionally less sequence specificity. Although the OP-Cu test lanes 9 & 10 are rather over-exposed (Figure 46), I

believe that OP-Cu provides a clearer and more definitive footprint. For this reason, I chose OP-Cu to extend my MNase data in investigating nucleosome positioning over the BLG gene.

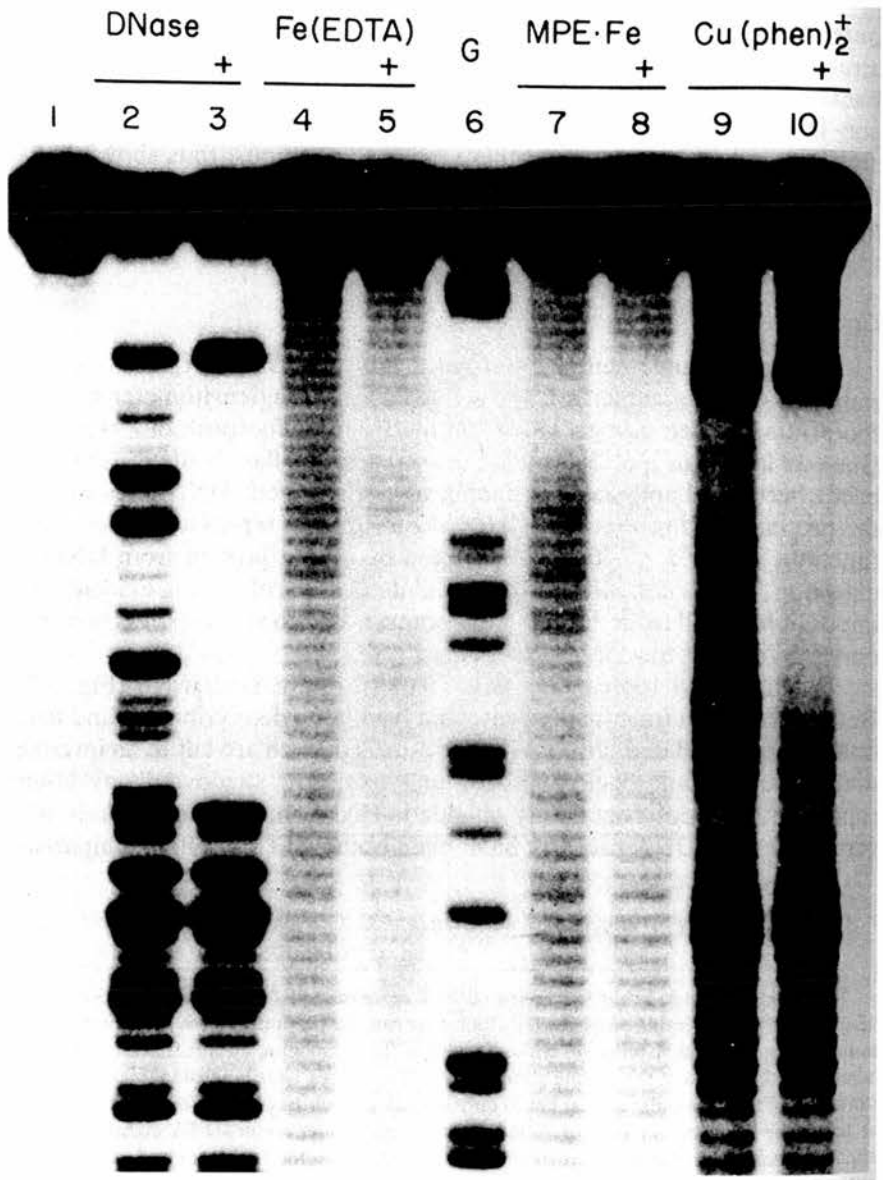



**Figure 46** Comparison of footprints of λ repressor made by four different reagents

DNA containing the $O_R$I operator sequence, labelled on the 3' end of the coding strand, was allowed to react with each of the four DNA-cutting reagents in the presence (lanes marked with +) or absence of λ repressor (90nM). Lane 1, untreated DNA. Lanes 2 and 3, products of DNA cut by DNase I. Lanes 4 and 5, products of DNA cut by hydroxyl radical produced by iron (II) EDTA reagent. Lane 6, products of the Maxim-Gilbert G-specific sequencing reaction. Lanes 7 and 8, products of DNA cut by MPE-Fe(II). Lanes 9 and 10, products of DNA cut by OP-Cu. Taken from Tullis et al, 1987.

# Chapter 4

# Chromatin Analysis of the Ovine BLG Gene using Cuprous Phenanthroline

## 4.1 Introduction

Owing to the limitations of micrococcal nuclease in examining chromatin structure (reviewed in **3.3.3**), I decided to use an alternative probe. Cuprous 1,10-phenanthroline (OP-Cu) is a chemical nuclease which yields fragments closely resembling those of MNase, except that the exonucleic action and 'pushing' of nucleosomes is not observed (Godde *et al.*, 1992; Cartwright *et al.*, 1982) (Figure 47). OP-Cu forms a tetrahedral complex consisting of a copper ion which associates with two phenanthroline molecules (Figure 48). It binds the minor groove of DNA through non-covalent interactions. Addition of 3-mercaptopropionic acid oxidises the complex to form a copper-oxo species. This generates a hydroxy-radical which reacts with the C1 hydrogen of the deoxyribose sugar, leading to its excision in the form of 5-methylenefuranone and therefore a break in the DNA strand (Sigman *et al.*, 1990b). Although alternatives exist, this reaction predominates.

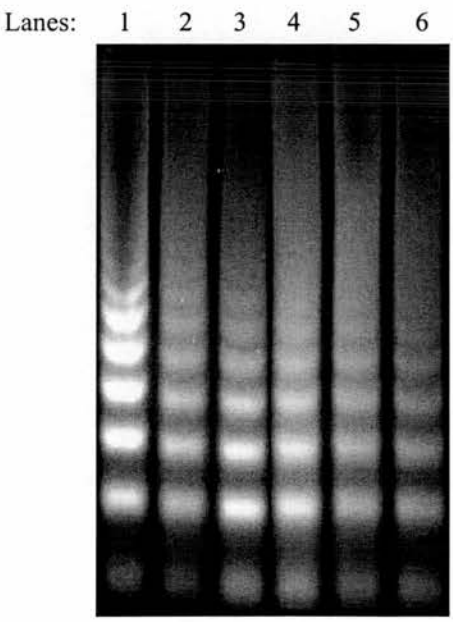



**Figure 47**

A comparison between MNase and OP-Cu digestion time-courses. Lanes 1 to 3 show increasing digestion time with MNase and a concomitant decrease in fragment sizes caused by its exonucleic activity. Lanes 4 to 6 show increasing digestion time with OP-Cu with consistent fragment sizes. Note: these two time-courses are taken from different experiments solely to illustrate this point.

The complex has a strong preference for linker DNA and very low sequence specificity, making it an ideal probe for nucleosome positioning (Yoon *et al.*, 1990). The small size of the molecule allows footprinting with higher resolution than with nucleases, which are sterically hindered to cleave close to histone-DNA contacts. Its reactivity with DNA is related to the accessibility and geometry of the minor groove and therefore it does display a limited amount of sequence specificity. This can be influenced by protein-DNA contacts and local distortions in the DNA structure (Papavassiliou, 1994). However, this specificity is much less severe than that of MNase (Quivy *et al.*, 1996).

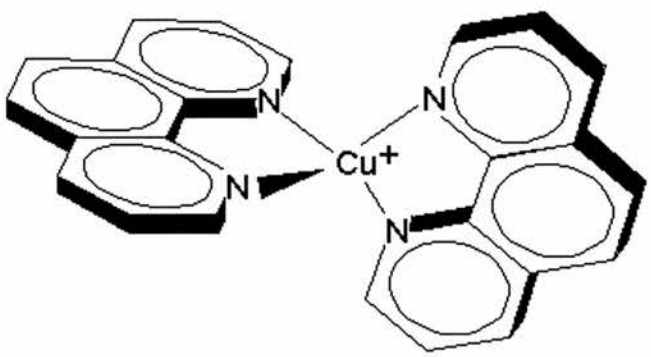



**Figure 48** The cuprous phenanthroline complex

## 4.2 Results

Using the same approach as that taken with MNase, sheep mammary and liver nuclei were digested with OP-Cu for 0, 10, 20, 30, 40 and 60 minutes to produce a broad distribution of partial digests. After purification, the DNA was digested with a specific restriction enzyme. The resulting fragments were separated on an agarose gel by electrophoresis and transferred to a nylon membrane via Southern blotting. Indirect end-labelling the fragments which abut the restriction enzyme site on the BLG gene reveals if and where nucleosomes are positioned.

In this way, I mapped the ovine BLG locus from -4.0kb to +5.5kb. The full time course experiments are shown in Appendix III, figures A1 to A16. From this primary data, I selected representative time points from mammary and liver digests in order to

compare them directly on a composite gel. These were electrophoresed with the appropriate controls and are displayed in figures 50- 57.

### 4.2.1 Controls

Similar controls to those in the MNase/restriction digested blots were used (chapter 3). The positive control was sheep genomic DNA digested with a restriction enzyme whose site abuts the probe (see chapter 3, figure 26). The mouse genomic DNA negative control was digested with the same restriction enzyme to generate a similar pattern of fragment sizes. The naked DNA control, which tests for OP-Cu sequence specificity, was digested with OP-Cu and the reaction terminated with neocuproine at various time points (Figure 49). A digestion time of 75 seconds produced the most representative smear of DNA fragments in relation to the test lanes. Further digestion, with the restriction enzyme whose site abuts the probe, enabled any sequence specific digestion sites to be pinpointed on the gene.

### 4.2.2 Probes Utilised

All eight probes from the MNase experiments (chapter 3, figure 26; excluding –2700UD in the repetitive region) were used to produce a contiguous nucleosome map over the entire gene (Table 7).

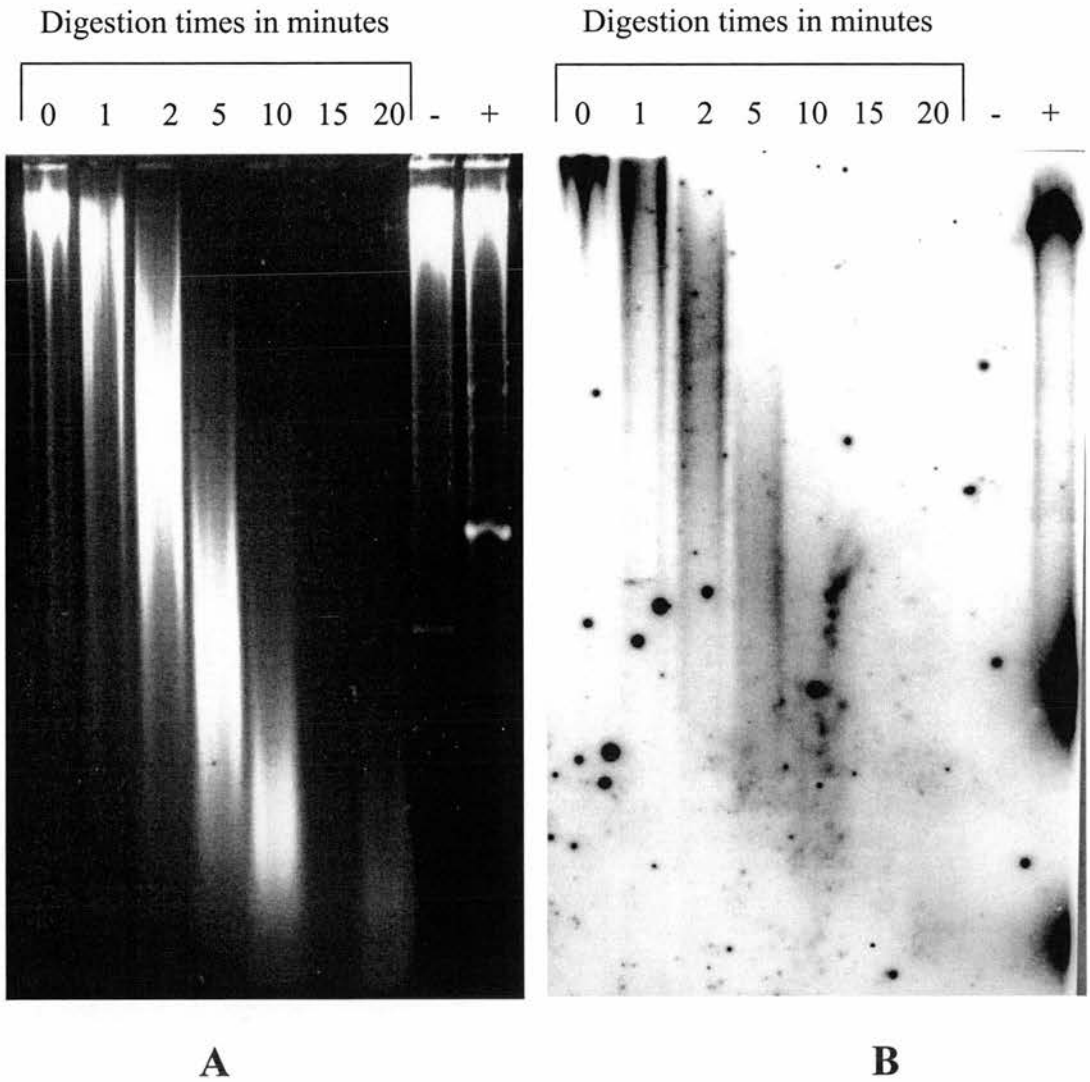
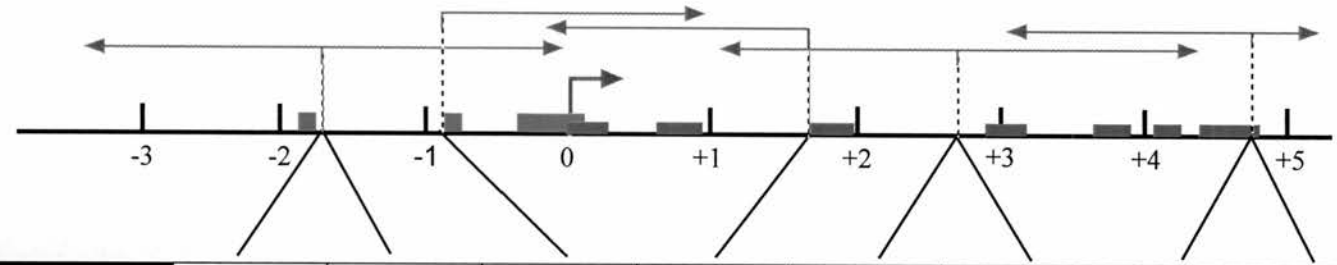



**Figure 49** The OP-Cu Naked DNA Positive Control

Sheep naked genomic DNA was digested with OP-Cu over a period of 20 minutes. Aliquots were taken at several time points during the reaction and the fragments generated in each were separated by agarose electrophoresis (A). This gel was Southern blotted and hybridised with the -1700DOWN probe to examine the smear of fragments (B). The negative control (-) is mouse genomic DNA and the positive control (+) is sheep genomic DNA. Both controls were digested with ***Bam*HI**. The two strong signals at the bottom of the positive control lane exposure were due to intense signals from the marker in the adjacent lane (not shown).




| | -1700UP | -1700DOWN | -900DOWN | +1700UP | +2700UP | +2700DOWN | +4700UP | +4700DOWN |
|---|---|---|---|---|---|---|---|---|
| Liver Titration | Fig. A1 | Fig. A3 | Fig. A5 | Fig. A7 | Fig. A9 | Fig. A11 | Fig. A13 | Fig. A15 |
| Mammary Titration | Fig. A2 | Fig. A4 | Fig. A6 | Fig. A8 | Fig. A10 | Fig. A12 | Fig. A14 | Fig. A16 |
| Composites | Figure 50 | Figure 51 | Figure 52 | Figure 53 | Figure 54 | Figure 55 | Figure 56 | Figure 57 |

**Table 7** Cuprous Phenanthroline Blots Spanning the Ovine BLG Gene

Titrations of liver and mammary nuclei with OP-Cu, and a comparison of both (composites), are indicated in the left-most column. The blots generated are listed as figures in columns to the right in relation to the probes used. Above the table, the probe positions are indicated in relation to the BLG gene (exons are blue; DNase I hypersensitive sites are red; the transcriptional start site is a green arrow). The region of the gene mapped by each probe is indicated by a purple arrow, e.g. for the -1700UP probe the arrow points upstream of the gene from position -1.7kb. The figure numbers for the complete mammary and liver titrations are prefixed with an 'A' to represent Appendix III where they are bound.

### 4.2.3 OP-Cu Map of the Ovine BLG Gene

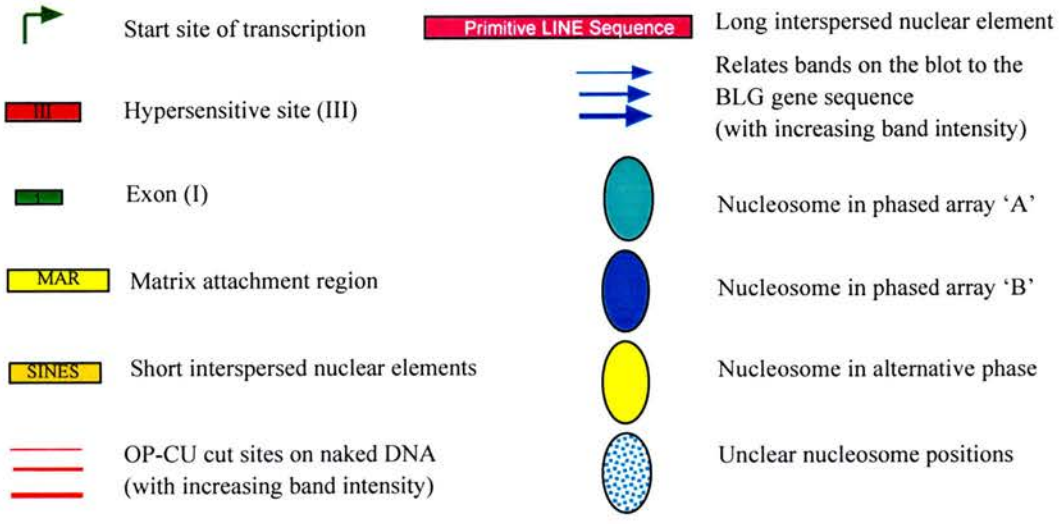
Two sets of data have been generated. The first set contains the full time course experiments for both mammary and liver chromatin which are bound in Appendix III (figures A1 to A16). These include all time points taken: at 0, 10, 20, 30, 40 and 60 minutes (time points 0, 1, 2, 3, 4 and 5 respectively). The second set are representative of the first and directly compare liver and mammary OP-Cu digestion sites on the same gel over the same region of the BLG gene. I have loaded two time points on the composite gels from each titration for both liver and mammary tissues. Unless otherwise stated, these are the points taken at 20 minutes (2) and 30 minutes (3). With each experiment there is a small discussion of the data generated.

All of the following experiments are explained diagrammatically in a standard format: (A) shows an ethidium-stained gel of the DNA before transfer to nylon membrane and indirect end-labelling. The resulting exposure is shown in (B) and the positions of the bands in relation to the BLG gene in (C). The mapping direction is always from the bottom of the gel, indicated by a red arrow. (D) shows nucleosome positions on the gene which would accommodate the banding data. As banding patterns differ over the promoter region in liver and mammary, the two figures spanning this region have (D) as liver-specific bands and (E) as mammary-specific bands. The legend to these diagrams is below.

The gels have all been loaded in the same order. The outside lanes (1 and 10) contain sizing markers together with negative control DNA. Owing to the large amount of genomic DNA loaded in the test lanes, they electrophorese at a slower rate in relation to the small amounts of plasmid DNA in the marker lane. Negative control DNA was added to overcome this problem which would otherwise lead to erroneous band sizing. Lanes 2 and 3 contain the OP-Cu / restriction enzyme digested time points from mammary nuclei. Lane 4 contains mammary time point 2, but only OP-Cu digested to show the nucleosome repeat. Lane 5 is the naked DNA control to test for OP-Cu sequence specific digestion. Lane 6 is as lane 4 for liver nuclei. Lanes 7 and 8 are as 2 and 3, for liver nuclei. Lane 9 is the positive control. All band sizes and their positions

on BLG are summarised in table 8. The complete nucleosome map generated from the following experiments can be seen at the end of this chapter as a foldout (Figure 58)

It uses the same legend as below and is in the same format as the following experiments.




**Legend to Figures**

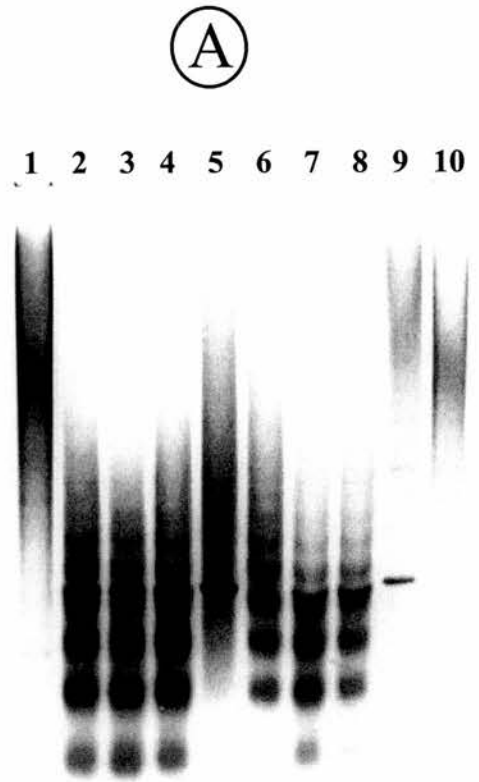
| -1700UP | | -1700DOWN | | -900DOWN | | +1700UP | |
|---|---|---|---|---|---|---|---|
| Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene |
| 170 | -1870 | 180 | -1520 | 260 | -605 | 250 | +1410 |
| 250 | -1950 | 300 | -1400 | 430 | -435 | 300 | +1360 |
| 350 | -2050 | 360 | -1340 | 610 | -255 | 360 | +1300 |
| 430 | -2130 | 480 | -1220 | 790 | -75 | 400 | +1260 |
| 540 | -2240 | 540 | -1160 | 970 | +105 | 470 | +1190 |
| 620 | -2320 | 720 | -980 | 1150 | +285 | 560 | +1100 |
| 710 | -2410 | 900 | -800 | 1330 | +465 | 640 | +1020 |
| 800 | -2500 | 1080 | -620 | | | 740 | +920 |
| 900 | -2600 | 1260 | -440 | Mammary specific digestion sites | | 1020 | +640 |
| 980 | -2680 | 1440 | -260 | | | 1200 | +460 |
| 1080 | -2780 | 1620 | -80 | 380 | -485 | 1380 | +280 |
| 1160 | -2860 | 1800 | +100 | 545 | -320 | | |
| 1260 | -2960 | 1980 | +280 | | | | |
| 1340 | -3040 | 2160 | +460 | | | | |
| 1520 | -3220 | 2340 | +640 | | | | |
| 1700 | -3400 | Mammary specific digestion sites | | | | | |
| 1950 | -3650 | | | | | | |
| 2130 | -3830 | 1220 | -480 | | | | |
| 2310 | -4010 | 1400 | -300 | | | | |

| +2700UP | | +2700DOWN | | +4700UP | | +4700DOWN | |
|---|---|---|---|---|---|---|---|
| Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene | Fragment size | Position on gene |
| 300 | +2440 | 280 | +3020 | 195 | +4560 | 180 | +4935 |
| 380 | +2360 | 380 | +3120 | 285 | +4470 | 260 | +5015 |
| 470 | +2270 | 460 | +3200 | 375 | +4380 | 360 | +5115 |
| 560 | +2180 | 640 | +3380 | 450 | +4305 | 450 | +5205 |
| 650 | +2090 | 830 | +3570 | 540 | +4215 | 520 | +5275 |
| 740 | +2000 | 1010 | +3750 | 620 | +4135 | 700 | +5455 |
| 830 | +1910 | 1190 | +3930 | 660 | +4095 | | |
| 920 | +1820 | 1370 | +4110 | 820 | +3935 | | |
| 1000 | +1740 | 1550 | +4290 | 1000 | +3755 | | |
| 1090 | +1650 | 1730 | +4470 | 1180 | +3575 | | |
| 1280 | +1460 | | | 1360 | +3395 | | |
| 1400 | +1260 | | | 1540 | +3215 | | |
| 1930 | +810 | | | 1720 | +3035 | | |
| 2100 | +640 | | | | | | |
| 2280 | +460 | | | | | | |
| 2460 | +280 | | | | | | |
| 2640 | +100 | | | | | | |

**Table 8** A summary of the OP-Cu digestion sites throughout the sheep BLG gene

**Figure 50 Nucleosome Mapping from -1.7kb to Upstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 4
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with OP-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.

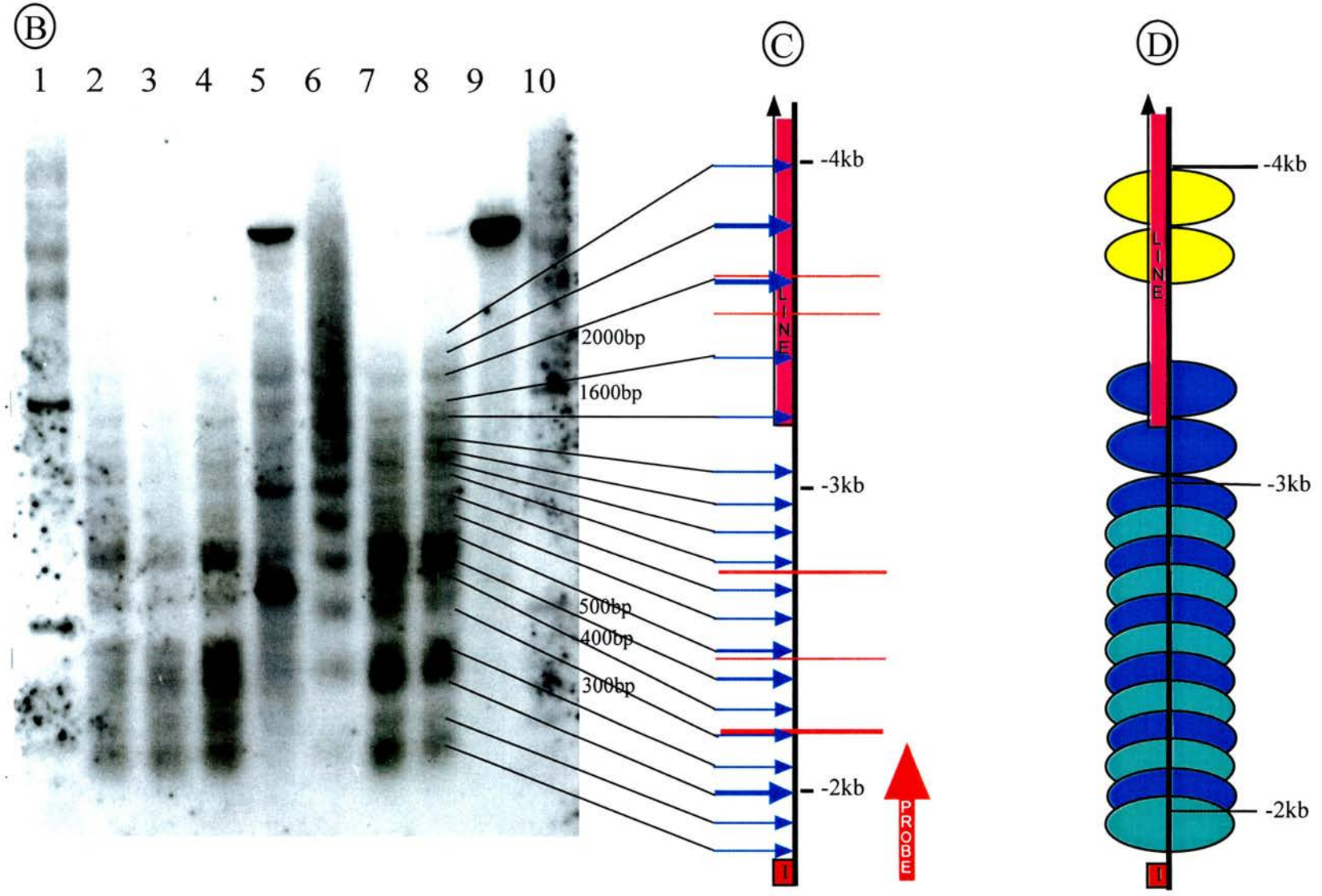
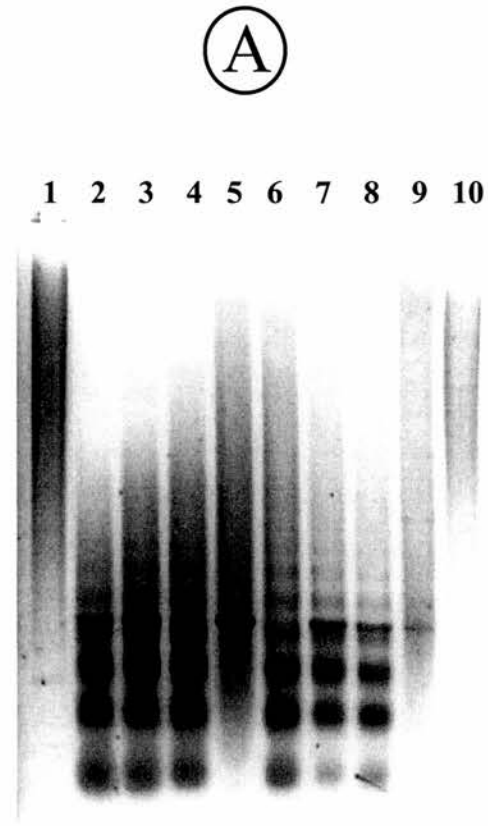

B
1 2 3 4 5 6 7 8 9 10
C
-4kb
LINE
2000bp
1600bp
-3kb
500bp
400bp
300bp
-2kb
PROBE
I
D
-4kb
LINE
-3kb
-2kb
I

**Figure 51 Nucleosome Mapping from -1.7kb to Downstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver chromatin
(E) Diagram of nucleosome positions on the gene in mammary chromatin




**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5µg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with OP-Cu only time point 2
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with *OP*-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.

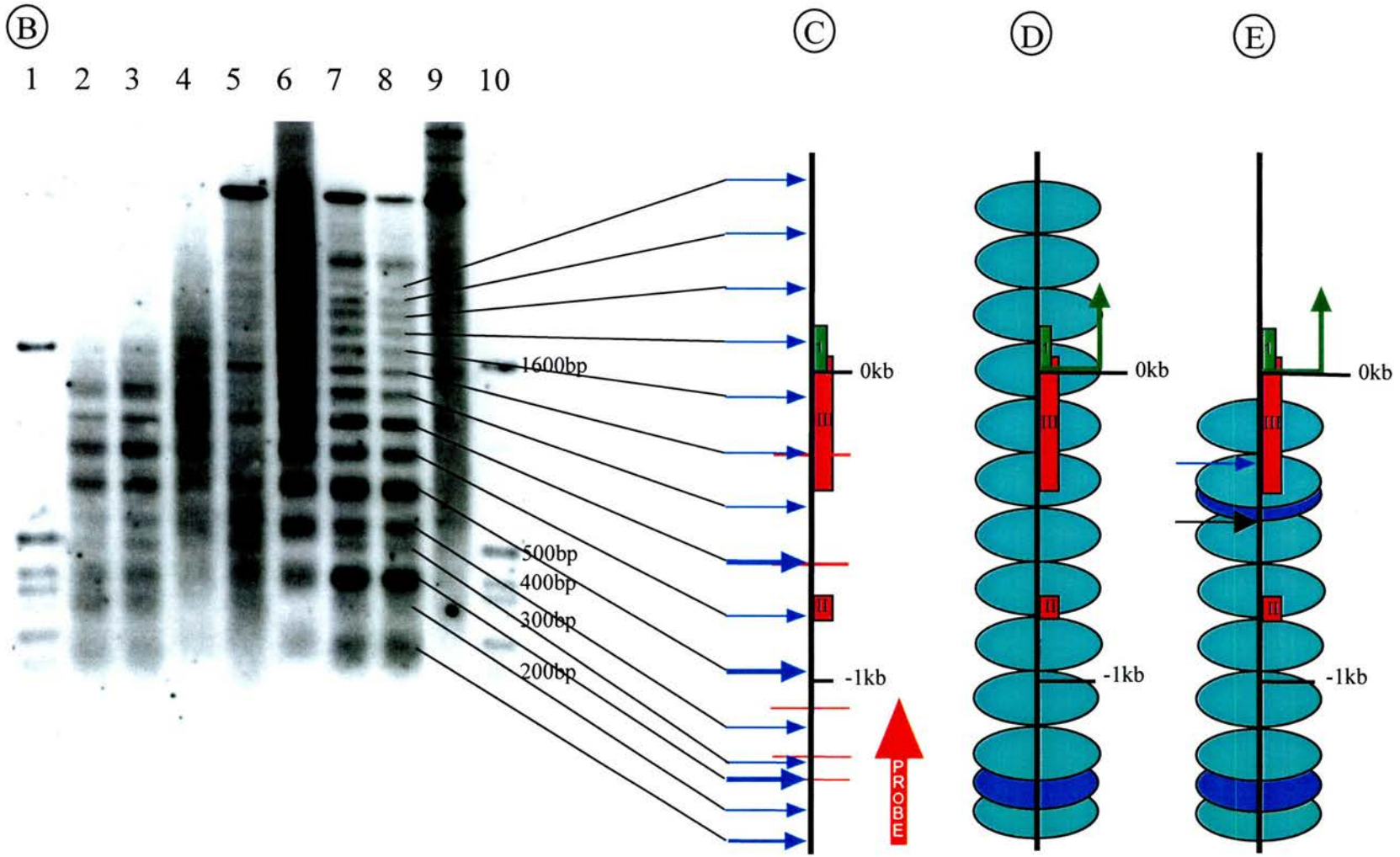
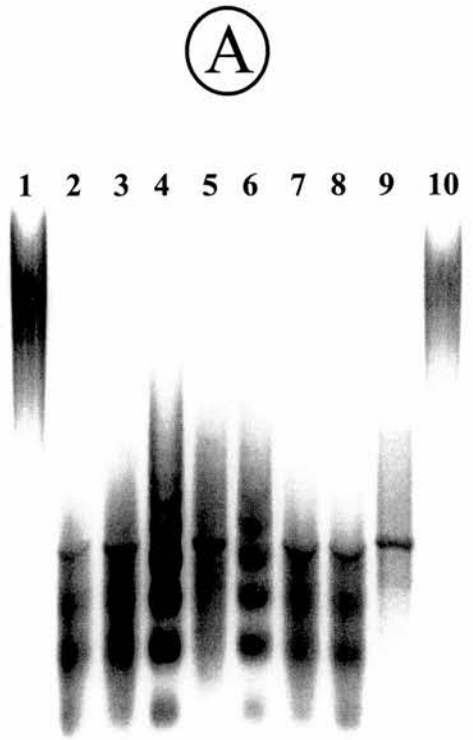

B
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
400bp
300bp
200bp
C
0kb
-1kb
PROBE
D
0kb
-1kb
E
0kb
-1kb

**Figure 52 Nucleosome Mapping from -0.9kb to Downstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver chromatin
(E) Diagram of nucleosome positions on the gene in mammary chromatin




**Lanes:**

1. Mouse genomic DNA digested with ***Rsa*I** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Rsa*I** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Rsa*I** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with ***Rsa*I** plus OP-Cu; time point 4
5. Sheep naked genomic DNA positive control digested with ***Rsa*I** plus OP-Cu
6. Sheep liver nuclei digested with *OP*-Cu only time point 2
7. Sheep liver nuclei digested with ***Rsa*I** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Rsa*I** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control Time point 0 digested with ***Rsa*I** only
10. As lane 1.

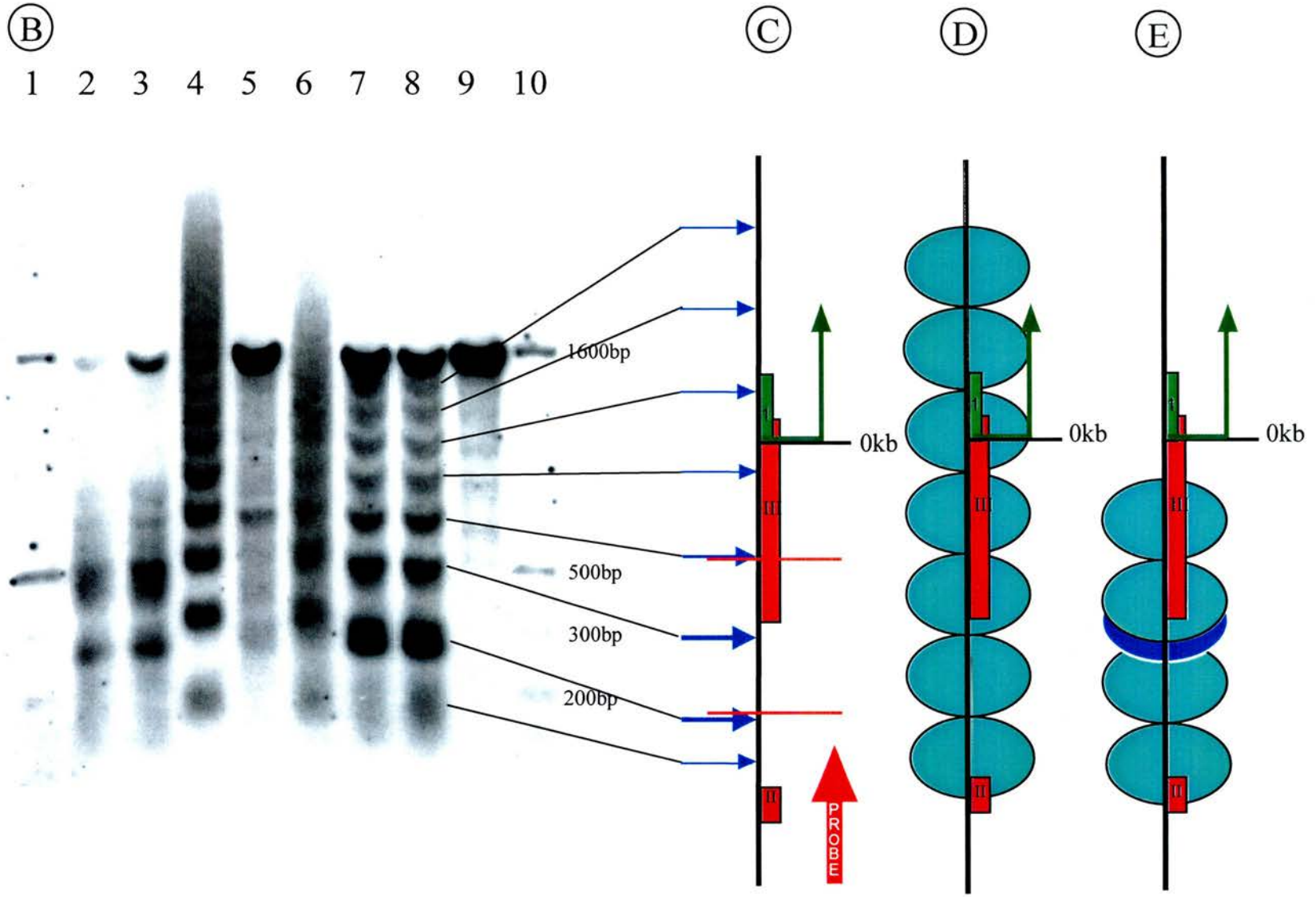
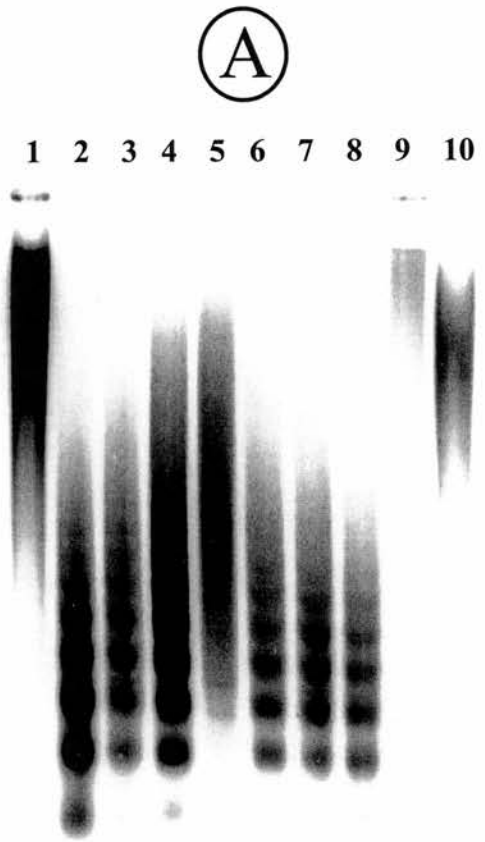

B
C
D
E
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
300bp
200bp
0kb
0kb
0kb
I
II
III
PROBE

**Figure 53 Nucleosome Mapping from +1.7kb to Upstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Mouse genomic DNA digested with ***Pvu*II** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Pvu*II** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Pvu*II** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with OP-Cu only time point 2
5. Sheep naked genomic DNA positive control digested with ***Pvu*II** plus OP-Cu
6. Sheep liver nuclei digested with *OP*-Cu only time point 2
7. Sheep liver nuclei digested with ***Pvu*II** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Pvu*II** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Pvu*II** only
10. As lane 1.

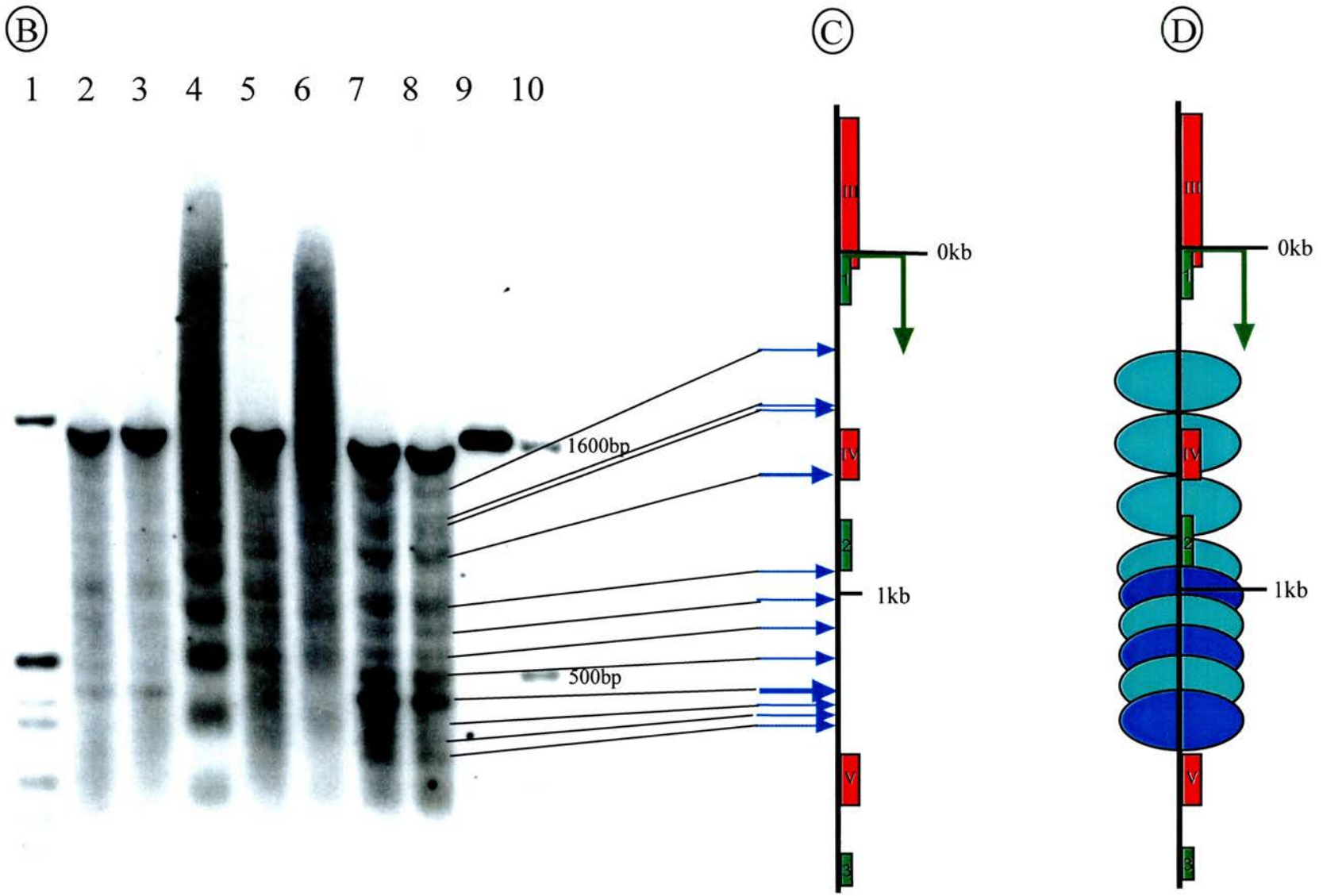
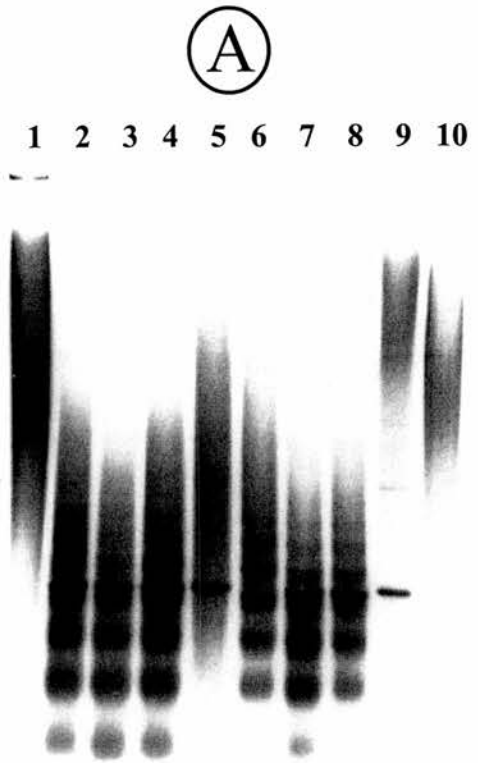

B
C
D
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
0kb
1kb
III
I
IV
II
V
0kb
1kb

**Figure 54 Nucleosome Mapping from +2.7kb to Upstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 4
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with OP-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.

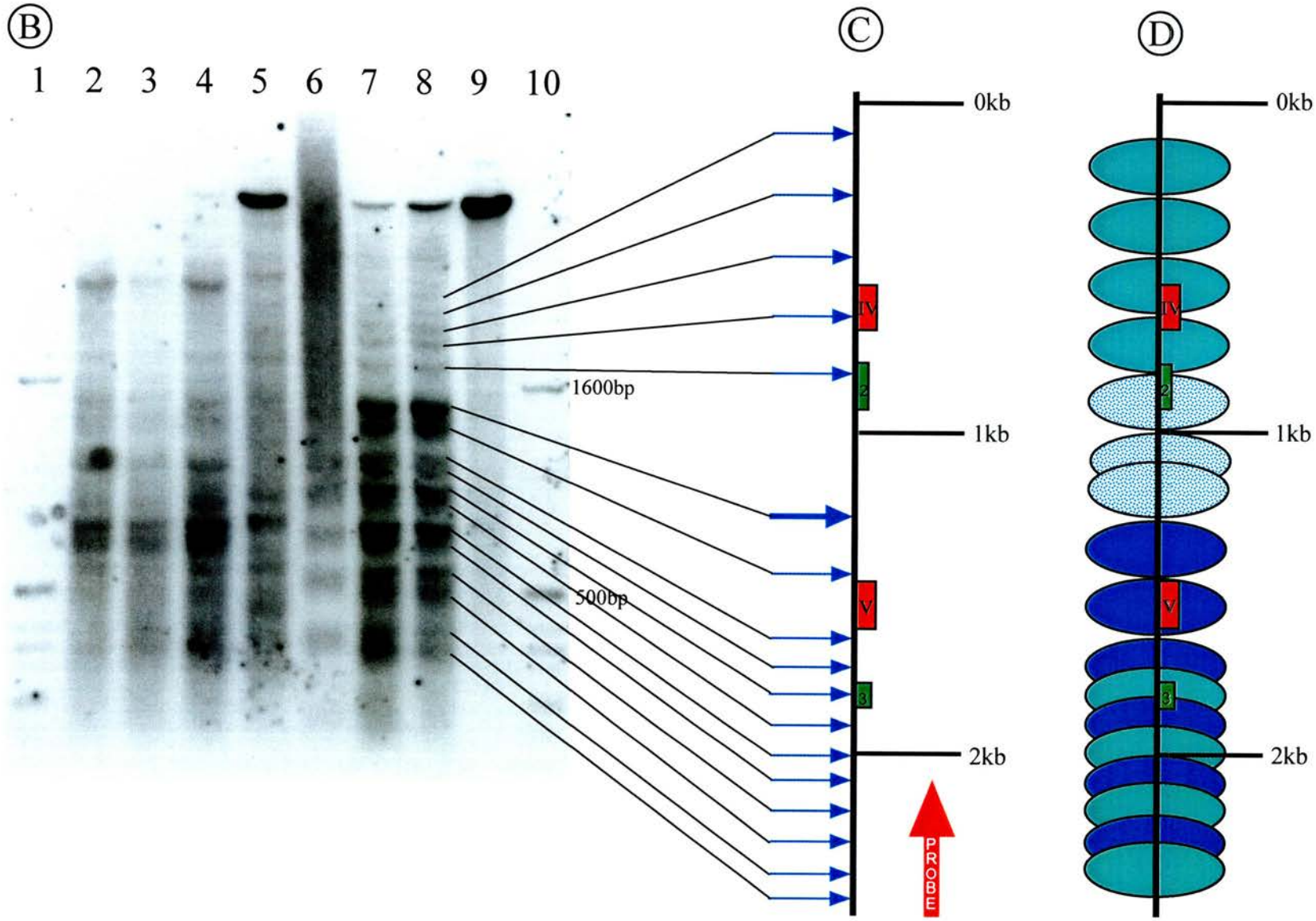

B
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
C
0kb
1kb
2kb
IV
2
V
3
PROBE
D
0kb
1kb
2kb
IV
2
V
3

**Figure 55 Nucleosome Mapping from +2.7kb to Downstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene


**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with OP-Cu only time point 2
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with OP-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.

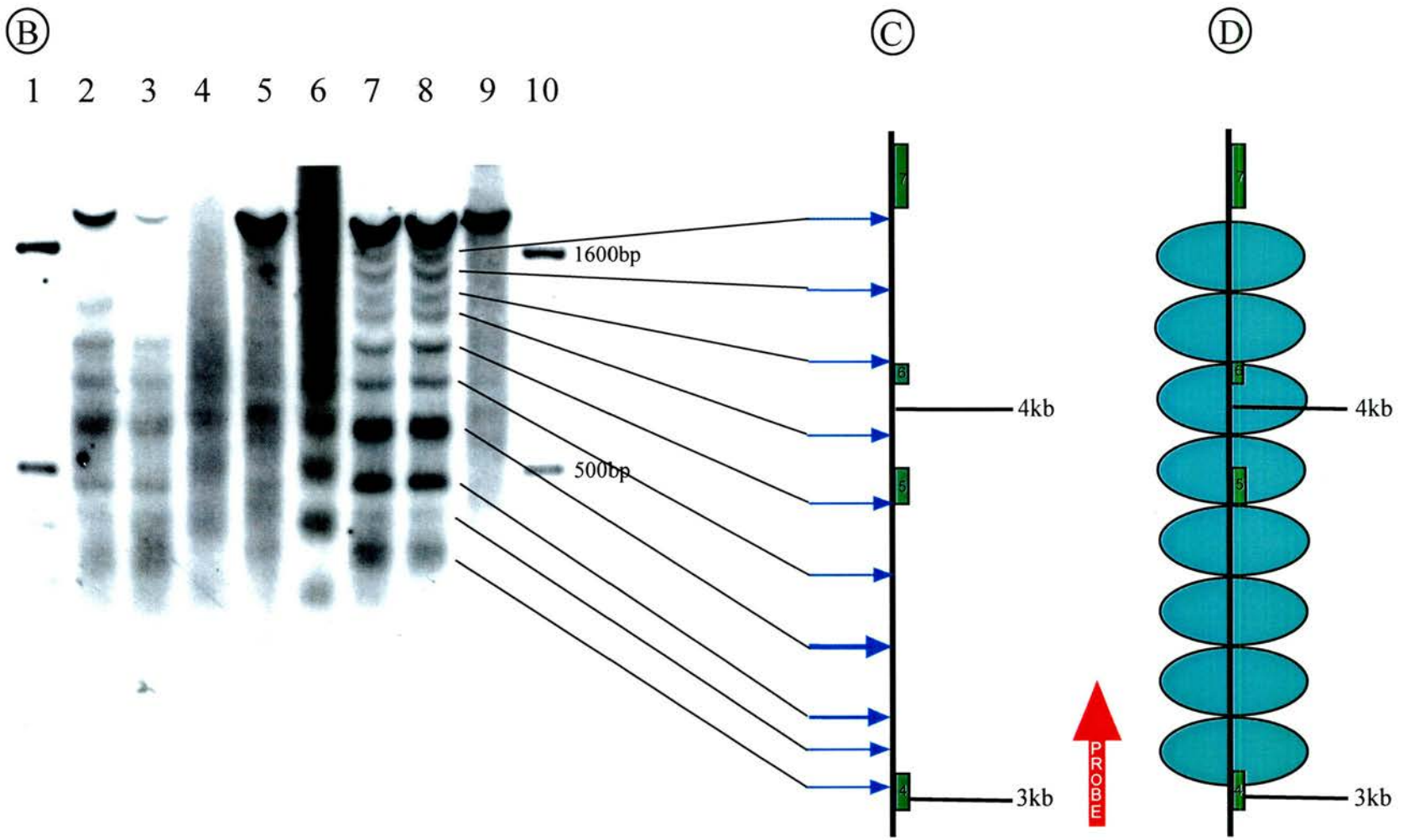
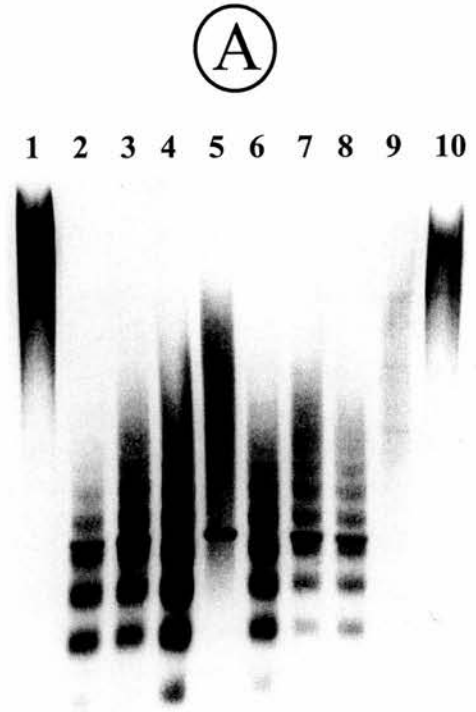

B
C
D
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
4kb
3kb
PROBE
4kb
3kb

**Figure 56 Nucleosome Mapping from +4.7kb to Upstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with OP-Cu only time point 2
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with OP-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.

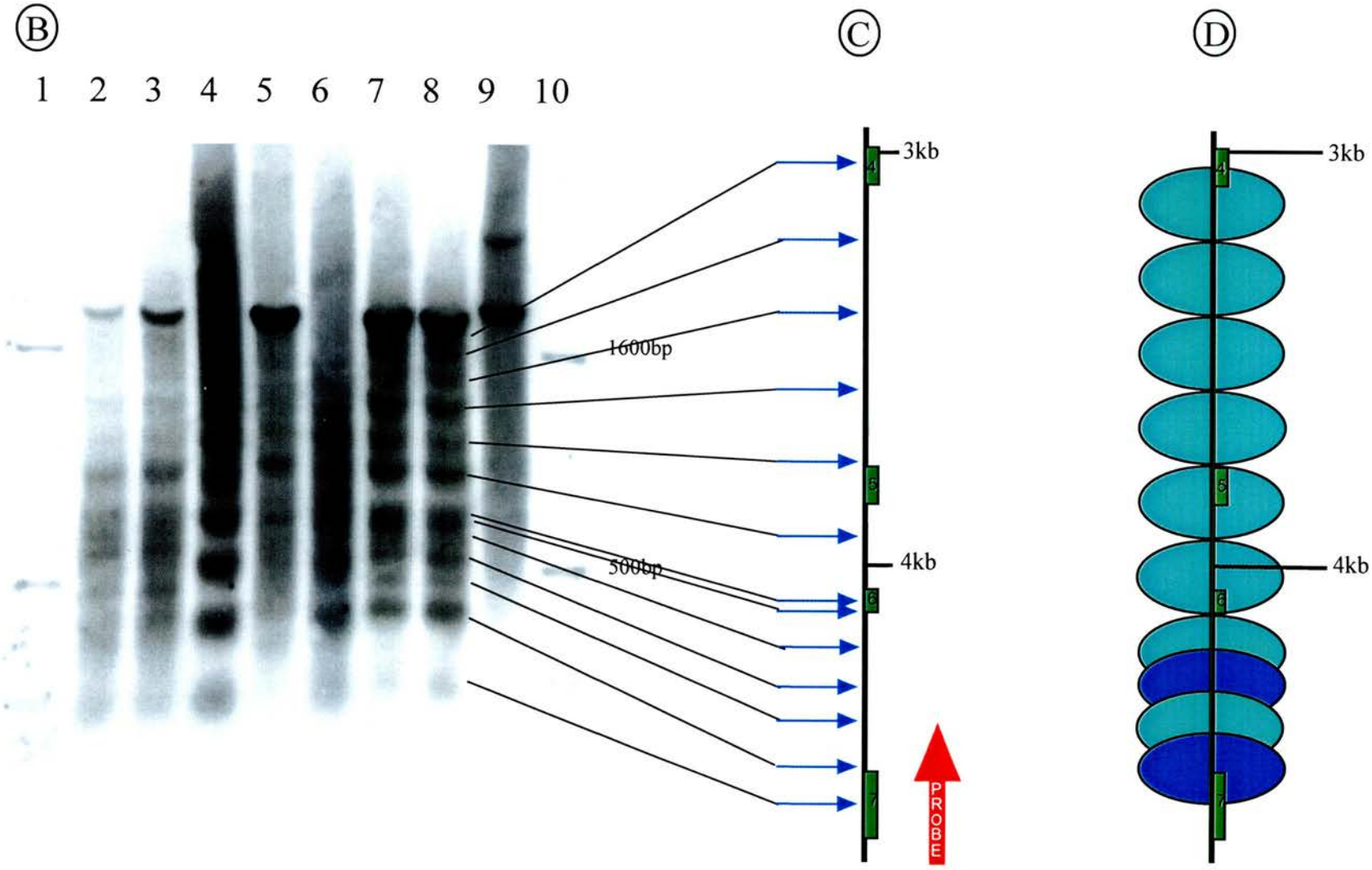
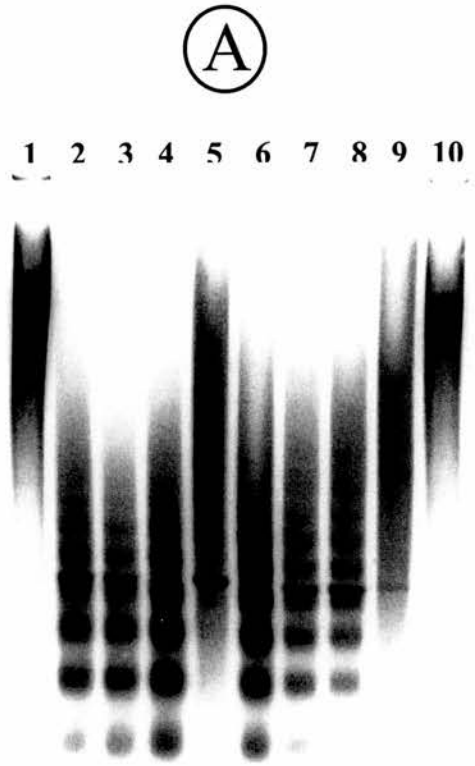

B
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
C
3kb
4kb
PROBE
D
3kb
4kb

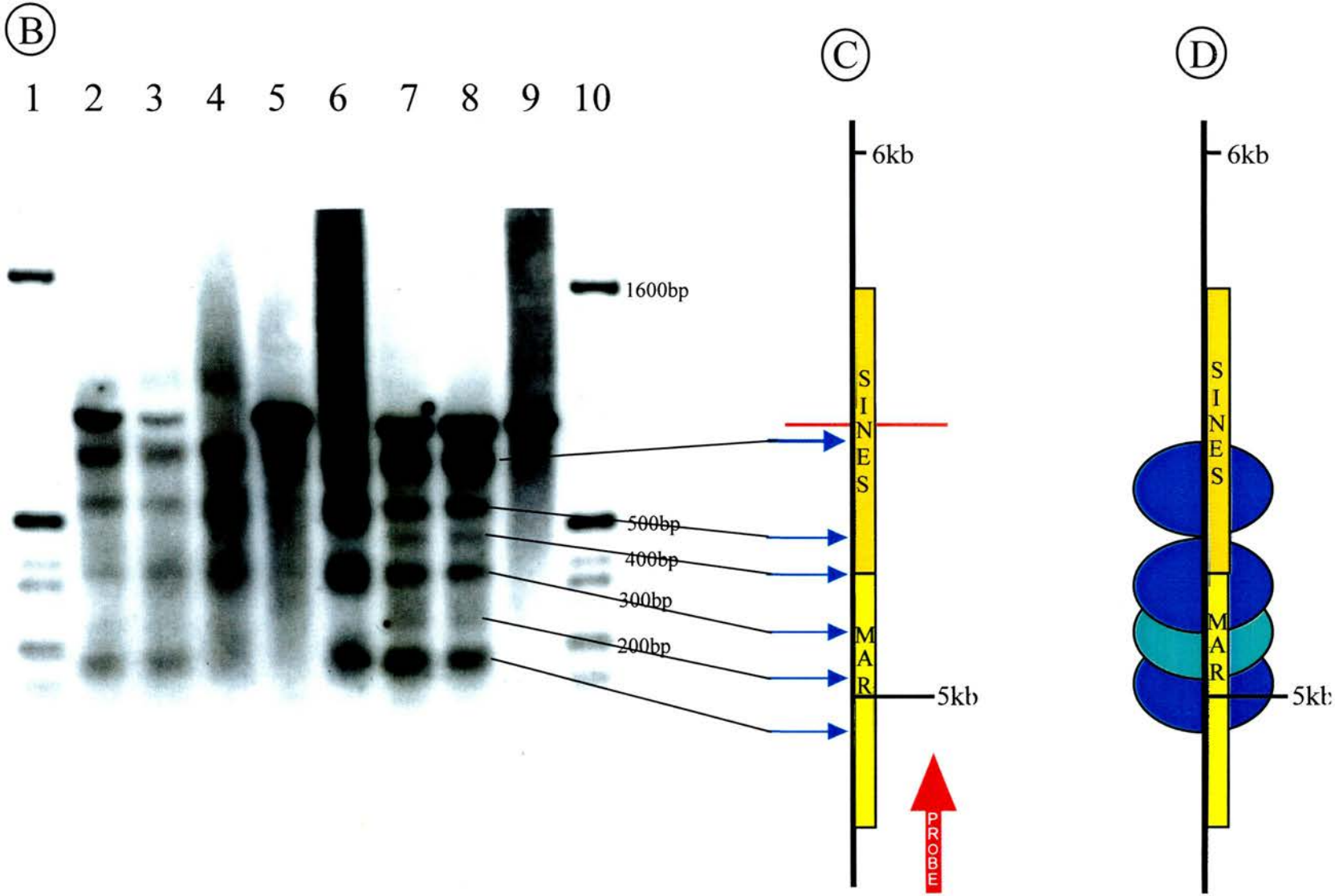
**Figure 57 Nucleosome Mapping from +4.7kb to Downstream of the Ovine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene




**Lanes:**

1. Mouse genomic DNA digested with ***Bam*HI** plus 5μg of 1kb ladder (Boehringer Mannheim)
2. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
3. Sheep mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
4. Sheep mammary nuclei digested with OP-Cu only time point 2
5. Sheep naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
6. Sheep liver nuclei digested with OP-Cu only time point 2
7. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
8. Sheep liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
9. Sheep genomic DNA positive control digested with ***Bam*HI** only
10. As lane 1.


B
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
400bp
300bp
200bp
C
6kb
SINES
MAR
5kb
PROBE
D
6kb
SINES
MAR
5kb

**Composite from -1.7kb to Upstream of BLG (Figure 50)**

This experiment maps the gene from -1700bp to -4814bp, generating a genomic fragment of 3.1kb in size. By accident, lane 4 is time point 4 (40mins) digested with OP-Cu and ***Bam*HI** rather than time point 2 only digested with OP-Cu.

From -2kb to -3kb of the gene there appear to be two alternative nucleosome phases which I have denoted 'A' for light blue and 'B' for dark blue. Phase 'B' continues just inside the long interspersed nuclear element (LINE) sequence before there is a gap of approximately 250bp from -3.4kb to -3.65kb. This is more clearly demonstrated in the full time course in liver (fig.A2). Here you can see the gap before the two prominent bands at the top of the gel. There are two very faint bands in this gap which may indicate binding of non-nucleosomal proteins. Moreover, in fig.A2 you can also see a large uncut region further upstream of the last mapped band at -4010bp.

The region from -3.2kb to -4.2kb has weak homology to reverse transcriptase and hence may be a LINE. If this is the case, further LINE sequences will stretch further upstream out-with the mapped region. It is feasible that LINE-specific factors are binding within these undigested regions or that the region has been folded in such a nature so as to exclude OP-Cu digestion.

One very strong OP-Cu sequence specific digestion site exists at -2280bp, next to DNase hypersensitive site I, and another slightly weaker at -2750bp. Neither of these lie in the linker region between nucleosomes. Unlike MNase, this sequence specificity is much less frequent and is clearly defined.

**Composite from -1.7kb to Downstream of BLG (Figure 51)**

The genomic band at the top of the picture is 4.4kb in size. It encompasses the region mapped from -1700bp to +2740bp.

The alternating array of nucleosomes ends at -1340bp with a strong digestion site. Similar strongly digesting sites are present at -980bp, -620bp and at -260bp, enclosing two nucleosomes each. This might indicate alternative nucleosome positions a very short distance apart, so that two bands fuse to generate one thick, more intense band. Alternatively, they may be indicative of a higher order structural formation (see

1.4). Interestingly, the naked DNA control parallels these highly digesting sites. Rather than solely indicating sequence preference of OP-Cu, these regions may have a more open conformation or may be structurally modified which allows greater access.

In the liver, a regular array of 'A' phased nucleosomes extends over the promoter into the coding region. However, this does not happen in the mammary where the regular phasing stops at -620bp and is replaced with several digestion sites in close proximity. This breakdown in the array is most likely due to the binding of transcription factors over the promoter, which are necessary for transcription to occur. Although they may not remove the nucleosomes, they may alter the accessibility of OP-Cu to nucleosome-bound DNA, making it possible to digest DNA further into the nucleosomes.

After -180bp no specific digestion sites are generated in the mammary. While the gene is actively transcribing during digestion, nucleosomes are constantly shuffled from in front of to behind the polymerase complex (Studitsky *et al.*, 1994; Studitsky *et al.*, 1995). This continual movement will generate a random digestion pattern, evident as a smear (the full titration in fig.A4 also shows an absence of specific digestion sites).

**Composite from -0.9kb to Downstream of BLG (Figure 52)**

The genomic band at the top of the picture is 1.75kb in size from -865bp to +882bp.

This experiment maps the promoter region at closer range than the last experiment and provides evidence of the same nucleosomal array seen in figure 51, probing in a downstream direction from –1.7kb. Moreover, this correlates well with the other experiments, in which the nucleosome phasing is in line with this.

**Composite from +1.7kb to Upstream of BLG (Figure 53)**

The region mapped spans from +1663bp to +38bp, generating a genomic band at the top of the picture of 1.6kb in size.

After +640bp at DNase I hypersensitive site II, the array returns to its previous state, alternating between 'A' and 'B' phases. This is exactly in phase with the pattern

which occurred upstream between -1160bp and -3040bp, thirteen nucleosomes away. However, the pattern appears to be slightly more complex now. There is no cut site at +820bp which would digest on the upstream side of the last nucleosome in the 'A' array before it enters the alternating array state. Moreover, additional digestion sits are present at +1410bp and +1300bp which do not fit with the nucleosome model I have portrayed. There is no obvious nucleosomal model which could account for these additional digestion sites while all eight histone proteins are bound to form a nucleosome. Additional factors or a more relaxed DNA conformation around the nucleosome might produce these sites.

Hypersensitive sites IV and V are only present in the sheep mammary gland before day 110 of pregnancy. They are present in the virgin tissue but not at any stage in the liver. Therefore, they cannot directly account for this complex array formation. However, this formation may be a consequence of them if they set the array initially.

Like the previous experiment, there are no obvious specific digestion sites in the mammary chromatin. The very weak digestion sites which can be seen could be due to non-expressing nuclei from connective tissue which is present at low concentration throughout the mammary gland (chapter 3).

**Composite from +2.7kb to Upstream of BLG (Figure 54)**

This experiment maps up to and over the previous region from +2740bp to -1700bp and has interesting similarities and differences. The genomic band at the top of the picture is 4.4kb in size. Both maps are identical except over the region between +820bp and +1360bp. An extra 'A'-phase nucleosome exists between +1360bp and +1190bp in the previous map which is not obvious here. However, this could be obscured by the intensely digested site at +1260bp. This site appears to demarcate the region of altered nucleosome phasing up to +820bp, when the regular phasing re-appears.

In the mammary gland all bands are again very weak, with the possible exception of two at +2000bp and +2090bp. The reason for this stronger digestion is unclear, but

could possibly be due to higher order chromatin formation, alternative linker histones or polymerase pausing. Near the top of the gel another strongly digesting site is present at about –100. This correlates with the multiple cut sites specific to the mammary chromatin within HSIII (see figures 51 and 52). Although not illustrated in the diagram, the 'A' phase in the liver can be seen to extend for at least another three nucleosomes past +100bp, again correlating with previous data.

**Composite from +2.7kb to Downstream of BLG (Figure 55)**

The genomic band at the top of the picture is 2kb in size which maps from +2740bp to +4750bp.

As the map extends outwards in an upstream direction from the ***Bam*HI** site at +2.7kb, there is a gap in map between +2440bp and +3020bp. However, there is reason to believe that both 'A' and 'B' nucleosome phases continue up until +3120bp. In this map there is a band at +3120bp and possibly one at +2940bp (a 180bp fragment). These would roughly be in line with the 'B' array which extends to +2360bp in the last experiment (Figure 54). At +3200bp and +3380bp there are very intense digestion sites, possibly indicating the end of the 'B' array with a strongly defined 'A' phase nucleosome. This array continues on its own up to the limit of this experiment at +4470bp.

The mammary chromatin displays much the same intensity of banding pattern at this point as in the liver. Although less sample has been loaded, it is obvious that the same sites are digested in the mammary as in the liver and with similar relative intensities. No DNase I hypersensitive site has been detected over this region.

**Composite from +4.7kb to Upstream of the BLG Gene (Figure 56)**

The genomic band at the top of the picture is 2kb in size which maps from +4750bp to +2740bp.

The 'A'-phase array digestion sites overlap and agree with the previous experiment. The 'B' array re-appears at +4215bp, but is not obvious in the previous experiment as it is at the upper extreme. Faint sub-bands may be present in the last

experiment at this position if viewed closely. This time the 'B' array does not match up precisely with the last time it was present at +2360bp (50bp difference). However, this was ten nucleosomes away; an increase of 6bp each to 186bp per nucleosome would account for this discrepancy indicating that there may be some interaction between these long separated arrays. Alternatively, this could be a separate array, quite different from the previous one. Interestingly, phase 'B' becomes dominant over phase 'A'.

Again, the mammary banding pattern is becoming increasingly stronger, especially after +3935bp where it is virtually identical to that in liver. It is unclear why the phasing should re-appear at this point, before the end of the transcription unit. One explanation could be that the polymerase complex slows as it approaches the end of the transcribed domain. This would mean that the nucleosomes were shifting less frequently than those further upstream and give rise to a more defined phased array.

**Composite from +4.7kb to Downstream of BLG (Figure 57)**

The genomic band at the top of the picture is 900bp in size from which maps the gene from +4750bp to +5650bp.

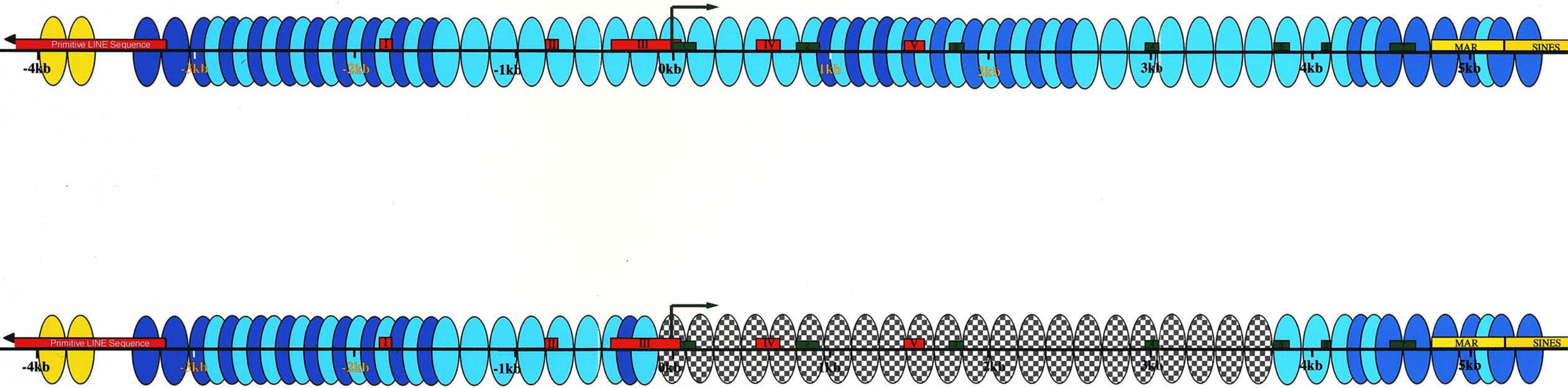
As the same ***Bam*HI** restriction enzyme site from the last experiment was used to map downstream in this experiment, a gap is present in the map from +4560bp to +4935bp. However, both 'A' and 'B' arrays persist in the same alignment from the last experiment and phase 'B' is still dominant. It is not clear whether array 'A' continues on its own further downstream out-with the mapping region, or if 'B' is also persists. There is a strong band beneath the genomic band which may be a doublet, indicating that array 'B' is also present. Alternatively, it could indicate the termination of the 'A' phase. Previous maps have shown that when two arrays overlap and one array ends, the other, persistent array has much stronger digestion sites over the first nucleosome.

Although this map extends across a matrix attachment region (within +4750bp to +5650bp) and SINEs (from +5202bp to +5732bp), the chromatin structure does not appear to be radically modified or remodelled.

**Figure 58** Nucleosome Structure over the Ovine BLG Gene (Foldout)

The inactive gene in liver chromatin (top) and active gene in mammary chromatin (bottom) are represented using the same legend as before. Indeterminate nucleosome positions over the transcribed region in the active gene are represented by stippled ovals. Note the alternative array structure over the proximal promoter region in the active gene.

Nucleosomal Structure Over the Ovine BLG Gene



## 4.3 Comparison with MNase Data

In the previous chapter, selected regions of the BLG gene were mapped using MNase as a probe for the nucleosome alignment. In particular, the regions mapped were probed with the +2700DOWN probe in mammary and liver chromatin, the +4700DOWN probe in mammary chromatin and the +4700UP probe in liver chromatin. Although I found that MNase had severe limitations as a probe for chromatin structure, it would be interesting to compare the maps generated with MNase and OP-Cu.

Therefore, the diagrammatic representation of all three regions are compared in figure 59 with the +2700DOWN probe, figure 60 with the +4700DOWN probe and figure 61 with the +4700UP probe.

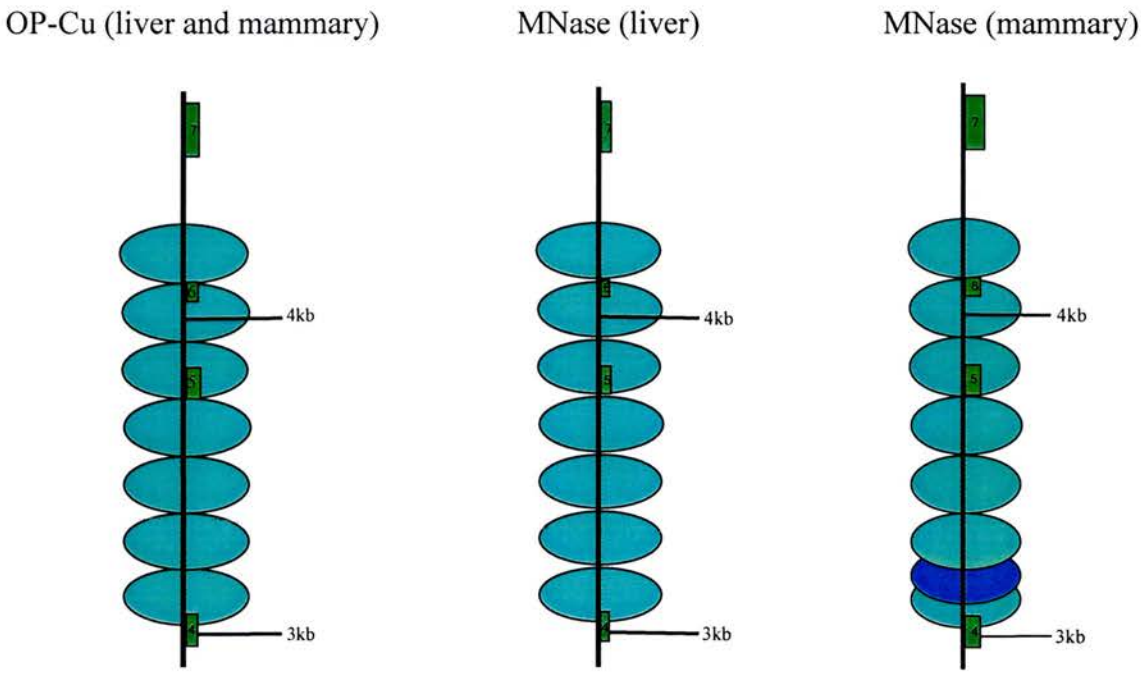
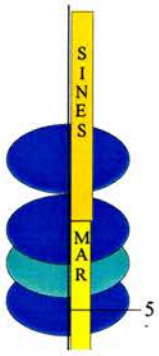



**Figure 59** Comparison Between MNase and OP-Cu with the +2700DOWN Probe

OP-Cu (liver and mammary)




MNase (mammary)

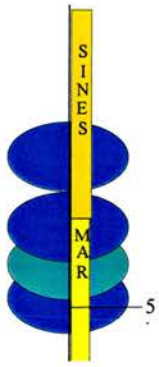
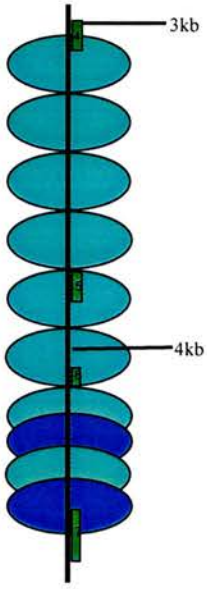



**Figure 60** Comparison Between MNase and OP-Cu with the +4700DOWN Probe

OP-Cu (liver and mammary)




MNase (liver)

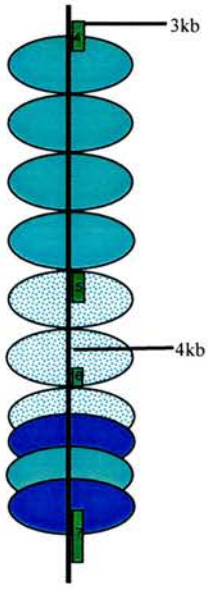



**Figure 61** Comparison Between MNase and OP-Cu with the +4700UP Probe

Both MNase and OP-Cu probes produce very similar maps of the nucleosome positions over the BLG gene. The nucleosome repeat length is the same in both, which indicates that digesting the chromatin either alters the nucleosomes equally, or more likely does not affect them at all so that the resultant map reflects a real phenomenon. Moreover, the nucleosome positions are the same in both, which strengthens this argument and shows

that although MNase possesses higher sequence specificity than OP-Cu, it still functions reasonably as a probe for nucleosomal positioning.

There are two noticeable differences between the maps. Firstly, although the nucleosome positions are the same in both cases, there is an extra alternatively phased nucleosome, proximal to the probe in figure 59. This is only present in mammary chromatin when tested with MNase. The mammary lanes from both digests are compared directly in figure 62. From this direct comparison, it appears that a single digestion site exists in the MNase digest where one does not in the OP-Cu digest (indicated by a red arrow). This represents the 3′ side of the nucleosome. A digestion site exists in on the 5′ side in both MNase and OP-Cu digests. Therefore, the only difference in this respect is a single digestion site. This may be caused by sequence-specific digestion. It is also interesting to note that relative intensity of the bands is much more even in the OP-Cu digestion than in the MNase digestion. This probably reflects on the relative lack of sequence specificity of OP-Cu. The banding pattern is also much more tightly defined with OP-Cu, making interpretation easier.

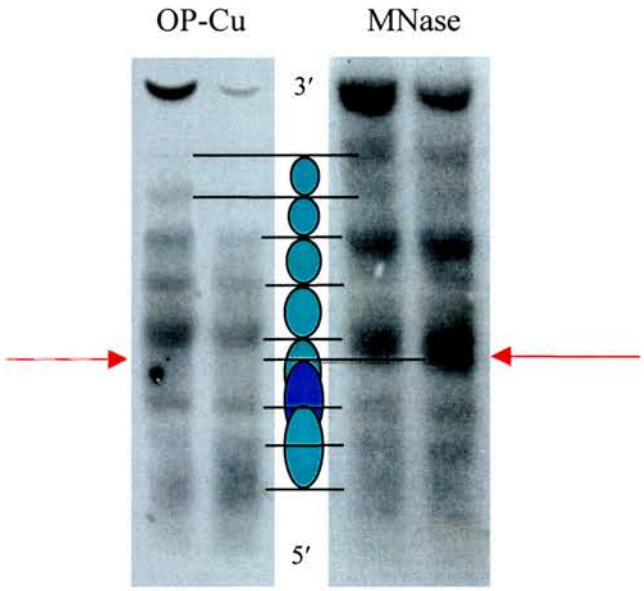



**Figure 62** A Direct Comparison of the Nucleosome Positions in OP-Cu and MNase Mammary Chromatin Digestion

The +2700DOWN probe was used in both cases to elucidate nucleosome positions from a ***Bam*HI** site at +2742bp in a downstream direction. Both images were taken from the original figures in chapters 3 & 4. Between images, ovals are drawn to indicate the positions of nucleosomes with respect to the bands. The red arrows indicate the discrepancy between results.

The second difference is that interpreting the data of the +4700UP probing of liver chromatin with MNase was difficult in the middle of the probed region (see stippled nucleosomes in figure 63 and figure 44 in chapter 3). When aligned with the OP-Cu digest of the same region, it is obvious that the same banding pattern is present in the MNase digestion. However, the bands are not so clearly defined.

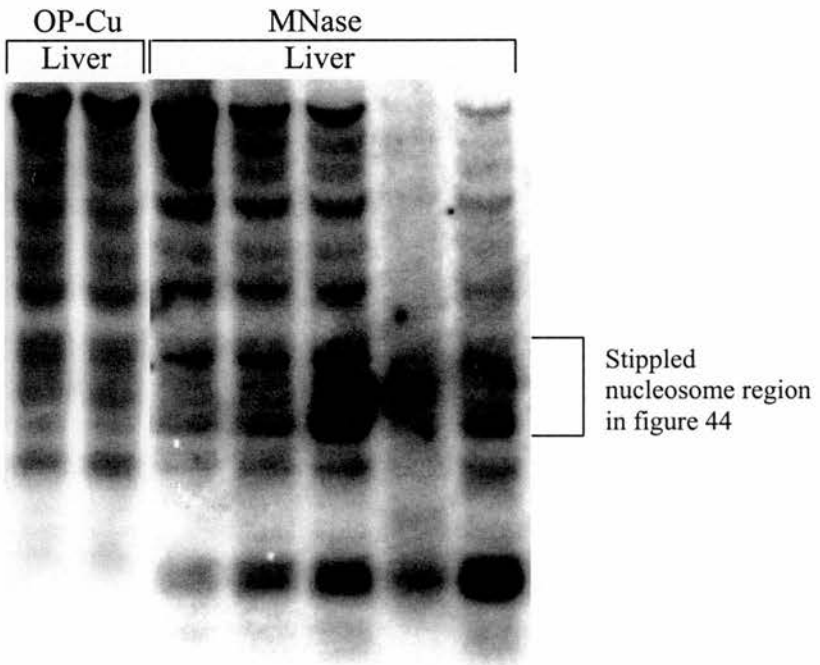



**Figure 63** A Direct Comparison of the Nucleosome Positions in OP-Cu and MNase Liver Chromatin Digestion

The +4700UP probe was used in both cases to elucidate nucleosome positions from a ***Bam*HI** site at +4755bp in an upstream direction. Both images were taken from the original figures in chapters 3 & 4. A box surrounds the region which was difficult to interpret in the MNase digest.

In summary, both probes generated almost identical maps of nucleosome positions over the BLG regions which have been compared. Only small discrepancies between the maps exist and in addition, most OP-Cu digests provided a clearer and more definitive map.

## 4.4 Comparison with Another Long-Range Nucleosome Map

As far as I am aware, this is the only published long-range nucleosome map of a pol II transcribed gene analysed from tissue. The closest work was performed by the Bonifer laboratory who study the chicken lysozyme gene (Huber *et al.*, 1996). They mapped approximately 7kb of 5' sequence upstream from the transcriptional start site in a variety of cell lines. These cell lines represent different stages of macrophage development, where the gene is expressed, and erythroid cell lines where it is not. Although they also employed an indirect end-labelling procedure to elucidate the nucleosome structure, their results are substantially different from mine.

Most prominently, they find that only a small number of sequences contain positioned nucleosomes. All of these are associated with DNase I hypersensitive site positions. The hypersensitive sites do not have to be present for this positioning to occur. Distinct positions are present in the inactive gene at the same regions as those in the active gene. It is difficult to tell whether the positions vary between the different cell types as results are aligned from different experiments to a marker lane from a different experiment.

The majority of sequences did not display nucleosome positioning in the chicken lysozyme gene. These are visible as a very light smear in the Southern blot data, completely lacking in any distinct banding pattern. However, nucleosomes may still adopt specific positions over these regions in macrophages, as cell lines frequently have a different nucleosome repeat length from the endogenous tissue (van Holde, 1988). Moreover, they employed MNase and do not control for sequence specific digestion. I have shown that this enzyme could generate a spurious banding pattern and also be refractive to DNA digestion in some linker regions. Using OP-Cu, I found distinct positions throughout the gene, with the only exception being in the active gene, where there were no strong positions between the transcriptional start site and exon five (figures 51, 52, 53, 54 & 56).

## 4.5 Summary

The ovine BLG gene comprises two distinct nucleosome phases which I have named 'A', represented in light blue, and 'B', represented as dark blue (see Legend). In a given population of nuclei and at specific locations on the gene they exist together, for instance between -2000 and -3000. Presumably, phase 'A' predominates in some nuclei and phase 'B' in others unless they oscillate between states. In addition, they can exist separately, for instance phase 'A' is only present between -1000 and +1000. Phase 'B' is only present by itself in small areas, for instance between -3040 and -3400.

The only point on the gene where the arrays differ is at about -4kb. Two nucleosomes (indicated yellow in figure 50) appear to exist in a separate array, divided from the rest of the gene by a 250bp stretch of undigested DNA (see figure A1 in Appendix III). This point may define the end of the gene domain. A DNase I general sensitivity assay could detect if this were the case.

The entire array is clearly defined in the liver where the gene is inactive. However, in the mammary where the gene is active, the pattern stops over the promoter region at HSIII with a unique pattern of digestion sites. There are no clearly defined digestion sites over the start of the transcribed region, but they gradually become more obvious towards the end, becoming equivalent to those in the liver chromatin around exon 5.

In BLG, the only distinct difference in chromatin pattern between the active and inactive gene is over the promoter region at HSIII. Obviously, this excludes the transcribed region where no chromatin pattern can be seen in the active gene. The promoter is particularly well mapped in figure 52 using the -0.9kb probe where *Rsa*I digestion sites at -865 and +882 delimit this area. In the next chapter I shall investigate this region using the *in vitro* monomer extension technique to find out the precise sites of sequence-dependent nucleosome positions. As the DNA sequence is the primary determinant of nucleosome positioning, this map should correlate with this *in vivo* map. In addition to corroborating the *in vivo* map, this will demonstrate whether some nucleosomes are more strongly positioned than others. A particularly strongly positioned nucleosome may be important in initially defining the array.

# Chapter 5

# Sequence-Dependent Nucleosome Positioning on the Ovine BLG Promoter *in vitro*

## 5.1 Introduction

The DNA sequence is thought to be a major determinant of nucleosome positioning (see **1.6.3**). For this reason, many groups have reconstituted histones on to short, end-labelled linear DNA fragments *in vitro* to examine this effect (Drew *et al.*, 1985; Hayes *et al.*, 1990; Roberts *et al.*, 1995; Fitzgerald *et al.*, 1998). These methods produce high-resolution nucleosome maps by digesting the unprotected DNA with MNase, DNase or exonuclease III. Analysis by polyacrylamide gel electrophoresis reveals an undigested footprint where a nucleosome sequesters the DNA.

Such methods are sufficient to map short regions of DNA, but do not lend themselves to mapping longer stretches of sequence. Recently, the monomer extension technique has been developed by Allan and co-workers which has been used to produce a long-range nucleosome positioning map spanning 8kb of the chicken β- and ε-globin gene locus (Yenidunya *et al.*, 1994; Davey *et al.*, 1995). This technique maps the precise positions adopted by core histone octamers reconstituted on to long DNA sequences. MNase is employed to nick the unprotected DNA and then progressively digest it until it reaches a nucleosome boundary. Large sequences of up to 2kb have been mapped in this way in a single experiment. I have employed this method to map the nucleosome positions *in vitro* over the ovine BLG promoter region.

A significant body of evidence suggests that nucleosomes can adopt well defined locations on the DNA sequence in both reconstituted systems and within the nucleus (Bock *et al.*, 1984; Thoma *et al.*, 1985; Drew *et al.*, 1987). Moreover, these locations are frequently the same in both instances (Adroer *et al.*, 1998). Therefore, the *in vitro* map of the BLG promoter should complement my *in vivo* data (chapters 3 & 4) by providing

a more accurate map of the nucleosome positions. Moreover, as the *in vitro* data is quantitative as well as qualitative, it indicates the relative positioning affinities of nucleosomes. A more strongly positioned nucleosome may have a more important role in either setting the local nucleosomal array and / or in modulating gene activity.

It should also be born in mind that the nucleosome repeat lengths can differ between tissues (van Holde, 1988) and gene activity state (De Ambrosis *et al.*, 1987) in the same organism. Therefore, although the DNA sequence appears to be a major determinant of nucleosome positioning, it cannot be the only one (Blank *et al.*, 1996; Liu *et al.*, 1997). Other factors such as: electrostatic interactions (Blank *et al.*, 1995), chromatin folding (Blank *et al.*, 1996), high mobility group proteins (Tremethick *et al.*, 1996), linker histones (Sun *et al.*, 1990; Jeong *et al.*, 1991), core histone variants (Leonardson *et al.*, 1989), non-histone protein complexes (Ito *et al.*, 1997) and even diet (Castro *et al.*, 1986) can modify nucleosome spacing. Hence, an *in vitro* map should only be interpreted in relation to an *in vivo* one.

## 5.2 The Monomer Extension Technique

The technique comprises five steps (Figure 65):

1. The region to be mapped is cloned in to pBluescript II KS (-) in both orientations to examine the nucleosome positioning on both strands.
2. Single stranded DNA is prepared from both constructs with the assistance of helper phage.
3. Core histones are reconstituted on to one of the dsDNA constructs to form nucleosomes. These must be at a concentration of approximately one nucleosome per 500bp of construct DNA in order to test solely for DNA sequence-directed positioning signals. These reconstitutes are then digested with MNase to release only nucleosome-protected 146bp fragments of DNA. These 'monomer' fragments are purified and $^{32}P$ end-labelled.
4. The monomer fragments of DNA are annealled to one of the ssDNA constructs and extended with the Klenow fragment of DNA polymerase to form dsDNA.

This extension reaction proceeds and eventually reaches a key restriction enzyme site, unique to the construct. While still ssDNA, this site cannot be utilised by the corresponding restriction enzyme which is present in the reaction mix. However, once the polymerase makes the site dsDNA, the DNA is cleaved. This site is used as a reference point in order to position the monomer fragments which had annealed to the ssDNA construct (Figure 65).

5. The labelled fragments are separated by polyacrylamide gel electrophoresis and visualised by phosphorimager and photographic film. The bands correspond to the distance from the restriction enzyme site to the far-end of the annealed monomer fragment (Figure 64 & Figure 65). These bands be sized by using the radiolabelled markers on the same gel and can be quantitated relative to one-another.

The positions of the nucleosomes can be calculated by simply subtracting the size of each band from the location of the restriction enzyme site used in the assay. This tells us the location of the terminal end of the nucleosome (Figure 64). To pinpoint the centre of the nucleosome 73nt must be added (half of one 146bp nucleosome width). It is accurate to within +/-1nt at the base of the gel, but may vary by up to 25nt at the top.

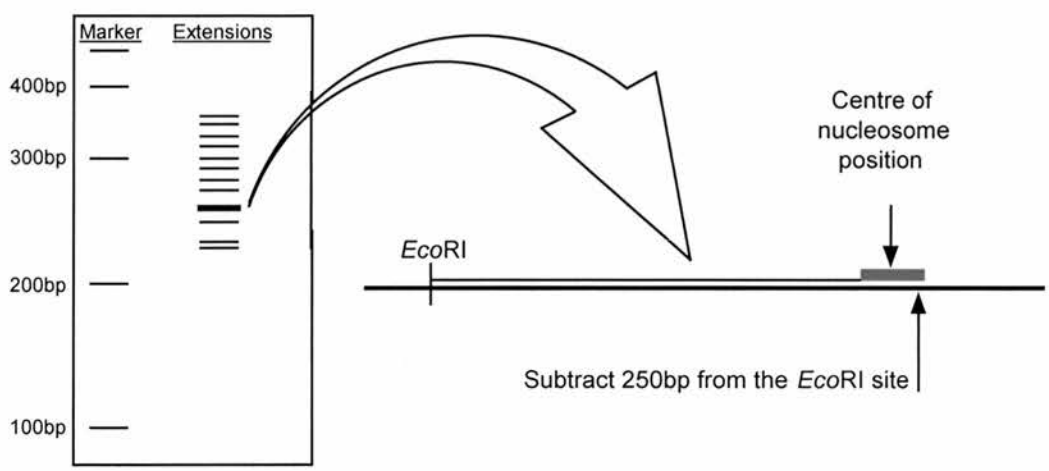



**Figure 64** Relating band size to nucleosome position

A strongly positioned nucleosome is evident in this diagram at 250bp. This distance upstream from the ***Eco*RI** site is the terminal end of the nucleosome on the DNA sequence. 73bp further in marks the centre of the nucleosome. The nucleosome-protected monomer DNA fragment isolated earlier is indicated as a red box.

As the Klenow fragment of DNA polymerase synthesises double-stranded DNA away from the 146bp monomer DNA fragment, it may reach DNA sequences which destabilise the polymerisation reaction. If the reaction terminates prematurely and consistently at a specific location before the restriction enzyme site, a band will appear on the phosphorimage of the resulting polyacrylamide gel. This band does not correspond to the position of a nucleosome, but reflects a sequence-specific termination reaction.

In order to control for this possibility, a control reaction is run in parallel which substitutes water for the restriction enzyme. In this extension reaction, any bands generated can be excluded from the test lanes which contain the restriction enzyme. Moreover, as the DNA polymerase reaction terminates prematurely in a sequence-independent manner, the smear of fragments generated can be subtracted from the test lanes to reduce the background.

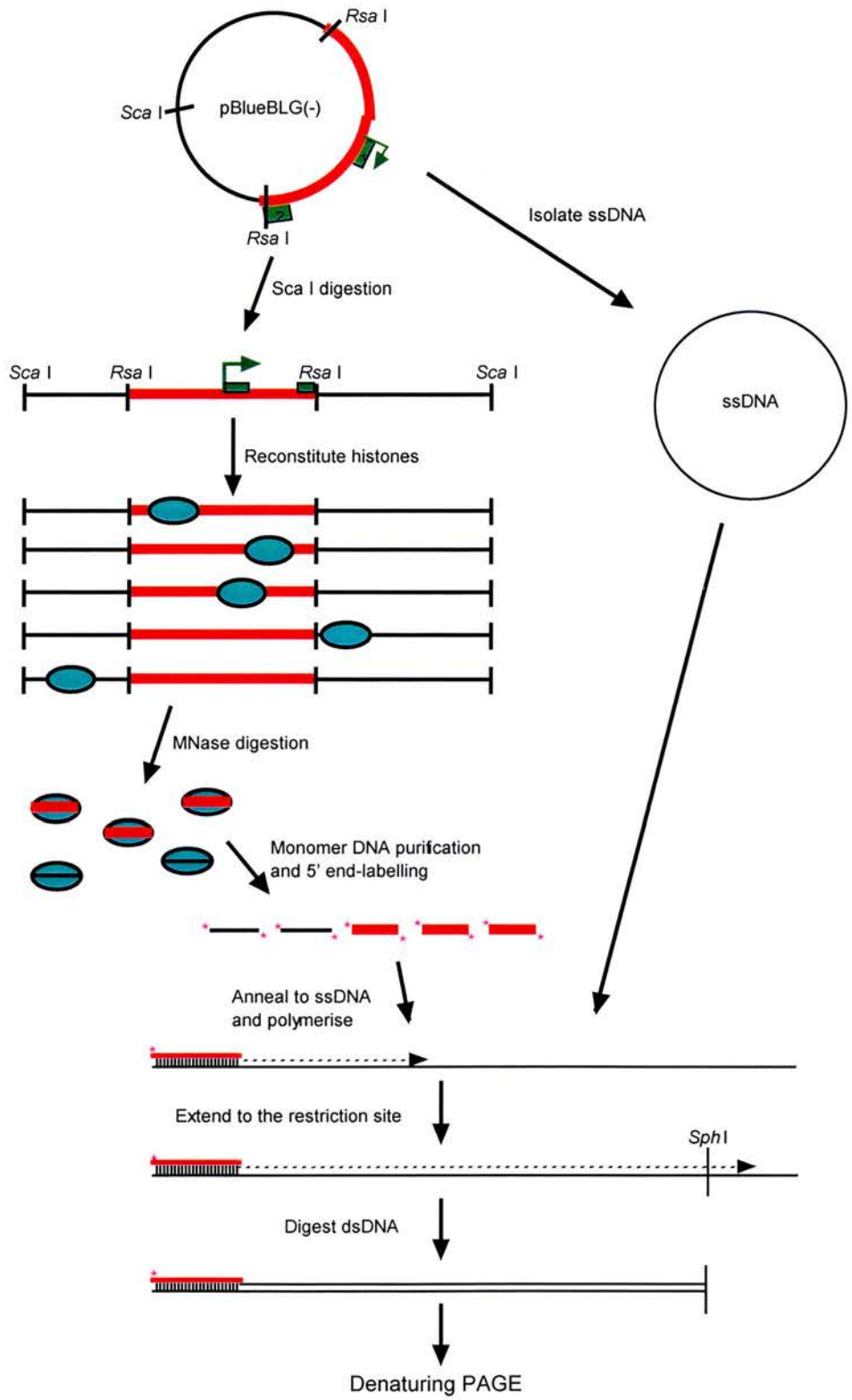



**Figure 65** Monomer Extension Technique

## 5.3 Results

Two constructs were generated which contained the 1746bp ***RsaI*** fragment of the ovine BLG gene promoter. They span from -864bp to +882bp with respect to the transcriptional start site and were inserted in to pBluescript II KS (-) via blunt-ended ligation into the ***Eco*RV** site in both orientations (Figure 66).

ssDNA was isolated from the pBlueBLG(-) construct using helper phage and PEG-mediated precipitation. Heated and unheated samples were electrophoresed through a 2% agarose gel to test for quality and dsDNA contamination (Figure 67).

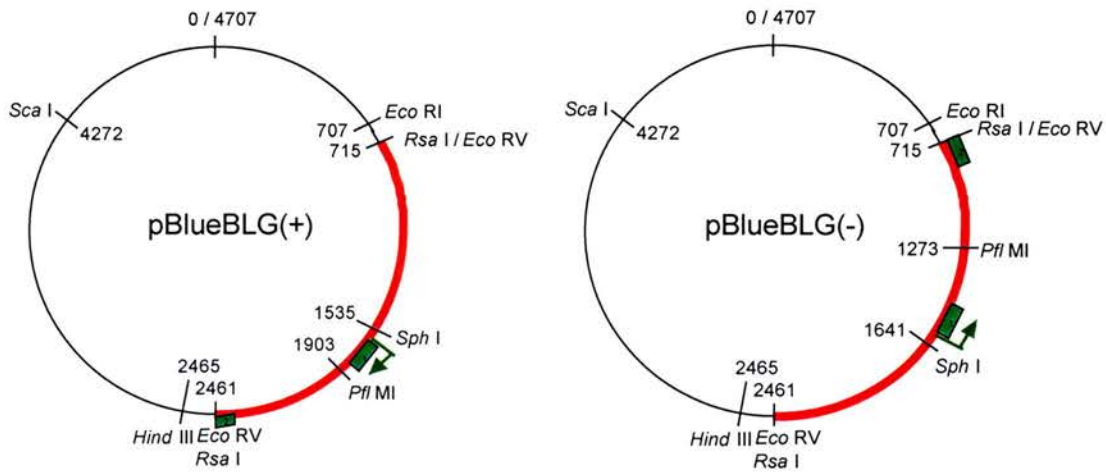



**Figure 66** Constructs Containing the Ovine BLG Promoter Region

The insert of 1746bp (red) has been inserted in both orientations. The sequence of pBlueBLG(+) decreases in line with that of the plasmid sequence. The insert is in the opposite orientation in pBlueBLG(-). The start site is indicated as a green arrow and relevant restriction enzyme sites are listed.


**Figure 67** Heated and Unheated Samples of pBlueBLG ssDNA

Five different isolates of pBlueBLG(-) were tested. All proved to be of high quality and free from dsDNA contamination. The upper band in each lane is phage ssDNA commonly found in these preparations and does not interfere with subsequent steps. 'Heated' means an additional step of 95°C for 5 minutes directly before loading. 'M' is the 1kb marker lane.

The monomer DNA was prepared by reconstituting of histones on to pBlueBLG(-), digesting it with MNase and isolating the protected 146bp fragments. These were checked for quality and concentration on an agarose gel (Figure 68).

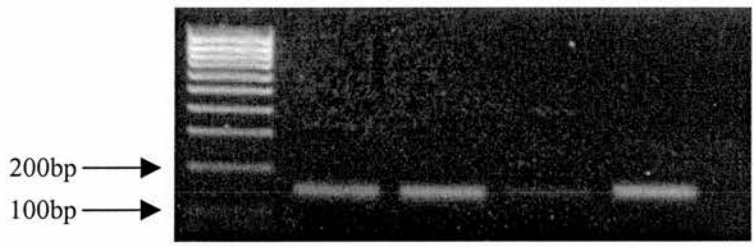



**Figure 68** pBlueBLG Monomer DNA Fragments

All monomer fragments are high quality and do not contain contaminating dimer or trimer nucleosome fragments.

The limit of nucleosome mapping for a single restriction enzyme is approximately 800bp. As the BLG insert is 1.7kb in size, several restriction enzymes were used to map the full region. These were ***Eco*RI**, ***Pfl*MI** and ***Sph*I** (Figure 66). ***Hind*III** was also used as a control as it maps nucleosome positions on the vector and can be used to compare experiments, for instance, between opposite strands. The polyacrylamide gel of the resulting reactions can be seen in figure 71.

### 5.3.1 Data Analysis

The relative intensity of each band was calculated using *Aida* software (Figure 72). It correlates the position of each band in millimetres to its relative intensity, measured in arbitrary units. Therefore, these peaks are related to their position on the image, rather than to their position in base pairs on the gene. In order to position these peaks relative to the gene sequence, we must look to the markers.

The sizes of all the bands in the marker lanes is known precisely. Using the *Sigma-plot* program, the size of these bands can be related to their position on the image. This relationship takes the form of a sixth order polynomial equation (Figure 73). Using this equation, all the bands generated from the monomer extension reactions were calculated in nucleotides and could therefore be related to the gene sequence. Bands

which overlap between extension reactions were used to normalise the relative band intensities between lanes. The background was subtracted from all ordinates by using the undigested control lane (lane 12).

The position of each band relative to the transcriptional start site was calculated by subtracting its size from the position of the restriction enzyme site used. For example, there is a prominent band in Figure 74 in lanes 3 & 4, indicated by red arrow. Using the polynomial equation, its size was calculated to be 590bp. The position of the ***Pfl*MI** restriction enzyme used in this extension reaction is +334bp with respect to the transcriptional start site (Figure 69). Therefore, the position of this digestion site on the gene is 334-590= -256bp. The centre of the nucleosome was determined by adding 73bp to give its centre at -183bp.

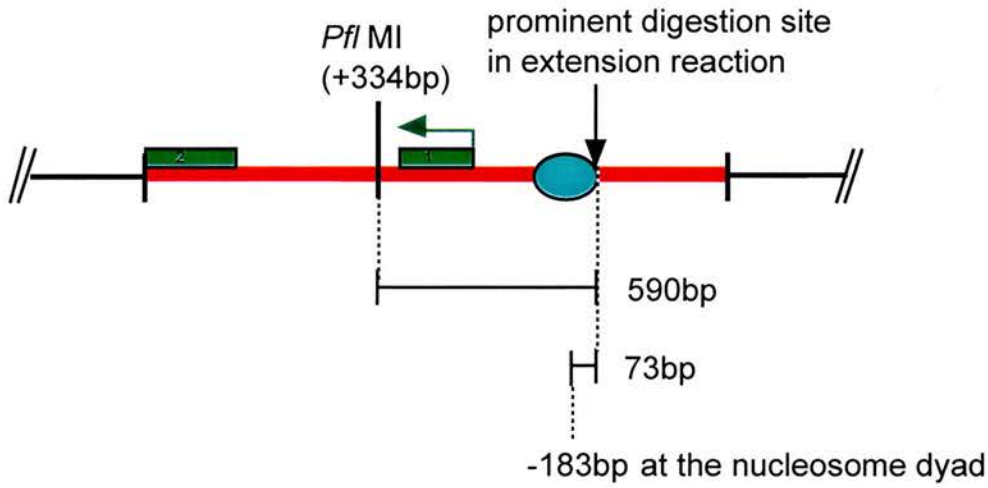



**Figure 69** Calculating the Position of a Monomer Extension Band Relative to the Gene
The end of the nucleosome is 590bp upstream from the ***Pfl*MI** digestion site. The centre of the nucleosome is 73bp further in, at -183bp relative to the transcriptional start site. The BLG insert is red, the exons are green boxes and the positioned nucleosome is blue.

The bands positions in the extension reactions containing restriction enzymes ***Sph*I** and ***Eco*RI** were calculated in the same way (Figure 70). The only exception was that 8bp of plasmid sequence also had to be accounted for in the ***Eco*RI** the digest (from 707 to 715 in figure 66). These results were plotted with their position on the BLG gene versus the relative band intensities (Figure 74).

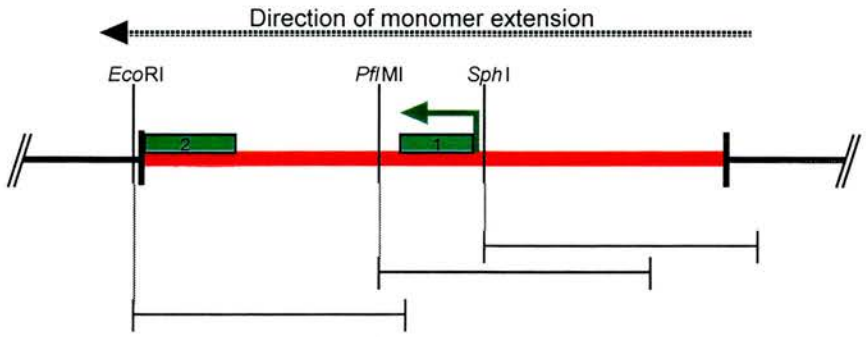



**Figure 70** Monomer Extension Mapping of the pBlueBLG(-) construct

The direction of the monomer extension reaction is indicated. The restriction enzyme in the reaction cleaves the DNA at the positions indicated. The maximum mapping distance of each restriction enzyme is up to 800bp of DNA, indicated below.

The lanes of the corresponding polyacrylamide gel overleaf are as follows:

1. Marker lane: phage lambda DNA digested with ***Dde*I** and end-labelled
2. Marker lane: phage lambda DNA digested with ***Hinf*I** and end-labelled
3. pBlueBLG(-) monomer extension digested with ***Pfl*MI**
4. pBlueBLG(-) monomer extension digested with ***Pfl*MI**
   (duplicate experiment to test for consistency)
5. pBlueBLG(-) monomer extension digested with ***Sph*I**
6. Marker lane: M13 phage sequenced for cytosine nucleotides
7. Marker lane: M13 phage sequenced for thymidine nucleotides
8. Marker lane: phage lambda DNA digested with ***Dde*I** and end-labelled
9. Marker lane: phage lambda DNA digested with ***Hinf*I** and end-labelled
10. pBlueBLG(-) monomer extension digested with ***Eco*RI**
11. pBlueBLG(-) monomer extension digested with ***Hind*III**
12. pBlueBLG(-) monomer extension control (undigested)
13. Marker lane: M13 phage sequenced for cytosine nucleotides
14. Marker lane: M13 phage sequenced for thymidine nucleotides

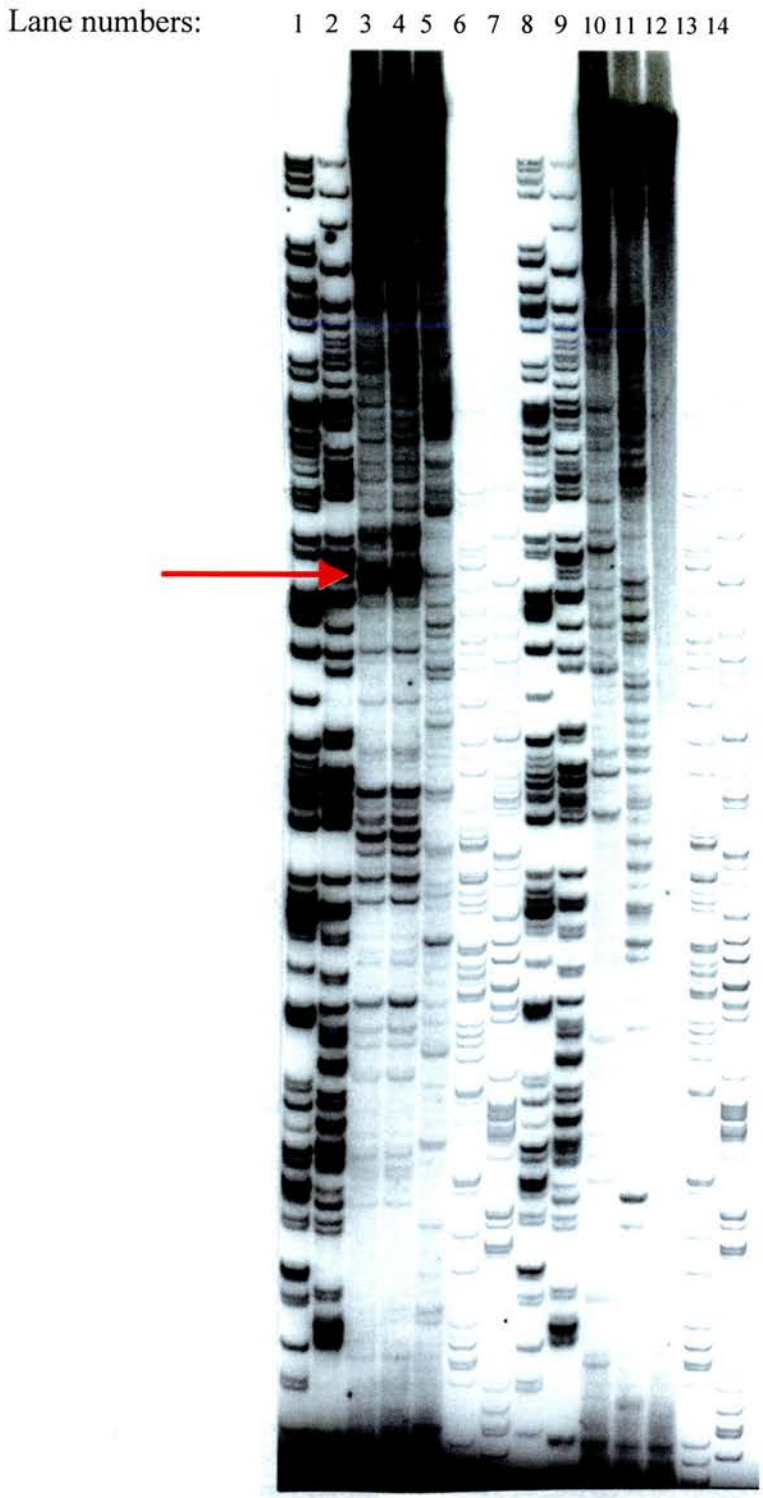



**Figure 71** Photographic Image of the pBlueBLG(-) Monomer Extension Reactions

Lanes are indicated above; see text for reference to the red arrow

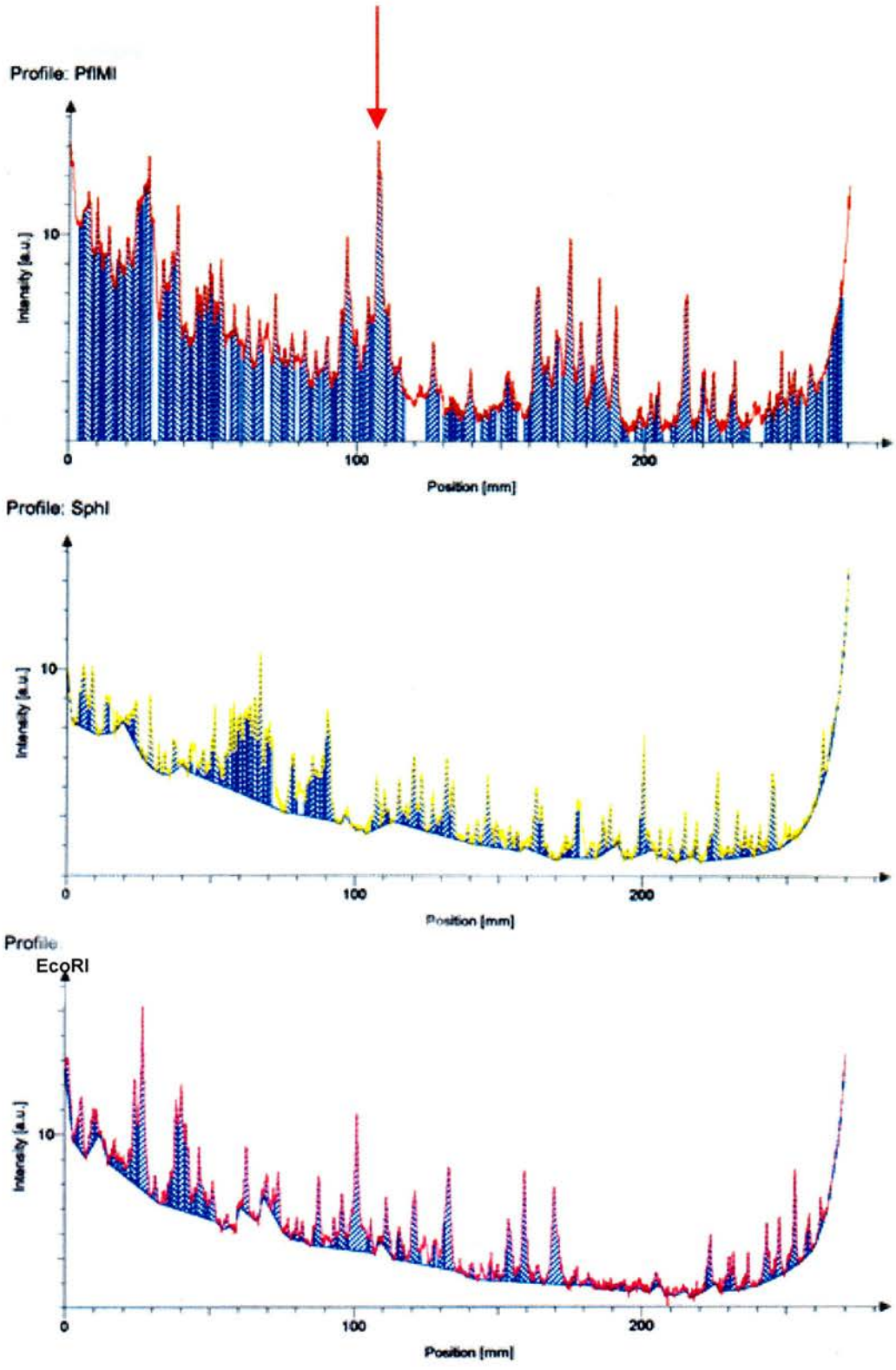



**Figure 72** Position of Each Band in Relation to its Intensity
A densitometer scan of figure 71, lanes 4 (***Pfl*MI**), 5 (***Sph*I**) and 10 (***Eco*RI**). The background has not been subtracted at this point

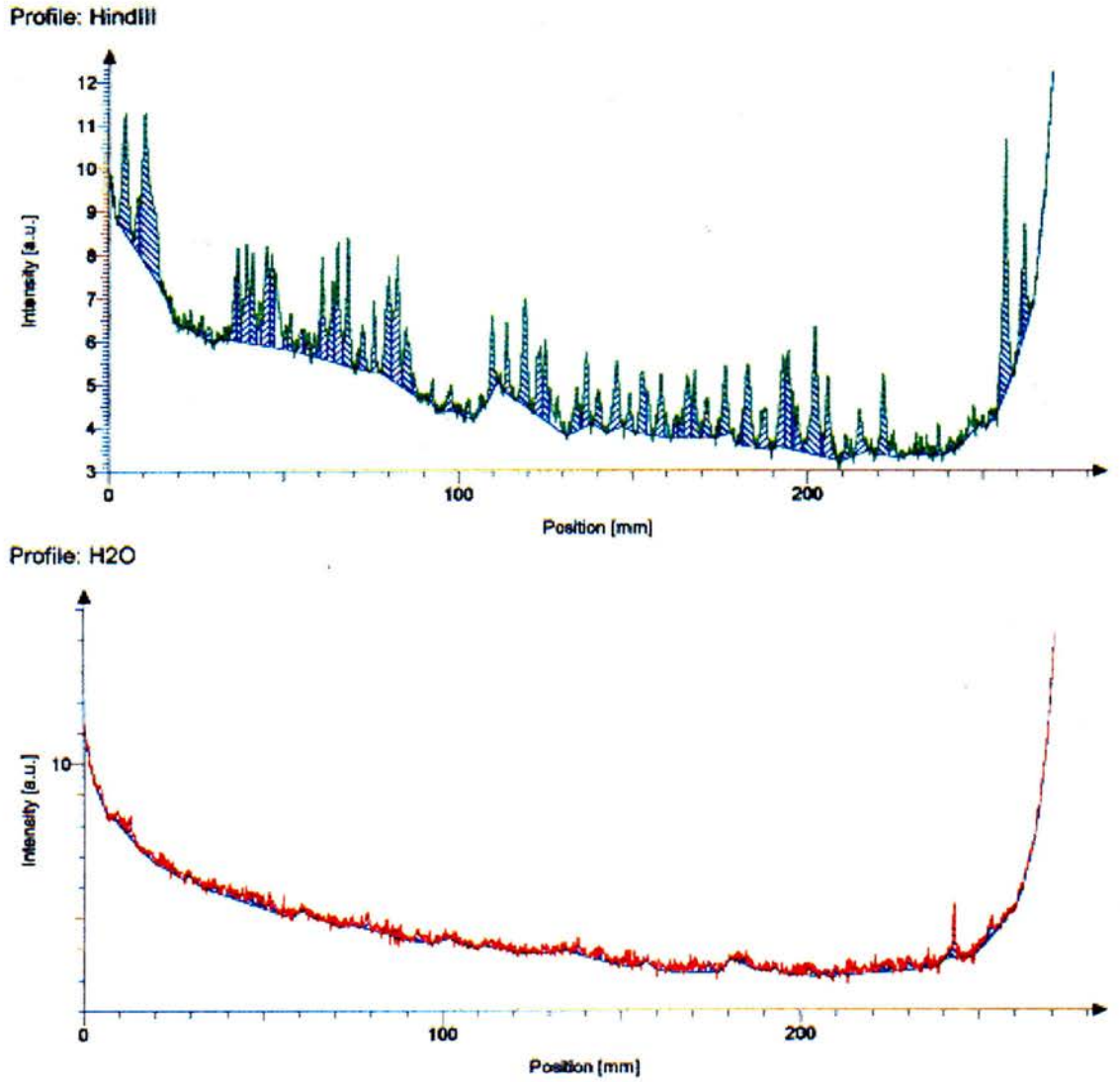



**Figure 72 cont.** Position of Each Band in Relation to its Intensity

A densitometer scan of figure 71, lanes 11 (***Hind*III**) and 12 (***$H_2O$*** control).

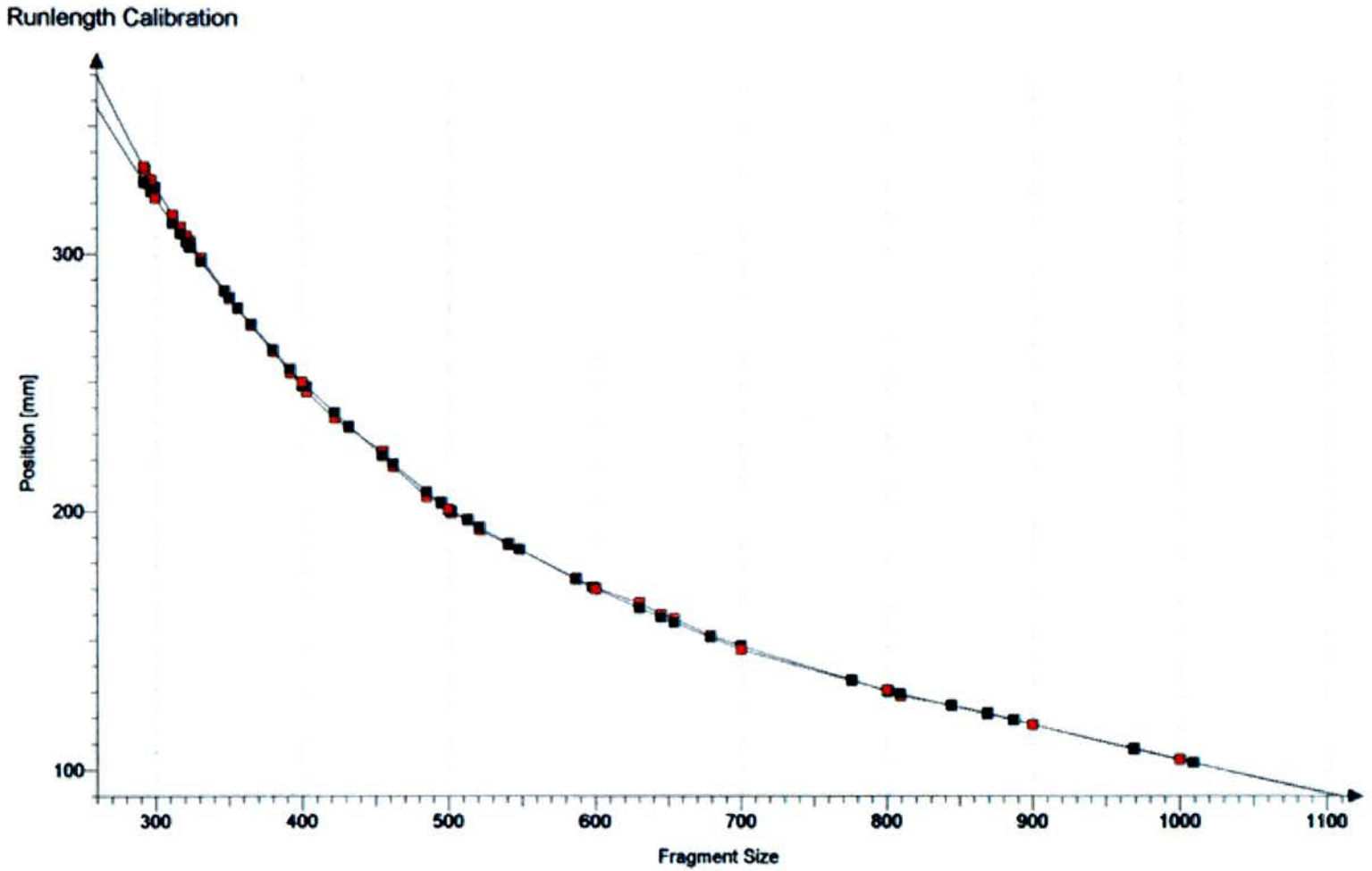



**Figure 73** Relationship between Fragment Size and Position

Red squares relate to the ***Dde*I** Lambda marker and back squares to the ***Hinf*I** Lambda marker

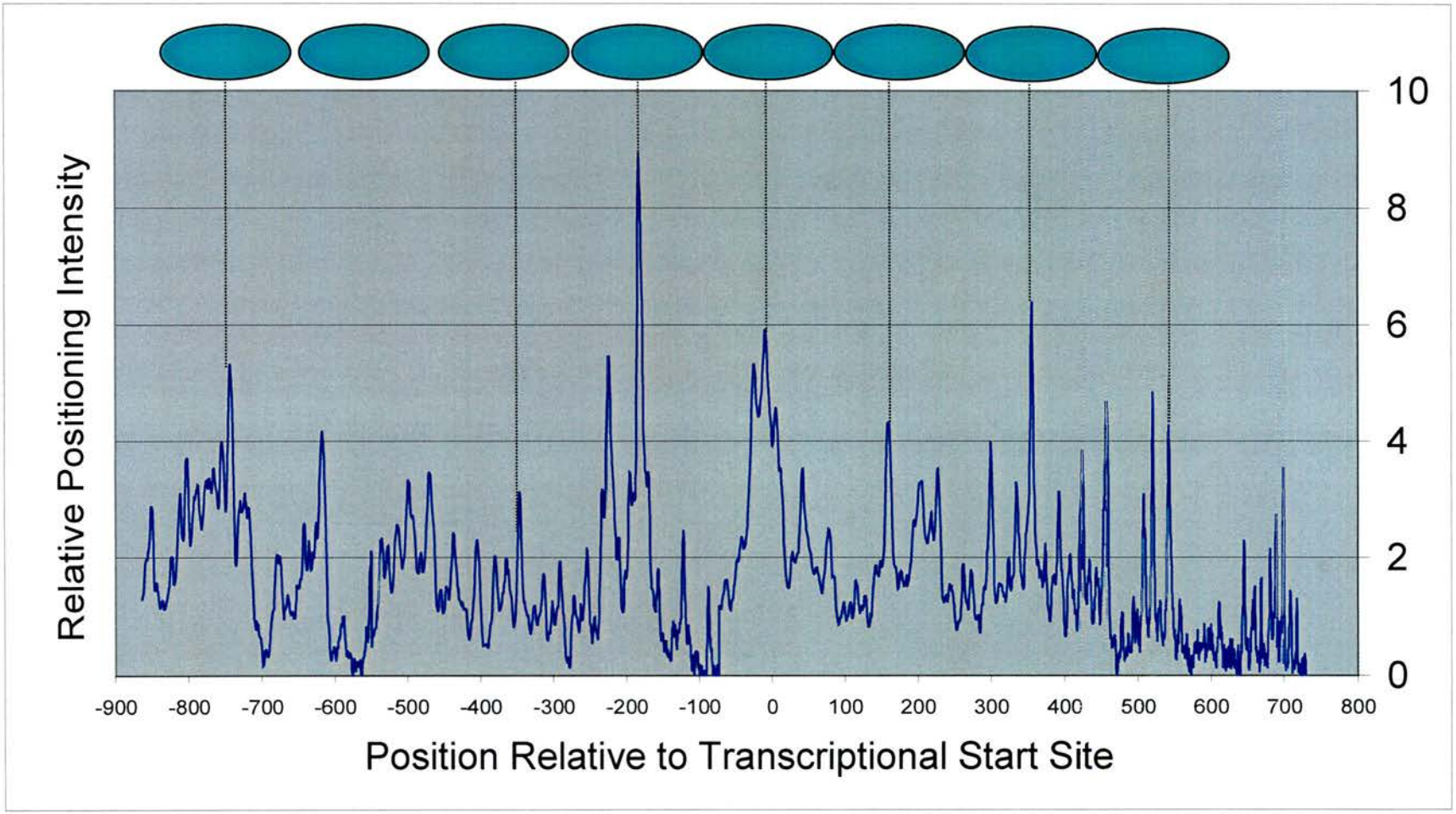



**Figure 74** Nucleosome Map of the *in vitro* Nucleosome Positions over the Sheep BLG Promoter Region

Blue ovals above the peaks indicate the position of nucleosome in liver chromatin *in vivo*

## 5.4 Discussion

The most striking feature of this map is the strong positioning signal at -183bp (red arrow). This would position the core particle between -110bp and -256bp and the full repeat length of 180bp between -93bp and -273bp (Figure 75). There is a smaller peak upstream of it at -224bp which may indicate an alternative position. The next two most intense peaks are at -10bp and +356bp. These would align very well with this strongly positioned nucleosome in a putative 180bp-repeat array. In accordance with this array are peaks at -743bp, +161bp and +542bp. There is a peak at -617bp which does not align with this array and several clustered next to the +542bp peak, notably +520bp and +458bp. However, owing to the net stabilisation energy from the former peaks, the most stable *array* on the insert might be formed by nucleosomes coloured light blue (Figure 75). Obviously, in its natural environment, long range effects and additional proteins could have a more pronounced influence on this short phased array.

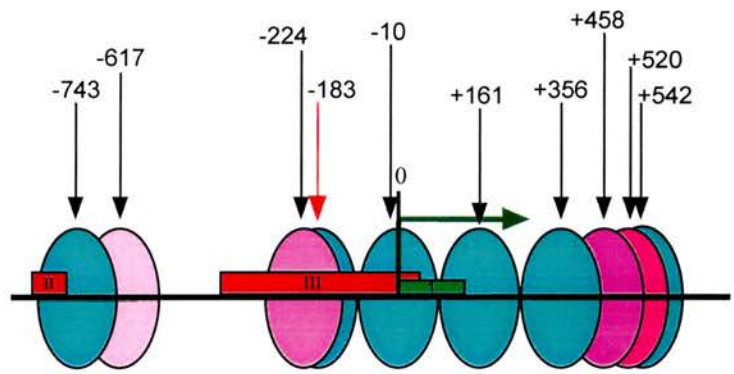



**Figure 75** Strong *in vitro* Nucleosome Positions over the Ovine BLG Promoter

Positions which correlate with a 180bp-array are coloured light blue. Those which do not are coloured shades of magenta. The start site (green arrow), exon 1 (green box) and DNase I hypersensitive sites (red box) are indicated. The exact position of the centre of each nucleosome is marked above.

The strongly positioned nucleosome at -183 shows an interesting relationship with the Stat binding sites which it encompasses (Figure 76). Two of the sites are positioned exactly at the terminal ends of the nucleosome at -93 and -278. The third is at -210, approximately 30bp from the dyad axis.

Stat5 alone is not sufficient to induce full activation (see **1.12** & **1.13**), however, perhaps the location of its binding sites in the linker regions assist its binding to the DNA, which may induce nucleosome remodelling (Pfitzner *et al.*, 1998). Nucleosome remodelling could make the NF1 sites more accessible, leading to full activation. Alternatively, these positions on the nucleosome may enable all three Stat-bound dimers to interact with each other, possibly mediating a co-operative action. If this were the case, the nucleosome might also have to be rotationally positioned to allow the Stat5 dimer access to the middle site within the nucleosome. It is unknown whether Stat5 requires nucleosome removal or remodelling to bind its cognate sequence or whether, like the glucocorticoid receptor, it can bind within a rotationally positioned nucleosome when its cognate site faces away from the histone proteins. However, from its structure and indirect evidence, it probably does not bind within a nucleosome (see **1.13**).

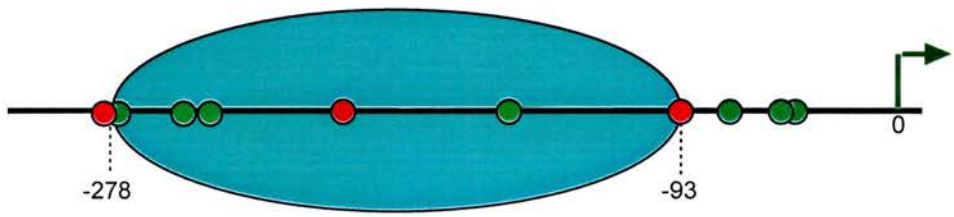



**Figure 76** Position of Stat and NF1 Sites on the Strongly Positioned Nucleosome
The exact positions are: NF1: -51, -55, -77, -159, -254, -266, -272 and -379 Stat: -93, -210 and -278

The nucleosome array present *in vivo* reflects this hypothetical array formed by positioned nucleosomes *in vitro*. In both mammary and liver chromatin the array over the promoter region (chapter 4, figure 58) matches the light-blue coloured array in figure 74. This is compelling evidence that the underlying DNA sequence over the promoter region of BLG is a major determinant of nucleosome positioning within this localised domain. This could result from a number of strong, in-phase positioning signals. Moreover, this strengthens previous evidence that the monomer extension technique is a valid method of studying nucleosome positions *in vitro*.

# Chapter 6

# A 16kb BLG Transgene Organises a Normal Nucleosomal Array over the Promoter

## 6.1 Introduction

The same nucleosomal array was present over the promoter region of BLG *in vitro* as that *in vivo* (chapter 5). This would suggest that the promoter itself contains sufficient information to form the correct array structure. However, it is feasible that sequences out-with the local domain are important in augmenting this structure. Frequently, it has been shown that regions many kilobases away are vital for correct gene activation (Pinkert *et al.*, 1987; Montoliu *et al.*, 1996). These distal elements may be involved in modifying the chromatin structure of the promoter region prior to gene activation as well as directly activating the gene via transcription factor interactions. I wished to test if long-range interactions were important in BLG gene activation and as an initial step I examined the BLG14 transgene.

The BLG14 mouse line contains about 5 copies of a BLG transgene comprising 4.2kb of 5' sequence, a 4.7kb transcription unit and 7.3kb of 3' flanking sequence. This line expresses BLG in a stable and copy number related manner, homogeneously in secretory epithelial cells throughout the mammary gland (Dobie *et al.*, 1996). Analysis of this BLG transgene in liver and lactating mammary tissues should answer this question, since it comprises a finite domain which cannot be influenced by distal elements from its natural setting.

## 6.2 Results

Nuclei were isolated by standard procedures from liver and mammary tissues taken from a BLG14 mouse at mid-lactation (chapter 2). They were partially digested with OP-Cu and aliquots removed at various time points (**2.2.6**). After purification, the DNA was digested with ***Rsa*I** to generate a 1.7kb genomic fragment. Only time points 2 and 3 from

each tissue were loaded on an agarose gel together with the appropriate controls. These time points contained representative fragments from the full time-coarse and could be compared directly.

After electrophoresis, the DNA fragments were transferred to a nylon membrane via Southern blotting. Fragments which abut the ***RsaI*** site on the BLG gene were indirect end-labelled using -900DOWN, the same probe used when probing the equivalent digest of sheep genomic DNA (figures 26 & 52). Radioactively labelled fragments were visualised on photographic film. Nucleosome positions have been related diagrammatically to the transcriptional start site of BLG in figure 77 and their positions are summarised in table 9.

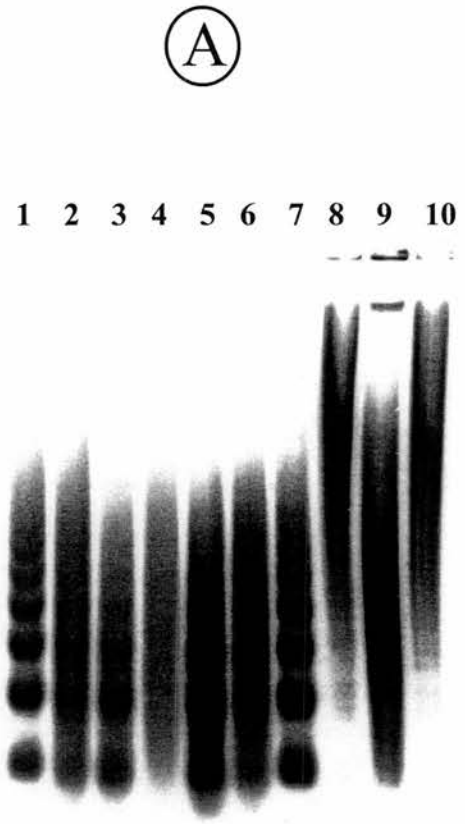
| | | | | | | |
|---|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 260 | 380 | 430 | 545 | 610 |
| | Position on Gene | -604 | -485 | -434 | -320 | -254 |
| Liver Chromatin | Band sizes | 260 | | 430 | | 610 |
| | Position on Gene | -604 | | -434 | | -254 |

| | | | | | |
|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 790 | | | |
| | Position on Gene | -74 | | | |
| Liver Chromatin | Band sizes | 790 | 970 | 1150 | 1330 |
| | Position on Gene | -74 | +106 | +286 | +466 |

**Table 9** Positions of OP-Cu Digestion Sites on the BLG14 Gene Promoter

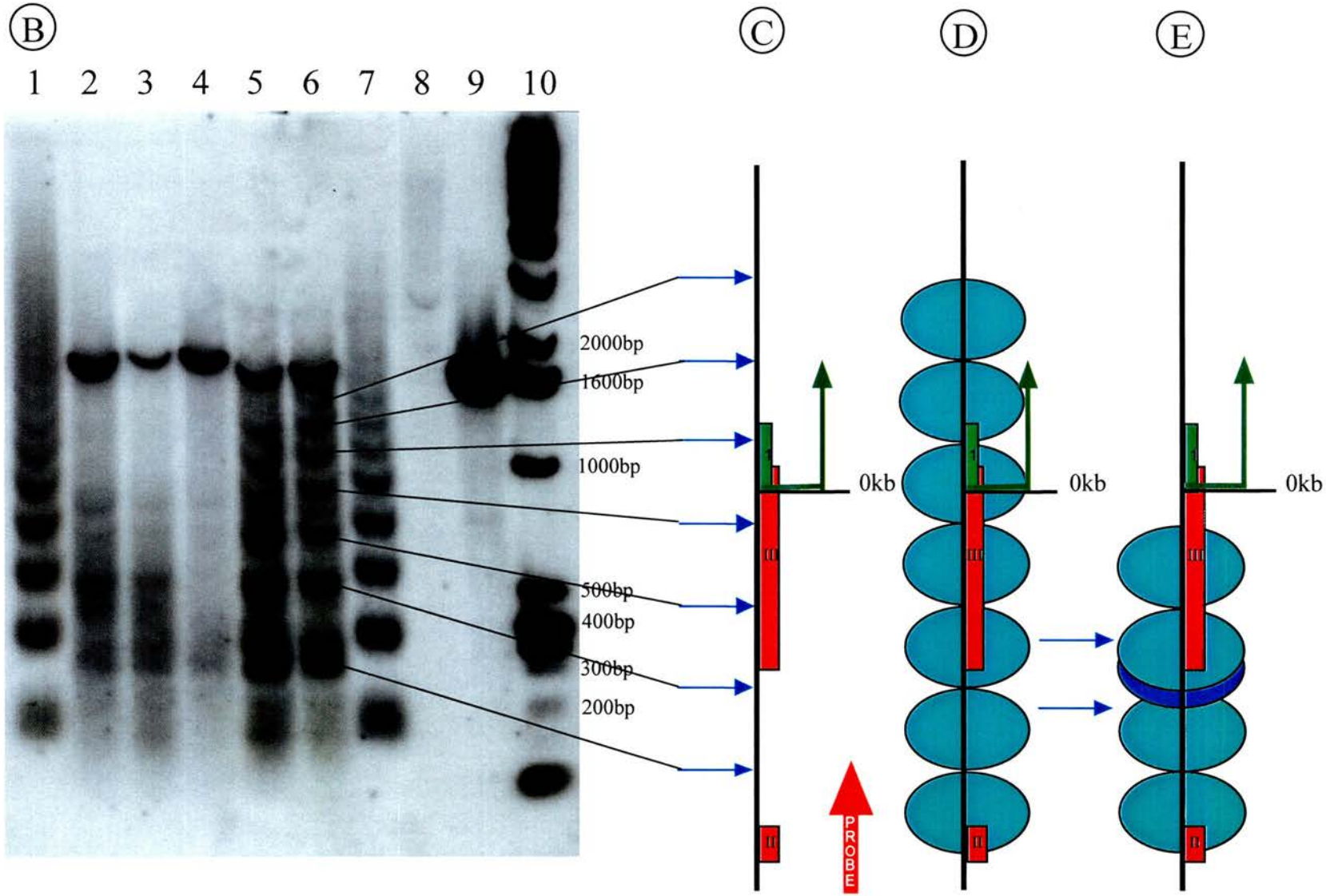
**Figure 77 Nucleosome Mapping from -0.9kb to Downstream of the BLG14 Transgene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver chromatin
(E) Diagram of nucleosome positions on the gene in mammary chromatin




**Lanes:**

1 BLG14 mammary nuclei digested with *OP*-Cu only time point 2
2. BLG14 mammary nuclei digested with ***Rsa*I** plus OP-Cu; time point 2
3. BLG14 mammary nuclei digested with ***Rsa*I** plus OP-Cu; time point 3
4. BLG14 naked genomic DNA positive control digested with ***Rsa*I** plus OP-Cu
5. BLG14 liver nuclei digested with ***Rsa*I** plus OP-Cu; time point 2
6. BLG14 liver nuclei digested with ***Rsa*I** plus OP-Cu; time point 3
7. BLG14 liver nuclei digested with *OP*-Cu only time point 2
8. Mouse genomic DNA negative control digested with ***Rsa*I** only
9. BLG14 genomic DNA positive control digested with ***Rsa*I** only
10. Mouse genomic DNA digested with ***Rsa*I** plus 5μg of 1kb ladder (Boehringer Mannheim)


B
1 2 3 4 5 6 7 8 9 10
2000bp
1600bp
1000bp
500bp
400bp
300bp
200bp
C
D
E
0kb
0kb
0kb
I
II
III
PROBE

## 6.3 Discussion

The nucleosomal pattern in the promoter region of the BLG14 transgene appears essentially identical to that of the gene in its natural environment in sheep (cf. figure 52 in chapter 4). In the liver chromatin, there is a regular and clearly defined array over the transcriptional start site. In the active gene in mammary tissue, there appears to be an alternatively positioned nucleosome near the transcriptional start site. At this point only a smear is evident, although the nucleosomes are probably still present.

Lane 1 (Figure 77) was probed with OP-Cu digested mouse mammary DNA which was not digested with ***Rsa*I**. At least ten nucleosome-repeat bands are evident which do not appear to drop in their relative intensity at any point. A drop in intensity might indicate that nucleosomes were absent on one side of the probe. Therefore, I suggest that the nucleosomes within the coding region are in a constant state of flux owing to polymerase read-though. Moreover, they appear to move as an entire array, rather than individually, since a regular repeat structure is still evident.

Like the corresponding Southern blot of sheep genomic DNA in chapter 4 (figure 52), the nucleosome positions align with the rest of those in the BLG gene. Therefore, within the limits of this assay, this experiment shows that the chromatin structure of the promoter region of the 16kb transgene in mouse line BLG14 possesses an essentially identical nucleosome structure to that of the gene in its natural setting in mammary and liver chromatin in the sheep genome. This means that all sequences necessary to induce the changes in the nucleosomal alignment on gene activation are present within the confines of this transgene. As the nucleosome repeat length in these tissues is so similar in both sheep and mouse, it is difficult to tell whether the transgene repeat length differs from the average repeat length in mouse chromatin.

It is possible, but highly unlikely, that endogenous mouse sequences out-with the transgene could compensate for an absence of essential distal elements in the BLG gene. It is particularly unlikely since these results reflect those not only from the gene in its natural setting (chapter 4), but also from the *in vitro* results (chapter 5). Completing the same experiment in a number of other lines, which contain the same BLG transgene at a different integration site, could effectively discount this possibility. Indeed, testing

smaller and smaller transgenes, e.g. only the BLG promoter region, might identify which regions are capable of setting this nucleosomal structure.

Therefore, this evidence and the *in vitro* evidence in the previous chapter strongly suggest that it is the local sequences within the BLG gene which are sufficient to reorganise the chromatin structure to induce gene activation. This is probably mediated through the recruitment of trans-acting factors which interact with the underlying DNA sequence.

# Chapter 7

# *In vivo* and *in vitro* Investigation of Additional Milk Protein Gene Promoters which are Stat5 Regulated

## 7.1 Introduction

In the previous chapter, the Stat5 GAS binding sites showed an interesting relationship with the *in vitro* nucleosome positions on the ovine BLG promoter region. Two of the three binding sites are located in the linker regions on either side of the nucleosome positioned at -180bp on the sheep gene. The third is located 30bp away from the dyad axis of that nucleosome. If this were significant in relation to Stat5-mediated gene regulation, a similar relationship might be found in the highly related caprine BLG gene.

Other milk protein genes, which are also Stat5 regulated, might have a similar relationship with positioned nucleosomes over their promoter regions. If this were the case, it could suggest that positioned nucleosomes have a direct bearing on Stat-mediated gene regulation.

To test this hypothesis, the position of nucleosomes over the promoter region of a number of other Stat5-regulated genes was analysed (Figure 78). The highly related caprine BLG gene was tested initially, since a strong deviation from the nucleosome positions in the ovine BLG gene would realistically negate this hypothesis. The following genes were also tested: goat β-casein, mouse β-casein, mouse α-lactalbumin and human α-lactalbumin. A correlation over such a broad range of genes and/or species would lend weight to the hypothesis.

All gene promoter regions were cloned in both orientations in to pBluescript II KS(-) to examine the nucleosome positions *in vitro*. In addition, with the exception of the human gene, they were analysed *in vivo* in nuclei extracted from tissue. My efforts of persuasion in the neighbouring neo-natal clinics were not sufficient in order to acquire

healthy, fresh, lactating human mammary tissue. Neoplastic or cancerous tissue could have been obtained, but it is usually highly degraded, fixed or non-lactating and may well have an altered chromatin configuration (Patel *et al.*, 1997).

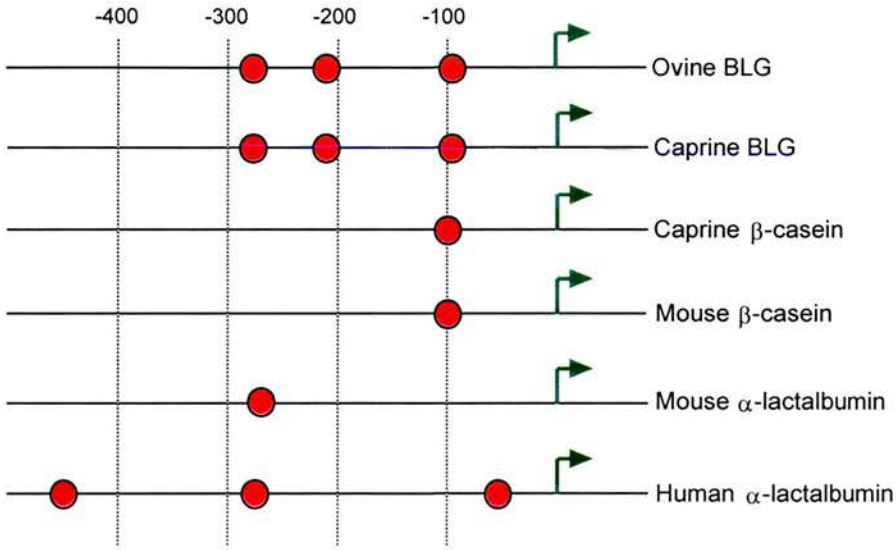



**Figure 78** Positions of Stat5-Binding Sites on the Genes Analysed

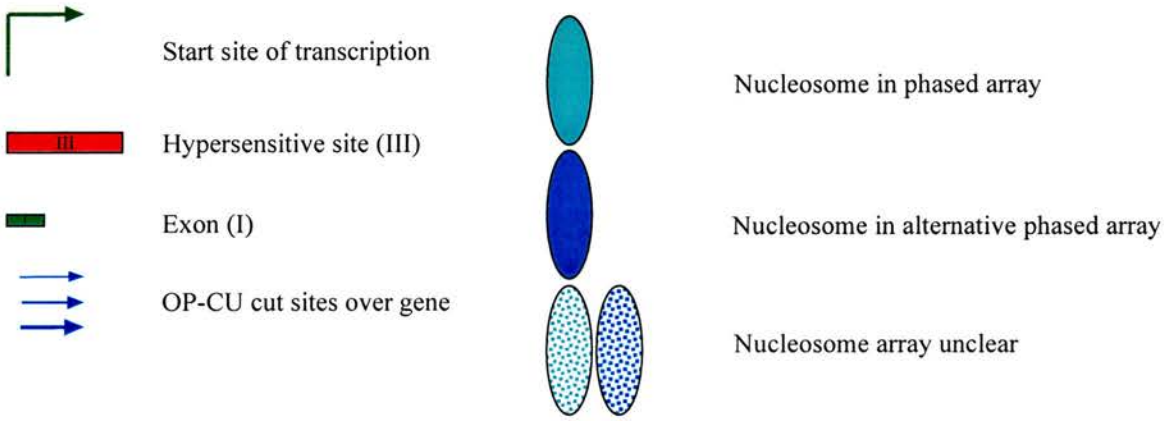
Stat site positions (red) are as follows: ovine & caprine BLG -93, -210 & -278; caprine β-casein -98; mouse β-casein -98; mouse α-lactalbumin -267; human α-lactalbumin -62, -278 & -457. Distance in base pairs is indicated above & the transcriptional start site is displayed as a green arrow.

## 7.2 Results

I shall present the data in the following manner:

1. Analysis of the aforementioned gene promoters in the following order: the caprine BLG gene (**7.3**), the caprine β-casein gene (**7.4**), the mouse β-casein gene (**7.5**), the mouse α-lactalbumin gene (**7.6**) and the human α-lactalbumin gene (**7.7**).
2. Each section shall begin with a description of the experimental design, followed by the *in vitro* analysis and then the *in vivo* analysis.
3. At the end of each section there will be a discussion relating to these findings.
4. At the end of the chapter, I shall present an overall discussion (**7.8**).

The legend used to represent the *in vivo* data in figures 84, 91, 99 and 106 is the same as that used in chapters 3 and 4 (see legend).




**Legend** *In vivo* data figures

## 7.3 Caprine β-Lactoglobulin

### 7.3.1 In vitro analysis

The promoter region from -727 (***Bsr*BI**) to +31 (***Pvu*II**) was cloned into the ***Eco*RV** site of pBluescript II KS (-) by blunt ended ligation to create pK'BLG (D) (Figure 79). The insert was cloned in the opposite orientation by digesting with ***Bss*HII** which cleaves on either side of the multiple cloning region. After re-ligation, pK'BLG (R) was isolated. "K' " stands for caprine and "D" and "R" for "direct" and "reverse" orientations, respectively.

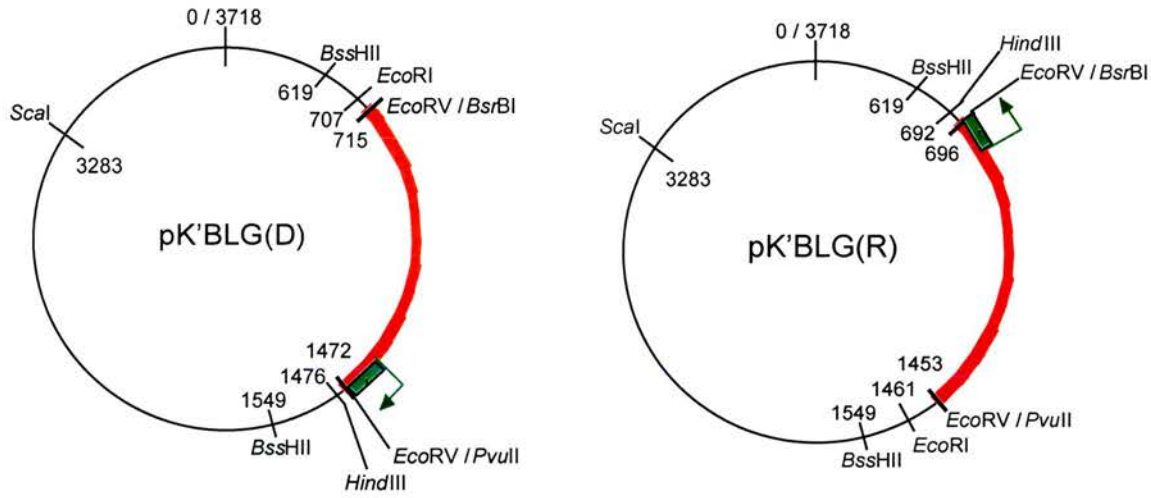



**Figure 79** Constructs Containing the Caprine BLG Promoter Region in Both Orientations

Notable restriction enzyme sites are indicated next to their position on the plasmid. The first exon (green box) is positioned next to the transcriptional start site (green arrow).

The core particle DNA fragments and ssDNA were isolated as before (**2.4** & **5.2**) (Figure 80). The restriction enzyme used in the pK'BLG (D) monomer extension reactions was ***Eco*RI** and for the pK'BLG (R) reaction ***Hind*III** was used. These produce extension fragments which span the insert. Once electrophoresed through a polyacrylamide gel the banding patterns were visualised on a phosphorimaging screen (Figure 81).

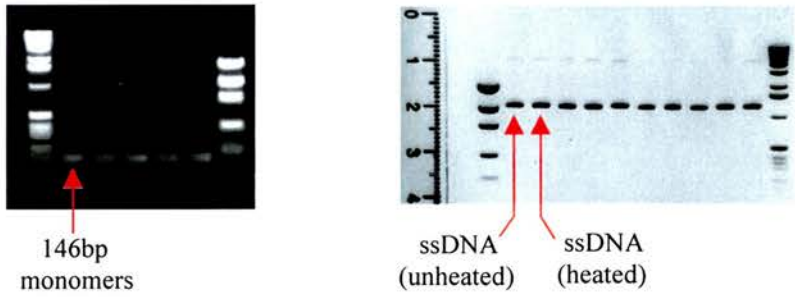



**Figure 80** Monomer DNA Fragments and ssDNA from K'BLG Constructs

The band sizes were correlated with their position on the image using *Aida* and *Sigma-plot* software as before (chapter 5). Once the background had been subtracted, the peaks were plotted with respect to their position on the caprine BLG promoter for both pK'BLG (D) (Figure 82) and pK'BLG (R) (Figure 83).

Lanes:

1. 100bp marker
2. Lambda ***Hinf*I** marker
3. Lambda ***Dde*I** marker
4. K'BLG (D) & ***Eco*RI**
5. K'BLG (D) & $H_2O$
6. K'BLG (R) & ***Hind*III**
7. K'BLG (R) & $H_2O$
8. 100bp marker

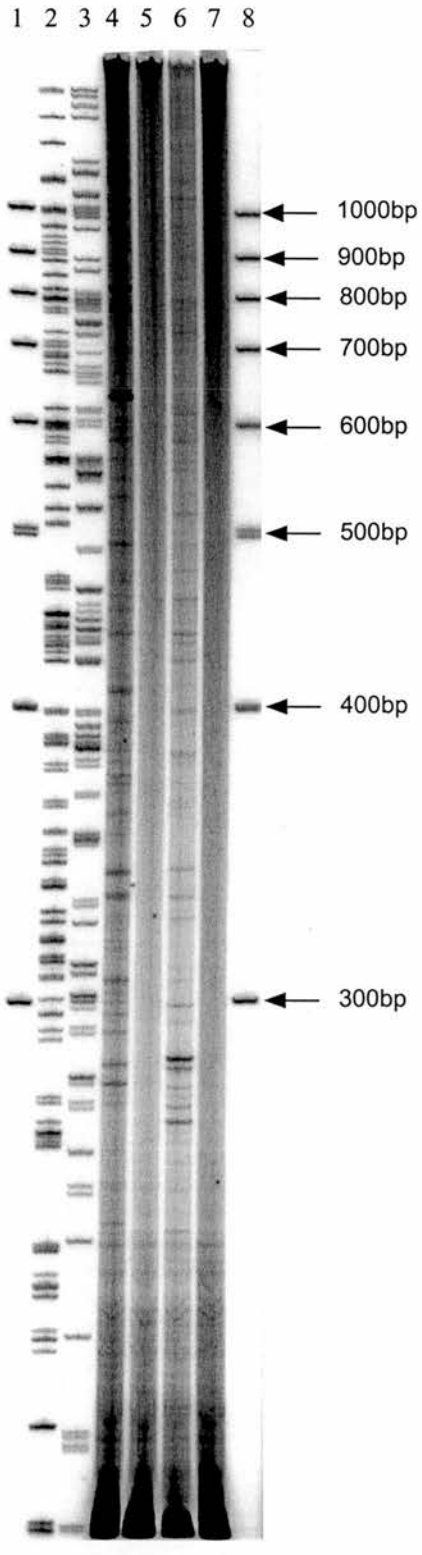



**Figure 81** Monomer Extension Reactions of the pK'BLG (D) and (R) Constructs

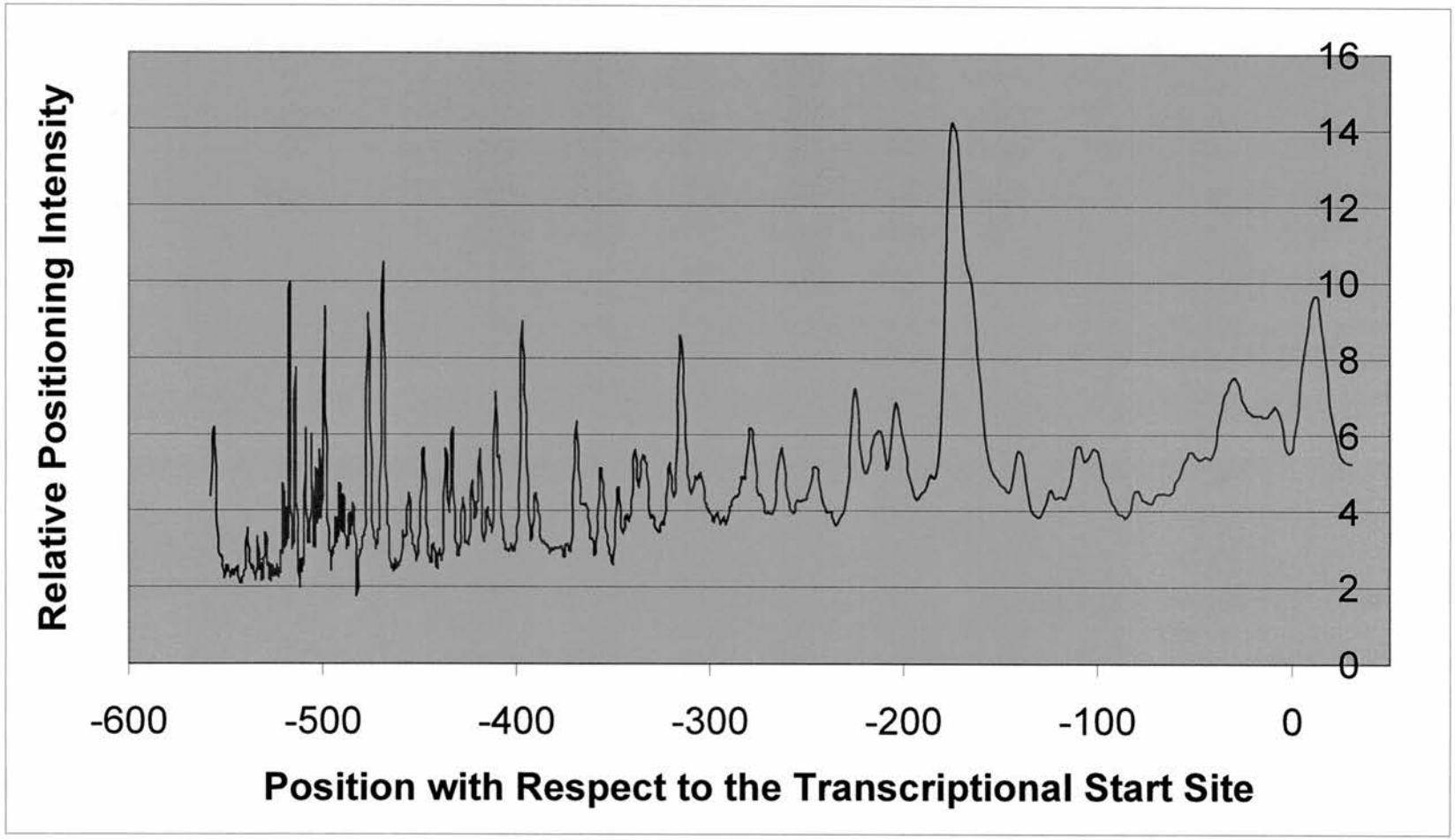



**Figure 82** Nucleosome Positions Over the Caprine BLG Promoter Region in pK'BLG (D)

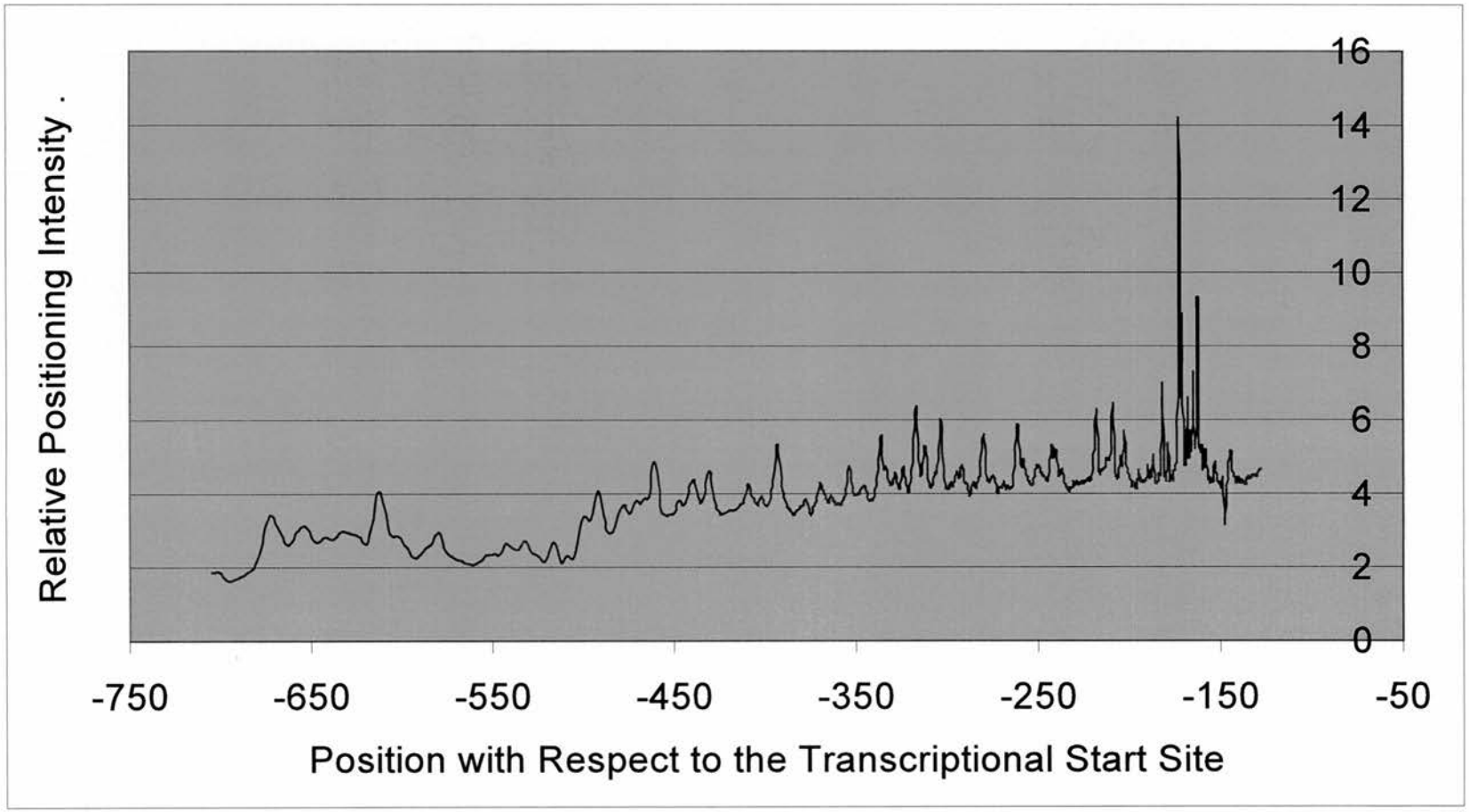



**Figure 83** Nucleosome Positions Over the Caprine BLG Promoter Region in pK'BLG(R)

### 7.3.2 In vivo analysis

Goat liver and lactating mammary nuclei were isolated by standard procedures (**2.2**). They were digested with OP-Cu in the same manner, and for the same time, as those of sheep nuclei (**4.2**). The time points are the same as follows:

time point 0: 0 minutes OP-Cu digestion
time point 1: 10 minutes OP-Cu digestion
time point 2: 20 minutes OP-Cu digestion
time point 3: 30 minutes OP-Cu digestion
time point 4: 40 minutes OP-Cu digestion
time point 5: 60 minutes OP-Cu digestion

After protein extraction, the DNA was further digested with ***Bam*HI**, which cuts the BLG promoter at -1.7kb with respect to the transcriptional start site. Primers were designed to abut this site on its downstream side and amplify a region of 247bp adjacent to it. This fragment was purified and end-labelled to act as a probe for the Southern blotted membrane.

The radiolabelled fragments were visualised on the membrane using photographic film and the band sizes were calculated in relation to the markers (Table 10). A diagrammatic representation of the banding pattern (Figure 84) shows the OP-Cu digestion sites in relation to the gene sequence.

Like the ovine BLG gene, only time points 2 and 3 for both mammary and liver titrations were electrophoresed on the composite gel so that the active and inactive promoter regions could be compared directly.

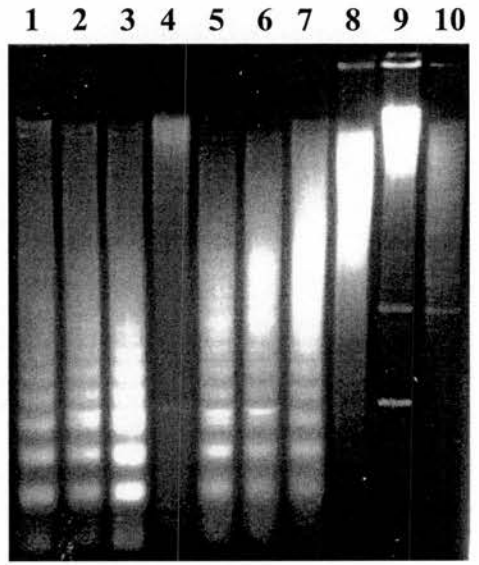
| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 180 | 310 | 370 | 480 | 530 | 660 | 710 | 900 | 1070 | 1250 | 1190 |
| | Position on Gene | -1.52 | -1.37 | -1.33 | -1.22 | -1.17 | -1.04 | -0.99 | -0.80 | -0.63 | -0.45 | -0.51 |
| Mammary Chromatin | Band sizes | 1370 | 1430 | | | | | | | | | |
| | Position on Gene | -0.21 | -0.27 | | | | | | | | | |
| Liver Chromatin | Band sizes | 180 | 370 | 530 | 710 | 900 | 1070 | 1250 | 1430 | 1610 | 1790 | 1970 |
| | Position on Gene | -1.52 | -1.33 | -1.17 | -0.99 | -0.80 | -0.63 | -0.45 | -0.27 | -0.09 | +0.09 | +0.27 |

**Table 10** Positions of OP-Cu Digestion Sites on the Caprine BLG Gene Promoter

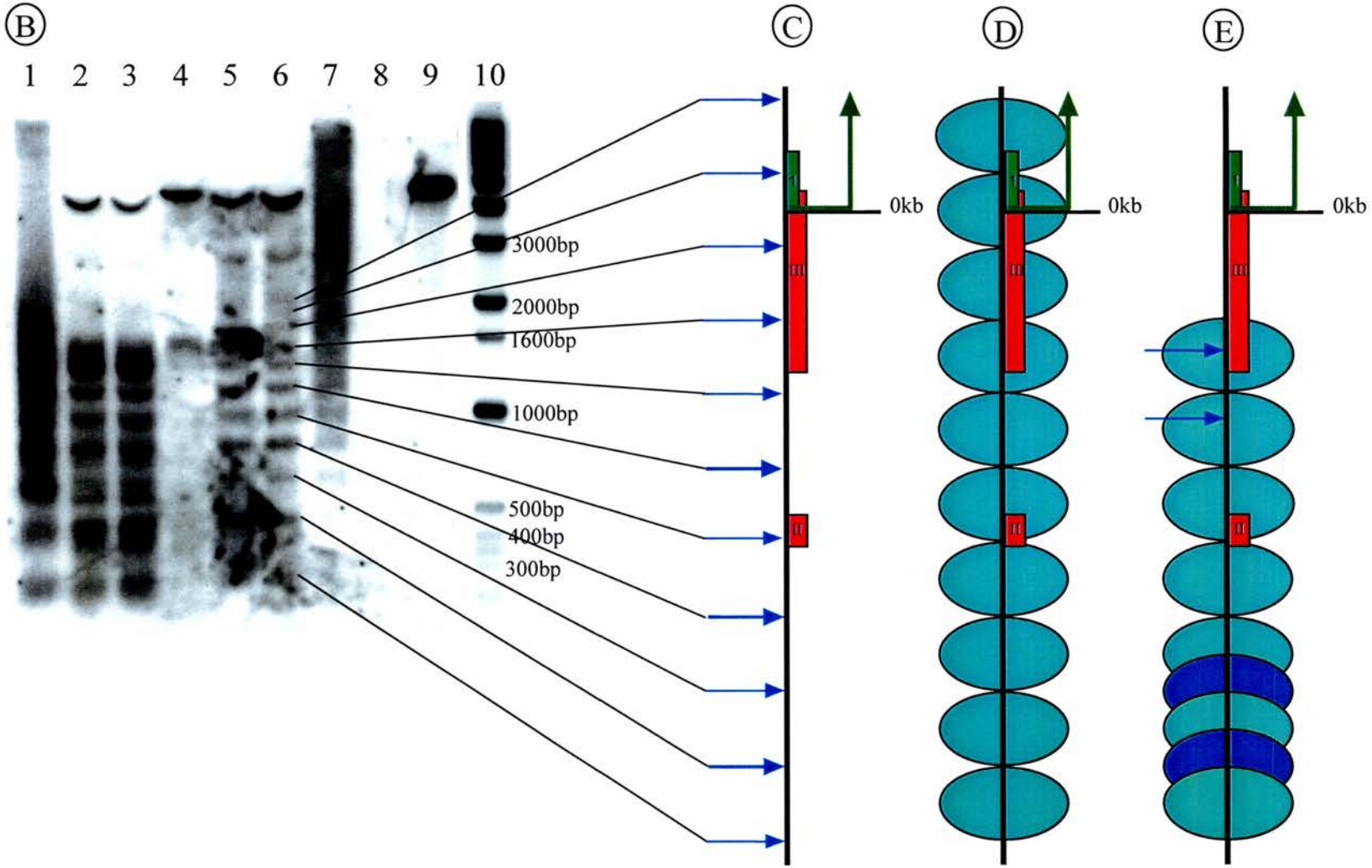
**Figure 84 Nucleosome Mapping from -1.7kb to Downstream of the Caprine BLG Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver nuclei
(E) Diagram of nucleosome positions on the gene in mammary nuclei


**Lanes:**




1. Caprine mammary nuclei digested with *OP*-Cu only time point 2
2. Caprine mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
3. Caprine mammary nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
4. Caprine naked genomic DNA positive control digested with ***Bam*HI** plus OP-Cu
5. Caprine liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 2
6. Caprine liver nuclei digested with ***Bam*HI** plus OP-Cu; time point 3
7. Caprine liver nuclei digested with *OP*-Cu only time point 2
8. Mouse genomic DNA negative control digested with ***Bam*HI** only
9. Caprine genomic DNA positive control digested with ***Bam*HI** only
10. Mouse genomic DNA digested with ***Bam*HI** plus 5µg of 1kb ladder (Boehringer Mannheim)


B
C
D
E
1 2 3 4 5 6 7 8 9 10
3000bp
2000bp
1600bp
1000bp
500bp
400bp
300bp
0kb
0kb
0kb
I
II
III

### 7.3.3 Discussion

The *in vitro* data indicates a strongly positioned nucleosome at –174bp with respect to the transcriptional start site. There is an additional sub-band close to this site at –165bp which may indicate an alternative position. Most importantly this is the strongest position in this map and it correlates almost exactly with the same position on the ovine BLG map.

There are additional peaks at positions +9bp, -316bp and –397bp with a cluster further upstream at positions –469bp, -477bp, -499bp and –518bp. These are most obvious in pK'BLG (R) where they are more intense than their counterparts in pK'BLG (D). This probably because of a decrease in the relative intensity of the –174bp band, so that in actual fact they are not as intense as they appear. The relative decrease in intensity of this band may be caused by the high background experienced at the top of the gel (Figure 81). Although in the figure the band is very prominent, the high background may have diluted this relative intensity when subtracted from the $H_2O$ control lane. Therefore, I postulate that the only strong positioning signal is the one present at -174bp.

This position is reflected in the *in vivo* analysis in liver chromatin (Figure 84). There is a regular array of nucleosomes extending from –1520bp until the limit of the assay at +270bp. It is feasible that this entire array could be positioned from the strongly positioned nucleosome at –174bp. This array reflects that of the ovine BLG gene, except that there is no distinct alternative array. This alternative array is indicated by dark-blue coloured nucleosomes in the ovine BLG gene. In the ovine gene, it starts at approximately -3kb and continues for about 1800bp until just short of the HS II site at -800bp. There are faint indications of bands in this region in the goat BLG gene which may be indicative of this very weak alternative array.

In goat mammary chromatin, the same array seen in liver chromatin stops at –270bp with respect to the BLG transcriptional start site. No obvious bands exist downstream of this point, possibly because the nucleosomes are in a constant state of flux as the polymerase repeatedly initiates transcription.

Like the sheep mammary BLG gene, the nucleosomal array is disrupted at about –630bp. A number of additional digestion sites exist between this point and –270bp, but their exact positions are unclear as these bands are tightly compressed. Since the transcription factor binding sites are highly conserved between goat and sheep BLG genes, these digestion sites are probably the same in both cases, indicating the presence of a possible alternative array.

This alternative array seen further upstream in the sheep BLG gene is also obvious in the same region in the goat mammary BLG gene. In the goat gene it starts at –1040bp and extends in an upstream direction to at least –1370bp (Figure 84; nucleosomes coloured dark blue). This differs slightly from the sheep gene where it starts at –1220bp, one nucleosome further upstream. It is intriguing that this switch in nucleosome phases occurs at almost the same position in both species.

In summary, the most strongly positioned nucleosome *in vitro* exists at -174bp with respect to the transcriptional start site for both sheep and goat BLG. This position is also occupied *in vivo* and may direct the array of phased nucleosomes which surround it. The Stat5 binding sites and this strongly positioned nucleosome are conserved in goat and sheep BLG genes (Figure 85). Therefore, caprine BLG data supports the hypothesis that the position of this nucleosome may in some way augment Stat5-mediated induction of the BLG gene.

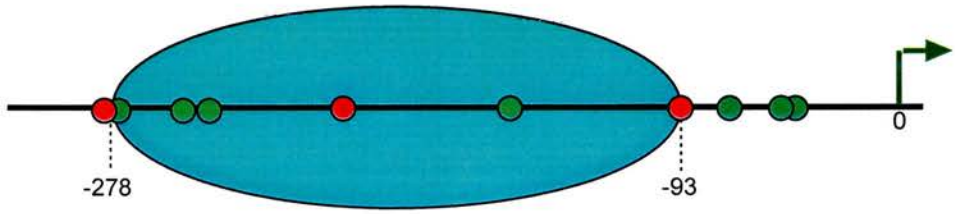



**Figure 85** Position of Stat and NF1 Sites on the Strongly Positioned Nucleosome

The exact positions are: NF1: -51, -55, -77, -159, -254, -266, -272 and -379 Stat: -93, -210 and -278

## 7.4 Caprine β-Casein

### 7.4.1 In vitro analysis

The caprine β-casein promoter region from -309 (***Bgl* II**) to +139 (***Bgl* II**) was cloned into pBluescript II KS (-) using a ligation compatible ***Bam*HI** site. It inserted in both orientations to produce pK'bcas (D) and pK'bcas (R) (Figure 86).

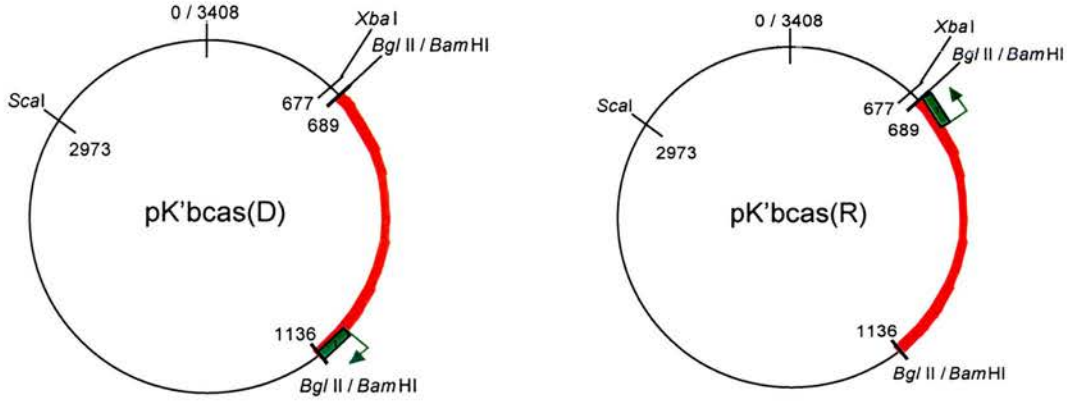



**Figure 86** Constructs Containing the Caprine β-Casein Promoter Region in Both Orientations

The monomer DNA fragments and the ssDNA were isolated as before and displayed no contamination when checked on a 2% agarose gel (Figure 87).

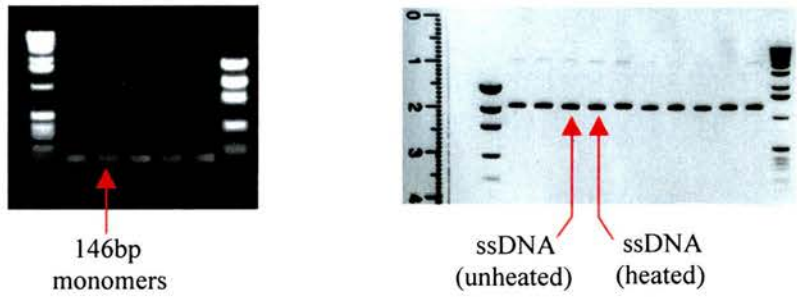



**Figure 87** Caprine β-Casein monomer DNA fragments and ssDNA from pK'bcas (D) and pK'bcas (R)

The monomer extension reactions were performed with the ***Xba*I** restriction enzyme to test for nucleosome positions on the insert, and $H_2O$ in the undigested control (Figure 88). After data analysis, the peaks were plotted in relation to their position on the gene for both constructs pK'bcas (D) (Figure 89) and pK'bcas (R) (Figure 90).

Lanes:

1. 100bp marker
2. Lambda ***Hinf*I** marker
3. Lambda ***Dde*I** marker
4. K'bcas (D) & ***Xba*I**
5. K'bcas (D) & $H_2O$
6. K'bcas (R) & ***Xba*I**
7. K'bcas (R) & $H_2O$
8. 100bp marker

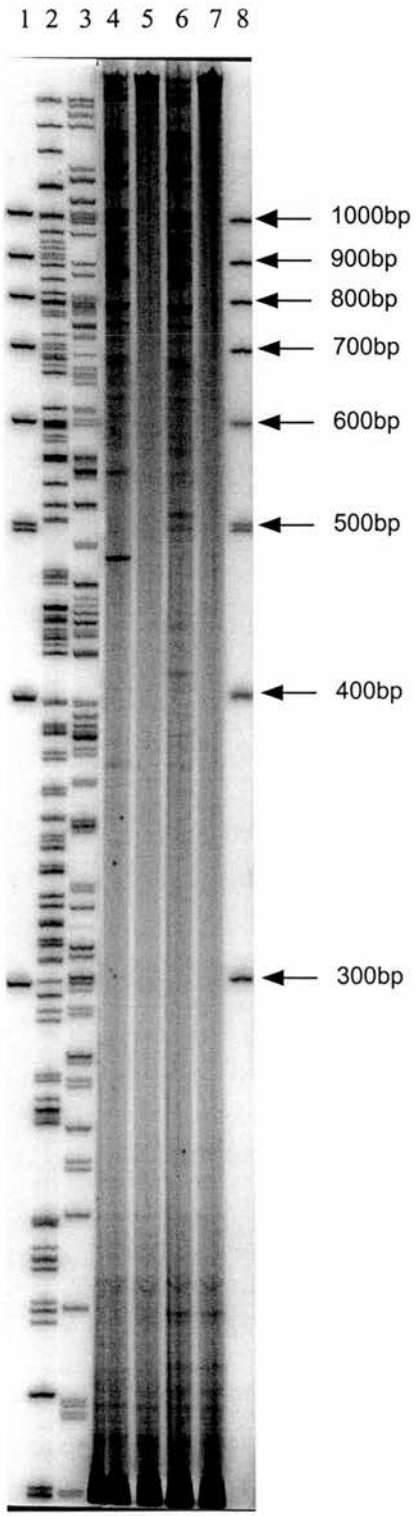



**Figure 88** Monomer Extension Reactions of the pK'bcas Constructs

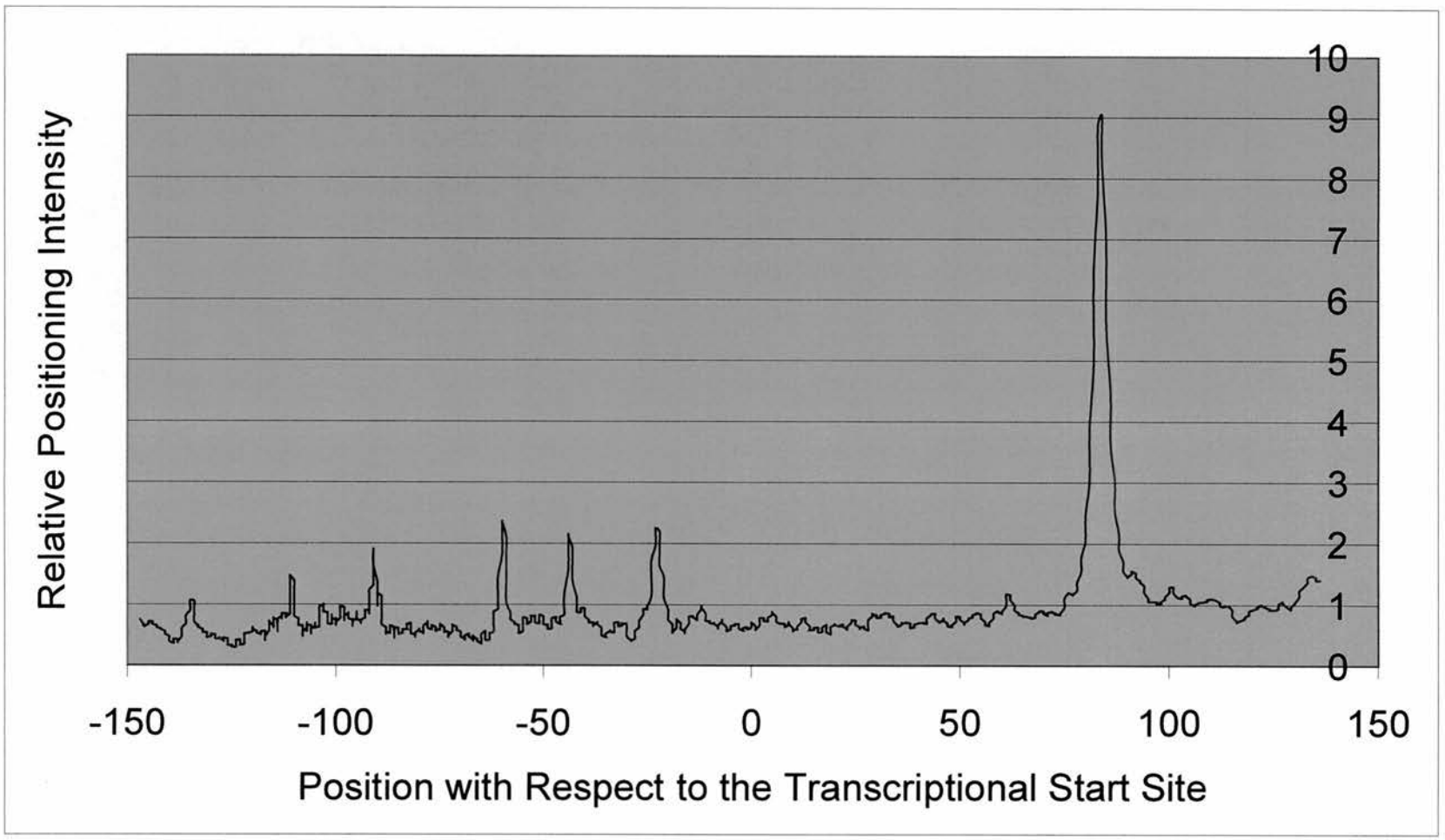



**Figure 89** Map of the *in vitro* Nucleosome Positions over the Caprine β-Casein Promoter Region in pK'bcas (D)

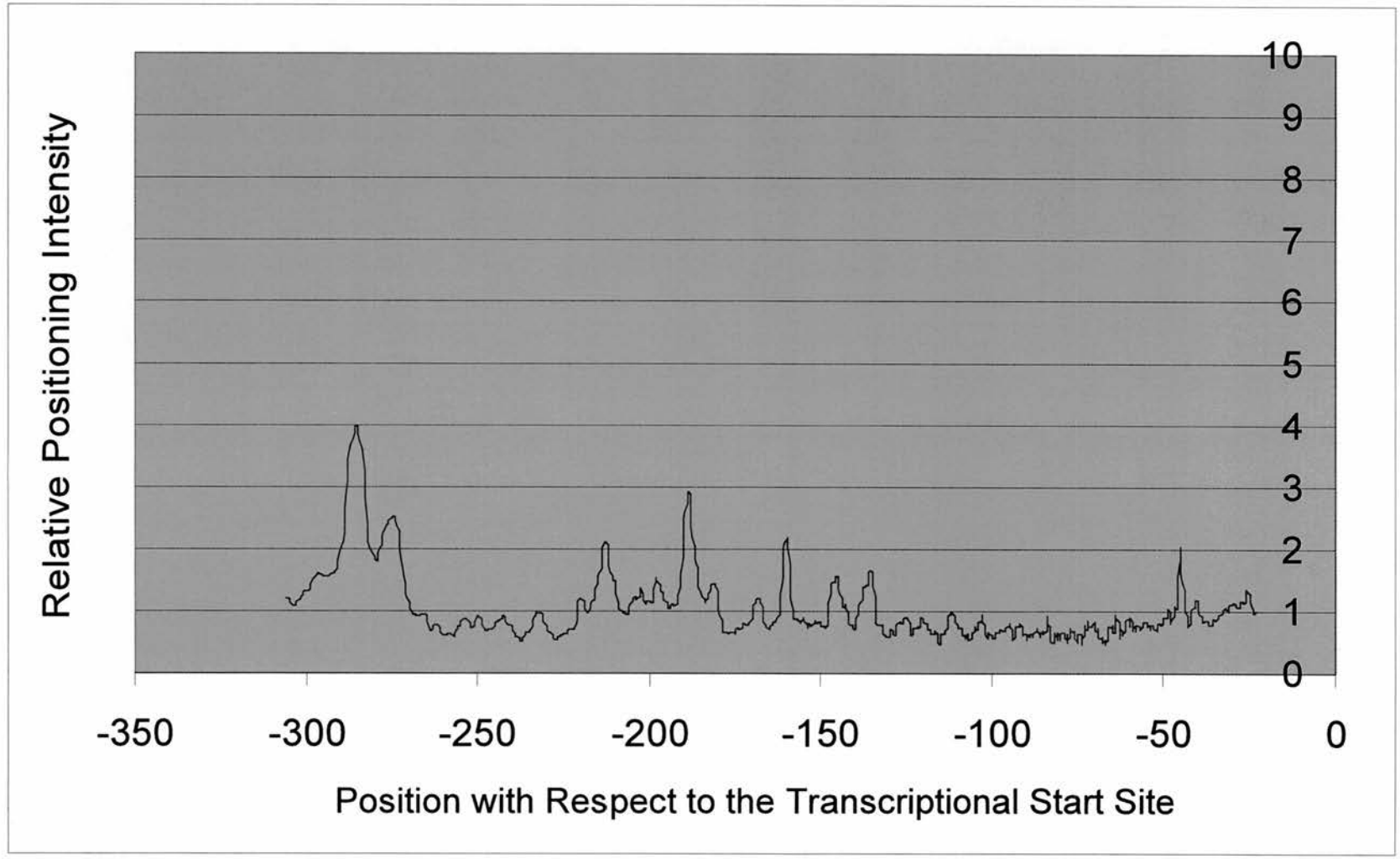



**Figure 90** Map of the *in vitro* Nucleosome Positions over the Caprine β-Casein Promoter Region in pK'bcas (R)

### 7.4.2 *In vivo* analysis

The same OP-Cu digests of goat mammary and liver nuclei generated previously to probe for caprine BLG were used here. However, instead of subsequently digesting the DNA time points with ***Bam*HI**, they were digested with ***Hind*III**. This digests the goat β-casein promoter at –1762bp, generating a genomic fragment spanning the promoter region in excess of 2kb. Primer pairs were designed to amplify a 279bp region abutting the downstream end of the ***Hind*III** site. This was used as the $^{32}$PdCTP end-labelled probe specific to the caprine β-casein gene. If there are any positioned nucleosomes over the goat β-casein promoter region, this probe would highlight them as footprints, bounded on either side by distinct bands.

The samples loaded in the 1.5% agarose gel were in the same order as the previous Southern blot experiments. Again, only time points 2 and 3 from both liver and mammary nuclei titrations were loaded in order to compare the active and inactive states directly. The resulting radiolabelled fragments on the membrane were visualised on photographic film. A diagram of the banding pattern has been drawn to indicate the possible positions of nucleosomes which would correlate best with this data (Figure 91). The positions of the digestion sites on the gene are indicated in table 11.

Although the liver chromatin sample was obscured by non-specific background in regions further up the membrane, nucleosome positioning in bands smaller than 1130bp is still very clear.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 170 | 260 | 350 | 430 | 520 | 590 | 680 | 770 |
| | Position on Gene | -1592 | -1502 | -1412 | -1332 | -1242 | -1172 | -1082 | -992 |
| Mammary Chromatin | Band sizes | 860 | 950 | 1130 | 1310 | 1400 | 1490 | 1580 | 1670 |
| | Position on Gene | -902 | -812 | -632 | -452 | -362 | -272 | -182 | -92 |
| Mammary Chromatin | Band sizes | 2110 | | | | | | | |
| | Position on Gene | +348 | | | | | | | |
| Liver Chromatin | Band sizes | 350 | 520 | 680 | 860 | 1040 | | | |
| | Position on Gene | -1412 | -1242 | -1082 | -902 | -722 | | | |

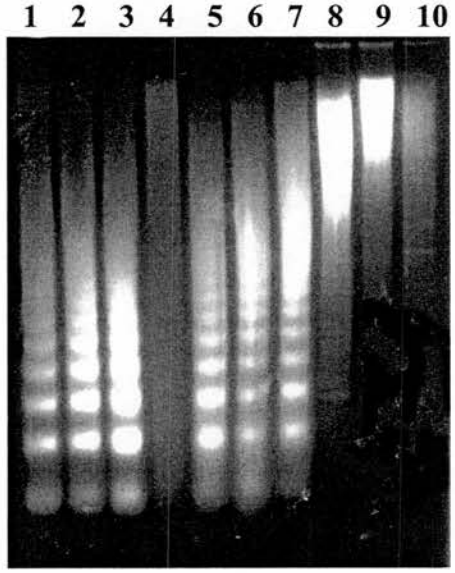
**Table 11** Positions of OP-Cu Digestion Sites on the Caprine β-Casein Gene Promoter

**Figure 91 Nucleosome Mapping from -1.76kb to Downstream of the Caprine β-Casein Gene**

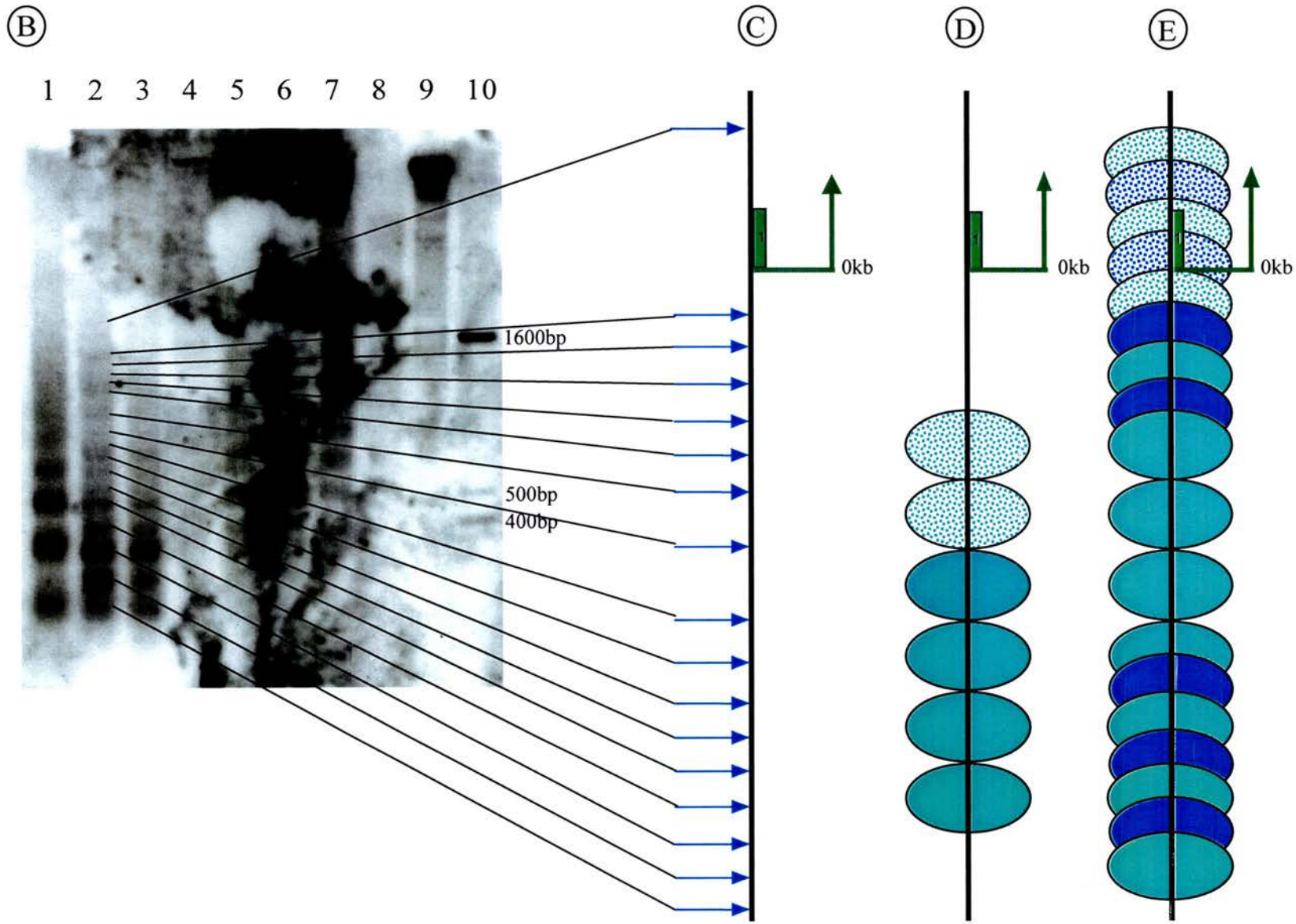
(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver nuclei
(E) Diagram of nucleosome positions on the gene in mammary nuclei




**Lanes:**




1. Caprine mammary nuclei digested with OP-Cu only time point 2
2. Caprine mammary nuclei digested with ***Hind*III** plus OP-Cu; time point 3
3. Caprine mammary nuclei digested with ***Hind*III** plus OP-Cu; time point 2
4. Caprine naked genomic DNA positive control digested with ***Hind*III** plus OP-Cu
5. Caprine liver nuclei digested with ***Hind*III** plus OP-Cu; time point 2
6. Caprine liver nuclei digested with ***Hind*III** plus OP-Cu; time point 3
7. Caprine liver nuclei digested with *OP*-Cu only time point 2
8. Mouse genomic DNA negative control digested with ***Hind*III** only
9. Caprine genomic DNA positive control digested with ***Hind*III** only
10. Mouse genomic DNA digested with ***Hind*III** plus 5μg of 1kb ladder (Boehringer Mannheim)


B
C
D
E
1 2 3 4 5 6 7 8 9 10
1600bp
500bp
400bp
0kb
0kb
0kb

### 7.4.3 Discussion

There is a very strongly positioned nucleosome at +83bp in the *in vitro* map. It is three times the intensity of all other peaks, except the one at -295bp. This peak is under half its intensity, but interestingly it is quite broad, spanning to -284bp, which would be in phase with a nucleosome array positioned over the +83bp site (Figure 92). Within this array, the Stat5 site is situated directly in the centre of a nucleosome (Figure 92), raising the possibility that this may exclude Stat5 access to its cognate site *in vivo*.

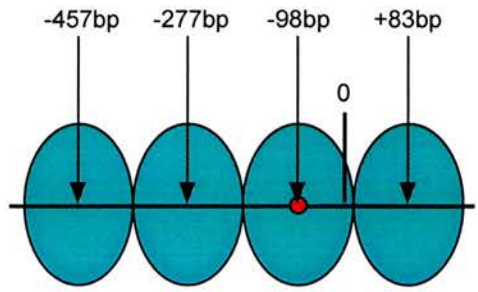



**Figure 92** Nucleosome Array Extrapolated from *in vitro* Data
The Stat5 site at -98bp is indicated by a red circle

A complex array of nucleosomes can be visualised over the promoter region *in vivo* (Figure 91). The liver sample banding patterns are rather unclear above the 1130bp band owing to non-specific hybridisation. However, a clear array is present up until that point which is in phase with the *in vitro* array (Figure 92). This array places the Stat5 binding sites in the centre of a nucleosome.

The nucleosome array in mammary chromatin can be visualised much further over the gene, extending to over 2kb in size (Table 11). Two alternative nucleosome arrays extend between -1592bp and -812bp, exactly out of phase with a 90bp difference. After this point the array indicated by light-blue coloured nucleosome appears to predominate until -452bp where both arrays resume. Between -92bp and +348bp it is very difficult to distinguish individual bands as they are so tightly packed (Figure 91).

There are two principal differences between the nucleosome arrays in the ovine and caprine BLG genes and this gene. Firstly, bands are present over the mammary β-casein gene transcriptional start site, whereas, in both BLG genes there are none. This

could indicate that the goat β-casein gene is not as actively transcribed as the BLG genes. Secondly, an alternative nucleosomal array permeates the β-casein gene in mammary chromatin but not liver chromatin. Conversely, in the BLG genes the nucleosome structure in mammary and liver chromatin is very similar. This may be instrumental in the regulation of goat β-casein discussed below.

The nucleosome array, indicated in Figure 91 by light-blue coloured nucleosomes, positions the Stat5 binding site in the centre of a nucleosome. This array is evident in both liver and mammary chromatin and also in the *in vitro* assay. However, the alternative array in mammary chromatin, indicated by dark-blue coloured nucleosomes, is 90bp out of phase with the light-blue coloured one. This would position the Stat5 binding site directly in the linker region between nucleosomes (Figure 93).

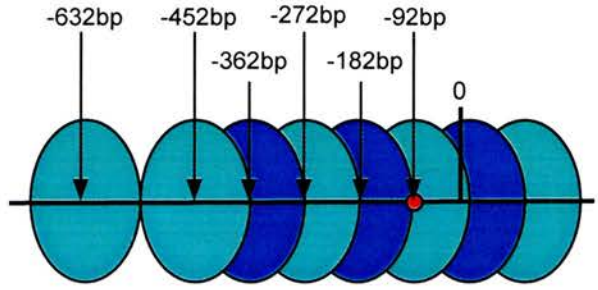



**Figure 93** Nucleosome Array in the Caprine Mammary β-Casein Gene
Positions relative to the transcriptional start site and the Stat5 site (red) are indicated.

The light-blue coloured array may be the more stable array, owing to the strong positioning signal at +83bp. This is present in the inactive gene in liver chromatin and may stabilise the inactive state by excluding transcription factors. The dominant array could shift to the alternative dark-blue coloured array when the gene is activated, allowing Stat5 to possibly bind its cognate site more easily in the linker region. It is unknown whether Stat5 can bind within a nucleosome (**1.13.1**). The mechanism by which the phasing could shift between arrays is unclear, but could be due to long range sequence dependent interactions. A modification to the array further away could generate enough local stability to overcome the positioning signal at +83bp. Alternatively, a chromatin remodelling complex may interact directly with the gene, shifting and stabilising the alternative array structure.

## 7.5 Mouse β-Casein

### 7.5.1 *In vitro* analysis

The mouse β-casein promoter region was excised from -545 (***Nsi*I**) to +95 (***Nsi*I**) with respect to the transcriptional start site. Using the ligation-compatible ***Pst*I** site in the vector, it was inserted into pBluescript II KS (-) in both orientations (Figure 94).

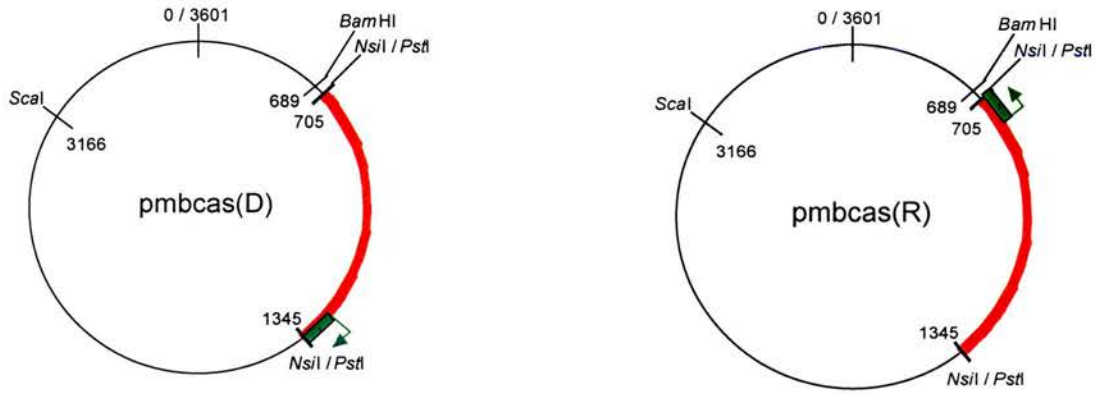



**Figure 94** Constructs Generated which Contain the Mouse β-Casein Promoter Region in both Orientations

The isolated monomer DNA fragments were electrophoresed through an agarose gel to check for contaminating dimer fragments or degraded DNA beneath the 146bp band. The ssDNA was tested in a similar manner for dsDNA contamination or degradation. Both proved to be of sufficient quality to proceed with the monomer extension assay (Figure 95).

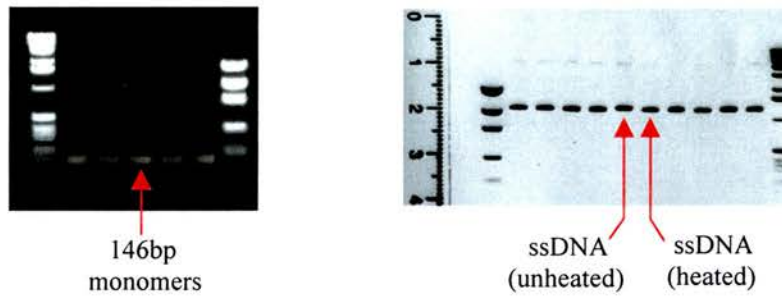



**Figure 95** Monomer DNA fragments and ssDNA from pmbcas (D) and (R)

Each construct was tested for nucleosome positioning by conducting the monomer extension reaction in the presence of ***Bam*HI**, which acts as a reference point in order to

locate the nucleosome binding sites. In addition, each reaction was also completed in the absence of restriction enzyme in case the DNA polymerisation reaction terminated prematurely at any specific point, generating fragments which do not relate to nucleosome positions (Figure 96).

The banding patterns generated were analysed using *Aida* and *Sigma-plot* software and the intensity of each position was plotted in relation to the centre of that nucleosome (as chapter 5). The nucleosome positions on constructs pK'bcas(D) (Figure 97) and pK'bcas(R) (Figure 98) were plotted in relation to the transcriptional start site of the gene.

### 7.5.2 *In vivo* analysis

Mouse liver and mammary tissues were isolated from a (CBA x C57Bl/6)F1 non-transgenic mouse and the nuclei isolated by standard procedures (**2.5**). They were digested with OP-Cu as before, with aliquots taken at the following time points:

time point 0: 0 minutes OP-Cu digestion
time point 1: 10 minutes OP-Cu digestion
time point 2: 20 minutes OP-Cu digestion
time point 3: 30 minutes OP-Cu digestion
time point 4: 40 minutes OP-Cu digestion
time point 5: 60 minutes OP-Cu digestion

As in previous experiments, only time points 2 and 3 were loaded from both tissues, so that they could be compared directly on the same blot. These time points were further digested with restriction enzyme ***Xmn*I** which digests the gene at +1349bp within the transcribed region. A primer pair was designed to amplify a 324bp region abutting the upstream side of the ***Xmn*I** site for use in the indirect end-labelling reaction. The results are displayed in the same format as in previous experiments (Figure 99) (Table 12).

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 110 | 290 | 470 | 650 | 830 | |
| | Position on Gene | +1240 | +1060 | +880 | +700 | +520 | |
| Liver Chromatin | Band sizes | 110 | 290 | 470 | 650 | 830 | 1010 |
| | Position on Gene | +1240 | +1060 | +880 | +700 | +520 | +340 |

**Table 12** Positions of OP-Cu Digestion Sites on the Mouse β-casein Gene Promoter

Lanes:

1. Lambda ***Dde*I** marker
2. Lambda ***Hinf*I** marker
3. mbcas (D) & ***Bam*HI**
4. mbcas (D) & $H_2O$
5. mbcas (R) & ***Bam*HI**
6. mbcas (R) & $H_2O$
7. Lambda ***Dde*I** marker
8. Lambda ***Hinf*I** marker

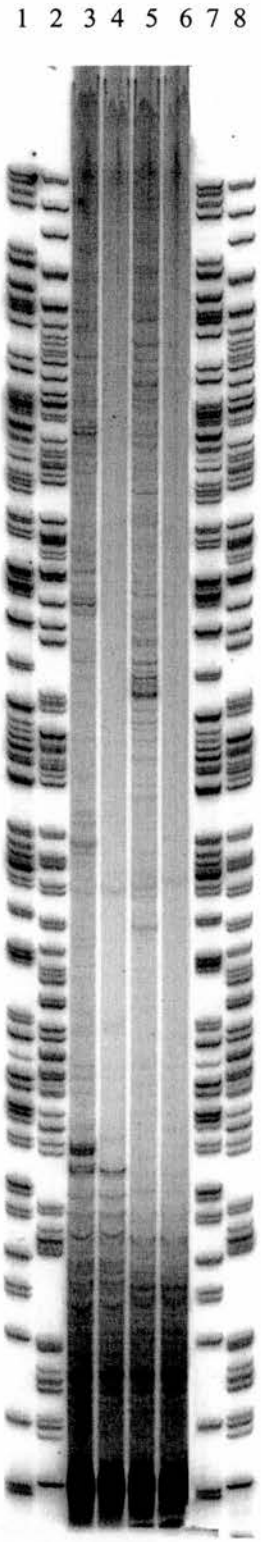



**Figure 96** Monomer Extension Reactions of the mbcas (D) and (R) Constructs

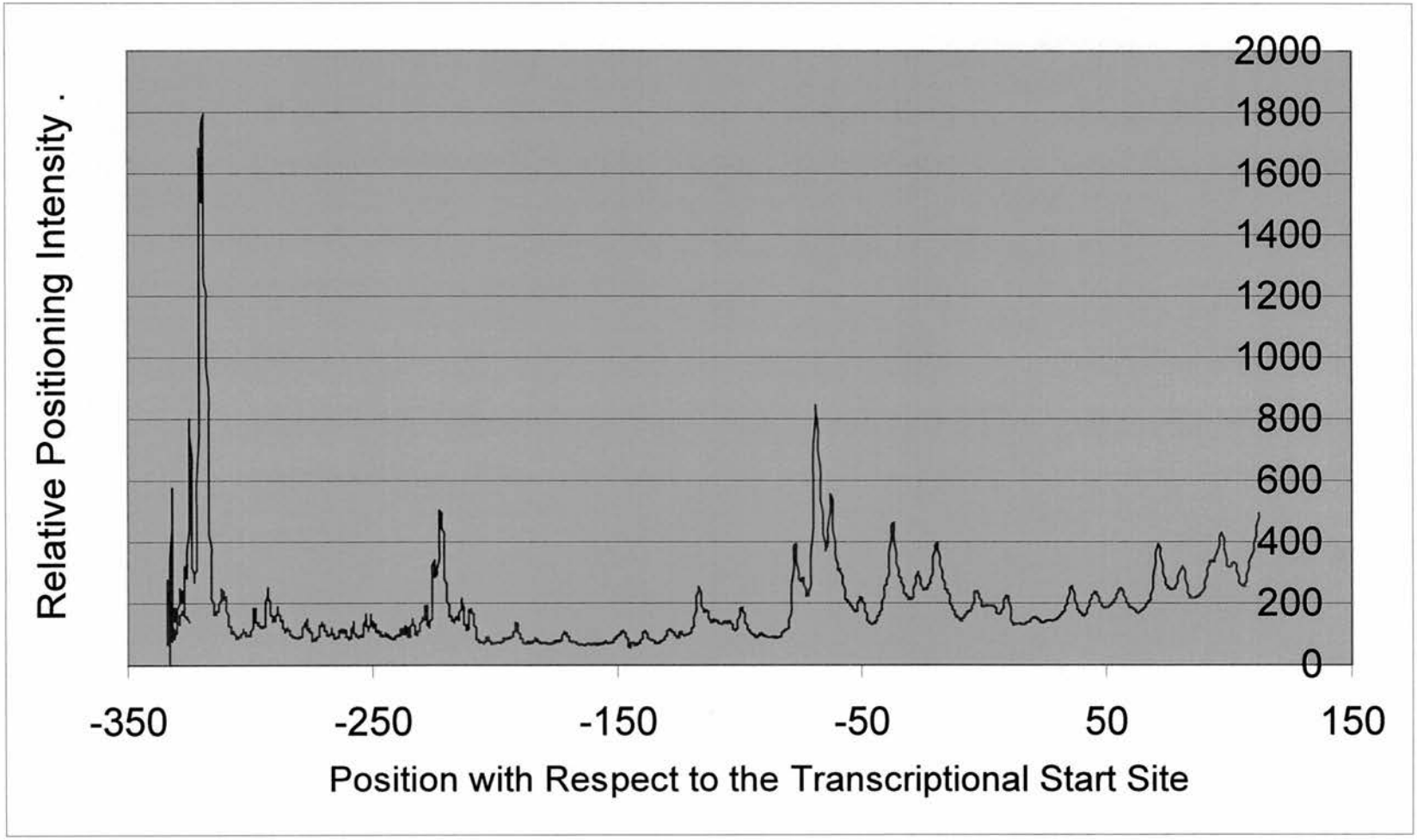



**Figure 97** Map of the In Vitro Nucleosome Positions over the Mouse β-casein Promoter Region in the pmbcas (D) Construct

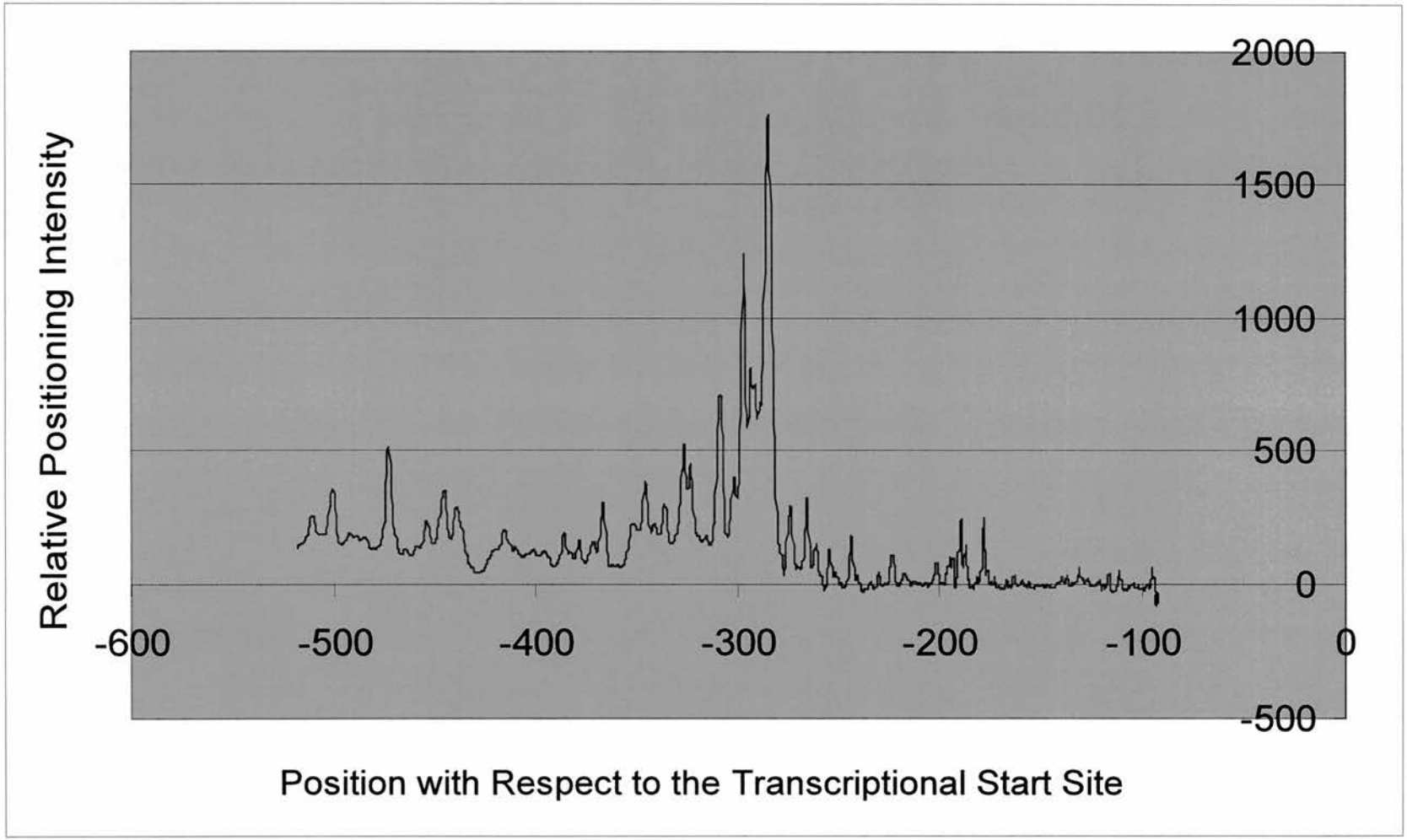


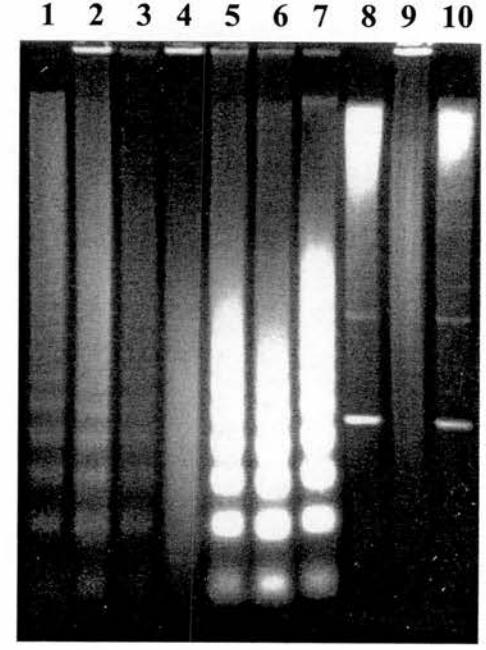

**Figure 98** Map of the In Vitro Nucleosome Positions over the Mouse β-casein Promoter Region in the pmbcas (R) Construct

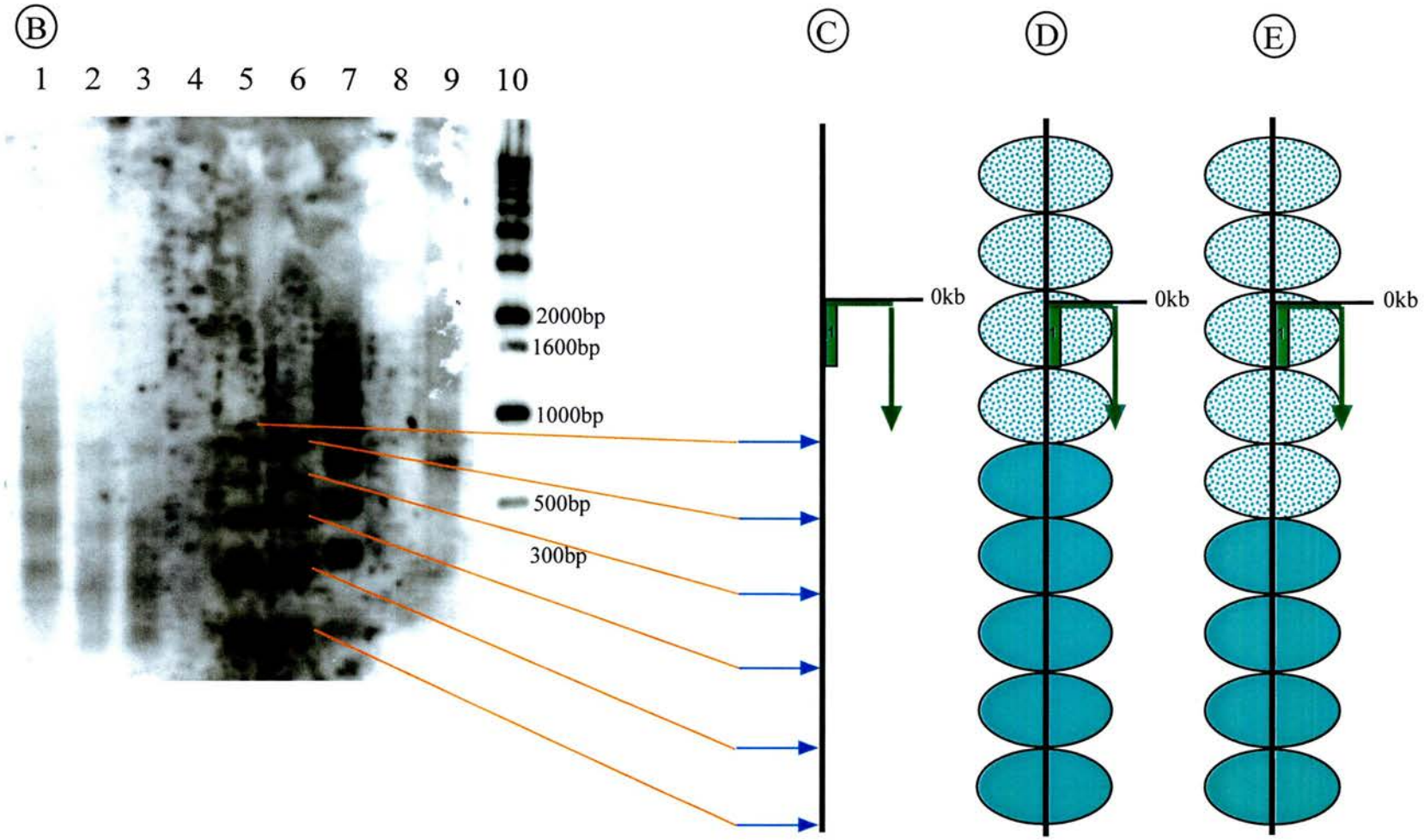
**Figure 99 Nucleosome Mapping from 1.34kb to Upstream of the Mouse β-Casein Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the gene
(D) Diagram of nucleosome positions on the gene in liver nuclei
(E) Diagram of nucleosome positions on the gene in mammary nuclei


**Lanes:**

1. Mouse mammary nuclei digested with *OP*-Cu only time point 2
2. Mouse mammary nuclei digested with ***Xmn*I** plus OP-Cu; time point 3
3. Mouse mammary nuclei digested with ***Xmn*I** plus OP-Cu; time point 2
4. Mouse naked genomic DNA positive control digested with ***Xmn*I** plus OP-Cu
5. Mouse liver nuclei digested with ***Xmn*I** plus OP-Cu; time point 2
6. Mouse liver nuclei digested with ***Xmn*I** plus OP-Cu; time point 3
7. Mouse liver nuclei digested with *OP*-Cu only time point 2
8. Porcine genomic DNA negative control digested with ***Xmn*I** only
9. Mouse genomic DNA positive control digested with ***Xmn*I** only
10. Porcine genomic DNA digested with ***Xmn*I** plus 5μg of 1kb ladder (Boehringer Mannheim)




B
1 2 3 4 5 6 7 8 9 10
2000bp
1600bp
1000bp
500bp
300bp
C
0kb
D
0kb
E
0kb

### 7.5.3 Discussion

The dominant positioning signal *in vitro* has its dyad centred at -284bp, with a substantial sub-band at -297bp (Figure 98). The repeat length of this major nucleosome position spans from -194bp to -374bp (Figure 100). An array which extends out from this strongly positioned nucleosome will align the Stat5 site in the centre of an adjacent nucleosome. This is very similar to the goat β-casein gene, whose Stat5 site is also located at the centre of a nucleosome at the same position.

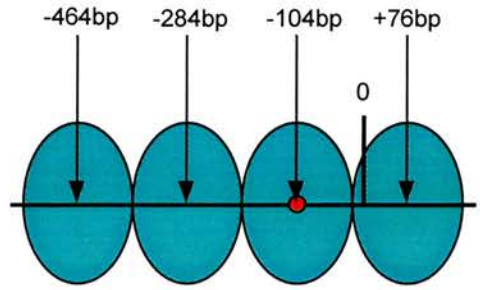



**Figure 100** Nucleosome Array in the Mouse Mammary β-Casein Gene
Positions relative to the transcriptional start site and the Stat5 site (red at -99bp) are indicated.

The *in vivo* map extends from +1240bp to +340bp (Table 12) in liver chromatin with only a single phased array evident (Figure 99). This array does not extend over the region mapped *in vitro*, but if extrapolated both *in vivo* and *in vitro* arrays do align. Therefore, in liver chromatin, the Stat5 site is probably aligned to the centre of a nucleosome.

The array is identical in the mammary gene chromatin, but can only be visualised up to +520bp with respect to the transcriptional start site. Like the goat β-casein gene, an alternative array may exist over the promoter region which allows Stat5 to bind its cognate site within the linker region. This is quite feasible since the array which is evident in the mouse β-casein gene aligns almost exactly with one in the goat β-casein gene. The Stat5 site in both genes is present at the same position and in both cases is positioned at the centre of a nucleosome in this array. Further investigation upstream of the promoter region should reveal if this is the case.

## 7.6 Mouse α-Lactalbumin

### 7.6.1 *In vitro* analysis

The mouse α-lactalbumin promoter region was cloned from -561 (***Bam*HI**) to +50 (***Bam*HI**) and inserted into the pBluescript II KS (-) vector at the ***Bam*HI** site. This created pmalac(R). This phagemid was digested with ***Bss*HII** and re-ligated to generate pmalac(R) in the opposite orientation (Figure 101).

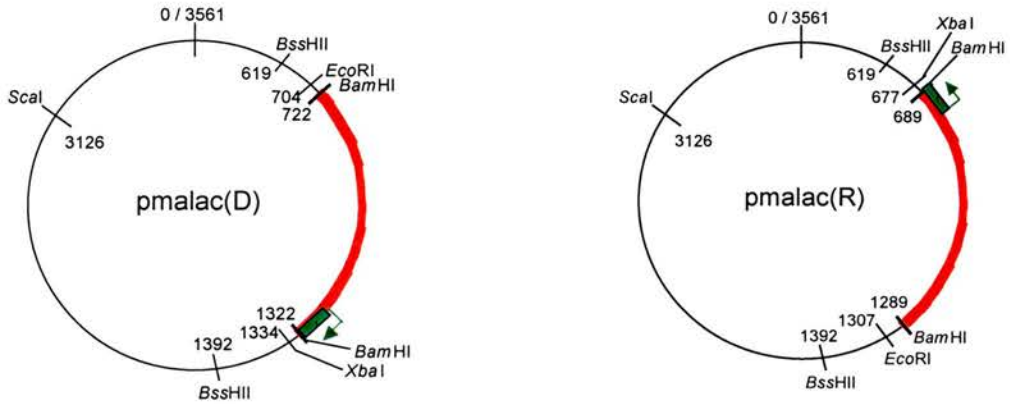



**Figure 101** Constructs Generated which Contain the Mouse α-Lactalbumin Promoter Region in both Orientations

The monomer DNA fragments and ssDNA generated from both constructs were checked on an agarose gel before proceeding to the extension reaction step (Figure 102).

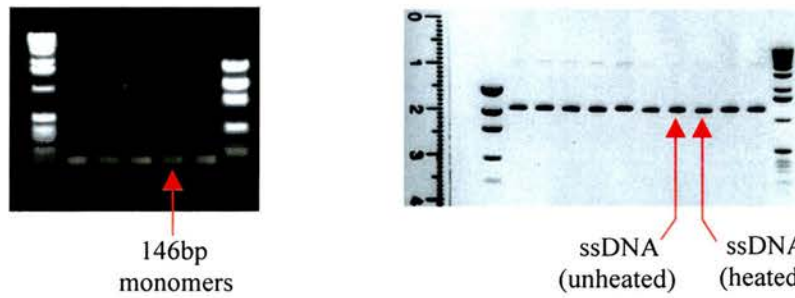
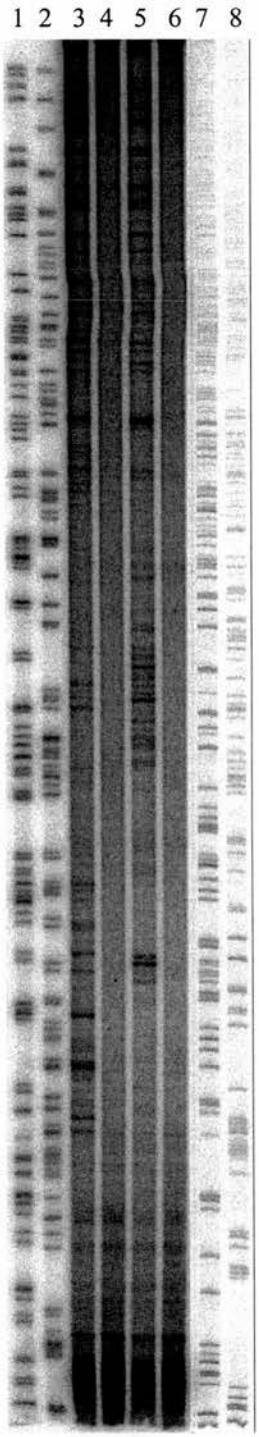



**Figure 102** Monomer DNA fragments and ssDNA from pmalac (D) and (R)

***Eco*RI** was included in the pmalac(D) monomer extension reaction and ***Xba*I** in the pmalac(R) reaction to act as a reference point in order to locate nucleosome positions on the constructs. Each reaction was also conducted with water in place of the restriction enzyme as a control (Figure 103).




Lanes:

1. Lambda ***Dde*I** marker
2. Lambda ***Hinf*I** marker
3. malac (D) & ***Eco*RI**
4. malac (D) & $H_2O$
5. malac (R) & ***Xba*I**
6. malac (R) & $H_2O$
7. M13 'C' nucleotide marker
8. M13 'T' nucleotide marker

**Figure 103** Monomer Extension Reactions of the pmalac (D) and (R) Constructs

The monomer extension bands were sized using the M13 sequencing markers and the bacteriophage lambda digestion markers with the *Sigma-plot* software package. In addition, the bands were quantified with the *Aida* densitometer software. The resulting peaks, which represent the relative nucleosome positioning intensities, were plotted with respect to the centre of each nucleosome relative to the start site of the gene. pmalac(D) (Figure 104) and pmalac(R) (Figure 105) maps overlap between -250bp and -200bp. This is most obvious at the large peak at approximately -200bp.

### 7.6.2 Mouse α-Lactalbumin: in vivo analysis

The mouse mammary and liver OP-Cu digestion time points generated in the previous experiment for the mouse β-casein gene were also used here. They were further digested with restriction enzyme ***Xba*I** which digests the mouse α-lactalbumin gene at +1257bp within the transcribed region. A primer pair was designed to amplify a 378bp region abutting the upstream side of the ***Xba*I** site for use in the subsequent indirect end-labelling reaction.

The radioactively labelled membrane was exposed to photographic film to reveal any banding pattern over the OP-Cu & ***Xba*I** digested mouse α-lactalbumin promoter region. This could indicate the presence of positioned nucleosomes. The band sizes were calculated in relation to the marker lane and are shown in table 13. Possible nucleosome positions over the promoter region are shown in relation to the photographic image of the membrane (Figure 106).

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Mammary Chromatin | Band sizes | 360 | 540 | 720 | 900 | 980 | 1080 | 1160 | 1260 | 1440 |
| | Position on Gene | +897 | +717 | +537 | +357 | +277 | +177 | +97 | -3 | -183 |
| Liver Chromatin | Band sizes | 180 | 360 | 540 | 720 | 900 | 1080 | 1260 | 1440 | 1620 |
| | Position on Gene | +1077 | +897 | +717 | +537 | +357 | +177 | -3 | -183 | -363 |

**Table 13** Positions of Digestion Sites on the Mouse α-Lactalbumin Gene Promoter

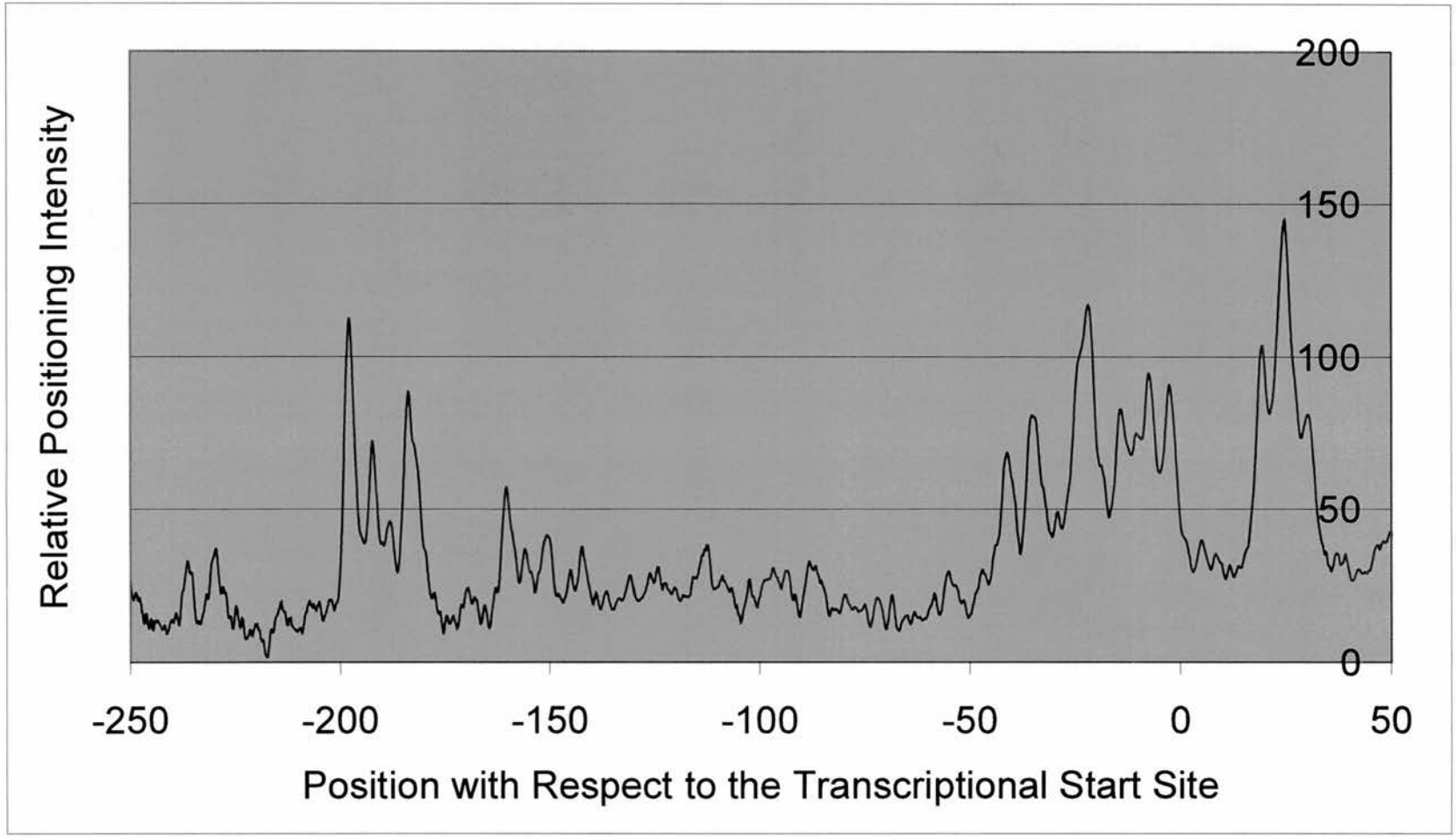



**Figure 104** Map of the *In Vitro* Nucleosome Positions over the Mouse α-lactalbumin Promoter in the pmalac(D) Construct

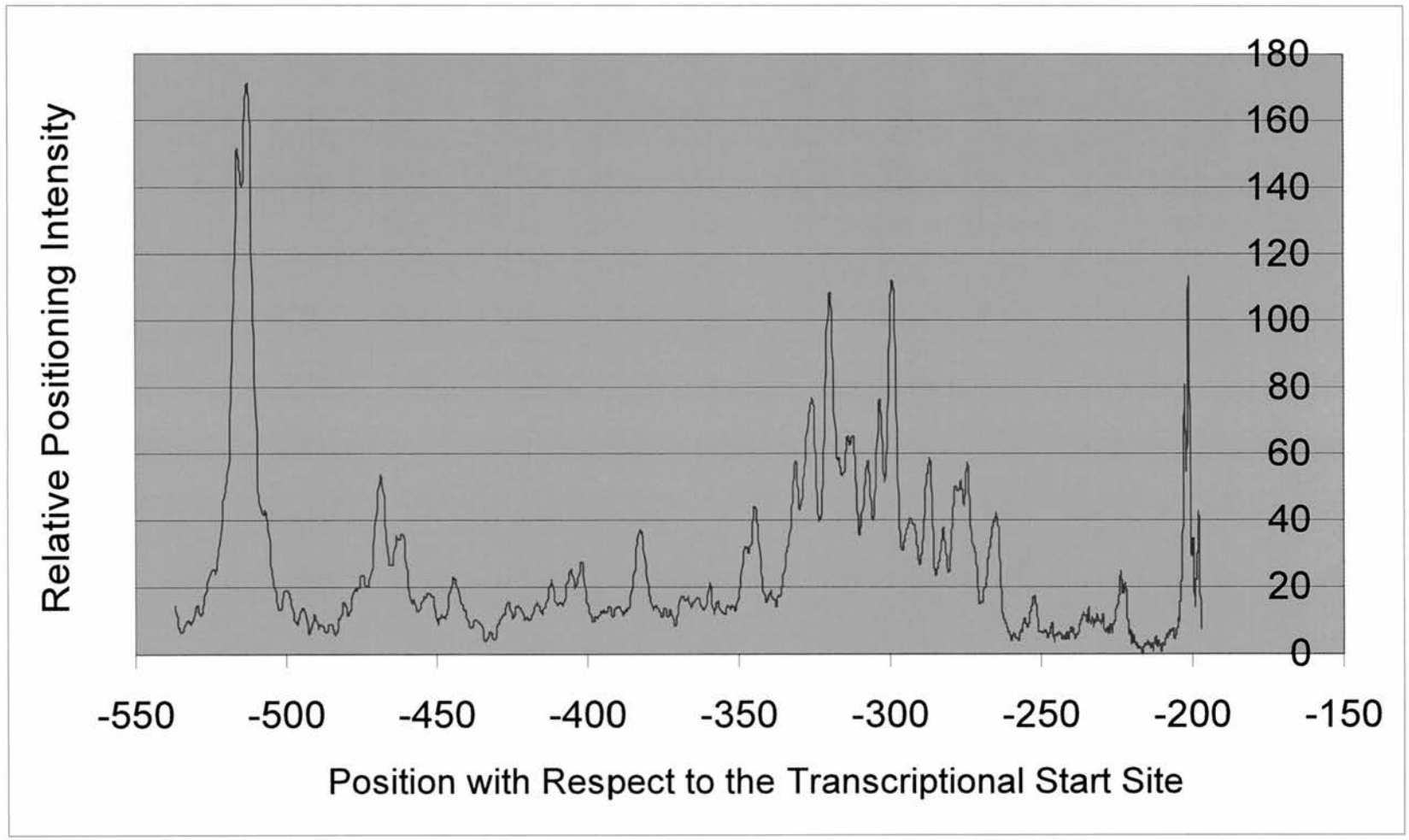


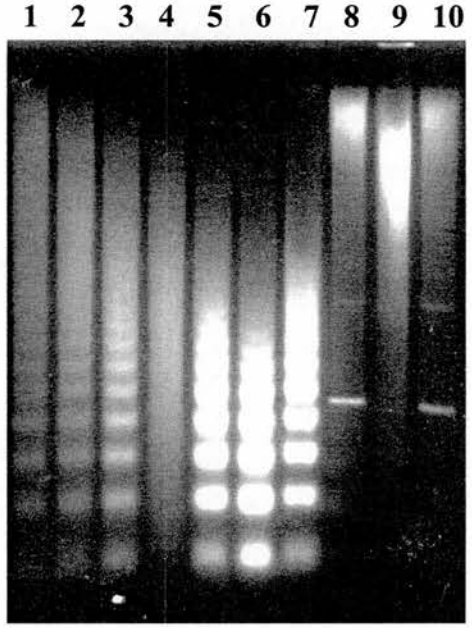

**Figure 105** Map of the *In Vitro* Nucleosome Positions over the Mouse α-lactalbumin Promoter in the pmalac(R) Construct

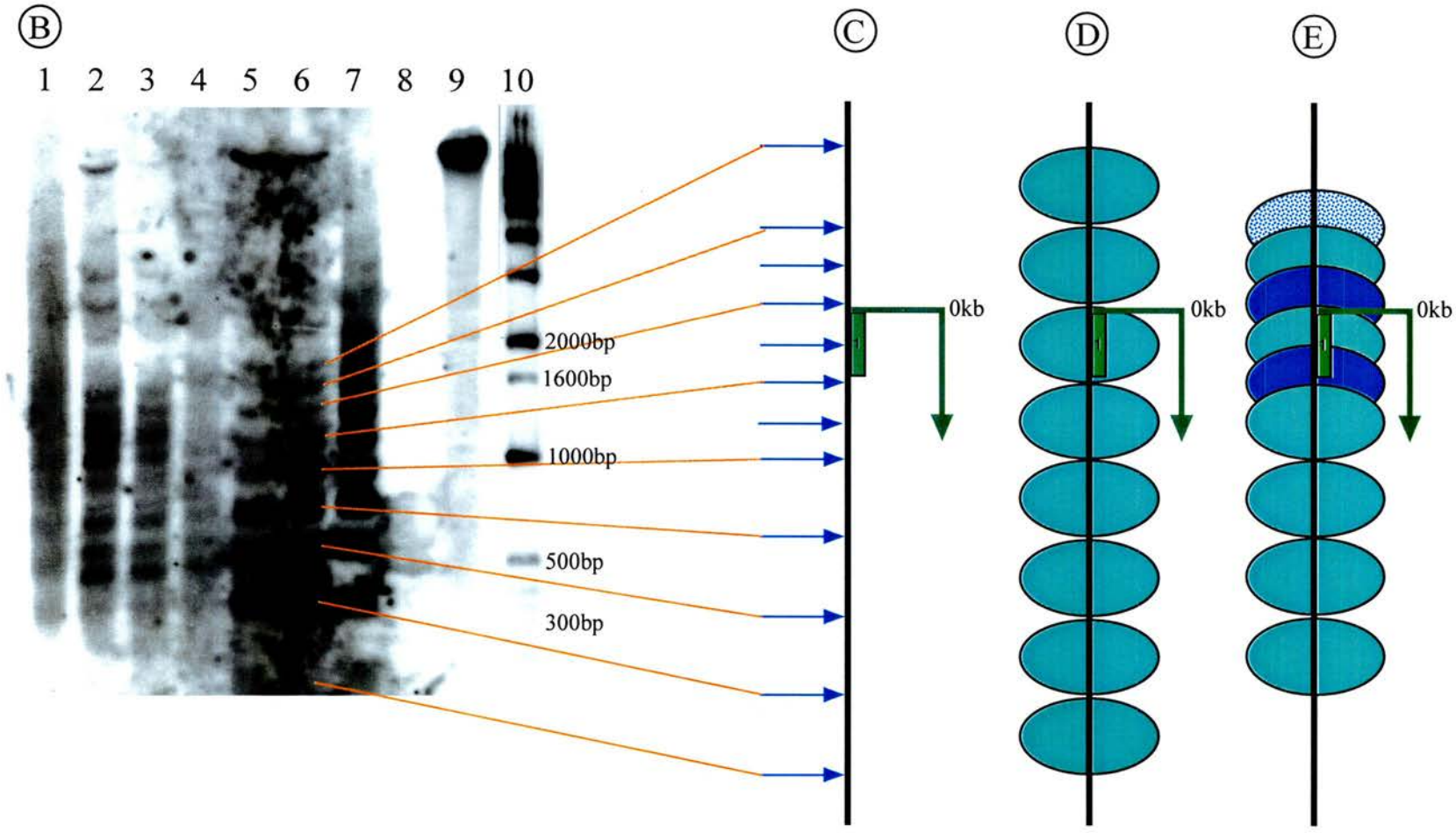
**Figure 106 Nucleosome Mapping from +1.26kb to Upstream of the Mouse α-Lactalbumin Gene**

(A) Ethidium-stained gel before blotting
(B) Exposure of indirect end-labelled Southern blot (sizes indicated)
(C) OP-Cu digestion sites in relation to the BLG gene
(D) Diagram of nucleosome positions on the gene in liver nuclei
(E) Diagram of nucleosome positions on the gene in mammary nuclei


**Lanes:**

1. Mouse mammary nuclei digested with *OP*-Cu only time point 2
2. Mouse mammary nuclei digested with ***Xba*I** plus OP-Cu; time point 3
3. Mouse mammary nuclei digested with ***Xba*I** plus OP-Cu; time point 2
4. Mouse naked genomic DNA positive control digested with ***Xba*I** plus OP-Cu
5. Mouse liver nuclei digested with ***Xba*I** plus OP-Cu; time point 2
6. Mouse liver nuclei digested with ***Xba*I** plus OP-Cu; time point 3
7. Mouse liver nuclei digested with *OP*-Cu only time point 2
5. Porcine genomic DNA negative control digested with ***Xba*I** only
9. Mouse genomic DNA positive control digested with ***Xba*I** only
10. Porcine genomic DNA digested with ***Xba*I** plus 5μg of 1kb ladder (Boehringer Mannheim)




B
1 2 3 4 5 6 7 8 9 10
2000bp
1600bp
1000bp
500bp
300bp
C
0kb
D
0kb
E
0kb

### 7.6.3 Discussion

The distribution of peaks in this analysis is rather more complex than in the previous genes. The most prominent peak in the *in vitro* analysis of the promoter region is positioned at -512bp. In addition to this, there are several clusters of peaks, distributed around a major peak at: -319bp, -298bp, -200bp, -184bp, -22bp, -7bp and +24bp. These can be aligned into the following three nucleosome arrays (Figure 107):

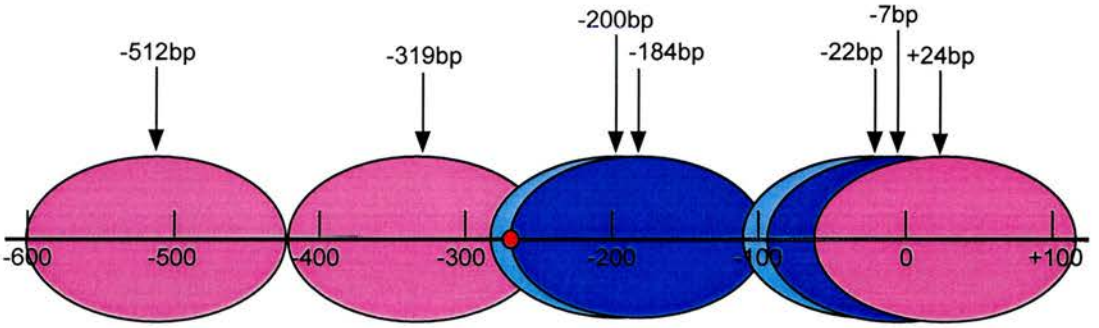



**Figure 107** Nucleosome Array Extrapolated from *in vitro* Data of the Mouse α-Lactalbumin Gene Positions are relative to the transcriptional start site; the Stat5 site at –267bp is indicated by a red circle.

Only the array represented by dark-blue coloured nucleosomes aligns with the *in vivo* data (Figure 107). *In vitro* these are centrally positioned at -184bp and -7bp and *in vivo* the positions are at +7bp, +187bp and if extrapolated, -173bp. Although these positions vary by approximately 12bp, this is within the error margin.

Only the active gene in mammary chromatin displays this dark-blue coloured nucleosomal array. It starts at +277bp and extends over the transcriptional start site to the limit of the assay at -83bp. It is unknown how far upstream it persists. Notably, this array positions the Stat5 site in the linker region between nucleosomes. *In vivo* there is also an alternative light-blue coloured array in the active gene, which is approximately 90bp out of phase with the former array. Therefore, this array positions the Stat5 site close to the dyad of the nucleosome.

This light-blue coloured array is the only one evident in the inactive gene in liver chromatin. This array may deny Stat5 access to its cognate site when the gene is not actively transcribing. This array might be present in mammary chromatin when the gene is not highly active.

## 7.7 Human α-Lactalbumin

### 7.7.1 *In vitro* analysis

The gene promoter region was cloned from -570 (***Acc*I**) to +113 (***Pvu*II**) and inserted in one orientation into pBluescript II KS (-) between the ***Acc*I** and ***Eco*RV** sites to produce phalac(R). This construct was digested with ***Bss*HII** and re-ligated to produce phalac(R), which contains the insert in the opposite orientation (Figure 108).

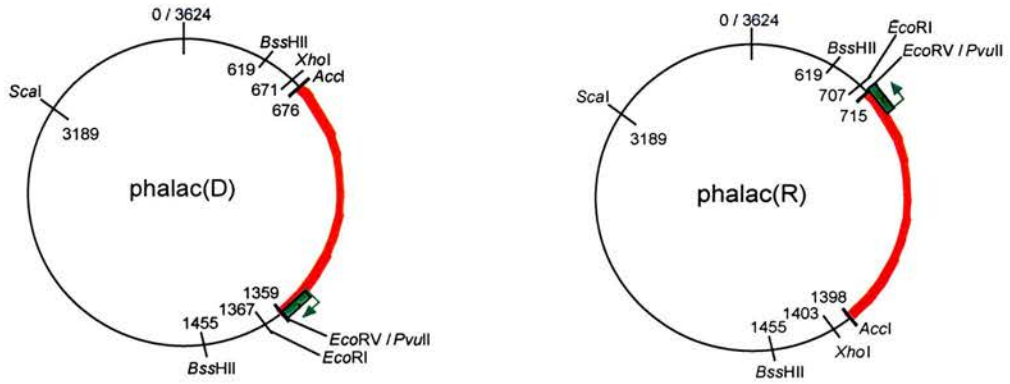



**Figure 108** Constructs Generated which Contain the Human α-Lactalbumin Promoter Region in both Orientations

The monomer DNA fragments and the ssDNA isolated from both constructs (Figure 109) was of sufficient quality to proceed with the monomer extension reaction. ***Xho*I** was included in the phalac(D) extension reaction and ***Eco*RI** in the phalac(R) extension reaction. These act as a reference point in sizing the resulting fragments (Figure 110).

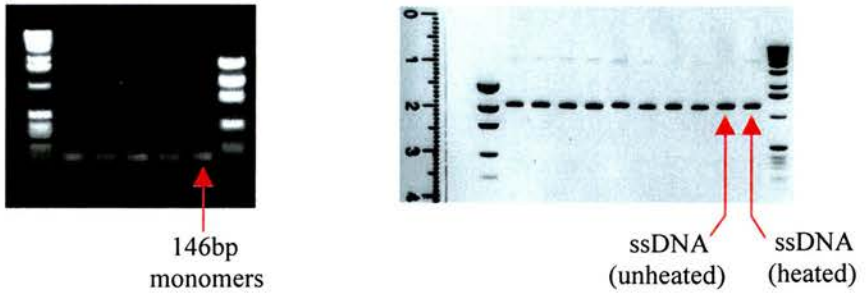



**Figure 109** 146bp Monomer DNA Fragments and ssDNA from phalac(D) and phalac(R) Constructs

The positions and relative intensities of the bands within the insert were plotted relative to each other for both the phalac(D) (Figure 111) and the phalac(R) (Figure 112) constructs.

Lanes:

1. 100bp marker
2. halac (D) & ***Xho*I**
3. halac (D) & $H_2O$
4. halac (R) & ***Eco*RI**
5. halac (R) & $H_2O$
6. Lambda ***Dde*I** marker
7. Lambda ***Hinf*I** marker
8. 100bp marker

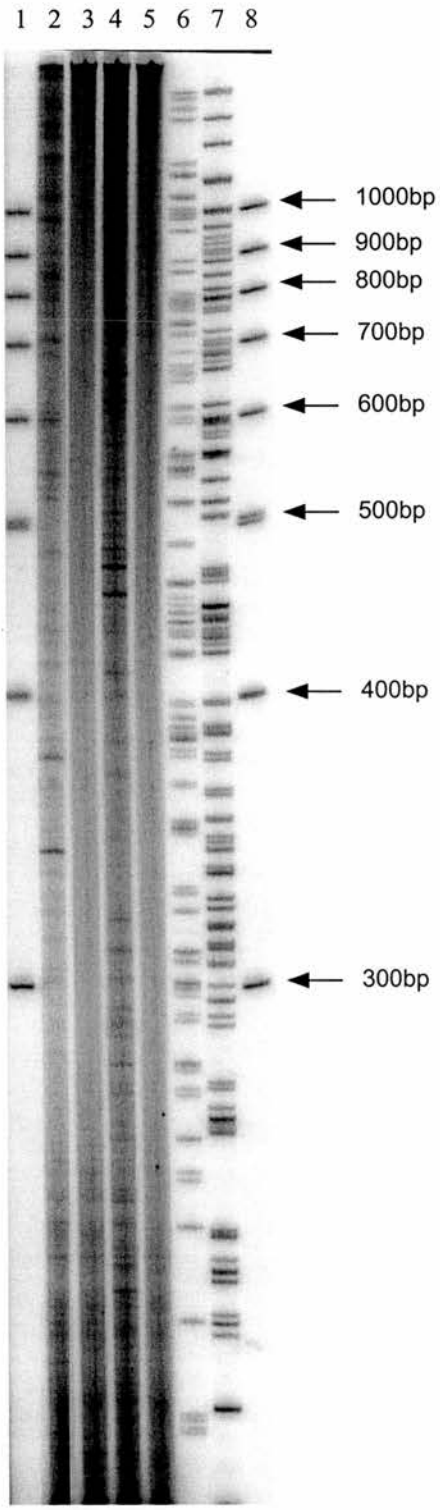



**Figure 110** Monomer Extension Reactions of the halac (D) and (R) Constructs

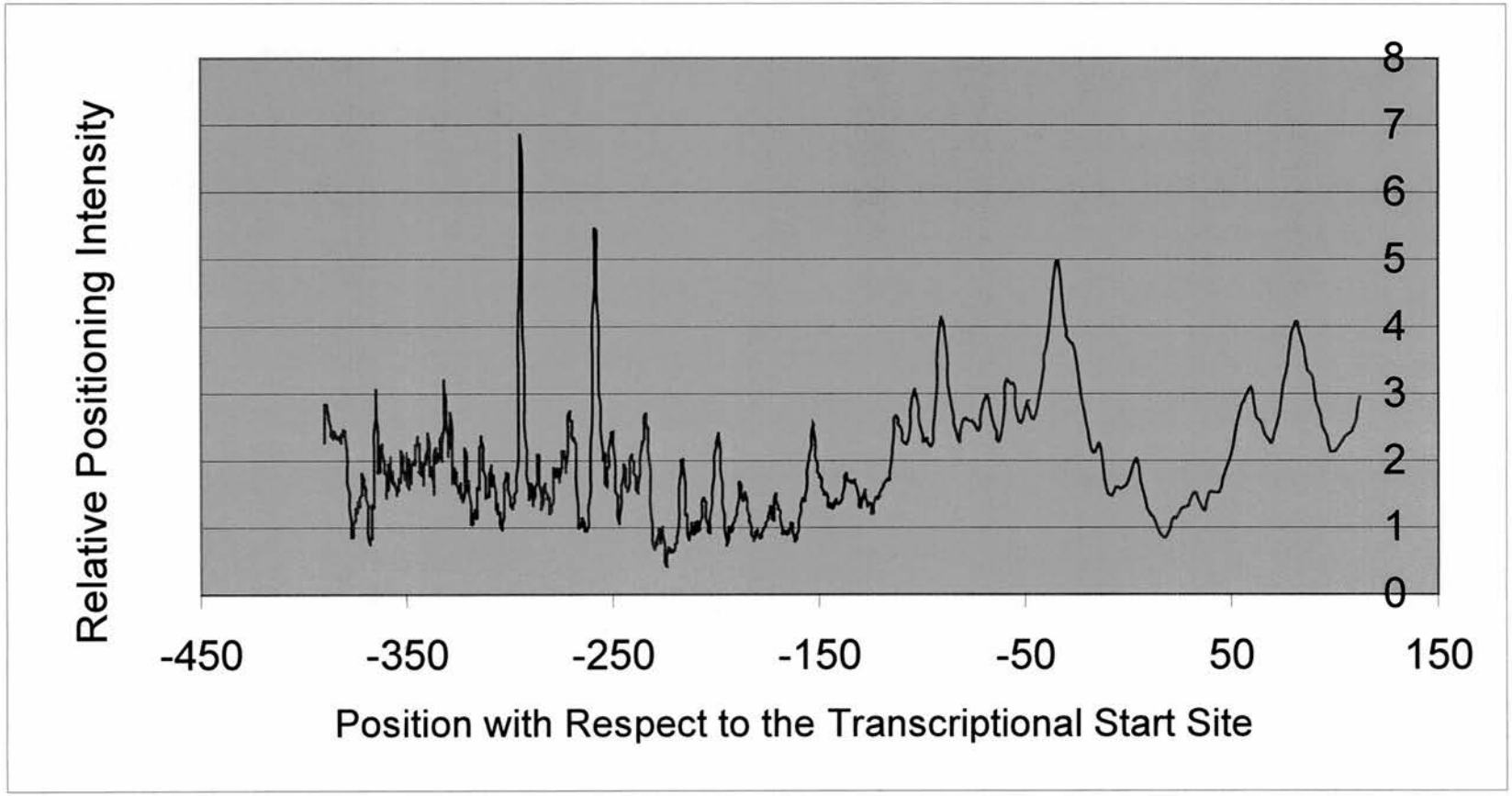



**Figure 111** Map of the *In Vitro* Nucleosome Positions over the Human α-lactalbumin Promoter in the phalac(D) Construct

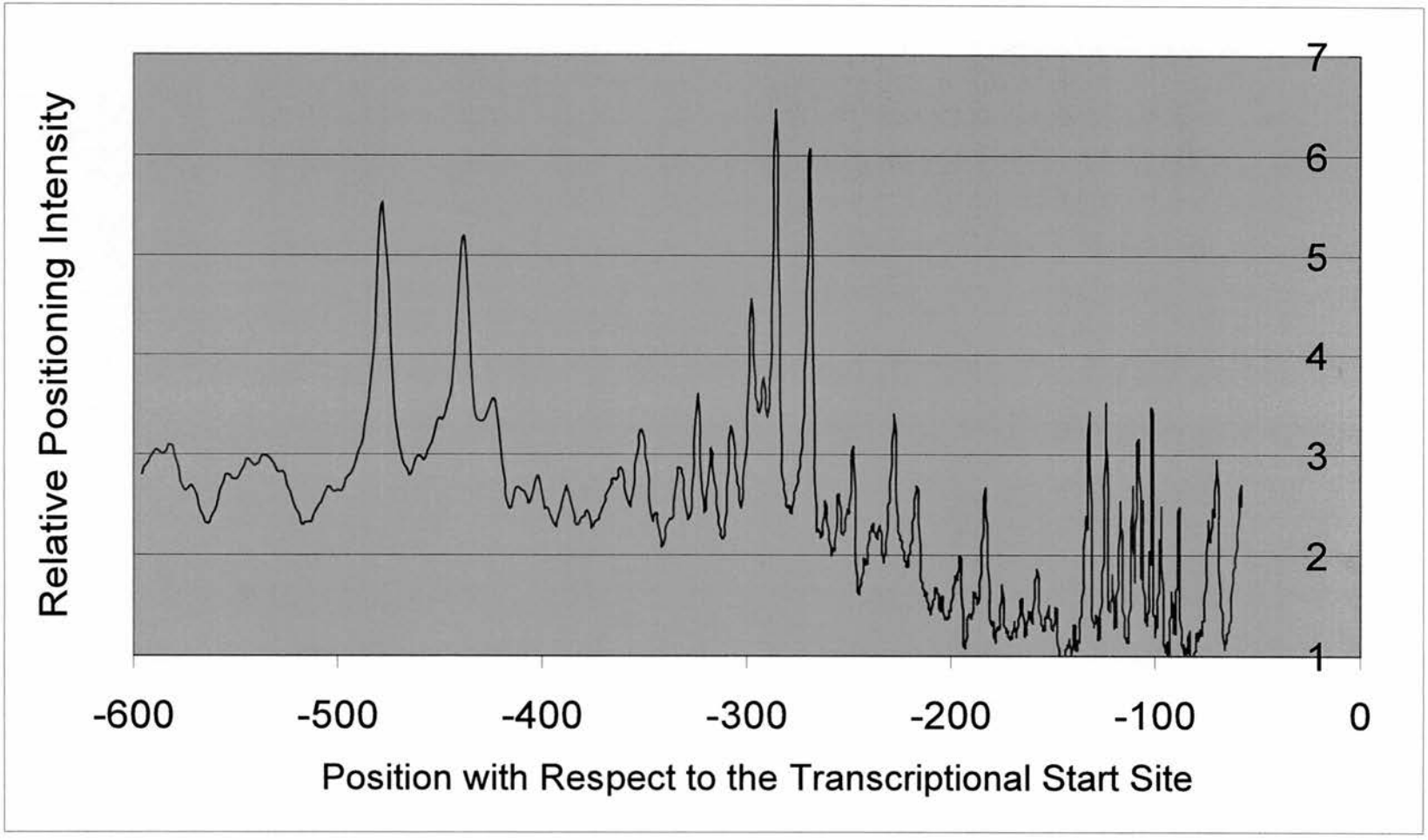



**Figure 112** Map of the *In Vitro* Nucleosome Positions over the Human α-lactalbumin Promoter in the phalac(R) Construct

### 7.7.2 Discussion

Only an *in vitro* map was generated over the human α-lactalbumin promoter region owing to the difficulty in obtaining tissue for *in vivo* work. The map displays several prominent peaks at positions: -478bp, -438bp, -290bp, -265bp, -91bp, -35bp and +81bp. Nucleosomes can be aligned to fit this series of peaks as follows (Figure 113).

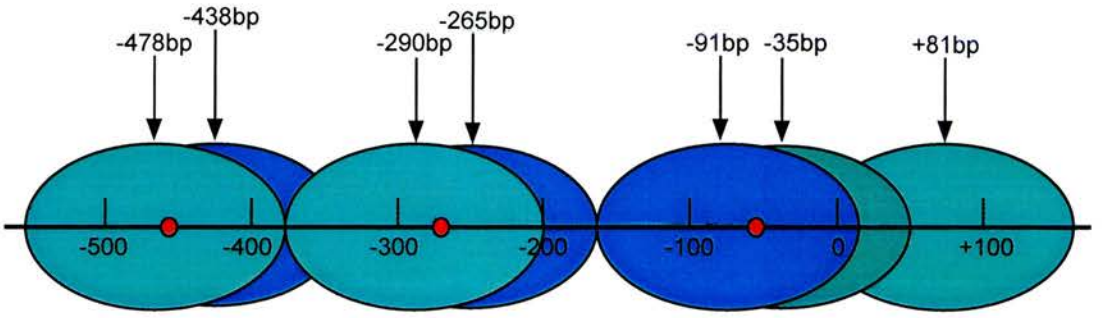



**Figure 113** Nucleosome Array Extrapolated from *in vitro* Data of the Human α-Lactalbumin Gene Positions are relative to the transcriptional start site; the Stat5 sites (at –62bp, -278bp & -457bp) are indicated by red circles.

The peaks at -478bp, -290bp and +81bp could form an array (coloured light-blue) which is consistent with a repeat length of approximately 190bp. In addition, the peaks at -438bp, -265bp and -91bp could also form an alternative array with a repeat length of approximately 180bp. The average repeat length in leukocytes is 185bp +/-2bp (Kahmann & Rake, 1993) and in HeLa cells approximately 191bp (Compton *et al.*, 1976; Tate & Philipson, 1979) . All Stat5 sites would be positioned close to the dyad axis in either of these arrays. The nucleosome position at -35bp does not appear to align with any other positions although it is quite prominent in the analysis (Figure 111).

If Stat5 preferentially binds its cognate site within the linker region between nucleosomes, this alignment may assist in inhibiting Stat5-mediated gene activation. In this case, chromatin remodelling might be required in order to allow Stat5 access to its site.

## 7.8 Summary

### 7.8.1 Positions of Stat5 Sites in vitro

Using the monomer extension technique, all of the aforementioned gene promoter regions were analysed for sequence-dependent nucleosome positioning. In each case, nucleosome arrays were calculated which fitted the data in a 180bp periodicity In addition, the human α-lactalbumin promoter region was also tested. In this case the data was fitted differently to accommodate the human nucleosome repeat (see above).

*The BLG Genes*

Both the ovine and caprine BLG genes displayed a single, strong positioning signal at about -180bp. This aligns with the array present in both the active and inactive genes and positions the outer Stat5 sites within the linker region on either side of a nucleosome. The active gene may possess an additional array which places only the middle site within the linker region. This change in array might lead to differential access to Stat5 binding sites.

*The β-Casein Genes*

The goat β-casein and mouse β-casein genes both displayed a single, strong positioning signal. In each case, this fitted with smaller peaks in the *in vitro* map to generate a putative array. Notably, in both cases this aligned with the array present in both the active and the inactive gene (Table 14). It is feasible that this is the default state, where the Stat5 site is hidden close to the nucleosomal dyad. Chromatin remodelling may switch array settings between the active and inactive states, allowing Stat5 to bind its site more easily within the linker region on gene activation.

*The α-Lactalbumin Genes*

The mouse α-lactalbumin gene does not possess one dominant positioning signal in its promoter region. The two strongest signals are at +24bp and -512bp which would position the Stat5 site in the linker region. However, as there are many conflicting

positioning signals, a "default" array might not be set locally and could extend from more dominant sites further away. In this way, the most stable array may actually be the one present in both the active and inactive genes. The alternative array in the active gene may be set by a chromatin remodelling mechanism.

Like the mouse α-lactalbumin gene, the human α-lactalbumin gene does not show any dominant positioning signal. There are two possible array which fit this data, both of which position the three Stat5 sites close to the nucleosomal dyad axes. Whether this is similar to the scenario of the goat and mouse β-casein genes or the mouse α-lactalbumin gene is unclear.

### 7.8.2 Positions of Stat5 Sites in vivo

Goat BLG, goat β-casein, mouse β-casein and mouse α-lactalbumin gene promoter regions were tested *in vivo* for the presence of positioned nucleosomes. In all these cases the liver and mammary chromatin is covered by precisely positioned arrays which have a repeat length of approximately 180bp.

Particular note was paid to the positions of the Stat5 sites on the genes in relation to the positioned nucleosomes. These appear to be located in either the linker region between nucleosomes or close to the dyad axis of a nucleosome (summarised in table 14). Interestingly, although there was only ever a single array in the inactive gene, there were usually two different arrays in the active gene. Moreover, the alternative array was usually located specifically around the region containing the Stat5 binding site(s). This alternative array was approximately 90bp out of alignment with the other array in these cases. This meant that the Stat5 site would be located in the centre of a nucleosome in one array and in the linker region in the other array (Figure 114).

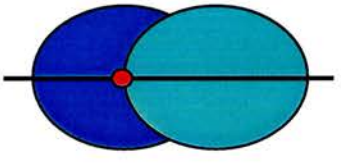


**Figure 114** Position of a Stat5 Site in Relation to the Two Nucleosomal Arrays
The dark-blue coloured array positions the Stat5 site (red) close to the dyad axis, whereas the light-blue coloured array places it within the linker region.

In sheep and goat BLG genes it is difficult to tell if an alternative array exists within the Stat5 region in the active gene. In both genes there is a single, clear array extending in a downstream direction towards the promoter until -630bp. At this point there are multiple OP-Cu digestion sites for approximately 540bp at which point no distinct bands can be seen. In the inactive gene in liver chromatin, the single, clear array extends over the entire promoter region.

| Gene Promoter | Tissue analysed | | *In vitro* array |
|---|---|---|---|
| | mammary | liver | |
| **Sheep BLG** | linker / dyad[#] | linker | linker |
| **Goat BLG** | linker / dyad[#] | linker | linker |
| **Goat β-casein** | linker / dyad | dyad | dyad |
| **Mouse β-casein** | linker? / dyad | dyad | dyad |
| **Mouse α-lactalbumin** | linker / dyad | dyad | linker[↗] |
| **Human α-lactalbumin** | Not tested | Not tested | dyad |

**Table 14** Position of Stat5 sites in relation to nucleosomes on the gene promoter

The positions of the Stat5 sites in relation to nucleosomes in the *in vitro* assay are based upon the most obvious array structure. (#) unclear; (?) requires further investigation; (↗) one candidate in a complex pattern.

Within this 540bp region, two alternative nucleosome arrays may exist when the gene is active. These arrays would fit with the OP-Cu digestion sites generated. This would allow the two outer Stat5 sites to be positioned either in the linker region in one array, or near the dyad axes in the other array. Conversely, the middle Stat5 site would suffer the opposite fate (Figure 115). If Stat5 binds preferentially within the linker region, this could lead to a novel form of gene regulation.

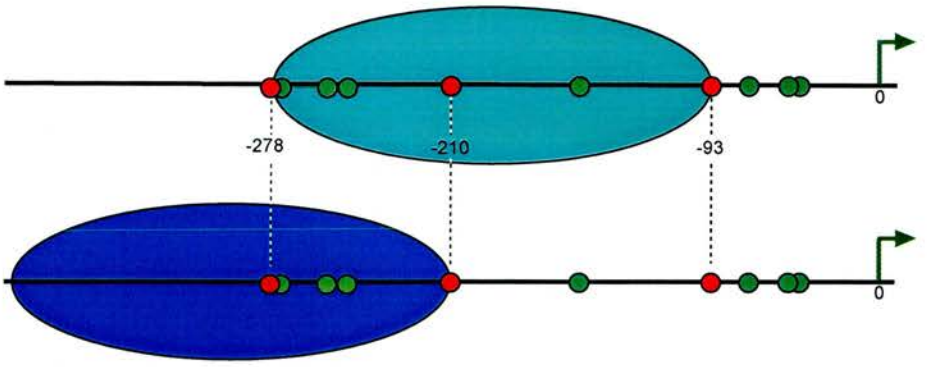



**Figure 115** Alternative Nucleosome Positions over the Ovine and Caprine BLG Promoter Regions in Relation to the Stat5 Sites NF1 sites are green circles; Stat5 sites are red circles

In the goat β-casein gene and the mouse α-lactalbumin gene, a single array positions the Stat5 site in the centre of a nucleosome in the inactive gene. This array also exists in the active gene, as does an alternative array, which positions the site within the linker region. This is quite different from the BLG genes, where the two outer Stat5 sites are positioned in the linker region in the active and inactive genes and may alternatively be close to the nucleosome dyad in the active gene.

Unfortunately, the map of the mouse β-casein gene does not extend as far as the promoter region and stops at +340bp with respect to the transcriptional start site. If this array is extrapolated, it positions the Stat5 site at the centre of a nucleosome. It is feasible that, like the previous genes, an alternative array exists in the active gene. This would require further investigation to be verified.

The nucleosome positions of all the promoter regions tested are summarised relative to one-another in Figure 116.

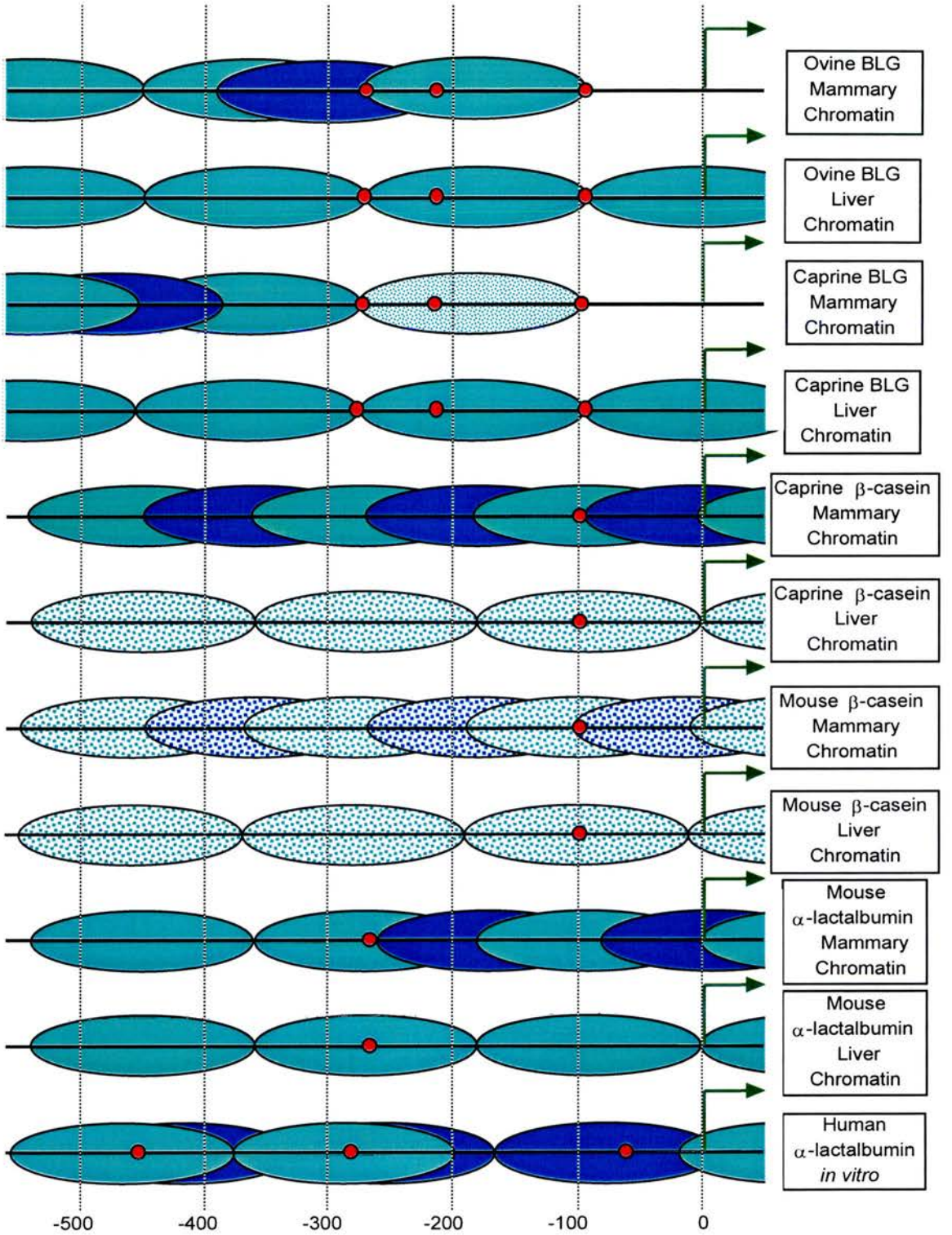



**Figure 116** Alignment of Nucleosome Positions of all Promoters Analysed

# Chapter 8

# Conclusions

## 8.1 Introduction

Chromatin is the functionally relevant template within which gene regulation must occur (Workman & Kingston, 1998). Nucleosomes can be shifted, structurally modified and locked into position by trans-acing factors in order to mediate gene activation in the correct temporal framework. This is frequently ignored in cell-free systems in which the DNA substrate is histone free, and also in tissue culture where the cell type has been derived from a tissue at a particular developmental stage. For example, the majority of cell culture work carried out on mammary gene regulation employs HC11 cells (Ball *et al.*, 1988). These were derived from COMMA-D1 cells (Danielson *et al.*, 1984) from a mid-pregnant mouse. Such cells cannot elucidate the full repertoire of developmental signals leading to a change in nucleosome array structure prior to gene activation. Moreover, nucleosome repeat length can vary between tissue and cell culture conditions (Compton *et al.*, 1976).

In order to understand how gene activity is controlled, a detailed knowledge of the chromatin encompassing the gene is required. This initial plan will then enable investigation of how chromatin is remodelled in association with gene expression status. I have presented an initial study, examining the array structure over the BLG gene in mammary and liver tissue from a mid-lactating Merino sheep. In addition, I have examined the promoter regions of the following genes, also in mammary and liver tissues: goat BLG, goat β-casein, mouse β -casein and mouse α-lactalbumin.

## 8.2 Nucleosomal Structure of the BLG Gene

There are limited studies with which to compare my data. No-one has published a long-range nucleosomal map of a pol II transcribed gene analysed from tissue-derived chromatin. I have presented a comparison with the lysozyme gene (**4.4**), which was

analysed in cell culture by the Bonifer laboratory. Their results differ markedly from those presented here, as the majority of nucleosomes were not positioned.

My study shows that there are distinct nucleosome positions throughout the BLG gene in its inactive state. Notably, there appear to be two types of phasing. The first comprises a single phased array and the second comprises two overlapping arrays (see chapter 4, figure 58). These two array types each span approximately ten nucleosome widths which alternate across the gene. It is interesting that a solenoid structure is composed of between five and six nucleosomes (McGhee *et al.*, 1980). Therefore, each ten nucleosome block might be encapsulated within two solenoid structures.

This phenomenon might also explain the overlapping alternatively phased arrays. There may be two particularly stable arrangements for nucleosome positioning over this ten nucleosome stretch. A two-solenoid array structure may form a distinct unit in which nucleosome positions can be locked into place. Indeed, this two-solenoid array structure might be packaged independently of the adjacent sequences such that nucleosomes within it are not influenced by the most stable array structure of the entire locus. Therefore, either array could exist within this two-solenoid structure.

If this hypothesis were true, it might be reflected in the mapping data. In this model, the linker length would vary in size at the junction between the two-solenoid array structures. This might result in some perturbation in the array structure at the junction (Figure 117). This could manifest itself as a heavily digested site (chapter 4; -1700DOWN), a heavily undigested site (chapter 4; +1700UP, +4700UP) or both (chapter 4, +2700UP). Therefore, my data may be compatible with this hypothesis.

The active gene displays the same nucleosome positions as the inactive gene, except over two regions. The first region, which spans the proximal promoter region, has an alternatively phased nucleosome next to the Stat5 sites (see chapter 4, figure 52). This alternative position might allow Stat5 access to the middle binding site on the BLG promoter at –210bp (see **5.4** and **7.8.2** for discussion). The second region is from –90bp to approximately exon 5 at +3.8kb. There are no clearly defined nucleosome positions over this region, perhaps because the nucleosomes are in a constant state of flux owing

to the movement of RNA polymerase. It is unclear why nucleosomes are visible after exon 5 to the end of the coding region.

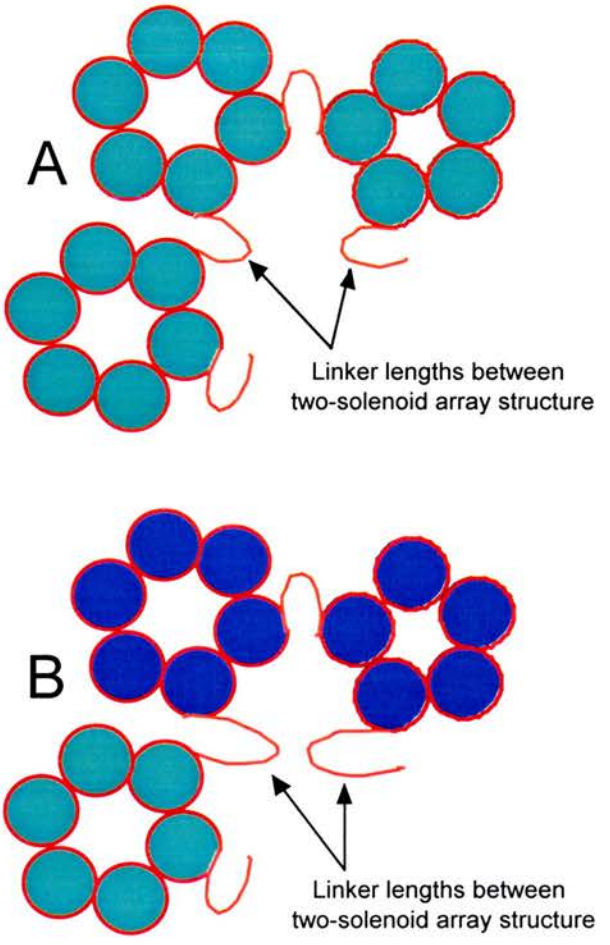



**Figure 117** Implications of the Two-Solenoid Array Model

A two solenoid structure can encapsulate approximately 10 or 11 nucleosomes (blue). DNA is coloured red. (A) The top two solenoids are in the same phase as the adjacent nucleosomal arrays, therefore the linker length between solenoids is the same. (B) The top two solenoids are in a different phase from the adjacent nucleosomal arrays, therefore, the linker sizes between the two-solenoid array structures would differ. This diagram indicates the solenoids as distinct entities to convey this point. Note that in the solenoid model, nucleosomes form a continuous spiral and are not flattened out as indicated here.

The next step in this analysis is to investigate the mechanisms involved in changing the nucleosome positions between the inactive and active gene states. Analysis of the nucleosome phasing at different times during mammary development would reveal the point at which any alteration in the phasing takes place. This may relate to the formation of a 'repressed', rather than 'silent' state, when the chromatin over the gene promoter is remodelled in preparation for activation. This remodelling may allow trans-acting factors easier access to their cognate sites in some instances (**1.7.1**).

In order to map nucleosome positions more accurately, the ligation mediated-polymerase chain reaction technique could be used. This would allow base pair resolution from OP-Cu digested chromatin.

## 8.3 Monomer Extension Data

Using the monomer extension technique, developed in Jim Allan's laboratory, I tested for sequence-dependent nucleosome positioning over the promoter region-transcriptional start site of the sheep BLG gene, goat BLG gene, goat β-casein gene, mouse β-casein gene, mouse α-lactalbumin gene and the human α-lactalbumin gene. The former five genes were also tested *in vivo* for nucleosome positioning in mammary and liver chromatin. In all cases, the *in vitro* array generated, whether it be from several or just one dominant position, correlated with the array in the inactive gene in liver chromatin. This may represent a 'default' array structure which is most stable. If this were the case, energy would be required to shift the array structure to that observed in the active gene. The only possible exception was the mouse α-lactalbumin gene where there was no obviously dominant array.

To extend this analysis, transgenic mice could be used to define the gene regions required for setting up the correct array structure. A number of mouse strains have already been generated which contain complete sheep BLG transgenes and others with selected deletions. I have already shown that the correct array structure was formed over the promoter region in transgenic mice containing an 11kb fragment spanning the BLG gene. Deletions and resections of this fragment should ellucidate which sequences are involved in setting the initial array and which, if any, are necessary for setting the

appropriate nucleosomal structure over the promoter region for gene activation. For instance, why does the inclusion of introns consistently give more representative gene expression levels in transgenic mice relative to cDNA constructs? Surely chromatin structure is a key factor, whether it be at a local level or as a higher order structure (Csordas, 1989).

## 8.4 Split-nucleosomes

Throughout this work I have suggested that two alternative array structures exist where digestion sites are approximately 90bp apart. There is, however, a more radical solution: nucleosome splitting (Thoma, 1991). This is a functional term for the observation of a 90bp periodicity in the array structure. It could represent the absence of the H2A/H2B dimers associated with each nucleosome or a completely different, possibly non-histone, protein complex. Only one group has reported this occurrence, which in itself is reason for caution! They digested the chromatin of the hsp82 gene in *S. cerevisiae* with MNase and DNase to reveal a 90bp cutting periodicity within the transcribed region of the gene (Lee & Garrard, 1991). Unfortunately, they did not conduct a control for sequence-specific digestion sites and therefore their data is rather inconclusive.

An absence of H2A/H2B dimers from the nucleosome octamer has been correlated with diverse effects on transcription (Chang & Luse, 1997; Santisteban *et al.*, 1997) and can relieve the requirement for chromatin remodelling in some instances (Kruger *et al.*, 1995; Recht & Osley, 1999). They have also been shown to associate with roughly 90bp of DNA (Puerta *et al.*,1993)

I have shown this 90bp cutting periodicity in regions both within and on either side of the transcribed region, in both active and inactive gene states. This is contrary to Lee's hypothesis that nucleosome splitting occurs due to the action of polymerase transcription. Notably, chromatin over the BLG gene which has only been digested with OP-Cu or MNase always displays a regular repeat pattern of approximately 180bp. Therefore, it is unlikely that split-nucleosomes exist at any point over the gene. Nevertheless, it is intriguing that the cutting periodicity is approximately 90bp when the regular repeat in the gene differs from 180bp.

**Models for Gene Activation**

I have formulated models for gene activation, with respect to the positioned nucleosomes I have elucidated, for the BLG, β-casein and α-lactalbumin genes. Since the β-casein gene is by far the most thoroughly studied of the three, I shall begin with it.

## 8.5 The β-Casein Gene

### 8.5.1 Nucleosome Positioning May Modulate Gene Activation

The promoter proximal regulatory region of the β-casein gene is highly related in many species, including goat and mouse. Both of these species contain a consensus binding site for Stat5 at -98bp and a partial consensus site at -140bp. There is a Yin-Yang 1 (YY1) binding site at -115bp and a CCAAT/Enhancer Binding Protein (C/EBP) site at -134bp. I propose a model for β-casein gene activation in which nucleosome positioning plays a key role (Figure 118).

### 8.5.2 The Silent State

Nucleosomes in the inactive gene in liver chromatin are phased in a tightly defined array downstream of the transcriptional start site and probably over the promoter region (chapter 6). Within this array, one of these nucleosomes is positioned at approximately -100bp, placing all of the above transcription factor binding sites close to the dyad axis. This could be the 'default' state of the array as the monomer extension data suggests that this is the most stable array over the promoter. This array could be essential in excluding ubiquitously expressed transcription factors from their sites within the β-casein promoter region.

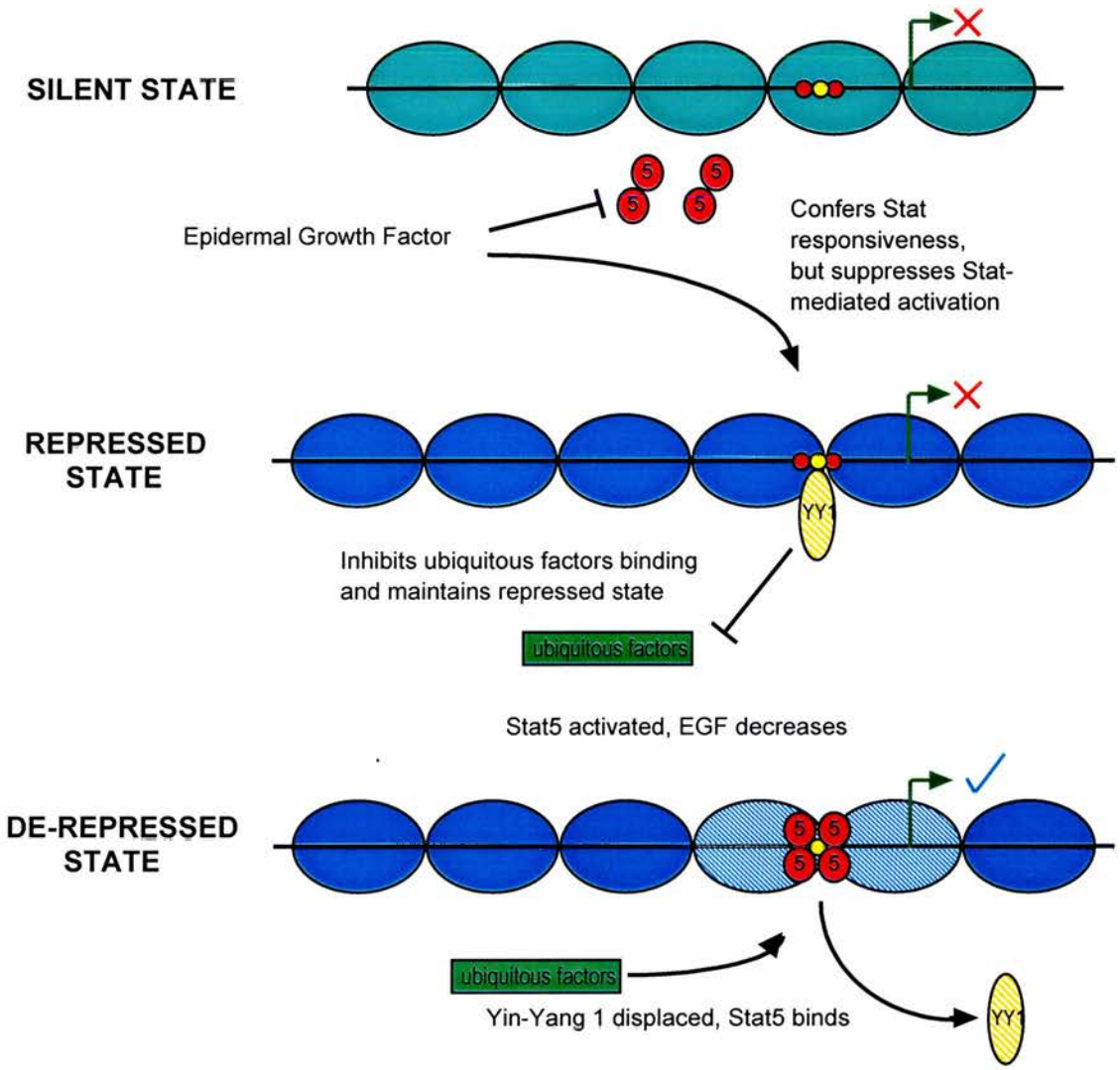



**Figure 118** β-Casein Gene Activation May Require Chromatin Remodelling

See text for details. Transcription factors are indicated as follows: Stat5, red; YY1, gold stripes; ubiquitous factors, green. Nucleosomes in the silent state are light blue, while those in the active state are dark blue. For more detail of the model see figure 119.

Various forms of C/EBP are present in a number of tissues, such as the retina (Ben-Or & Okret, 1993), adipocytes, white adipose tissue (MacDougald *et al.*, 1994) and hematopoietic cells (Yamaguchi *et al.*, 1999) which are able to bind the C/EBP cognate site. Although it is unknown if C/EBP binds to the silent β-casein gene, it is unlikely as no DNase I hypersensitive sites are present. Likewise, YY1 is present in many, indeed almost certainly all, tissues (Shi *et al.*, 1991; Johansson *et al.*, 1998). Both of these factors are bound to the β-casein promoter region when it is in the repressed state, just before gene activation (Figure 119). Inappropriate binding in other tissues could lead to gene activation and may be prevented by nucleosome-mediated exclusion.

### 8.5.3 The Repressed State

At some point between the silent and active states, I propose that the nucleosomal array over the β-casein gene promoter region shifts to position the aforementioned transcription factor binding sites within the linker region.

After conception, the levels of epidermal growth factor (EGF) rise in the mammary tissue which leads to cell proliferation. Notably, *in vitro* experiments show that cells respond optimally to lactogenic hormones and synthesise high levels of β-casein only if they have been kept previously in a medium containing EGF (Taverna *et al.*, 1991). Although EGF confers *competence* to lactogenic hormones, it prevents hormonal induction of β-casein gene expression when it is present in the culture medium of confluent cells simultaneously with the lactogenic hormones (Hynes *et al.*, 1990). Therefore, it is involved either directly or indirectly in establishing the repressed gene state, but must be removed for gene activation.

EGF can induce the activation and nuclear localisation of Stat5 dimers in mouse liver (Ruff-Jamison *et al.*, 1995). However, as the nucleosomal array is not re-organised in the liver, Stat5 probably cannot bind the β-casein gene promoter. This induction pathway does not occur in the mammary tissue. Interestingly, prolactin failed to dimerise Stat5 or mediate Stat5 binding activity when injected into mouse livers. Obviously, Stat5 induction is regulated very differently in the liver (Jahn *et al.*, 1997), which may explain its pleiotropic effects in different tissues.

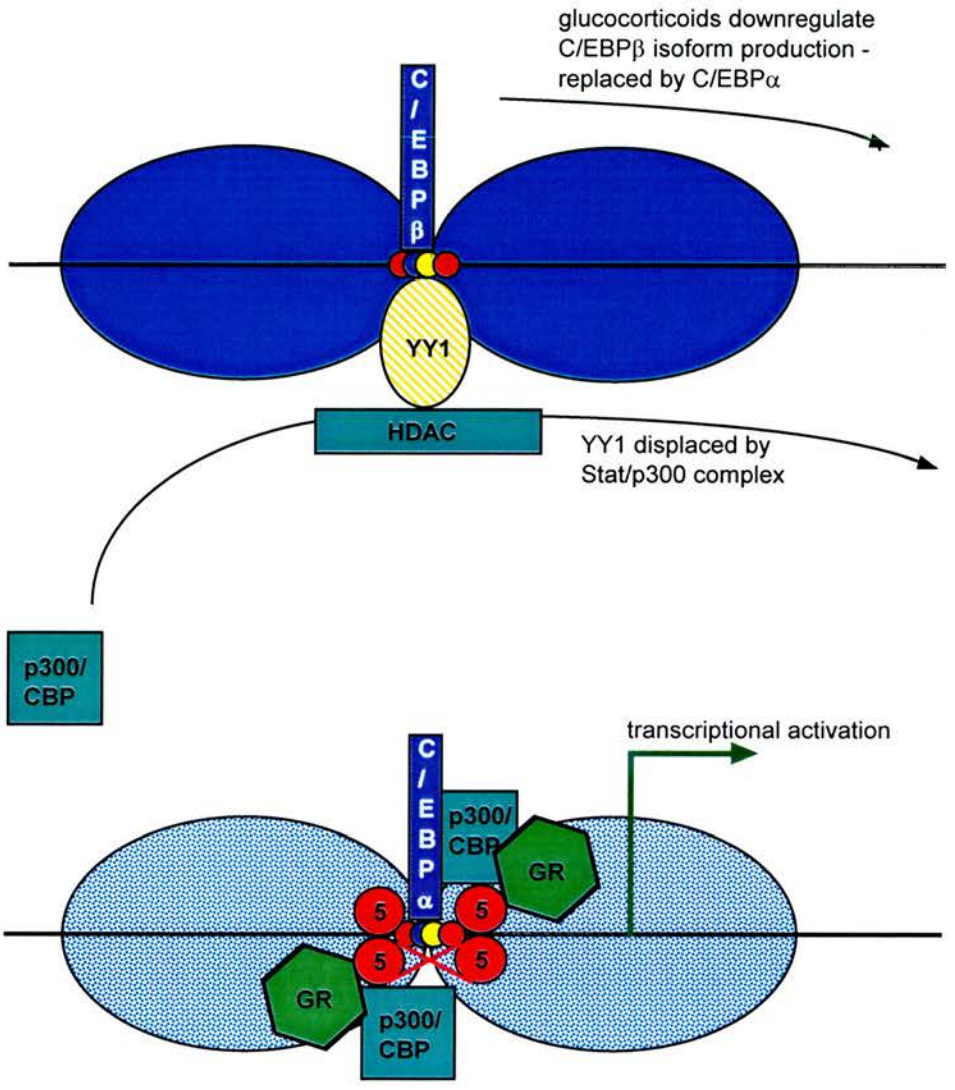



**Figure 119** A Model for β-Casein Gene Activation Over the Promoter Region

Nucleosomes (dark blue) may be remodelled (stippled dark blue) at some point between the repressed and active states in the process of gene activation. Once the YY1 (gold-striped)/C/EBPβ (purple)/HDAC (teal) complex is displaced by p300/CBP (teal)/Stat5 (red), the initiation complex can form. It comprises these displacement factors plus GR (green), C/EBPα (purple) and other histone acetylase associated factors (not shown). Possible Stat5 tetramerisation is indicated by a red cross.

In order to stabilise and maintain this repressed state, it is reasonable to assume that specific trans-acting factors bind this domain. The binding activity of YY1 is much higher in virgin and early pregnant mammary glands relative to that of lactating glands (Raught *et al.*, 1994). YY1 binds and repress β-casein gene activation in conjunction with another DNA binding factor, possibly C/EBPβ (Meier & Groner, 1994). YY-1 can also bind to HDAC1, 2 and 3. (Yang *et al.*, 1997). These proteins possesses histone deacetylase activity and represses transcription when tethered to a promoter. They are well conserved throughout evolution. Therefore, this complex may maintain the repressed state by deacetylating the surrounding histones, preventing other factors access to their binding sites (Brown *et al.*, 1998).

The other DNA-binding factor involved in maintaining this repressed state may be C/EBPβ. There are at least three isoforms of C/EBPβ in the mammary gland, all of which are developmentally regulated (Raught *et al.*, 1995). The protein levels of two particular isoforms, liver-enriched activating proteins (LAPs) and liver-enriched inhibiting protein (LIP) are elevated throughout pregnancy. As the names imply, LAPs can activate expression whereas LIP inhibits expression. During pregnancy a high LIP/LAP ratio is maintained which may aid in maintaining the repressed state.

### 8.5.4 The De-repressed State

Levels of EGF decrease from a maximum at day 10 of lactation to a minimum at parturition (Edery *et al.*, 1985). As the levels of EGF are decreasing, the concentration of prolactin is rising. This leads to a concurrent increase in Stat5 dimerisation and nuclear translocation, mediated through the Jak/Stat endocrine pathway. Stat5 dimers are able to displace YY1 from its binding on the β-casein gene promoter (Meier & Groner, 1994). This may be mediated via p300/CBP which interacts with both Stat5 (Pfitzner *et al.*, 1998) and C/EBPβ (Mink *et al.*, 1997) and has been shown to relieve YY1 transcriptional repression (Lee *et al.*, 1995).

During lactation, C/EBPβ is replaced by C/EBPα on the β-casein gene promoter (Raught *et al.*, 1995). This may be mediated via glucocorticoid receptor, as it down-regulates LIP C/EBPβ expression (Raught *et al.*, 1995). C/EBPα expression correlates

with cell proliferation arrest and terminal differentiation of mammary epithelial cells (Muller *et al.*, 1999) and may be regulated by cell-extra-cellular-matrix interactions (Rana *et al.*, 1994). C/EBPα is probably associated with stabilising and developing the promoter complex, rather than directly activating it, since in C/EBPα$^{-/-}$ mice appear to express β-casein at normal levels (Seagroves *et al.*, 1998). C/EBPα probably also associates with GR, like the C/EBPβ isoforms and C/EBPδ (Raught *et al.*, 1995).

Like prolactin, the level of other lactogenic hormones and steroids also increases during pregnancy and lactation. Hydrocortisone is one such steroid. It mediates glucocorticoid receptor (GR) dissociation from the heat-shock proteins with which GR associates in the cytoplasm (Pratt, 1993). GR interacts directly with Stat5 dimers in the cytoplasm (Cella *et al.*, 1998) and either one can translocate the other into the nucleus (Wyszomierski *et al.*, 1999).

This co-operation between GR and Stat5 dimers continues once Stat5 has bound to the β-casein gene promoter. The transactivation domain of Stat5 is very weak in relation to that of Stat6 (Moriggl *et al.*, 1997) or VP16 (Berchtold *et al.*, 1997). It is enhanced by its association with GR which also decreases its rate of dephosphorylation, leading to longer association of the Stat5 complex with the DNA (Wyszomierski *et al.*, 1999). This may explain why the Stat5 mRNA and protein levels drop dramatically during lactation without a concurrent decrease in β-casein gene expression (Kazansky *et al.*, 1995). Additional Stat5 protein may not be required if dimers are already stably associated with the β-casein gene. This stability may be strengthened by tetramerisation of the Stat5 proteins whose sites are four helical turns apart (see BLG gene below).

The Stat5/GR complex is able to recruit high levels of p300/CBP (Pfitzner *et al.*, 1998). p300 and CBP are similar in sequence and functionally interchangeable in most cases. They exhibit histone acetyltransferase activity and associate with P/CAF, which possesses the same enzymatic activity (Bannister & Kouzarides, 1996; Ogryzko *et al.*, 1996). Together they may make the promoter more accessible to transcription factors (Vettese-Dadey *et al.*, 1996), polymerase associated factors (Lee *et al.*, 1993) and relieve transcriptional repression (Ura *et al.*, 1997) which can lead to gene activation (Hebbes *et al.*, 1988).

This proximal promoter region is necessary and sufficient to confer hormonal regulation and gene activation in stably transfected cell lines. However, it is insufficient to confer accurate temporal and spatial regulation in transgenic mice (Cerdan *et al.*,1998) A larger 5' region of 3.8kb is required. Therefore, this model is incomplete and other regulatory regions are essential to encapsulate its full regulatory ensemble.

## 8.6 The α-Lactalbumin Gene

Like the β-casein genes, both murid (Lubon & Hennighausen, 1988) and human (Hall *et al.*, 1987) α-lactalbumin genes contain the 'milk box' consensus motif from -110bp to -140bp with respect to the transcriptional start site. However, unlike the β-casein genes, very little is known about what factors interact with the promoter region to mediate gene activation. The mouse α-lactalbumin gene contains a Stat5 site at -267bp and an incomplete one at -70bp (Vilotte & Soulier, 1992), approximately nineteen DNA helical turns away. The human α-lactalbumin gene contains Stat5 sites at -62bp, -278bp & -457bp. Only the former two sites are separated by almost a whole number of helical turns. Since Stat5 requires both sites to be in the same rotational phase in order to form a stable tetramer (see **8.6**), their relative orientation may be important. As both genes are dependent on prolactin for activation *in vivo* (Mercier & Vilotte, 1993), it is likely that Stat5 can interact with some, if not all, of these sites.

In the inactive mouse α-lactalbumin gene, nucleosomes are positioned in a tightly defined array, centred at +270bp, +90bp, -90bp, etc. This positions the Stat5 sites close to a nucleosome dyad in both instances. Like the β-casein genes, this may assist in inhibiting spurious gene activation. In the active gene, both this array and an alternative array exist. The alternative array is approximately 90bp out of phase, which positions the Stat5 sites in the linker region in both cases. Again, like the β-casein genes, this may be important in assisting Stat5 binding to the DNA and activating gene expression. The monomer extension data present a large number of peaks of similar intensity in different phases from one another. As there is no obvious dominant array, a DNA-induced 'default' array may be set by sequences out-with the local environment, perhaps synergistically with weak positioning sites within the promoter region.

The human α-lactalbumin gene was only tested *in vitro* using the monomer extension technique. Although many positioning sites exist, they appear to fall into one of two possible array structures. Like the β-casein gene results, either array positions all Stat5 sites close to nucleosome dyad axes. This could inhibit Stat5 binding and gene activation.

## 8.7 The β-Lactoglobulin Gene

The BLG nucleosomal array is the same in both the inactive and active states in both sheep and goats. They are centred at approximately -180bp, 0bp, +180bp, etc with respect to the transcriptional start site. This positions the Stat5 sites located at -93bp and -278bp within the linker region. This is a rather unexpected result as it differs substantially from both the β-casein and the α-lactalbumin genes, where the Stat5 sites are close to the dyad of a nucleosome when the gene is inactive. Indeed this is reflected in the *in vitro* data, where the DNA-induced nucleosome positions also centre the Stat5 sites close to a nucleosome dyad. Therefore, I suggest that BLG may be regulated in a different manner from the β-casein and α-lactalbumin genes, in relation to nucleosome positioning.

It appears that an alternative array exists over the promoter region in the active gene which would position the middle Stat5 site within the linker region. However, this would position the other two Stat5 sites close to a nucleosome dyad, possibly restricting their binding capacity. Alternatively, these additional digestion sites may be the result of nucleosome shifting by a SWI/SNF-type chromatin remodelling complex (Whitehouse *et al.*, 1999). Binding of the outer two Stat5 sites may recruit chromatin remodelling complexes, like in the β-casein gene, which could in turn make the domain more accessible to ubiquitous factors, such as NF1.

NF1 is essential for BLG activation and seven of its binding sites span the region encapsulated by the strongly positioned nucleosome at -180bp. As NF1 cannot bind within a nucleosome (see **1.13.1**), remodelling or shifting of this nucleosome may be a pre-requisite for gene activation. Indeed, this may be the reason why the nucleosomal array is positioned this way in the inactive gene. By positioning the outer Stat5 sites in

the linker region, most of the NF1 sites are inaccessible, encapsulated within the nucleosome. Therefore, only once the nucleosome was remodelled would the sites become accessible. Their exclusion may be more important than sequestering the Stat5 sites within the nucleosome at this stage. It is unknown if α-lactalbumin or β-casein genes are regulated by NF1 at the proximal promoter region. If they are not, this may account for the difference in nucleosome array patterns in the active and inactive genes. The most important aspect of gene suppression for these genes may be to inhibit Stats binding, while for BLG the most important aspect may be to exclude NF1.

An alternative hypothesis is that the position of the Stat5 sites within the linker region is not important. Rather, it is the position of other transcription factor sites within the linker region. An NF1 variant was found in the mammary gland which preferentially interacts with an NF1 site at –274bp, which overlaps the Stat5 binding site at -278bp (Watson *et al.*, 1991). This is positioned in the linker region. In addition, a site at -55bp forms an NF1-like complex with sheep mammary nuclear extracts. It is unclear what this complex is, but it is also positioned in the linker region. Mutation of the latter site in a transgene construct might indicate if this site is crucial to gene regulation. As the former site overlaps the Stat5 site, this approach may yield results which are difficult to interpret.

## 8.8 Stat5 Tetramerisation

The precise position of the outer Stat5 sites may be crucial for gene activation. The sites are exactly 185bp apart and since a nucleosome is positioned almost exactly between them, they must be very close to each other in three dimensional space (Figure 120). Moreover, with an average of 10.2bp/turn of the DNA helix within a nucleosome, the sites may also be facing the same side of the helix relative to each other. This could allow the Stat5 dimers to tetramerise. Presumably such an interaction could stabilise the transcription complex at the BLG promoter.

Stat tetramers were first described for Stat1 (Guyer *et al.*, 1995; Vinkemeier *et al.*, 1996) and Stat4 (Xu *et al.*, 1996). In both cases homodimers could interact via their

N-terminal domains to associate with equivalent homodimers. This was independent of the relative orientations of the binding sites, so long as the distance was not 5bp, which probably lead to stearic hindrance. These investigators concluded that since the dissociation rate was much faster when either site was mutated, this inferred co-operative stabilisation. Notably, one site was a 'weak' site and the other a 'strong' site in both cases. In BLG, the proximal Stat5 site is strong and the distal one is weak (Philp *et al.*, 1996).

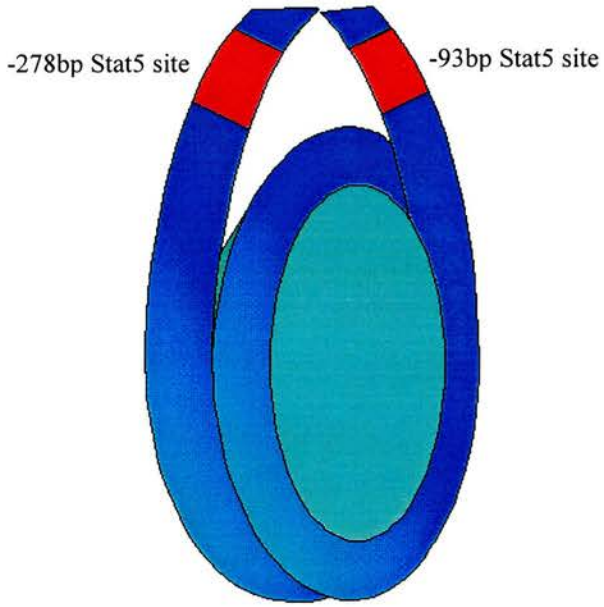



**Figure 120** Relative Position of the Outer Stat5 Sites on the BLG Promoter
The outer Stat5 sites (red) are located in the linker region on either side of the strongly positioned nucleosome (light blue) at -180bp with respect to the transcriptional start site.

The crystal structure of the Stat4 N-terminal domain showed that its hook-like structure can interact with another Stat4 monomer through an extensive interface. This is formed by polar interactions across one face of the hook (Vinkemeier *et al.*, 1998). Since the Stat N-terminal domains are highly conserved (Schindler & Darnell, 1995) (**1.13.1**), homodimers may be able to tetramerise with alternative homodimers.

Stat5 dimers can tetramerise on the interleukin-2-responsive enhancer on the IL-2Rα gene *in vitro* (Meyer *et al.*, 1997). Like Stats 1 and 4, one site was 'weak' and the other 'strong'. However, unlike Stats 1 and 4, Stat5 dimers could only bind when the two sites were in the same relative orientation, i.e. a whole number of helical turns apart.

Whether the position of the Stat5 sites on the face of the helix is important in a chromatin context remains to be proved. As interleukin-2 activates both Stat5a and b, it is unclear which Stat5 dimers interacted.

A recent paper has examined the relationship between nucleosome positioning and Stat5 binding on this gene (Rusterholz *et al.*, 1999). The binding region is only 51bp in length comprising the two Stat sites and an Elf-1 transcription factor site. Although Elf-1 is constitutively expressed (Davis *et al.*, 1996; Serdobova *et al.*, 1997), it cannot bind its cognate site. It is only once interleukin-2 (IL2) induces Stat5 dimerisation and both Stat5 binding sites are occupied that Elf-1 can bind its site. Together these factors are sufficient and necessary for IL2-mediated induction.

Notably, this enhancer sequence is enclosed within the centre of a nucleosome. It is highly likely that this is sufficient to prevent Elf-1 binding. It is only once the positioned nucleosome is shifted and both Stat5 dimers bind, and probably tetramerise, that Elf-1 gains access to its cognate site. Therefore, nucleosome positioning may be a critical parameter in the regulation of this gene.

As indicated above, the outer Stat5 binding sites on the BLG proximal promoter are in the same orientation, eighteen helical turns from each other. The strongly positioned nucleosome should bring them extremely close together. This Stat5 tetramerisation mechanism is quite distinct from that of the IL-2Rα gene where the Stat5 sites are only 20bp apart, requiring the positioned nucleosome to be removed or shifted. I propose that in BLG the nucleosome is an integral part of the tetramerisation mechanism, bringing the sites together in a similar manner to the heat shock elements of the *Drosophila* hsp26 gene (**1.7.2**). An obvious next step would be to reconstitute a nucleosome onto a fragment containing this region, adding phosphorylated Stat5a protein, and seeing if Stat dimerisation and tetramerisation occurred.

## 8.9 Nucleosome-Mediated Stat5 Tetramerisation and Additional Factor Access

It is unclear how this model would accommodate binding of the middle Stat5 site and the NF1 sites. The middle Stat5 site is 6.67 helical turns from the distal Stat5 site, and

11.5 turns from the proximal site. Since it is facing the opposite side of the helix and is within the nucleosome, unlike the outer sites, it probably does not bind Stat5 in this context. Similarly, the NF1 sites are all located within the nucleosome making their binding unfavourable.

It is possible that only an H3/H4 tetramer is present on either side of position -180bp *in vivo*. This would correlate with the digestion pattern *in vivo* and an H3/H4 tetramer in some instances can bind NF1 (Spangenberg *et al.*, 1998). This may also allow Stat5 to bind the middle site. However, the H3/H4 tetramer appears to contain the nucleosome positioning information (Dong & van Holde, 1991). This would not correlate with the *in vitro* data which would place the strong positioning site in between the histone tetramers.

## 8.10 Is Stat5 Tetramerisation the Reason for the Specific Nucleosome Phasing?

Unlike the goat β-casein gene, mouse α-lactalbumin gene and possibly the mouse β-casein gene, the nucleosome phasing in the sheep and goat BLG genes does not alter to position Stat5 sites within the linker region on gene activation. The outer Stat5 sites are constantly positioned in the linker region on either side of a strongly positioned nucleosome. The main function of this nucleosome could be two-fold:

1. To exclude the binding of ubiquitously expressed NF1 in inappropriate cell types. This is essential for tissue- and stage-specific expression.
2. To mediate Stat5a tetramerisation on gene activation. How this is prevented in the inactive gene is unclear, but may be mediated through higher-order structural compaction or a repressive linker protein (Kandolf, 1994; Vermaak *et al.*, 1998).

How these two points can be reconciled is unclear, since Stat5 tetramerisation and NF1 binding would appear to be mutually exclusive. However, this should be an interesting topic for future research.

# Appendix I

## Conversion Table for rpm into RCF

Relative Centrifugal Force (RCF) refers to the force during centrifugation that moves the particulate outwards from the centre of rotation. This force is proportional to the rpm squared and rotor radius. It is calculated using the following formula:

$$RCF = 11.17r\,(n/1000)2$$

Where:

r = the radius in centimeters from the centre of the rotor to the point in the tube where the RCF value is required.

n = the rotor speed in rpm

Here are some examples of frequently used speeds on the following centrifuge rotors:

| | | |
|---|---|---|
| **Desktop microcentrifuge** | 14,000rpm=14,230 RCF | (average with maximum volume) |
| **Sorval SS-34 Rotor** | 9,000rpm = 6,315 RCF | (average with maximum volume) |
| **Sorval F16 Rotor** | 6,000rpm = 3,470 RCF | (average with maximum volume) |
| **Beckman JA14 Rotor** | 7,000rpm = 5,200 RCF | (average with maximum volume) |
| **Beckman JA20 Rotor** | 7,000rpm = 3,010 RCF | (average with maximum volume) |
| **Jouan CR3000 Swing-Bucket Rotor** | 1,000rpm = 142 RCF | (average with maximum volume) |

# Appendix II

**Cuprous Phenanthroline Digested Ovine Liver and Mammary DNA**

The following sixteen indirectly end-labelled Southern blots are an addendum to chapter 4. These contain all six time points of the partial digestion of sheep chromatin with OP-Cu. The loading in all gels is the same and is as follows:

**Lane**

**1** **Time point 0** (0 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**2** **Time point 1** (10 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**3** **Time point 2** (20 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**4** **Time point 3** (30 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**5** **Time point 4** (40 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**6** **Time point 6** (60 minutes OP-Cu digestion); digested to completion with a restriction enzyme indicated below the picture

**n** **Naked DNA control** (75 seconds OP-Cu digestion of protein-free DNA); digested to completion with a restriction enzyme indicated below the picture

**-** **Negative control** Mouse DNA digested to completion with a restriction enzyme indicated below the picture

**+** **Positive control** Sheep DNA digested to completion with a restriction enzyme indicated below the picture

**m** **Marker** Either Böhringer Mannheim 1kb ladder or sheep genomic DNA digested only with OP-Cu (liver or mammary chromatin used depending on gel).

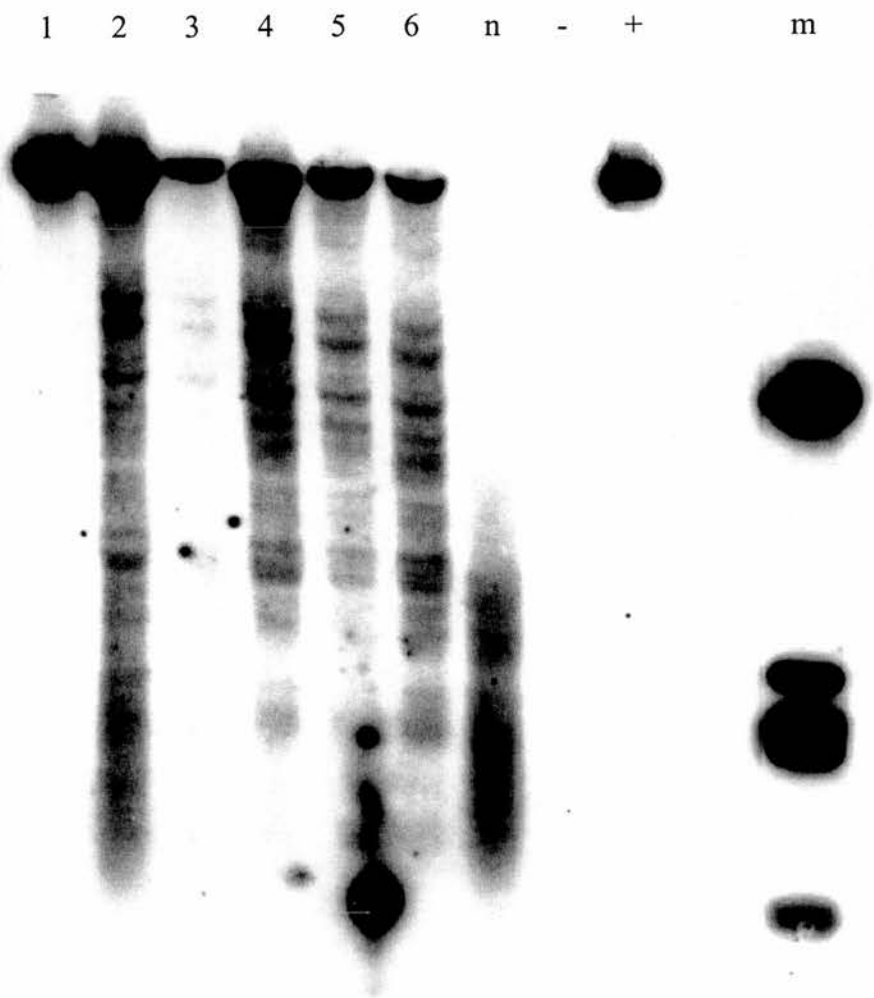



**Figure A1**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with -1700UP probe.

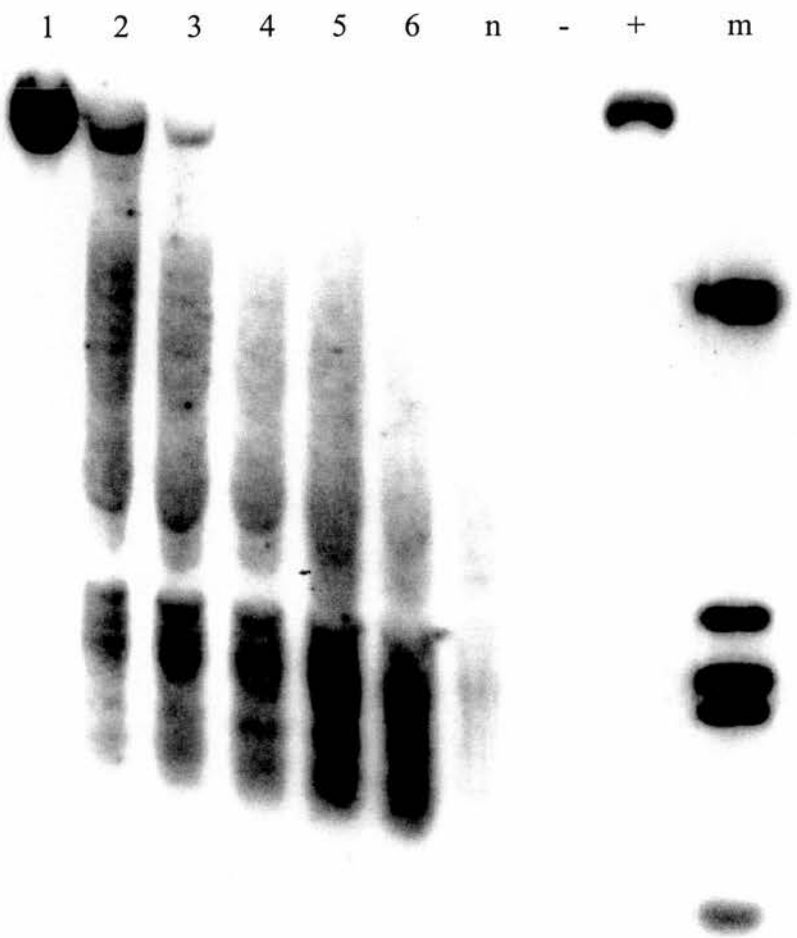



**Figure A2**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with -1700UP probe.

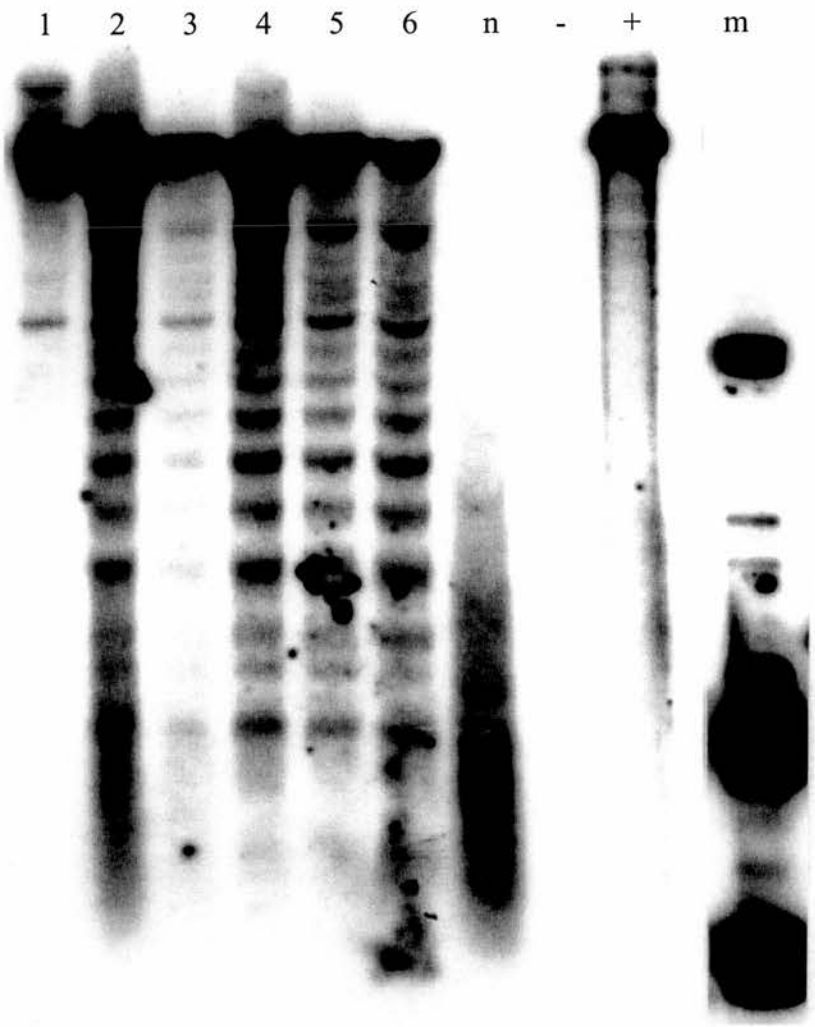



**Figure A3**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with –1700DOWN probe.

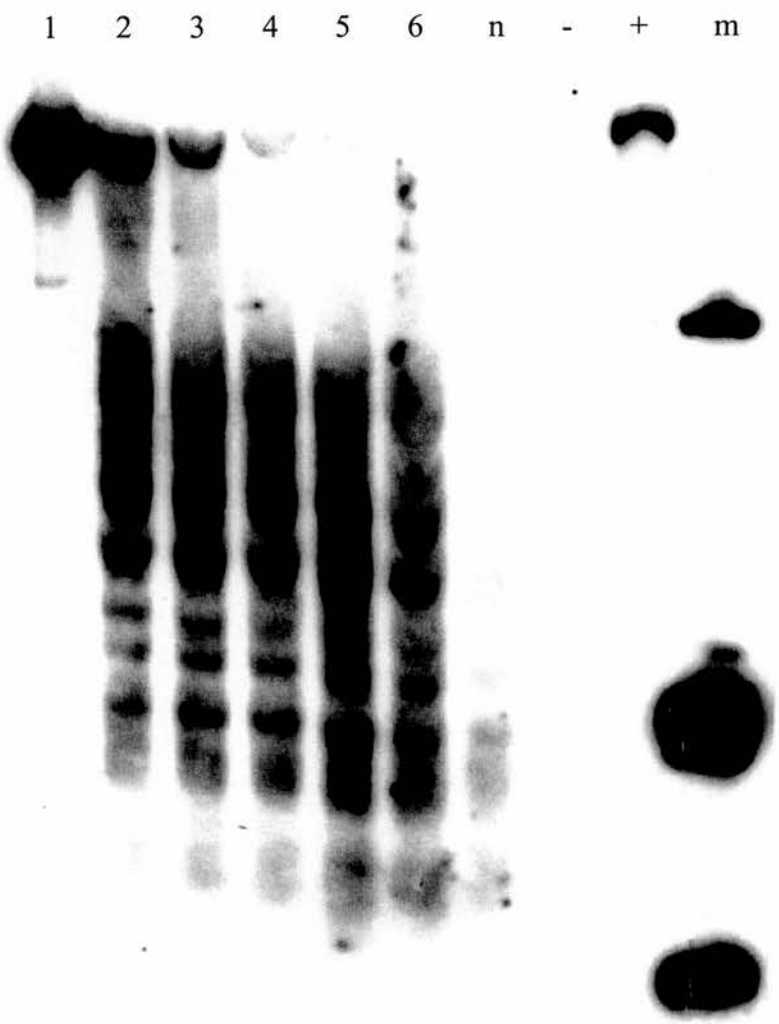



**Figure A4**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with –1700DOWN probe.

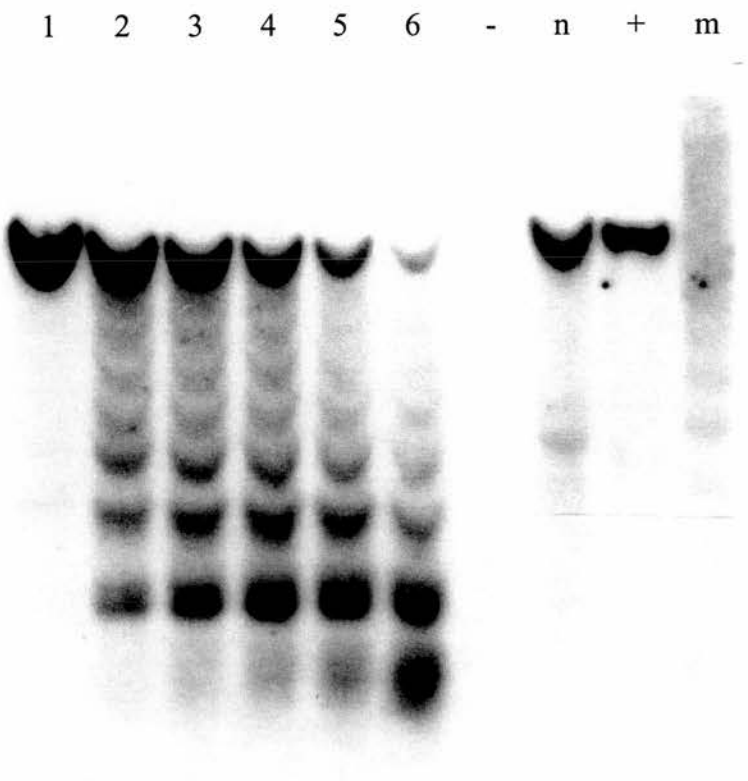



**Figure A5**

Liver Nuclei digestion with OP-Cu (partial) and ***Rsa*I** (complete).
Probed with –900DOWN probe.

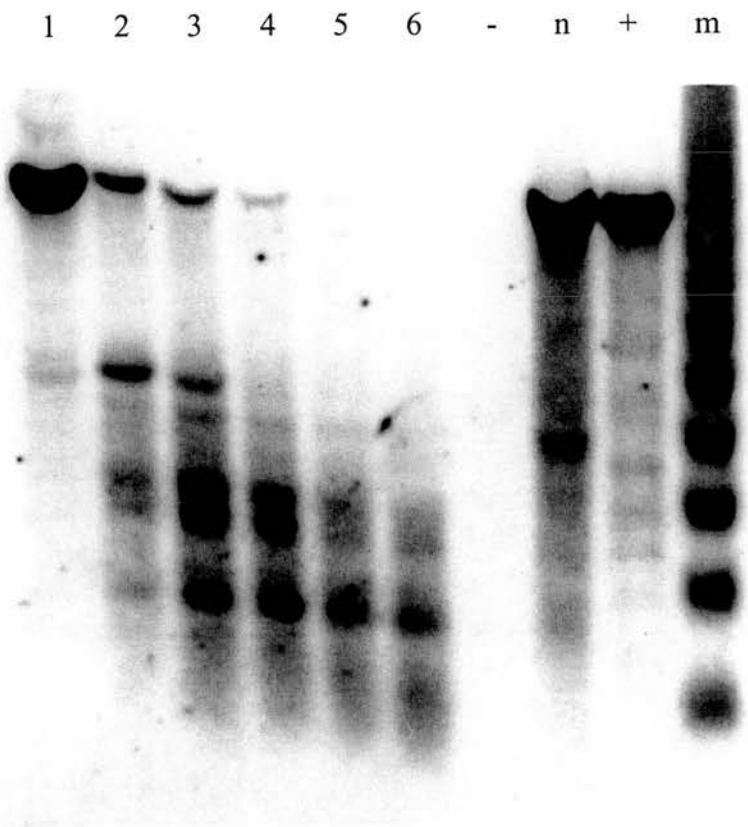



**Figure A6**

Mammary Nuclei digestion with OP-Cu (partial) and ***Rsa*I** (complete).
Probed with –900DOWN probe.

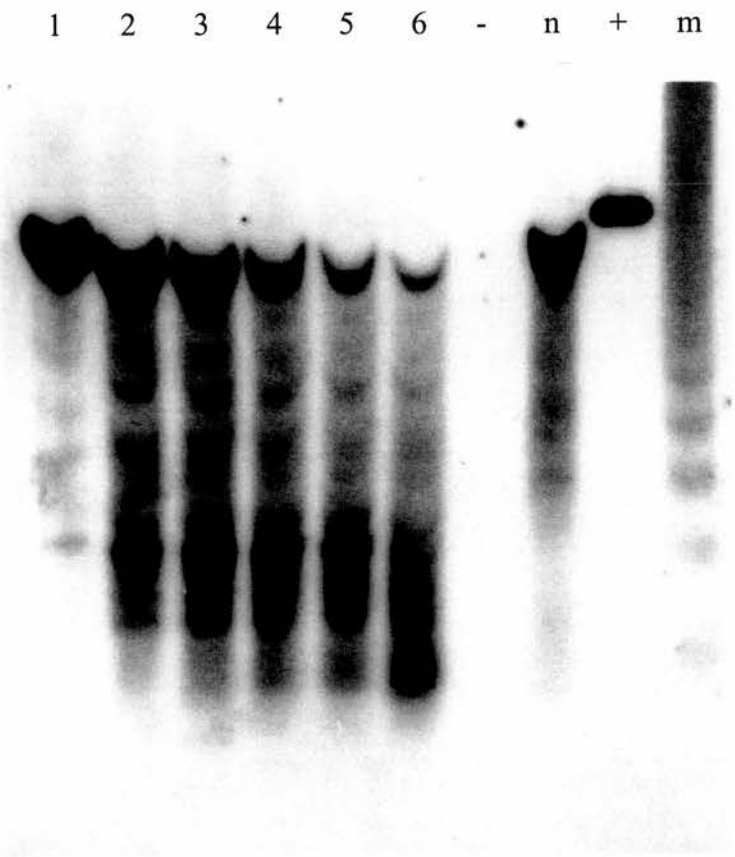



**Figure A7**

Liver Nuclei digestion with OP-Cu (partial) and ***Pvu*II** (complete).
Probed with +1700UP probe.

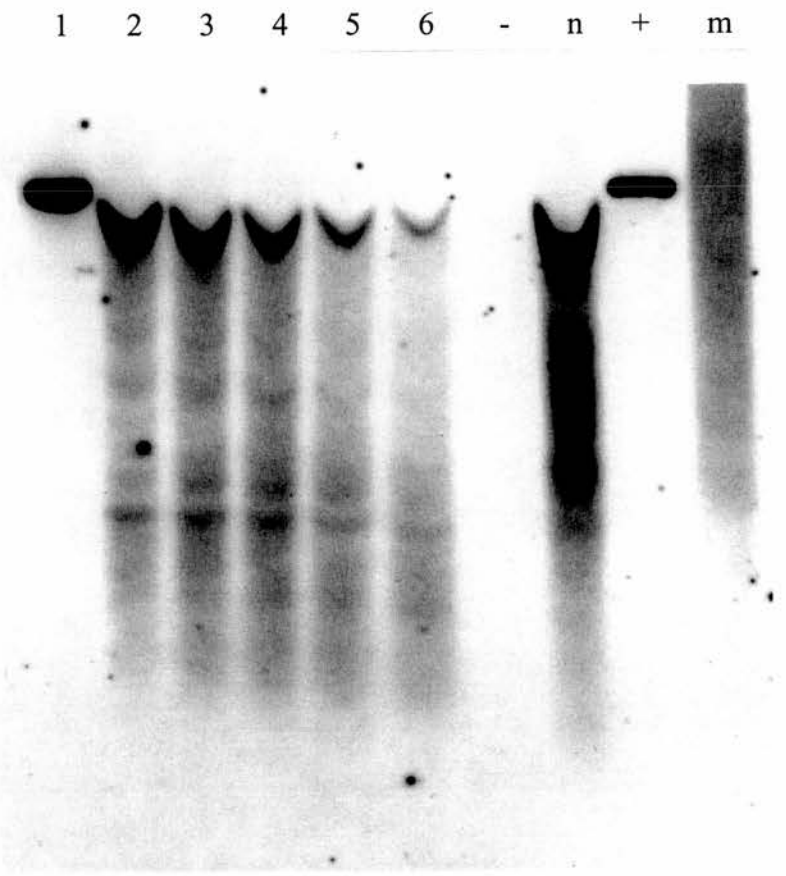



**Figure A8**

Mammary Nuclei digestion with OP-Cu (partial) and ***Pvu*II** (complete).
Probed with +1700UP probe.

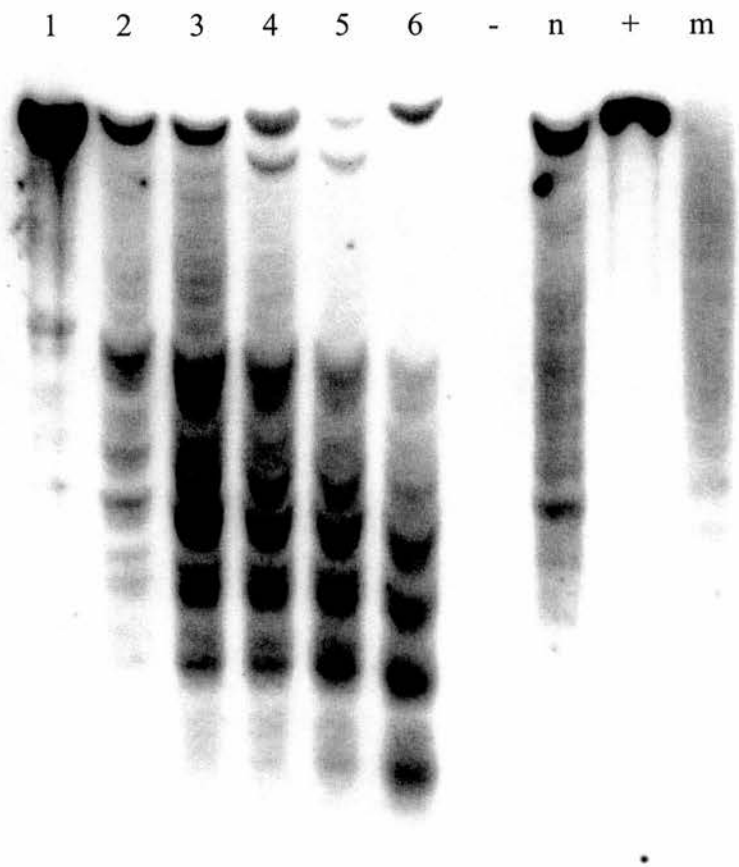



**Figure A9**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +2700UP probe.

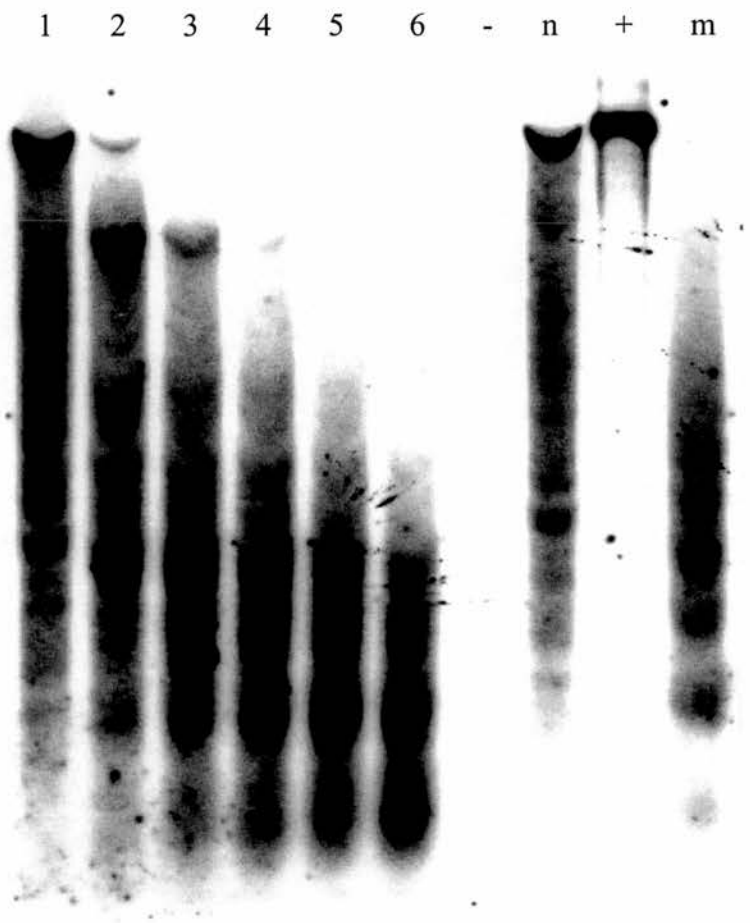



**Figure A10**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +2700UP probe.

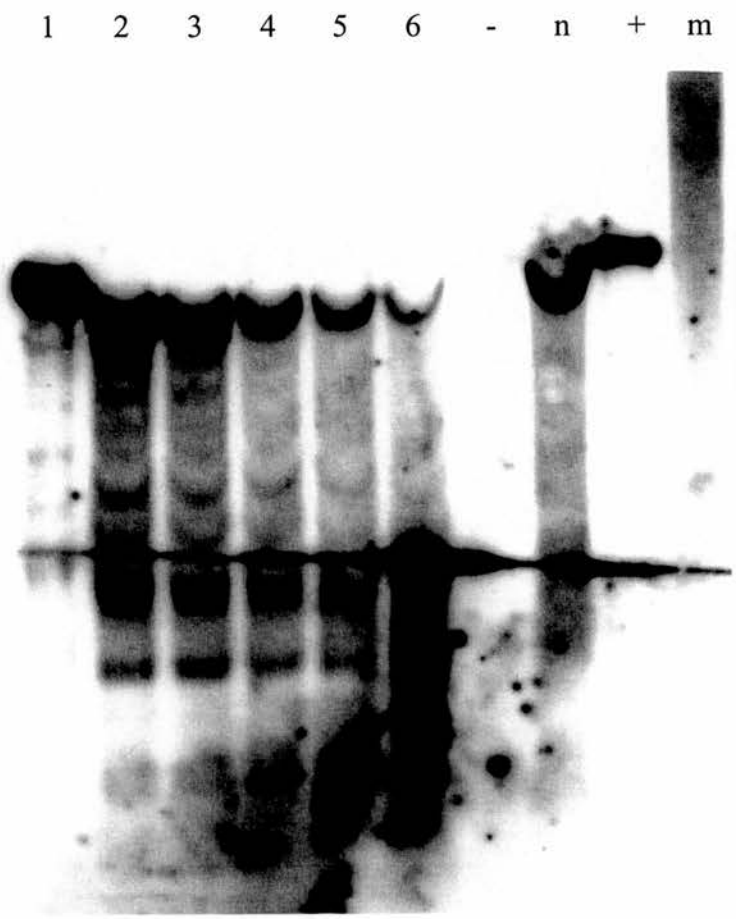



**Figure A11**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +2700DOWN probe.

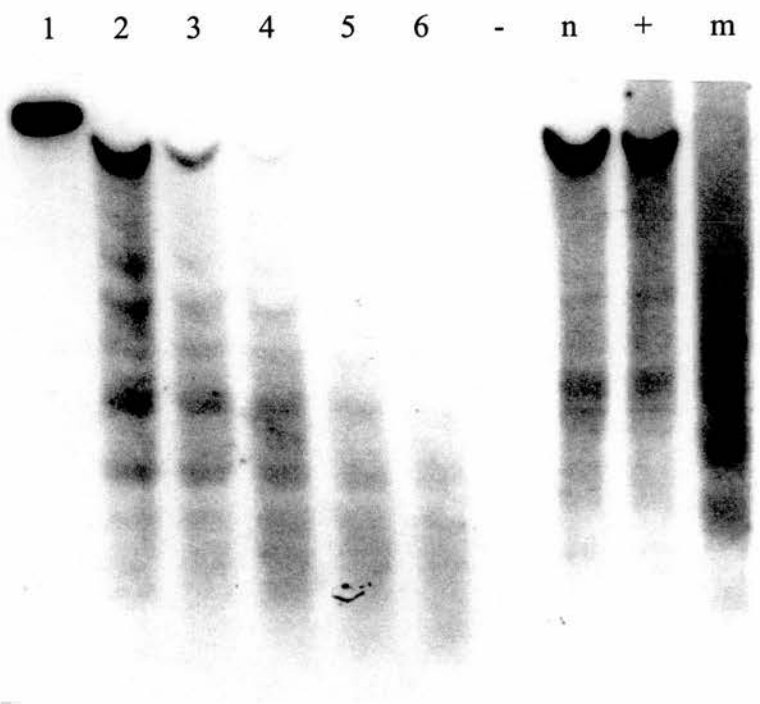



**Figure A12**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +2700DOWN probe.

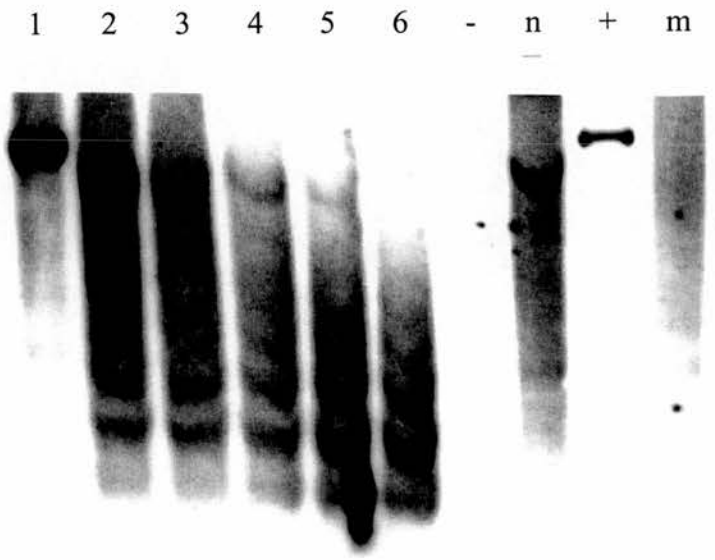



**Figure A13**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +4700UP probe.

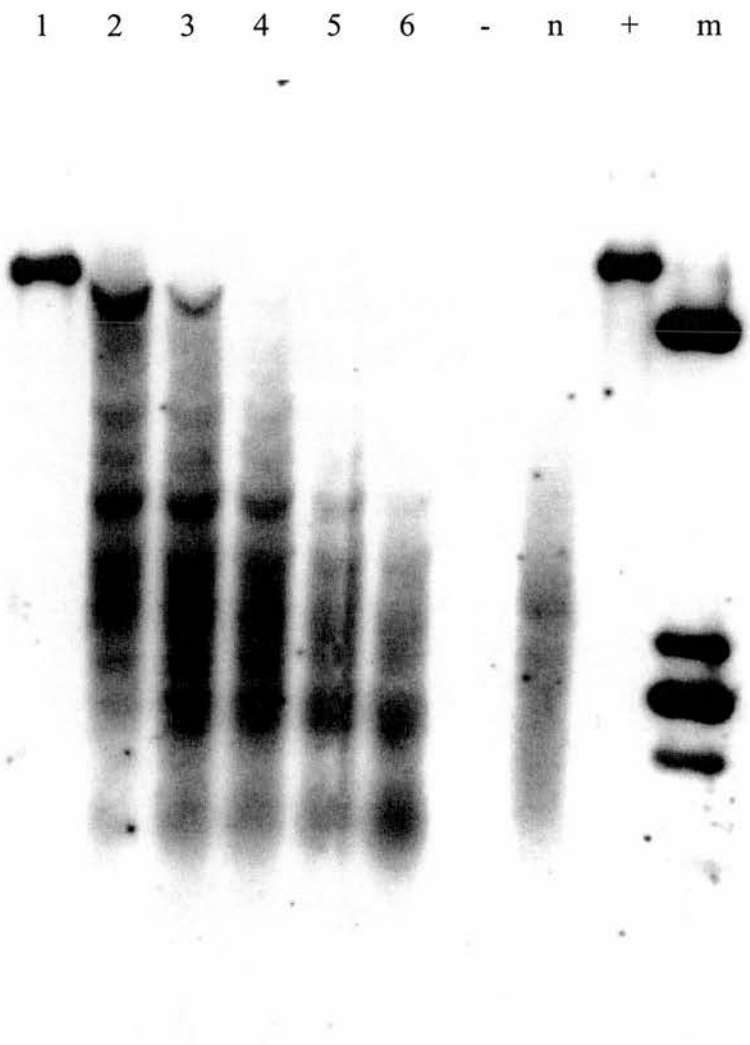



**Figure A14**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +4700UP probe.

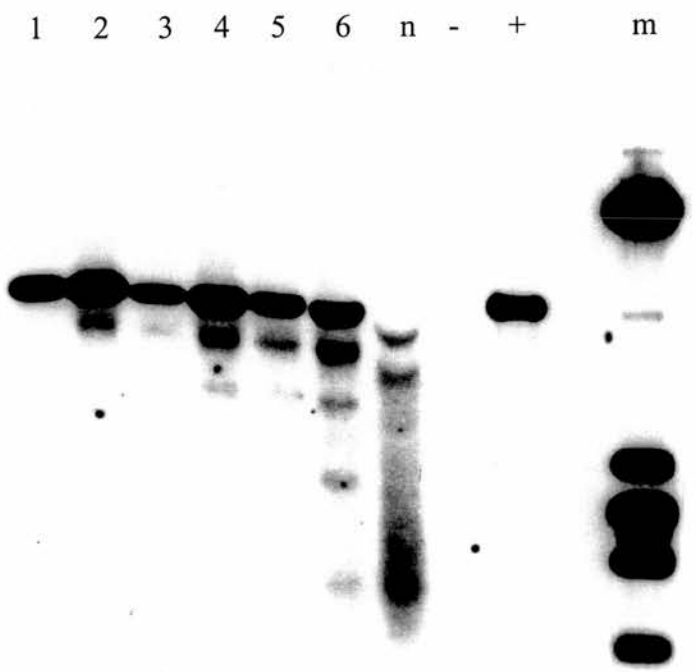



**Figure A15**

Liver Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +4700DOWN probe.

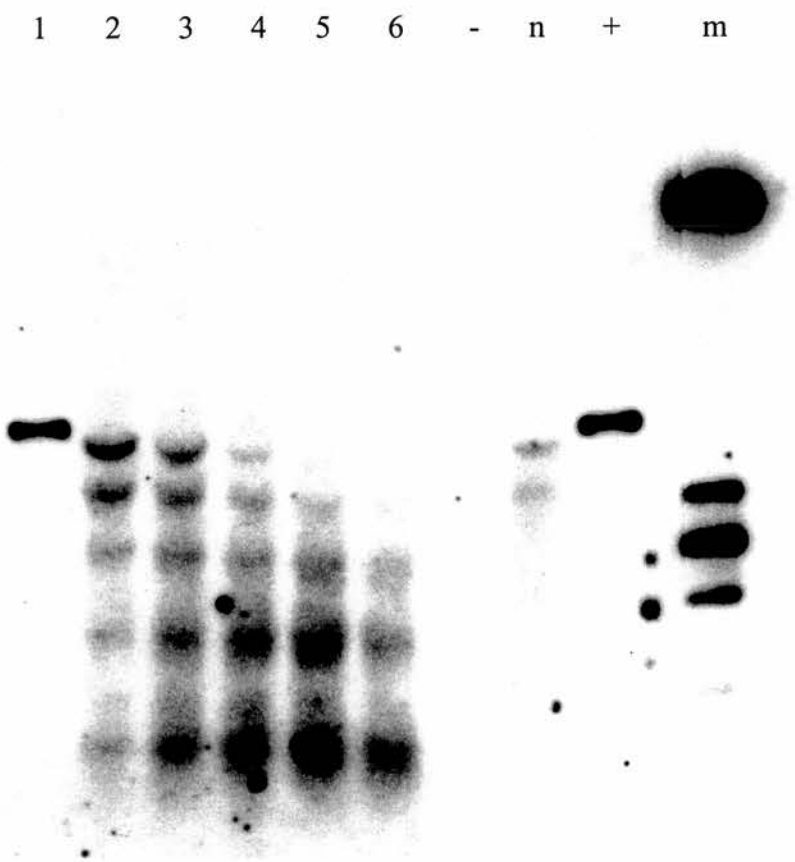



**Figure A16**

Mammary Nuclei digestion with OP-Cu (partial) and ***Bam*HI** (complete).
Probed with +4700UP probe.

# Reference List


Adachi, Y., Kas, E., and Laemmli, U. K. (1989)
**Preferential, cooperative binding of DNA topoisomerase II to scaffold- associated regions** EMBO J *8* p.3997-4006

Adroer, R. and Oliva, R. (1998)
**Nucleosome positioning in the rat protamine 1 gene in vivo and in vitro** Biochim Biophys Acta *1442* p.252-260

Ait, Si Ali, Ramirez, S., Barre, F. X., Dkhissi, F., Magnaghi-Jaulin, L., Girault, J. A., Robin, P., Knibiehler, M., Pritchard, L. L., Ducommun, B., Trouche, D., and Harel-Bellan, A. (1998)
**Histone acetyltransferase activity of CBP is controlled by cycle- dependent kinases and oncoprotein E1A** Nature *396* p.184-186

Alexander, M., Heppel, L. A., and Hurwitz, J. (1961)
**The Purification and Properties of Micrococcal Nuclease** J.Biol.Chem. *236* p.3014-3019

Ali, S. and Clark, A. J. (1988)
**Characterization of the gene encoding ovine beta-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins** J Mol Biol *199* p.415-426

Allan, J., Harborne, N., Rau, D. C., and Gould, H. (1982)
**Participation of core histone "tails" in the stabilization of the chromatin solenoid** J Cell Biol *93* p.285-297

Althaus, F. R., Hofferer, L., Kleczkowska, H. E., Malanga, M., Naegeli, H., Panzeter, P. L., and Realini, C. A. (1994)
**Histone shuttling by poly ADP-ribosylation** Mol Cell Biochem *138* p.53-59

Archer, T. K., Lefebvre, P., Wolford, R. G., and Hager, G. L. (1992)
**Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation** Science *255* p.1573-1576

Archibald, A. L., McClenaghan, M., Hornsey, V., Simons, J. P., and Clark, A. J. (1990)
**High-level expression of biologically active human alpha 1-antitrypsin in the milk of transgenic mice** Proc Natl Acad Sci U S A *87* p.5178-5182

Avery, O. T., MacLeod, C. M., and McCarthy, M. (1944)
**Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III** J.Exp.Med. *79* p.137-158

Ball, R. K., Friis, R. R., Schoenenberger, C. A., Doppler, W., and Groner, B. (1988)
**Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line** EMBO J *7* p.2089-2095

Bannister, A. J. and Kouzarides, T. (1996)
**The CBP co-activator is a histone acetyltransferase** Nature *384* p.641-643

Barahmand-Pour, F., Meinke, A., Groner, B., and Decker, T. (1998)
**Jak2-Stat5 interactions analyzed in yeast** J Biol Chem *273* p.12567-12575

Barlev, N. A., Candau, R., Wang, L., Darpino, P., Silverman, N., and Berger, S. L. (1995)
**Characterization of physical interactions of the putative transcriptional adaptor, ADA2, with acidic activation domains and TATA- binding protein** J Biol Chem *270* p.19337-19344

Bartsch, J., Truss, M., Bode, J., and Beato, M. (1996)
**Moderate increase in histone acetylation activates the mouse mammary tumor virus promoter and remodels its nucleosome structure** Proc Natl Acad Sci U S A *93* p.10741-10746

Beato, M., Herrlich, P., and Schutz, G. (1995)
**Steroid hormone receptors: many actors in search of a plot** Cell *83* p.851-857

Beato, M. and Eisfeld, K. (1997)
**Transcription factor access to chromatin** Nucleic Acids Res *25* p.3559-3563

Becker, P. B. (1994)
**The establishment of active promoters in chromatin** Bioessays *16* p.541-547
Becker, S., Groner, B., and Muller, C. W. (1998)
**Three-dimensional structure of the Stat3 beta homodimer bound to DNA** Nature *394* p.145-151
Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., and Woodcock, C. L. (1998)
**Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin** Proc Natl Acad Sci U S A *95* p.14173-14178
Belmont, A. S., Sedat, J. W., and Agard, D. A. (1987)
**A three-dimensional approach to mitotic chromosome structure: evidence for a complex hierarchical organization** J Cell Biol *105* p.77-92
Belmont, A. S., Braunfeld, M. B., Sedat, J. W., and Agard, D. A. (1989)
**Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro** Chromosoma *98* p.129-143
Ben-Or, S. and Okret, S. (1993)
**Involvement of a C/EBP-like protein in the acquisition of responsiveness to glucocorticoid hormones during chick neural retina development** Mol Cell Biol *13* p.331-340
Benezra, R., Cantor, C. R., and Axel, R. (1986)
**Nucleosomes are phased along the mouse beta-major globin gene in erythroid and nonerythroid cells** Cell *44* p.697-704
Bengal, E., Flores, O., Krauskopf, A., Reinberg, D., and Aloni, Y. (1991)
**Role of the mammalian transcription factors IIF, IIS, and IIX during elongation by RNA polymerase II** Mol Cell Biol *11* p.1195-1206
Benkirane, M., Chun, R. F., Xiao, H., Ogryzko, V. V., Howard, B. H., Nakatani, Y., and Jeang, K. T. (1998)
**Activation of integrated provirus requires histone acetyltransferase. p300 and P/CAF are coactivators for HIV-1 Tat** J Biol Chem *273* p.24898-24905
Benore-Parsons, M. and Ayoub, M. A. (1997)
**Presence of RNase A causes aberrant DNA band shifts** Biotechniques *23* p.128-131
Berchtold, S., Moriggl, R., Gouilleux, F., Silvennoinen, O., Beisenherz, C., Pfitzner, E., Wissler, M., Stocklin, E., and Groner, B. (1997)
**Cytokine receptor-independent, constitutively active variants of STAT5** J Biol Chem *272* p.30237-30243
Bestor, T. H. (1998)
**The host defence function of genomic methylation patterns** Novartis Found Symp *214* p.187-195
Bi, X. and Broach, J. R. (1997)
**DNA in transcriptionally silent chromatin assumes a distinct topology that is sensitive to cell cycle progression** Mol Cell Biol *17* p.7077-7087
Bishop, J. O. (1996)
**Chromosomal insertion of foreign DNA** Reprod Nutr Dev *36* p.607-618
Bjorklund, S., Almouzni, G., Davidson, I., Nightingale, K. P., and Weiss, K. (1999)
**Global transcription regulators of eukaryotes** Cell *96* p.759-767
Blank, T. A. and Becker, P. B. (1995)
**Electrostatic mechanism of nucleosome spacing** J Mol Biol *252* p.305-313
Blank, T. A. and Becker, P. B. (1996)
**The effect of nucleosome phasing sequences and DNA topology on nucleosome spacing** J Mol Biol *260* p.1-8
Blomquist, P., Li, Q., and Wrange, O. (1996)
**The affinity of nuclear factor 1 for its DNA site is drastically reduced by nucleosome**

**organization irrespective of its rotational or translational position** J Biol Chem *271* p.153-159

Bock, H., Abler, S., Zhang, X. Y., Fritton, H., and Igo-Kemenes, T. (1984)
**Positioning of nucleosomes in satellite I-containing chromatin of rat liver** J Mol Biol *176* p.131-154

Bohm, L. and Crane-Robinson, C. (1984)
**Proteases as structural probes for chromatin: the domain structure of histones** Biosci Rep *4* p.365-386

Bonifer, C., Huber, M. C., Jagle, U., Faust, N., and Sippel, A. E. (1996)
**Prerequisites for tissue specific and position independent expression of a gene locus in transgenic mice** J Mol Med *74* p.663-671

Bonner, J. and Ts'o, P. (1964)
**The Nucleohistones** Holden Day Inc. San Francisco 1st ed.

Bordas, J., Perez-Grau, L., Koch, M. H., Vega, M. C., and Nave, C. (1986)
**The superstructure of chromatin and its condensation mechanism. II. Theoretical analysis of the X-ray scattering patterns and model calculations** Eur Biophys J *13* p.175-185

Bradbury, E. M. (1992)
**Reversible histone modifications and the chromosome cell cycle** Bioessays *14* p.9-16

Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R. (1996)
**Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern** Mol Cell Biol *16* p.4349-4356

Bresnick, E. H., John, S., Berard, D. S., Lefebvre, P., and Hager, G. L. (1990)
**Glucocorticoid receptor-dependent disruption of a specific nucleosome on the mouse mammary tumor virus promoter is prevented by sodium butyrate** Proc Natl Acad Sci U S A *87* p.3977-3981

Bresnick, E. H., Bustin, M., Marsaud, V., Richard-Foy, H., and Hager, G. L. (1992)
**The transcriptionally-active MMTV promoter is depleted of histone H1** Nucleic Acids Res *20* p.273-278

Brinster, R. L., Allen, J. M., Behringer, R. R., Gelinas, R. E., and Palmiter, R. D. (1988)
**Introns increase transcriptional efficiency in transgenic mice** Proc Natl Acad Sci U S A *85* p.836-840

Brown, J. L., Mucci, D., Whiteley, M., Dirksen, M. L., and Kassis, J. A. (1998)
**The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1** Mol Cell *1* p.1057-1064

Brown, S. A., Imbalzano, A. N., and Kingston, R. E. (1996)
**Activator-dependent regulation of transcriptional pausing on nucleosomal templates** Genes Dev *10* p.1479-1490

Burdon, T. G., Maitland, K. A., Clark, A. J., Wallace, R., and Watson, C. J. (1994)
**Regulation of the sheep beta-lactoglobulin gene by lactogenic hormones is mediated by a transcription factor that binds an interferon-gamma activation site-related element** Mol Endocrinol *8* p.1528-1536

Cairns, B. R., Kim, Y. J., Sayre, M. H., Laurent, B. C., and Kornberg, R. D. (1994)
**A multi-subunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast** Proc Natl Acad Sci U S A *91* p.1950-1954

Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., and Kornberg, R. D. (1996)
**RSC, an essential, abundant chromatin-remodeling complex** Cell *87* p.1249-1260

Cairns, B. R. (1998)
**Chromatin remodeling machines: similar motors, ulterior motives** Trends Biochem Sci *23* p.20-25

Calladine, C. R. and Drew, H. R. (1992)
**Understanding DNA** 2nd ed. Academic Press, San Diego

Candido, E. P. and Dixon, G. H. (1972)
**Acetylation of histones in different cell types from developing trout testis** J Biol Chem *247* p.5506-5510
Caplan, A., Kimura, T., Gould, H., and Allan, J. (1987)
**Perturbation of chromatin structure in the region of the adult beta- globin gene in chicken erythrocyte chromatin** J Mol Biol *193* p.57-70
Cartwright, I. L. and Elgin, S. C. (1982)
**Analysis of chromatin structure and DNA sequence organization: use of the 1,10-phenanthroline-cuprous complex** Nucleic Acids Res *10* p.5835-5852
Cartwright, I. L., Hertzberg, R. P., Dervan, P. B., and Elgin, S. C. (1983)
**Cleavage of chromatin with methidiumpropyl-EDTA . iron(II)** Proc Natl Acad Sci U S A *80* p.3213-3217
Castro, C. E., Armstrong-Major, J., and Ramirez, M. E. (1986)
**Diet-mediated alteration of chromatin structure** Fed Proc *45* p.2394-2398
Cattini, P. A. and Allan, J. (1988)
**Locating the folded domain of H5 histone in chicken erythrocyte higher- order fiber** J Histochem Cytochem *36* p.425-432
Cella, N., Groner, B., and Hynes, N. E. (1998)
**Characterization of Stat5a and Stat5b homodimers and heterodimers and their association with the glucocortiocoid receptor in mammary cells** Mol Cell Biol *18* p.1783-1792
Cerdan, M. G., Young, J. I., Zino, E., Falzone, T. L., Otero, V., Torres, H. N., and Rubinstein, M. (1998)
**Accurate spatial and temporal transgene expression driven by a 3.8- kilobase promoter of the bovine beta-casein gene in the lactating mouse mammary gland** Mol Reprod Dev *49* p.236-245
Chang, C. H. and Luse, D. S. (1997)
**The H3/H4 tetramer blocks transcript elongation by RNA polymerase II in vitro** J Biol Chem *272* p.23427-23434
Chavez, S. and Beato, M. (1997)
**Nucleosome-mediated synergism between transcription factors on the mouse mammary tumor virus promoter** Proc Natl Acad Sci U S A *94* p.2885-2890
Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E. Jr, and Kuriyan, J. (1998)
**Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA** Cell *93* p.827-839
Cheng,T.H., Li,Y.C. and Gartenberg,M.R. (1998)
**Persistence of an alternate chromatin structure at silenced loci in the absence of silencers** Proc Natl Acad Sci U S A *95* p.5521-6.
Chiang, Y. C., Komarnitsky, P., Chase, D., and Denis, C. L. (1996)
**ADR1 activation domains contact the histone acetyltransferase GCN5 and the core transcriptional factor TFIIB** J Biol Chem *271* p.32359-32365
Clark, A. J., Ali, S., Archibald, A. L., Bessos, H., Brown, P., Harris, S., McClenaghan, M., Prowse, C., Simons, J. P., and Whitelaw, C. B. (1989)
**The molecular manipulation of milk composition** Genome *31* p.950-955
Clark, A. J., Cowper, A., Wallace, R., Wright, G., and Simons, J. P. (1992)
**Rescuing transgene expression by co-integration** Biotechnology (N Y) *10* p.1450-1454
Clark, A. J., Harold, G., and Yull, F. E. (1997)
**Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice** Nucleic Acids Res *25* p.1009-1014
Clark, D. J. and Kimura, T. (1990)
**Electrostatic mechanism of chromatin folding** J Mol Biol *211* p.883-896
Colot, V and Rossignol, J-L. (1999)
**Eukaryotic DNA methylatin as an evolutionary device** Bioessays *21* p.402-411
Compton, J. L., Bellard, M., and Chambon, P. (1976)
**Biochemical evidence of variability in the DNA repeat length in the chromatin of higher eukaryotes** Proc Natl Acad Sci U S A *73* p.4382-4386

Cook, J. L., Re, R. N., Giardina, J. F., Fontenot, F. E., Cheng, D. Y., and Alam, J. (1995)
**Distance constraints and stereospecific alignment requirements characteristic of p53 DNA-binding consensus sequence homologies** Oncogene *11* p.723-733
Cook, J. L., Zhang, Z., Alam, J., and Re, R. N. (1999)
**Effects of chromosomal integration site upon p53 interactions with DNA consensus sequence homologies** Oncogene *18* p.2373-2379
Cook, P. R. and Brazell, I. A. (1975)
**Supercoils in human DNA** J Cell Sci *19* p.261-279
Corona, D. F., Langst, G., Clapier, C. R., Bonte, E. J., Ferrari, S., Tamkun, J. W., and Becker, P. B. (1999)
**ISWI is an ATP-dependent nucleosome remodeling factor** Mol Cell *3* p.239-245
Cote, J., Peterson, C. L., and Workman, J. L. (1998)
**Perturbation of nucleosome core structure by the SWI/SNF complex persists after its detachment, enhancing subsequent transcription factor binding** Proc Natl Acad Sci U S A *95* p.4947-4952
Cotmore, S. F. and Tattersall, P. (1998)
**High-mobility group 1/2 proteins are essential for initiating rolling- circle-type DNA replication at a parvovirus hairpin origin** J Virol *72* p.8477-8484
Crippa, M. P., Trieschmann, L., Alfonso, P. J., Wolffe, A. P., and Bustin, M. (1993)
**Deposition of chromosomal protein HMG-17 during replication affects the nucleosomal ladder and transcriptional potential of nascent chromatin** EMBO J *12* p.3855-3864
Cross, S. H., Meehan, R. R., Nan, X., and Bird, A. (1997)
**A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins** Nat Genet *16* p.256-259
Csordas, A. (1989)
**A proposal for a possible role of nucleosome positioning in the evolutionary adjustment of introns** Int J Biochem *21* p.455-461
Cuatrecasas, P., Fuchs, S. and Anfinsen, C.B. (1967)
**Catalytic properties and specificity of the extracellular nuclease of Staphylococcus aureus** J Biol Chem *242* p.541-7
Cunningham, L. (1958)
**Natural 3'-deoxyribomononucleotides** J.Am.Chem.Soc. *80* p.2546-2549
Cusick, M. E., Lee, K. S., DePamphilis, M. L., and Wassarman, P. M. (1983)
**Structure of chromatin at deoxyribonucleic acid replication forks: nuclease hypersensitivity results from both prenucleosomal deoxyribonucleic acid and an immature chromatin structure** Biochemistry *22* p.3873-3884
Danielson, K. G., Oborn, C. J., Durban, E. M., Butel, J. S., and Medina, D. (1984)
**Epithelial mouse mammary cell line exhibiting normal morphogenesis in vivo and functional differentiation in vitro** Proc Natl Acad Sci U S A *81* p.3756-3760
Davey, C., Pennings, S., Meersseman, G., Wess, T. J., and Allan, J. (1995)
**Periodicity of strong nucleosome positioning sites around the chicken adult beta-globin gene may encode regularly spaced chromatin** Proc Natl Acad Sci U S A *92* p.11210-11214
Davey, C., Pennings, S., and Allan, J. (1997)
**CpG methylation remodels chromatin structure in vitro** J Mol Biol *267* p.276-288
Davies, N. and Lindsey, G. G. (1994)
**Histone H2B (and H2A) ubiquitination allows normal histone octamer and core particle reconstitution** Biochim Biophys Acta *1218* p.187-193
Davis, J. N. and Roussel, M. F. (1996)
**Cloning and expression of the murine Elf-1 cDNA** Gene *171* p.265-269
De Ambrosis, A., Ferrari, N., Bonassi, S., and Vidali, G. (1987)
**Nucleosomal repeat length in active and inactive genes** FEBS Lett *225* p.120-122
Dean, S. W., Tew, K. D., Clark, A. E., and Schein, P. S. (1985)
**DNA repeat length in chromatin from murine bone marrow and L1210 leukaemia cells** Br J Cancer *52* p.377-382

DeLange, R. J., Fambrough, D. M., Smith, E. L., and Bonner, J. (1969)
**Calf and pea histone IV. 3. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone** J Biol Chem *244* p.5669-5679

Demmer, J., Burdon, T. G., Djiane, J., Watson, C. J., and Clark, A. J. (1995)
**The proximal milk protein binding factor binding site is required for the prolactin responsiveness of the sheep beta-lactoglobulin promoter in Chinese hamster ovary cells** Mol Cell Endocrinol *107* p.113-121

Di Croce, L., Koop, R., Venditti, P., Westphal, H. M., Nightingale, K. P., Corona, D. F., Becker, P. B., and Beato, M. (1999)
**Two-step synergism between the progesterone receptor and the DNA- binding domain of nuclear factor 1 on MMTV minichromosomes** Mol Cell *4* p.45-54

Dimitrov, S. I., Russanova, V. R., and Pashev, I. G. (1987)
**The globular domain of histone H5 is internally located in the 30 nm chromatin fiber: an immunochemical study** EMBO J *6* p.2387-2392

Ding, H. F., Rimsky, S., Batson, S. C., Bustin, M., and Hansen, U. (1994)
**Stimulation of RNA polymerase II elongation by chromosomal protein HMG- 14** Science *265* p.796-799

Ding, H. F., Bustin, M., and Hansen, U. (1997)
**Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14** Mol Cell Biol *17* p.5843-5855

Dingwall, A. K., Beek, S. J., McCallum, C. M., Tamkun, J. W., Kalpana, G. V., Goff, S. P., and Scott, M. P. (1995)
**The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex** Mol Biol Cell *6* p.777-791

Dingwall, C., Lomonossoff, G. P., and Laskey, R. A. (1981)
**High sequence specificity of micrococcal nuclease** Nucleic Acids Res *9* p.2659-2673

Dobie, K. W., Lee, M., Fantes, J. A., Graham, E., Clark, A. J., Springbett, A., Lathe, R., and McClenaghan, M. (1996)
**Variegated transgene expression in mouse mammary gland is determined by the transgene integration locus** Proc Natl Acad Sci U S A *93* p.6659-6664

Dong, F. and van Holde, K. E. (1991)
**Nucleosome positioning is determined by the (H3-H4)2 tetramer** Proc Natl Acad Sci U S A *88* p.10596-10600

Doppler, W., Hock, W., Hofer, P., Groner, B., and Ball, R. K. (1990)
**Prolactin and glucocorticoid hormones control transcription of the beta- casein gene by kinetically distinct mechanisms** Mol Endocrinol *4* p.912-919

Dorer, D. R. and Henikoff, S. (1994)
**Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophila** Cell *77* p.993-1002

Drew, H. R. and Travers, A. A. (1985)
**DNA bending and its relation to nucleosome positioning** J Mol Biol *186* p.773-790

Drew, H. R. and Calladine, C. R. (1987)
**Sequence-specific positioning of core histones on an 860 base-pair DNA. Experiment and theory** J Mol Biol *195* p.143-173

Du, W., Thanos, D., and Maniatis, T. (1993)
**Mechanisms of transcriptional synergism between distinct virus- inducible enhancer elements** Cell *74* p.887-898

Duerre, J. A., Wallwork, J. C., Quick, D. P., and Ford, K. M. (1977)
**In vitro studies on the methylation of histones in rat brain nuclei** J Biol Chem *252* p.6981-6985

Edery, M., Pang, K., Larson, L., Colosi, T., and Nandi, S. (1985)
**Epidermal growth factor receptor levels in mouse mammary glands in various physiological states** Endocrinology *117* p.405-411

Edmondson, D. G., Smith, M. M., and Roth, S. Y. (1996)
**Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4** Genes Dev *10* p.1247-1259
Ehmann, A., Chafin, D., Lee, K. M., and Hayes, J. J. (1998)
**(1,4,7-trimethyl-1,4,7-triazacyclononane)iron (III)-mediated cleavage of DNA: detection of selected protein-DNA interactions** Nucleic Acids Res *26* p.2086-2091
Eisfeld, K., Candau, R., Truss, M., and Beato, M. (1997)
**Binding of NF1 to the MMTV promoter in nucleosomes: influence of rotational phasing, translational positioning and histone H1** Nucleic Acids Res *25* p.3733-3742
Elgin, S. C. (1981)
**DNAase I-hypersensitive sites of chromatin** Cell *27* p.413-415
Elgin, S. C. (1984)
**Anatomy of hypersensitive sites** Nature *309* p.213-214
Elgin, S. C. (1988)
**The formation and function of DNase I hypersensitive sites in the process of gene activation** J Biol Chem *263* p.19259-19262
Englander, E. W., Wolffe, A. P., and Howard, B. H. (1993)
**Nucleosome interactions with a human Alu element. Transcriptional repression and effects of template methylation** J Biol Chem *268* p.19565-19573
Fedor, M. J., Lue, N. F., and Kornberg, R. D. (1988)
**Statistical positioning of nucleosomes by specific protein-binding to an upstream activating sequence in yeast** J Mol Biol *204* p.109-127
Feinberg, A. P. and Vogelstein, B. (1983)
**A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity** Anal Biochem *132* p.6-13
Feinberg, A. P. and Vogelstein, B. (1984)
**"A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum** Anal Biochem *137* p.266-267
Felsenfeld, G. and McGhee, J. D. (1986)
**Structure of the 30 nm chromatin fiber** Cell *44* p.375-377
Fitzgerald, D. J. and Anderson, J. N. (1998)
**Unique translational positioning of nucleosomes on synthetic DNAs** Nucleic Acids Res *26* p.2526-2535
Flaus, A. and Richmond, T. J. (1998)
**Positioning and stability of nucleosomes on MMTV 3'LTR sequences** J Mol Biol *275* p.427-441
Fletcher, T. M. and Hansen, J. C. (1996)
**The nucleosomal array: structure/function relationships** Crit Rev Eukaryot Gene Expr *6* p.149-188
Folch, J. M., Coll, A., and Sanchez, A. (1994)
**Complete sequence of the caprine beta-lactoglobulin gene** J Dairy Sci *77* p.3493-3497
Forsyth, I. A. (1986)
**Variation among species in the endocrine control of mammary growth and function: the roles of prolactin, growth hormone, and placental lactogen** J Dairy Sci *69* p.886-903
Fragoso, G., John, S., Roberts, M. S., and Hager, G. L. (1995)
**Nucleosome positioning on the MMTV LTR results from the frequency- biased occupancy of multiple frames** Genes Dev *9* p.1933-1947
Funder, J. W. (1989)
**Hormonal regulation of gene expression** Biochem Soc Symp *55* p.105-114
Futterman, S. and Heller, J. (1972)
**The enhancement of fluorescence and the decreased susceptibility to enzymatic oxidation of retinol complexed with bovine serum albumin, - lactoglobulin, and the retinol-binding protein of human plasma** J Biol Chem *247* p.5168-5172

Garrick, D., Fiering, S., Martin, D. I., and Whitelaw, E. (1998)
**Repeat-induced gene silencing in mammals** Nat Genet *18* p.56-59
Gasser, S. M. and Laemmli, U. K. (1986)
**The organisation of chromatin loops: characterisation of a scaffold attachment site** EMBO J *5* p.511-518
Gaubatz, J., Ellis, M., and Chalkley, R. (1979)
**The structural organization of mouse chromatin as a function of age** Fed Proc *38* p.1973-1978
Gaye, P., Hue-Delahaie, D., Mercier, J. C., Soulier, S., Vilotte, J. L., and Furet, J. P. (1986)
**Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland** Biochimie *68* p.1097-1107
Geissler, E. N., Cheng, S. V., Gusella, J. F., and Housman, D. E. (1988)
**Genetic analysis of the dominant white-spotting (W) region on mouse chromosome 5: identification of cloned DNA markers near W** Proc Natl Acad Sci U S A *85* p.9635-9639
Georgakopoulos, T. and Thireos, G. (1992)
**Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription** EMBO J *11* p.4145-4152
Gerchman, S. E. and Ramakrishnan, V. (1987)
**Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy** Proc Natl Acad Sci U S A *84* p.7802-7806
Giles, R. H., Peters, D. J., and Breuning, M. H. (1998)
**Conjunction dysfunction: CBP/p300 in human disease** Trends Genet *14* p.178-183
Godde, J. S. and Widom, J. (1992)
**Chromatin structure of Schizosaccharomyces pombe. A nucleosome repeat length that is shorter than the chromatosomal DNA length** J Mol Biol *226* p.1009-1025
Godde, J. S., Kass, S. U., Hirst, M. C., and Wolffe, A. P. (1996)
**Nucleosome assembly on methylated CGG triplet repeats in the fragile X mental retardation gene 1 promoter** J Biol Chem *271* p.24325-24328
Goodman, H. S. and Rosen, J. M. (1990)
**Transcriptional analysis of the mouse beta-casein gene** Mol Endocrinol *4* p.1661-1670
Goodwin, G. H., Mathew, C. G., Wright, C. A., Venkov, C. D., and Johns, E. W. (1979)
**Analysis of the high mobility group proteins associated with salt- soluble nucleosomes** Nucleic Acids Res *7* p.1815-1835
Gottesfeld, J. M. (1980)
**Organization of 5S genes in chromatin of Xenopus laevis** Nucleic Acids Res *8* p.905-922
Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L., and Workman, J. L. (1997)
**Yeast Gcn5 functions in two multi-subunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex** Genes Dev *11* p.1640-1650
Groenen, A. M. and van der Poel, J. J. (1994)
**Regulation of Expression of Milk Protein Genes: A Review** Livestock Production Science *38* p.61-78
Gruenbaum, Y., Szyf, M., Cedar, H., and Razin, A. (1983)
**Methylation of replicating and post-replicated mouse L-cell DNA** Proc Natl Acad Sci U S A *80* p.4919-4921
Grunstein, M. (1997)
**Histone acetylation in chromatin structure and transcription** Nature *389* p.349-352
Gunjan, A. and Brown, D. T. (1999)
**Overproduction of histone H1 variants in vivo increases basal and induced activity of the mouse mammary tumor virus promoter** Nucleic Acids Res *27* p.3355-3363
Gurley, L. R., Tobey, R. A., Walters, R. A., Hildebrand, C. E., Hohmann, P. G., D'Anna, J. A., Barham, S. S., and Deavan, L. L. (1978)
**Cell Cycle Regulation** 1st ed. Academic Press, New York p.37-60

Guyer, N. B., Severns, C. W., Wong, P., Feghali, C. A., and Wright, T. M. (1995)
**IFN-gamma induces a p91/Stat1 alpha-related transcription factor with distinct activation and binding properties** J Immunol *155* p.3472-3480
Hall, L., Emery, D. C., Davies, M. S., Parker, D., and Craig, R. K. (1987)
**Organization and sequence of the human alpha-lactalbumin gene** Biochem J *242* p.735-742
Hambling, S. G., McAlpine, A. S., and Sawyer, L. (1992)
**Advances in Dairy Chemistry** *Vol.I, Proteins* p.141-190
Hamiche, A., Sandaltzopoulos, R., Gdula, D. A., and Wu, C. (1999)
**ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF** Cell *97* p.833-842
Han, Y., Watling, D., Rogers, N. C., and Stark, G. R. (1997)
**JAK2 and STAT5, but not JAK1 and STAT1, are required for prolactin- induced beta-lactoglobulin transcription** Mol Endocrinol *11* p.1180-1188
Harris, S., Ali, S., Anderson, S., Archibald, A. L., and Clark, A. J. (1988)
**Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene** Nucleic Acids Res *16* p.10379-10380
Harris, S., McClenaghan, M., Simons, J. P., Ali, S., and Clark, A. J. (1991)
**Developmental regulation of the sheep beta-lactoglobulin gene in the mammary gland of transgenic mice** Dev Genet *12* p.299-307
Hart, C. M. and Laemmli, U. K. (1998)
**Facilitation of chromatin dynamics by SARs** Curr Opin Genet Dev *8* p.519-525
Hartzog, G. A., Wada, T., Handa, H., and Winston, F. (1998)
**Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in Saccharomyces cerevisiae** Genes Dev *12* p.357-369
Hayes, H., Petit, E., Lemieux, N., and Dutrillaux, B. (1992)
**Chromosomal localization of the ovine beta-casein gene by non-isotopic in situ hybridization and R-banding** Cytogenet Cell Genet *61* p.286-288
Hayes, H., Petit, E., Bouniol, C., and Popescu, P. (1993)
**Localization of the alpha-S2-casein gene (CASAS2) to the homologous cattle, sheep, and goat chromosomes 4 by in situ hybridization** Cytogenet Cell Genet *64* p.281-285
Hayes, J. J., Tullius, T. D., and Wolffe, A. P. (1990)
**The structure of DNA in a nucleosome** Proc Natl Acad Sci U S A *87* p.7405-7409
Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988)
**A direct link between core histone acetylation and transcriptionally active chromatin** EMBO J *7* p.1395-1402
Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S. M., and Grunstein, M. (1995)
**Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast** Cell *80* p.583-592
Henikoff, S. (1998)
**Conspiracy of silence among repeated transgenes** Bioessays *20* p.532-535
Hertwig, O. and Jenaische Z. (1885)
**(Nuclein is the genetic material)** Naturwiss *18* p.276-318
Hertzberg, R. P. and Dervan, P. B. (1984)
**Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses** Biochemistry *23* p.3934-3945
Hoey, T. and Schindler, U. (1998)
**STAT structure and function in signaling** Curr Opin Genet Dev *8* p.582-587
Hohmann, P. (1983)
**Phosphorylation of H1 histones** Mol Cell Biochem *57* p.81-92
Holliday, R. (1987)
**The inheritance of epigenetic defects** Science *238* p.163-170
Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., and Bradbury, E. M. (1993)
**Studies of the DNA binding properties of histone H4 amino terminus. Thermal**

**denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA** J Biol Chem *268* p.305-314

Horz, W. and Altenburger, W. (1981)
**Sequence specific cleavage of DNA by micrococcal nuclease** Nucleic Acids Res *9* p.2643-2658

Horz, W., Fittler, F., and Zachau, H. G. (1983)
**Sequence specific cleavage of African green monkey alpha-satellite DNA by micrococcal nuclease** Nucleic Acids Res *11* p.4275-4285

Huber, M. C., Kruger, G., and Bonifer, C. (1996)
**Genomic position effects lead to an inefficient reorganization of nucleosomes in the 5'-regulatory region of the chicken lysozyme locus in transgenic mice** Nucleic Acids Res *24* p.1443-1452

Hynes, N. E., Taverna, D., Harwerth, I. M., Ciardiello, F., Salomon, D. S., Yamamoto, T., and Groner, B. (1990)
**Epidermal growth factor receptor, but not c-erbB-2, activation prevents lactogenic hormone induction of the beta-casein gene in mouse mammary epithelial cells** Mol Cell Biol *10* p.4027-4034

Isenberg, I. (1978)
**The Cell Nucleus** 1st ed. Vol.4 pp.135-154 Academic Press

Ito, T., Bulger, M., Pazin, M. J., Kobayashi, R., and Kadonaga, J. T. (1997)
**ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor** Cell *90* p.145-155

Izaurralde, E., Kas, E., and Laemmli, U. K. (1989)
**Highly preferential nucleation of histone H1 assembly on scaffold- associated regions** J Mol Biol *210* p.573-585

Jackson, D. A., Dickinson, P., and Cook, P. R. (1990)
**The size of chromatin loops in HeLa cells** EMBO J *9* p.567-571

Jahn, G. A., Daniel, N., Jolivet, G., Belair, L., Bole-Feysot, C., Kelly, P. A., and Djiane, J. (1997)
**In vivo study of prolactin (PRL) intracellular signalling during lactogenesis in the rat: JAK/STAT pathway is activated by PRL in the mammary gland but not in the liver** Biol Reprod *57* p.894-900

Jeddeloh, J. A., Stokes, T. L., and Richards, E. J. (1999)
**Maintenance of genetic methylation requires a SWI2/SNF2-like protein** Nat Genet *22* p.94-97

Jenness, R. (1986)
**Lactational performance of various mammalian species** J Dairy Sci *69* p.869-885

Jenster, G., Spencer, T. E., Burcin, M. M., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1997)
**Steroid receptor induction of gene transcription: a two-step model** Proc Natl Acad Sci U S A *94* p.7879-7884

Jeong, S. W., Lauderdale, J. D., and Stein, A. (1991)
**Chromatin assembly on plasmid DNA in vitro. Apparent spreading of nucleosome alignment from one region of pBR327 by histone H5** J Mol Biol *222* p.1131-1147

Johansson, E., Hjortsberg, K., and Thelander, L. (1998)
**Two YY-1-binding proximal elements regulate the promoter strength of the TATA-less mouse ribonucleotide reductase R1 gene** J Biol Chem *273* p.29816-29821

Johns, E. W., Phillips, D. M. P., Simpson, P., and Butler, J. A. V. (1960)
**The electrophoresis of histones in polyacrylamide gel and their quantitative determination** Biochem.J. *77* p.631-636

Johns, E. W. (1964)
**Studies on histones. Preparative methods for histone fractions from calf thymus.** Biochem.J. *92* p.55-59

Johnson, C. A., O'Neill, L. P., Mitchell, A., and Turner, B. M. (1998)
**Distinctive patterns of histone H4 acetylation are associated with defined sequence elements**

**within both heterochromatic and euchromatic regions of the human genome** Nucleic Acids Res *26* p.994-1001

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998)
**Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription** Nat Genet *19* p.187-191

Kahmann, N. H. and Rake, A. V. (1993)
**Altered nucleosome spacing associated with Down syndrome** Biochem Genet *31* p.207-214

Kandolf, H. (1994)
**The H1A histone variant is an in vivo repressor of oocyte-type 5S gene transcription in Xenopus laevis embryos** Proc Natl Acad Sci U S A *91* p.7257-7261

Kass, S. U., Pruss, D., and Wolffe, A. P. (1997a)
**How does DNA methylation repress transcription?** Trends Genet *13* p.444-449

Kass, S. U., Landsberger, N., and Wolffe, A. P. (1997b)
**DNA methylation directs a time-dependent repression of transcription initiation** Curr Biol *7* p.157-165

Katsani, K. R., Hajibagheri, M. A., and Verrijzer, C. P. (1999)
**Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology** EMBO J *18* p.698-708

Kawasaki, H., Eckner, R., Yao, T. P., Taira, K., Chiu, R., Livingston, D. M., and Yokoyama, K. K. (1998)
**Distinct roles of the co-activators p300 and CBP in retinoic-acid- induced F9-cell differentiation** Nature *393* p.284-289

Kazansky, A. V., Raught, B., Lindsey, S. M., Wang, Y. F., and Rosen, J. M. (1995)
**Regulation of mammary gland factor/Stat5a during mammary gland development** Mol Endocrinol *9* p.1598-1609

Kazansky, A. V., Kabotyanski, E. B., Wyszomierski, S. L., Mancini, M. A., and Rosen, J. M. (1999)
**Differential effects of prolactin and src/abl kinases on the nuclear translocation of STAT5B and STAT5A** J Biol Chem *274* p.22484-22492

Keene, M. A. and Elgin, S. C. (1981)
**Micrococcal nuclease as a probe of DNA sequence organization and chromatin structure** Cell *27* p.57-64

Keshet, I., Lieman-Hurwitz, J., and Cedar, H. (1986)
**DNA methylation affects the formation of active chromatin** Cell *44* p.535-543

Kladde, M. P. and Simpson, R. T. (1994)
**Positioned nucleosomes inhibit Dam methylation in vivo** Proc Natl Acad Sci U S A *91* p.1361-1365

Kold, H. J. and Braunitzer, G. (1983)
**The Primary Structure of β-Lactoglobulin, 2. Discussion and Genetic Aspects** Milchwissenschaft *38* p.70-72

Kornberg, R. D. (1974a)
**Chromatin structure: a repeating unit of histones and DNA** Science *184* p.868-871

Kornberg, R. D. and Thomas, J. O. (1974b)
**Chromatin structure; oligomers of the histones** Science *184* p.865-868

Kornberg, R. D. and Stryer, L. (1988)
**Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism** Nucleic Acids Res *16* p.6677-6690

Kornberg, R. D. and Lorch, Y. (1999)
**Chromatin-modifying and -remodeling complexes** Curr Opin Genet Dev *9* p.148-151

Kruger, G., Huber, M. C., and Bonifer, C. (1999)
**The -3.9 kb DNaseI hypersensitive site of the chicken lysozyme locus harbours an enhancer with unusual chromatin reorganizing activity** Gene *236* p.63-77

Kruger, W., Peterson, C. L., Sil, A., Coburn, C., Arents, G., Moudrianakis, E. N., and Herskowitz, I. (1995)

**Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription** Genes Dev *9* p.2770-2779

Kubista, M., Hagmar, P., Nielsen, P. E., and Norden, B. (1990)
**Reinterpretation of linear dichroism of chromatin supports a perpendicular linker orientation in the folded state** J Biomol Struct Dyn *8* p.37-54

Kumar, S., Clarke, A. R., Hooper, M. L., Horne, D. S., Law, A. J., Leaver, J., Springbett, A., Stevenson, E., and Simons, J. P. (1994)
**Milk composition and lactation of beta-casein-deficient mice** Proc Natl Acad Sci U S A *91* p.6138-6142

Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E., and Green, M. R. (1994)
**Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex** Nature *370* p.477-481

Laird, P. W., Zijderveld, A., Linders, K., Rudnicki, M. A., Jaenisch, R., and Berns, A. (1991)
**Simplified mammalian DNA isolation procedure** Nucleic Acids Res *19* p.4293

Langley, B., Vilotte, J. L., Stinnakre, M-G., Whitelaw, C. B., and L'Huillier, P. J. (1998)
**Rescue of an MMTV Transgene by Co-injection Reveals Novel Locus Control Properties of the Ovine β-Lactoglobulin Gene that Confer Locus Commitment to Heterogenious Tissues** Transgenic Res *7* p.205-212

Langst, G., Bonte, E. J., Corona, D. F., and Becker, P. B. (1999)
**Nucleosome movement by CHRAC and ISWI without disruption or trans- displacement of the histone octamer** Cell *97* p.843-852

Laskey, R. A., Mills, A. D., and Morris, N. R. (1977)
**Assembly of SV40 chromatin in a cell-free system from Xenopus eggs** Cell *10* p.237-243

Laybourn, P. J. and Kadonaga, J. T. (1992)
**Threshold phenomena and long-distance activation of transcription by RNA polymerase II** Science *257* p.1682-1685

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993)
**A positive role for histone acetylation in transcription factor access to nucleosomal DNA** Cell *72* p.73-84

Lee, J. S., Galvin, K. M., See, R. H., Eckner, R., Livingston, D., Moran, E., and Shi, Y. (1995)
**Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300** Genes Dev *9* p.1188-1198

Lee, M. S. and Garrard, W. T. (1991)
**Transcription-induced nucleosome 'splitting': an underlying structure for DNase I sensitive chromatin** EMBO J *10* p.607-615

Leonardson, K. E. and Levy, S. B. (1989)
**Chromatin reorganization during emergence of malignant Friend tumors: early changes in H2A and H2B variants and nucleosome repeat length** Exp Cell Res *180* p.209-219

LeRoy, G., Orphanides, G., Lane, W. S., and Reinberg, D. (1998)
**Requirement of RSF and FACT for transcription of chromatin templates in vitro** Science *282* p.1900-1904

Leuba, S. H., Zlatanova, J., and van Holde, K. (1993)
**On the location of histones H1 and H5 in the chromatin fiber. Studies with immobilized trypsin and chymotrypsin** J Mol Biol *229* p.917-929

Leuba, S. H., Yang, G., Robert, C., Samori, B., van Holde, K., Zlatanova, J., and Bustamante, C. (1994a)
**Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy** Proc Natl Acad Sci U S A *91* p.11621-11625

Leuba, S. H., Zlatanova, J., and van Holde, K. (1994b)
**On the location of linker DNA in the chromatin fiber. Studies with immobilized and soluble micrococcal nuclease** J Mol Biol *235* p.871-880

Leveziel, H., Metenier, L., Guerin, G., Cullen, P., Provot, C., Bertaud, M. and Mercier, J.C. (1991)
**Restriction fragment length polymorphism of ovine casein genes: close linkage between the alpha s1-, alpha s2-, beta- and kappa-casein loci** Anim Genet *22* p.1-10

Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992)

**Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA** Cell *69* p.905-14.

Lis, J. and Wu, C. (1993)
**Protein traffic on the heat shock promoter: parking, stalling, and trucking along** Cell *74* p.1-4

Liu, J. K., DiPersio, C. M., and Zaret, K. S. (1991)
**Extracellular signals that regulate liver transcription factors during hepatic differentiation in vitro** Mol Cell Biol *11* p.773-784

Liu, X., Robinson, G. W., Gouilleux, F., Groner, B., and Hennighausen, L. (1995)
**Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in prolactin signal transduction in mouse mammary tissue** Proc Natl Acad Sci U S A *92* p.8831-8835

Liu, X., Robinson, G. W., Wagner, K. U., Garrett, L., Wynshaw-Boris, A., and Hennighausen, L. (1997)
**Stat5a is mandatory for adult mammary gland development and lactogenesis** Genes Dev *11* p.179-186

Liu, X., Gallego, M. I., Smith, G. H., Robinson, G. W., and Hennighausen, L. (1998)
**Functional release of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b** Cell Growth Differ *9* p.795-803

Lohr, D. and Lopez, J. (1995)
**GAL4/GAL80-dependent nucleosome disruption/deposition on the upstream regions of the yeast GAL1-10 and GAL80 genes** J Biol Chem *270* p.27671-8

Lu, Q., Wallrath, L. L., and Elgin, S. C. (1995)
**The role of a positioned nucleosome at the Drosophila melanogaster hsp26 promoter** EMBO J *14* p.4738-4746

Lubon, H. and Hennighausen, L. (1988)
**Conserved region of the rat alpha-lactalbumin promoter is a target site for protein binding in vitro** Biochem J *256* p.391-396

Lubon, H., Pittius, C. W., and Hennighausen, L. (1989)
**In vitro transcription of the mouse whey acidic protein promoter is affected by upstream sequences** FEBS Lett *251* p.173-176

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997)
**Crystal structure of the nucleosome core particle at 2.8 A resolution** Nature *389* p.251-260

MacDougald, O. A., Cornelius, P., Lin, F. T., Chen, S. S., and Lane, M. D. (1994)
**Glucocorticoids reciprocally regulate expression of the CCAAT/enhancer- binding protein alpha and delta genes in 3T3-L1 adipocytes and white adipose tissue** J Biol Chem *269* p.19041-19047

Mahadevan, L. C., Willis, A. C., and Barratt, M. J. (1991)
**Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors** Cell *65* p.775-783

Marsden, M. P. and Laemmli, U. K. (1979)
**Metaphase chromosome structure: evidence for a radial loop model** Cell *17* p.849-858

Martens, J. A., Genereaux, J., Saleh, A., and Brandl, C. J. (1996)
**Transcriptional activation by yeast PDR1p is inhibited by its association with NGG1p/ADA3p** J Biol Chem *271* p.15884-15890

McClenaghan, M., Archibald, A. L., Harris, S., Simons, J. P., Whitelaw, C. B. A., Wilmut, I., and Clark, A. J. (1991)
**Production of Human alpha1-Antitrypsin in the Milk of Transgenic Sheep and Mice: Targetting Expression of cDNA Sequences to the Mammary Gland** Animal Biotechnology *2* p.161-176

McConkey, E. H., Menon, R., Williams, G., Baker, E., and Sutherland, G. R. (1996)
**Assignment of the gene for beta-casein (CSN2) to 4q13 --> q21 in humans and 3p13 --> p12 in chimpanzees** Cytogenet Cell Genet *72* p.60-62

McGhee, J. D., Rau, D. C., Charney, E., and Felsenfeld, G. (1980)
**Orientation of the nucleosome within the higher order structure of chromatin** Cell *22* p.87-96

McPherson, C. E., Shim, E. Y., Friedman, D. S., and Zaret, K. S. (1993)
**An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array** Cell *75* p.387-398
McPherson, C. E., Horowitz, R., Woodcock, C. L., Jiang, C., and Zaret, K. S. (1996)
**Nucleosome positioning properties of the albumin transcriptional enhancer** Nucleic Acids Res *24* p.397-404
Meehan, R. R., Lewis, J. D., and Bird, A. P. (1992)
**Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA** Nucleic Acids Res *20* p.5085-5092
Meersseman, G., Pennings, S., and Bradbury, E. M. (1992)
**Mobile nucleosomes-a general behavior** EMBO J *11* p.2951-2959
Meier, V. S. and Groner, B. (1994)
**The nuclear factor YY1 participates in repression of the beta-casein gene promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induction** Mol Cell Biol *14* p.128-137
Mercier, J. C. and Vilotte, J. L. (1993)
**Structure and function of milk protein genes** J Dairy Sci *76* p.3079-3098
Meyer, W. K., Reichenbach, P., Schindler, U., Soldaini, E., and Nabholz, M. (1997)
**Interaction of STAT5 dimers on two low affinity binding sites mediates interleukin 2 (IL 2) stimulation of IL-2 receptor alpha gene transcription** J Biol Chem *272* p.31821-31828
Miescher, F. (1874)
**(Isolation of salmon sperm DNA)** Ver.Naturforsch.Ges. *6* p.138-208
Mink, S., Haenig, B., and Klempnauer, K. H. (1997)
**Interaction and functional collaboration of p300 and C/EBPbeta** Mol Cell Biol *17* p.6609-6617
Montoliu, L., Umland, T., and Schutz, G. (1996)
**A locus control region at -12 kb of the tyrosinase gene** EMBO J *15* p.6026-6034
Moriggl, R., Berchtold, S., Friedrich, K., Standke, G. J., Kammer, W., Heim, M., Wissler, M., Stocklin, E., Gouilleux, F., and Groner, B. (1997)
**Comparison of the transactivation domains of Stat5 and Stat6 in lymphoid cells and mammary epithelial cells** Mol Cell Biol *17* p.3663-3678
Moss, D.E. and Fahrney, D.E. (1978)
**Phenylmethysulphonyl fluoride inhibits serine proteases** Biochem Pharmacol *27* pp.2693
Muller, C., Alunni-Fabbroni, M., Kowenz-Leutz, E., Mo, X., Tommasino, M., and Leutz, A. (1999)
**Separation of C/EBPalpha mediated proliferation arrest and differentiation pathways** Proc Natl Acad Sci U S A *96* p.7276-7281
Muyldermans, S. and Travers, A. A. (1994)
**DNA sequence organization in chromatosomes** J Mol Biol *235* p.855-870
Mymryk, J. S., Berard, D., Hager, G. L., and Archer, T. K. (1995)
**Mouse mammary tumor virus chromatin in human breast cancer cells is constitutively hypersensitive and exhibits steroid hormone independent loading of transcription factors in vivo** Mol Cell Biol *15* p.26-34
Nan, X., Campoy, F. J., and Bird, A. (1997)
**MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin** Cell *88* p.471-481
Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998)
**Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex** Nature *393* p.386-389
Nardacci, N. J., Lee, J. W., and McGuire, W. L. (1978)
**Differential regulation of alpha-lactalbumin and casein messenger RNA's in mammary tissue** Cancer Res *38* p.2694-2699
Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L., and Peterson, P. A. (1984)
**The three-dimensional structure of retinol-binding protein** EMBO J *3* p.1451-1454

Nickel, B. E., Allis, C. D., and Davie, J. R. (1989)
**Ubiquitinated histone H2B is preferentially located in transcriptionally active chromatin** Biochemistry *28* p.958-963
Nissen, M. S. and Reeves, R. (1995)
**Changes in superhelicity are introduced into closed circular DNA by binding of high mobility group protein I/Y** J Biol Chem *270* p.4355-4360
Noll, M. and Kornberg, R. D. (1977)
**Action of micrococcal nuclease on chromatin and the location of histone H1** J Mol Biol *109* p.393-404
Norton, V. G., Imai, B. S., Yau, P., and Bradbury, E. M. (1989)
**Histone acetylation reduces nucleosome core particle linking number change** Cell *57* p.449-457
O'Brien, T., Hardin, S., Greenleaf, A., and Lis, J. T. (1994)
**Phosphorylation of RNA polymerase II C terminal domain and transcriptional elongation** Nature *370* p.75-77
Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. (1996)
**The transcriptional coactivators p300 and CBP are histone acetyltransferases** Cell *87* p.953-959
Ono, M. and Oka, T. (1980)
**alpha-Lactalbumin-casein induction in virgin mouse mammary explants: dose dependent differential action of cortisol** Science *207* p.1367-1369
Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S., and Reinberg, D. (1998)
**FACT, a factor that facilitates transcript elongation through nucleosomes** Cell *92* p.105-116
Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., and Reinberg, D. (1999)
**The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins** Nature *400* p.284-288
Osborne, R., Howell, M., Clark, A. J., and Nicholas, K. R. (1995)
**Hormone-dependent expression of the ovine beta-lactoglobulin gene** J Dairy Res *62* p.321-329
Ostlund, Farrants AK, Blomquist, P., Kwon, H., and Wrange, O. (1997)
**Glucocorticoid receptor-glucocorticoid response element binding stimulates nucleosome disruption by the SWI/SNF complex** Mol Cell Biol *17* p.895-905
Ostrowski, M. C., Richard-Foy, H., Wolford, R. G., Berard, D. S., and Hager, G. L. (1983)
**Glucocorticoid regulation of transcription at an amplified, episomal promoter** Mol Cell Biol *3* p.2045-2057
Oudet, P., Gross-Bellard, M., and Chambon, P. (1975)
**Electron microscopic and biochemical evidence that chromatin structure is a repeating unit** Cell *4* p.281-300
Palmiter, R. D., Sandgren, E. P., Avarbock, M. R., Allen, D. D., and Brinster, R. L. (1991)
**Heterologous introns can enhance expression of transgenes in mice** Proc Natl Acad Sci U S A *88* p.478-482
Papavassiliou, A. G. (1994)
**1,10-Phenanthroline-copper ion nuclease footprinting of DNA-protein complexes in situ following mobility-shift electrophoresis assays** Methods Mol Biol *30* p.43-78
Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C., Findlay, J. B., Sivaprasadarao, R., Jones, T. A., Newcomer, M. E., and Kraulis, P. J. (1986)
**The structure of beta-lactoglobulin and its similarity to plasma retinol-binding protein** Nature *324* p.383-385
Patel, R. C., Lange, D., McConathy, W. J., Patel, Y. C., and Patel, S. C. (1997)
**Probing the structure of the ligand binding cavity of lipocalins by fluorescence spectroscopy** Protein Eng *10* p.621-625

Pelta, J., Livolant, F., and Sikorav, J. L. (1996)
**DNA aggregation induced by polyamines and cobalthexamine** J Biol Chem *271* p.5656-5662
Pfitzner, E., Jahne, R., Wissler, M., Stoecklin, E., and Groner, B. (1998)
**p300/CREB-binding protein enhances the prolactin-mediated transcriptional induction through direct interaction with the transactivation domain of Stat5, but does not participate in the Stat5- mediated suppression of the glucocorticoid response** Mol Endocrinol *12* p.1582-1593
Phillips, D. M. P. and Johns E. W. (1965)
**Interactions between histones and nucleic acids** Biochem.J. *94* p.127-130
Philp, J. A., Burdon, T. G., and Watson, C. J. (1996)
**Differential activation of STATs 3 and 5 during mammary gland development** FEBS Lett *396* p.77-80
Pinkert, C. A., Ornitz, D. M., Brinster, R. L., and Palmiter, R. D. (1987)
**An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice** Genes Dev *1* p.268-276
Pittius, C. W., Sankaran, L., Topper, Y. J., and Hennighausen, L. (1988)
**Comparison of the regulation of the whey acidic protein gene with that of a hybrid gene containing the whey acidic protein gene promoter in transgenic mice** Mol Endocrinol *2* p.1027-1032
Pollard, K. J. and Peterson, C. L. (1997)
**Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression** Mol Cell Biol *17* p.6212-6222
Pratt, W. B. (1993)
**The role of heat shock proteins in regulating the function, folding, and trafficking of the glucocorticoid receptor** J Biol Chem *268* p.21455-21458
Privat de Garilhe, M., Cunningham, L., Laurila, U. R., and Laskowski, M. (1957)
**Characterisation and properties of micrococcal nuclease** J.Biol.Chem. *224* p.751-
Prunell, A. and Kornberg, R. D. (1978)
**Relation of nucleosomes to DNA sequences** Cold Spring Harb Symp Quant Biol *42 Pt 1* p.103-108
Puerta, C., Hernandez, F., Gutierrez, C., Pineiro, M., Lopez-Alarcon, L., and Palacian, E. (1993)
**Efficient transcription of a DNA template associated with histone (H3.H4)2 tetramers** J Biol Chem *268* p.26663-26667
Quivy, J. P. and Becker, P. B. (1996)
**The architecture of the heat-inducible Drosophila hsp27 promoter in nuclei** J Mol Biol *256* p.249-263
Ramakrishnan, V. (1997)
**Histone H1 and chromatin higher-order structure** Crit Rev Eukaryot Gene Expr *7* p.215-230
Rana, B., Mischoulon, D., Xie, Y., Bucher, N. L., and Farmer, S. R. (1994)
**Cell-extracellular matrix interactions can regulate the switch between growth and differentiation in rat hepatocytes: reciprocal expression of C/EBP alpha and immediate-early growth response transcription factors** Mol Cell Biol *14* p.5858-5869
Raught, B., Khursheed, B., Kazansky, A., and Rosen, J. (1994)
**YY1 represses beta-casein gene expression by preventing the formation of a lactation-associated complex** Mol Cell Biol *14* p.1752-1763
Raught, B., Liao, W. S., and Rosen, J. M. (1995)
**Developmentally and hormonally regulated CCAAT/enhancer-binding protein isoforms influence beta-casein gene expression** Mol Endocrinol *9* p.1223-1232
Recht, J. and Osley, M. A. (1999)
**Mutations in both the structured domain and N-terminus of histone H2B bypass the requirement for Swi-Snf in yeast** EMBO J *18* p.229-240

Reeves, R. and Nissen, M. S. (1993)
**Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles** J Biol Chem *268* p.21137-21146
Reeves, R. and Wolffe, A. P. (1996)
**Substrate structure influences binding of the non-histone protein HMG- I(Y) to free nucleosomal DNA** Biochemistry *35* p.5063-5074
Richard-Foy, H. and Hager, G. L. (1987)
**Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter** EMBO J *6* p.2321-2328
Roberson, A. E., Wolffe, A. P., Hauser, L. J., and Olins, D. E. (1989)
**The 5S RNA gene minichromosome of Euplotes** Nucleic Acids Res *17* p.4699-4712
Roberts, B., DiTullio, P., Vitale, J., Hehir, K., and Gordon, K. (1992)
**Cloning of the goat beta-casein-encoding gene and expression in transgenic mice** Gene *121* p.255-262
Roberts, M. S., Fragoso, G., and Hager, G. L. (1995)
**Nucleosomes reconstituted in vitro on mouse mammary tumor virus B region DNA occupy multiple translational and rotational frames** Biochemistry *34* p.12470-12480
Ruff-Jamison, S., Chen, K., and Cohen, S. (1995)
**Epidermal growth factor induces the tyrosine phosphorylation and nuclear translocation of Stat 5 in mouse liver** Proc Natl Acad Sci U S A *92* p.4215-4218
Rundlett, S. E., Carmen, A. A., Suka, N., Turner, B. M., and Grunstein, M. (1998)
**Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3** Nature *392* p.831-835
Rusterholz, C., Henrioud, P. C., and Nabholz, M. (1999)
**Interleukin-2 (IL-2) regulates the accessibility of the IL-2-responsive enhancer in the IL-2 receptor alpha gene to transcription factors** Mol Cell Biol *19* p.2681-2689
Santisteban, M. S., Arents, G., Moudrianakis, E. N., and Smith, M. M. (1997)
**Histone octamer function in vivo: mutations in the dimer-tetramer interfaces disrupt both gene activation and repression** EMBO J *16* p.2493-2506
Schindler, C. and Darnell, J. E. Jr. (1995)
**Transcriptional responses to polypeptide ligands: the JAK-STAT pathway** Annu Rev Biochem *64* p.621-651
Schmidhauser, C., Bissell, M. J., Myers, C. A., and Casperson, G. F. (1990)
**Extracellular matrix and hormones transcriptionally regulate bovine beta-casein 5' sequences in stably transfected mouse mammary cells** Proc Natl Acad Sci U S A *87* p.9118-9122
Schmidhauser, C., Casperson, G. F., Myers, C. A., Sanzo, K. T., Bolten, S., and Bissell, M. J. (1992)
**A novel transcriptional enhancer is involved in the prolactin- and extracellular matrix-dependent regulation of beta-casein gene expression** Mol Biol Cell *3* p.699-709
Schmitt-Ney, M., Doppler, W., Ball, R. K., and Groner, B. (1991)
**Beta-casein gene promoter activity is regulated by the hormone-mediated relief of transcriptional repression and a mammary-gland-specific nuclear factor** Mol Cell Biol *11* p.3745-3755
Seagroves, T. N., Krnacik, S., Raught, B., Gay, J., Burgess-Beusse, B., Darlington, G. J., and Rosen, J. M. (1998)
**C/EBPbeta, but not C/EBPalpha, is essential for ductal morphogenesis, lobuloalveolar proliferation, and functional differentiation in the mouse mammary gland** Genes Dev *12* p.1917-1928
Sedat, J. and Manuelidis, L. (1978)
**A direct approach to the structure of eukaryotic chromosomes** Cold Spring Harb Symp Quant Biol *42 Pt 1* p.331-350
Seidel, H. M., Milocco, L. H., Lamb, P., Darnell, J. E. Jr, Stein, R. B., and Rosen, J. (1995)
**Spacing of palindromic half sites as a determinant of selective STAT (signal transducers**

**and activators of transcription) DNA binding and transcriptional activity** Proc Natl Acad Sci U S A *92* p.3041-3045

Serdobova, I., Pla, M., Reichenbach, P., Sperisen, P., Ghysdael, J., Wilson, A., Freeman, J., and Nabholz, M. (1997)
**Elf-1 contributes to the function of the complex interleukin (IL)-2- responsive enhancer in the mouse IL-2 receptor alpha gene** J Exp Med *185* p.1211-1221

Shi, Y., Seto, E., Chang, L. S., and Shenk, T. (1991)
**Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein** Cell *67* p.377-388

Shim, E. Y., Woodcock, C., and Zaret, K. S. (1998)
**Nucleosome positioning by the winged helix transcription factor HNF3** Genes Dev *12* p.5-10

Sigman, D. S. and Chen, C. H. (1990a)
**Chemical nucleases: new reagents in molecular biology** Annu Rev Biochem *59* p.207-236

Sigman, D. S. (1990b)
**Chemical nucleases** Biochemistry *29* p.9097-9105

Silverman, N., Agapite, J., and Guarente, L. (1994)
**Yeast ADA2 protein binds to the VP16 protein activation domain and activates transcription** Proc Natl Acad Sci U S A *91* p.11665-11668

Simmen, M. W., Leitgeb, S., Charlton, J., Jones, S. J., Harris, B. R., Clark, V. H., and Bird, A. (1999)
**Nonmethylated transposable elements and methylated genes in a chordate genome** Science *283* p.1164-1167

Simons, J. P. and Land, R. B. (1987)
**Transgenic livestock** J Reprod Fertil Suppl *34* p.237-250

Simons, J. P., Wilmut, I., Clark, A. J., Archibald, A. L., Bishop, J. O., and Lathe, R. (1988)
**Gene Transfer into Sheep** Biotechnology *6* p.179-183

Simpson, R. T. (1981)
**Modulation of nucleosome structure by histone subtypes in sea urchin embryos** Proc Natl Acad Sci U S A *78* p.6803-6807

Simpson, R. T. (1991)
**Nucleosome positioning: occurrence, mechanisms, and functional consequences** Prog Nucleic Acid Res Mol Biol *40* p.143-184

Simpson, R. T. (1999)
**In vivo methods to analyze chromatin structure** Curr Opin Genet Dev *9* p.225-229

Smith, M. R. and Lieberman, M. W. (1984)
**Nucleosome arrangement in alpha-satellite chromatin of African green monkey cells** Nucleic Acids Res *12* p.6493-6510

Soulier, S., Vilotte, J. L., Stinnakre, M. G., and Mercier, J. C. (1992)
**Expression analysis of ruminant alpha-lactalbumin in transgenic mice: developmental regulation and general location of important cis- regulatory elements** FEBS Lett *297* p.13-18

Spangenberg, C., Eisfeld, K., Stunkel, W., Luger, K., Flaus, A., Richmond, T. J., Truss, M., and Beato, M. (1998)
**The mouse mammary tumour virus promoter positioned on a tetramer of histones H3 and H4 binds nuclear factor 1 and OTF1** J Mol Biol *278* p.725-739

Stacey, A., Schnieke, A., Kerr, M., Scott, A., McKee, C., Cottingham, I., Binas, B., Wilde, C., and Colman, A. (1995)
**Lactation is disrupted by alpha-lactalbumin deficiency and can be restored by human alpha-lactalbumin gene replacement in mice** Proc Natl Acad Sci U S A *92* p.2835-2839

Stinnakre, M. G., Vilotte, J. L., Soulier, S., and Mercier, J. C. (1994)
**Creation and phenotypic analysis of alpha-lactalbumin-deficient mice** Proc Natl Acad Sci U S A *91* p.6544-6548

Streuli, C. H., Bailey, N., and Bissell, M. J. (1991)
**Control of mammary epithelial differentiation: basement membrane induces tissue specific**

**gene expression in the absence of cell-cell interaction and morphological polarity** J Cell Biol *115* p.1383-1395

Struhl, K. (1998)
**Histone acetylation and transcriptional regulatory mechanisms** Genes Dev *12* p.599-606

Struhl, K. (1999)
**Fundamentally different logic of gene regulation in eukaryotes and prokaryotes** Cell *98* p.1-4

Sun, J. M., Ali, Z., Lurz, R., and Ruiz-Carrillo, A. (1990)
**Replacement of histone H1 by H5 in vivo does not change the nucleosome repeat length of chromatin but increases its stability** EMBO J *9* p.1651-1658

Svaren, J. and Horz, W. (1997)
**Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast** Trends Biochem Sci *22* p.93-97

Takami, Y., Takeda, S., and Nakayama, T. (1997)
**An approximately half set of histone genes is enough for cell proliferation and a lack of several histone variants causes protein pattern changes in the DT40 chicken B cell line** J Mol Biol *265* p.394-408

Talasz, H., Sapojnikova, N., Helliger, W., Lindner, H., and Puschendorf, B. (1998)
**In vitro binding of H1 histone subtypes to nucleosomal organized mouse mammary tumor virus long terminal repeat promotor** J Biol Chem *273* p.32236-32243

Taniuchi, H., Anfinsen, C. B., and Sodja, A. (1967)
**The amino acid sequence of an extracellular nuclease of Staphylococcus aureus. 3. Complete amino acid sequence** J Biol Chem *242* p.4752-4758

Tate, V. E. and Philipson, L. (1979)
**Parental adenovirus DNA accumulates in nucleosome-like structures in infected cells** Nucleic Acids Res *6* p.2769-2785

Taverna, D., Groner, B., and Hynes, N. E. (1991)
**Epidermal growth factor receptor, platelet-derived growth factor receptor, and c-erbB-2 receptor activation all promote growth but have distinctive effects upon mouse mammary epithelial cell differentiation** Cell Growth Differ *2* p.145-154

Teglund, S., McKay, C., Schuetz, E., van Deursen, J. M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J. N. (1998)
**Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses** Cell *93* p.841-850

Thoma, F., Koller, T., and Klug, A. (1979)
**Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin** J Cell Biol *83* p.403-427

Thoma, F. and Simpson, R. T. (1985)
**Local protein-DNA interactions may determine nucleosome positions on yeast plasmids** Nature *315* p.250-252

Thoma, F. (1991)
**Structural changes in nucleosomes during transcription: strip, split or flip?** Trends Genet *7* p.175-177

Thoma, F. (1992)
**Nucleosome positioning** Biochim Biophys Acta *1130* p.1-19

Thomas, G. H. and Elgin, S. C. (1988)
**Protein/DNA architecture of the DNase I hypersensitive region of the Drosophila hsp26 promoter** EMBO J *7* p.2191-2201

Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., and Schreiber, S. L. (1998)
**Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex** Nature *395* p.917-921

Topper, Y. J. and Freeman, C. S. (1980)
**Multiple hormone interactions in the developmental biology of the mammary gland** Physiol Rev *60* p.1049-1106

Travers, A. (1999)
**The location of the linker histone on the nucleosome** Trends Biochem Sci *24* p.4-7
Tremethick, D. J. and Molloy, P. L. (1988)
**Effects of high mobility group proteins 1 and 2 on initiation and elongation of specific transcription by RNA polymerase II in vitro** Nucleic Acids Res *16* p.11107-11123
Tremethick, D. J. and Hyman, L. (1996)
**High mobility group protein 14 and 17 can prevent the close packing of nucleosomes by increasing the strength of protein contacts in the linker DNA** J Biol Chem *271* p.12009-12016
Trieschmann, L., Martin, B., and Bustin, M. (1998)
**The chromatin unfolding domain of chromosomal protein HMG-14 targets the N-terminal tail of histone H3 in nucleosomes** Proc Natl Acad Sci U S A *95* p.5468-5473
Truss, M., Bartsch, J., Schelbert, A., Hache, R. J., and Beato, M. (1995)
**Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo** EMBO J *14* p.1737-1751
Tsukiyama, T., Becker, P. B., and Wu, C. (1994)
**ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor** Nature *367* p.525-532
Tullius, T. D. and Dombroski, B. A. (1986)
**Hydroxyl radical "footprinting": high-resolution information about DNA- protein contacts and application to lambda repressor and Cro protein** Proc Natl Acad Sci U S A *83* p.5469-5473
Tullius, T. D., Dombroski, B. A., Churchill, M. E., and Kam, L. (1987)
**Hydroxyl radical footprinting: a high-resolution method for mapping protein-DNA contacts** Methods Enzymol *155* p.537-558
Turner, B. M., O'Neill, L. P., and Allan, I. M. (1989)
**Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies** FEBS Lett *253* p.141-145
Turner, B. M., Birley, A. J., and Lavender, J. (1992)
**Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei** Cell *69* p.375-384
Udy, G. B., Towers, R. P., Snell, R. G., Wilkins, R. J., Park, S. H., Ram, P. A., Waxman, D. J., and Davey, H. W. (1997)
**Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression** Proc Natl Acad Sci U S A *94* p.7239-7244
Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G., and Wolffe, A. P. (1997)
**Histone acetylation: influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression** EMBO J *16* p.2096-2107

van Holde, K.E., Shaw, B. R., Lohr, D., Herman, T. M., and Kovacic, R. T. (1975)
**Low-angle neutron scattering from chromatin subunit particles** Proc.Tenth FEBS Meeting *38* p.57-72
van Holde, K.E. (1988)
**Chromatin** Springer Series in Molecular Biology 1st ed.
van Holde, K.E. and Zlatanova, J. (1995)
**Chromatin higher order structure: chasing a mirage?** J Biol Chem *270* p.8373-8376
Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P. B. (1997)
**Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II** Nature *388* p.598-602
Vermaak, D., Steinbach, O. C., Dimitrov, S., Rupp, R. A. W., and Wolffe, A. P. (1998)
**The globular domain of histone H1 is sufficient to direct specific gene repression in early Xenopus embryos** Curr Biol *8* p.533-536

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane, Allis, C. D., and Workman, J. L. (1996)
**Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro** EMBO J *15* p.2508-2518
Vignali, M. and Workman, J. L. (1998)
**Location and function of linker histones** Nat Struct Biol *5* p.1025-1028
Vilotte, J. L. and Soulier, S. (1992)
**Isolation and characterization of the mouse alpha-lactalbumin-encoding gene: interspecies comparison, tissue- and stage-specific expression** Gene *119* p.287-292
Vinkemeier, U., Cohen, S. L., Moarefi, I., Chait, B. T., Kuriyan, J., and Darnell, J. E. Jr. (1996)
**DNA binding of in vitro activated Stat1 alpha, Stat1 beta and truncated Stat1: interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites** EMBO J *15* p.5616-5626
Vinkemeier, U., Moarefi, I., Darnell, J. E. Jr, and Kuriyan, J. (1998)
**Structure of the amino-terminal protein interaction domain of STAT-4** Science *279* p.1048-1052
Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F. and Wolffe, A.P. (1999a)
**Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation** Nat Genet *23* p.62-66.
Wade, P. A. and Wolffe, A. P. (1999b)
**Transcriptional regulation: SWItching circuitry** Curr Biol *9* p.R221-R224
Warburton, P. E., Cooke, C. A., Bourassa, S., Vafa, O., Sullivan, B. A., Stetten, G., Gimelli, G., Warburton, D., Tyler-Smith, C., Sullivan, K. F., Poirier, G. G., and Earnshaw, W. C. (1997)
**Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres** Curr Biol *7* p.901-904
Watson, C. J., Gordon, K. E., Robertson, M., and Clark, A. J. (1991)
**Interaction of DNA-binding proteins with a milk protein gene promoter in vitro: identification of a mammary gland-specific factor** Nucleic Acids Res *19* p.6603-6610
Watson, C. J. and Burdon, T. G. (1996)
**Prolactin signal transduction mechanisms in the mammary gland: the role of the Jak/Stat pathway** Rev Reprod *1* p.1-5
Webster, J., Wallace, R. M., Clark, A. J., and Whitelaw, C. B. (1995)
**Tissue-specific, temporally regulated expression mediated by the proximal ovine beta-lactoglobulin promoter in transgenic mice** Cell Mol Biol Res *41* p.11-15
Webster, J., Donofrio, G., Wallace, R., Clark, A. J., and Whitelaw, C. B. (1997)
**Intronic sequences modulate the sensitivity of beta-lactoglobulin transgenes to position effects** Gene *193* p.239-243
Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A., and Allis, C. D. (1999)
**Phosphorylation of histone H3 is required for proper chromosome condensation and segregation** Cell *97* p.99-109
Weintraub, H. (1984)
**Histone-H1-dependent chromatin superstructures and the suppression of gene activity** Cell *38* p.17-27
Weiss, K. and Simpson, R. T. (1998)
**High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating type locus HMLalpha** Mol Cell Biol *18* p.5392-5403
Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999)
**Nucleosome mobilization catalysed by the yeast SWI/SNF complex** Nature *400* p.784-787
Whitelaw, C. B., Archibald, A. L., Harris, S., McClenaghan, M., Simons, J. P., and Clark, A. J. (1991)
**Targeting expression to the mammary gland: intronic sequences can enhance the efficiency of gene expression in transgenic mice** Transgenic Res *1* p.3-13
Whitelaw, C. B., Harris, S., McClenaghan, M., Simons, J. P., and Clark, A. J. (1992)
**Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice** Biochem J *286* p.31-39

Whitelaw, C. B. and Webster, J. (1998)
**Temporal profiles of appearance of DNase I hypersensitive sites associated with the ovine beta-lactoglobulin gene differ in sheep and transgenic mice** Mol Gen Genet *257* p.649-654
Whitelaw, C. B., Grolli, S., Accornero, P., Donofrio, G., Farini, E., and Webster, J. (1999)
**Matrix attachment region regulates basal beta-lactoglobulin transgene expression** Gene (submited)
Widlund, H. R., Cao, H., Simonsson, S., Magnusson, E., Simonsson, T., Nielsen, P. E., Kahn, J. D., Crothers, D. M., and Kubista, M. (1997)
**Identification and characterization of genomic nucleosome-positioning sequences** J Mol Biol *267* p.807-817
Widom, J. (1986)
**Physicochemical studies of the folding of the 100 A nucleosome filament into the 300 A filament. Cation dependence** J Mol Biol *190* p.411-424
Widom, J. (1998)
**Chromatin structure: linking structure to function with histone H1** Curr Biol *8* p.R788-R791
Wilde, C. J., Clark, A. J., Kerr, M. A., Knight, C. H., McClenaghan, M., and Simons, J. P. (1992)
**Mammary development and milk secretion in transgenic mice expressing the sheep beta-lactoglobulin gene** Biochem J *284 ( Pt 3)* p.717-720
Wilkins, R. C. and Lis, J. T. (1997)
**Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation** Nucleic Acids Res *25* p.3963-3968
Wilkins, R. C. and Lis, J. T. (1999)
**DNA distortion and multimerization: novel functions of the glutamine- rich domain of GAGA factor** J Mol Biol *285* p.515-525
Wilks, A. F., Harpur, A. G., Kurban, R. R., Ralph, S. J., Zurcher, G., and Ziemiecki, A. (1991)
**Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase** Mol Cell Biol *11* p.2057-2065
Wilmut, I., Archibald, A. L., McClenaghan, M., Simons, J. P., Whitelaw, C. B., and Clark, A. J. (1991)
**Production of pharmaceutical proteins in milk** Experientia *47* p.905-912
Wilmut, I. and Whitelaw, C. B. (1994)
**Strategies for production of pharmaceutical proteins in milk** Reprod Fertil Dev *6* p.625-630
Wolffe, A. (1995)
**Chromatin: Structure and Function** 2nd ed. Academic Press, New York
Wolffe, A. P. and Pruss, D. (1996)
**Targeting chromatin disruption: Transcription regulators that acetylate histones** Cell *84* p.817-819
Woodcock, C. L., Frado, L. L., and Rattner, J. B. (1984)
**The higher-order structure of chromatin: evidence for a helical ribbon arrangement** J Cell Biol *99* p.42-52
Workman, J.L. and Kingston, R.E. (1998)
**Alteration of nucleosome structure as a mechanism of transcriptional regulation** Annu Rev Biochem *67* p.545-79
Wyszomierski, S. L., Yeh, J., and Rosen, J. M. (1999)
**Glucocorticoid receptor/signal transducer and activator of transcription 5 (STAT5) interactions enhance STAT5 activation by prolonging STAT5 DNA binding and tyrosine phosphorylation** Mol Endocrinol *13* p.330-343
Xu, X., Sun, Y. L., and Hoey, T. (1996)
**Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain [see comments]** Science *273* p.794-797
Yamaguchi, Y., Nishio, H., Kishi, K., Ackerman, S. J., and Suda, T. (1999)
**C/EBPbeta and GATA-1 Synergistically Regulate Activity of the Eosinophil Granule Major Basic Protein Promoter: Implication for C/EBPbeta Activity in Eosinophil Gene Expression** Blood *94* p.1429-1439

Yang, W. M., Yao, Y. L., Sun, J. M., Davie, J. R., and Seto, E. (1997)
**Isolation and characterization of cDNAs corresponding to an additional member of the human histone deacetylase gene family** J Biol Chem *272* p.28001-28007
Yenidunya, A., Davey, C., Clark, D., Felsenfeld, G., and Allan, J. (1994)
**Nucleosome positioning on chicken and human globin gene promoters in vitro. Novel mapping techniques** J Mol Biol *237* p.401-414
Yoda, K., Ando, S., Okuda, A., Kikuchi, A., and Okazaki, T. (1998)
**In vitro assembly of the CENP-B/alpha-satellite DNA/core histone complex: CENP-B causes nucleosome positioning** Genes Cells *3* p.533-548
Yoon, C., Kuwabara, M. D., Spassky, A., and Sigman, D. S. (1990)
**Sequence specificity of the deoxyribonuclease activity of 1,10- phenanthroline-copper ion** Biochemistry *29* p.2116-2121
Yoshimura, M. and Oka, T. (1989)
**Isolation and structural analysis of the mouse beta-casein gene** Gene *78* p.267-275
Yu, L. and Gorovsky, M. A. (1997)
**Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in Tetrahymena thermophila** Mol Cell Biol *17* p.6303-6310
Yull, F., Binas, B., Harold, G., Wallace, R., and Clark, A. J. (1997)
**Transgene rescue in the mammary gland is associated with transcription but does not require translation of BLG transgenes** Transgenic Res *6* p.11-17
Zlatanova, J., Leuba, S. H., and van Holde, K. (1998)
**Chromatin fiber structure: morphology, molecular determinants, structural transitions** Biophys J *74* p.2554-2566