



universität
wien

DISSERTATION

Titel der Dissertation

Transcriptional Regulation of the central MMTV LTR Promoter Complex

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer.nat.)

Verfasserin:	Eva Fuchs
Dissertationsgebiet (lt. Studienblatt):	Molekulare Biologie
Betreuer:	Univ.-Prof. Dr. Timothy Skern

Wien, im November 2010

Acknowledgements

First of all, I want to express my gratitude to Professor Tim Skern for taking over the responsibility as a supervisor and for keeping me motivating during the writing of the thesis.

I am thankful to Dr. Françoise Rouault for organizing an exit strategy after the (at least for us) unforeseen cessation of the project funding. Moreover, I want to thank her for critically reading the manuscript.

I am grateful Dr. Karin Schwertner-Komornyik for her straightforward attitude in solving work-related, yet non-scientific problems and the 'Frauenförderungstopf' of the University of Veterinary medicine for financing the final year of laboratory work.

I thank Professor Alexandra Lusser and her group for sharing their expertise in the field of chromatin and for the pleasant cooperation during my stay in Innsbruck.

I want to thank Reinhard Ertl for help with the qRT-PCR and the FACS-analysis.

Thanks to him and Cornelius Kasper for the coffee breaks which helped a lot in keeping me grounded during difficult times, when things did not run the way they should.

Most notably, I want to thank Markus Zojer for asking the right questions when I was stuck in the middle of nowhere and for his ability to put things into perspective.

*There is no comparison between that which is lost by not succeeding and that
which is lost by not trying.*
Francis Bacon

Summary

Mouse mammary tumor virus (MMTV) is a non-acute transforming *Betaretrovirus* that is transmitted by viral particles present in the mother's milk. Upon digestion, the virus infects and replicates in the cells of the lymphatic system until it finally reaches the mammary gland epithelial. There, the integrated virus remains silent until its major LTR promoter, P1, is activated by glucocorticoid hormones.

Uniquely among retroviruses, MMTV harbours a further promoter element within its LTR. The P2 promoter has been described as being enhanced in the presence of glucocorticoids. However, as yet, no transcriptional product of P2 has been related to its activity. To examine the regulation of P2 in more detail, we used a sensitive reporter gene assay based on luciferase expression. Unexpectedly, we detected another promoter during this analysis, P3, which is located immediately downstream of the P2 promoter. Both promoter elements appear to form a complex in the central part of the MMTV LTR.

We could show that both promoters are regulated by a downstream located enhancer. In contrast, only the activity of the P2 promoter is influenced by an upstream located negative regulatory element (NRE). Furthermore, we were able to relate the constitutive transcriptional activity originating from the P2/P3 promoter complex with a DNase hypersensitive site at the corresponding nucleosome. Additionally, we could demonstrate that this open chromatin formation is associated with elevated methylation levels at lysine 4 and lysine 36 of histone H3, respectively. Both modifications are predominantly found in transcriptionally active regions.

Independent of glucocorticoid stimulation, we were able to identify a doubly spliced message initiating from the complex. Moreover, we could demonstrate that the level of this transcript is regulated by the activities of histone deacetylases (HDACs). Incubation with the HDAC inhibitor trichostatin A (TSA) lead to a concentration dependent reduction of the doubly spliced messenger RNA, which was accompanied by increasing acetylation levels of the corresponding nucleosome. Congruently, we detected a tight association of the histone deacetylase 1 (HDAC1) to the region of the P2/P3 promoter complex.

Finally, we could show that the transcription factor CCAAT displacement protein (CDP) is a specific regulator of the promoter complex. By binding to the NRE, CDP downregulates the transcriptional activity of the P2 promoter. Thereby, the emergence of the doubly spliced message is intimately linked to early stages of cell differentiation and thus to the latency phase of the virus.

Zusammenfassung

Der zur Gruppe der *Betaretroviren* zählende Mouse mammary tumor virus (MMTV) verursacht nach einer längeren Latenzphase Tumoren in infizierten Tieren. Die Übertragung dieses Erregers erfolgt dabei durch virale Partikel, welche in die Muttermilch erwachsener weiblicher Tiere abgegeben werden. Im Verdauungstrakt der neugeborenen Mäuse befällt der Virus zunächst die Zellen des lymphatischen Gewebes; erst danach wird die Brustdrüse befallen. Die dort befindlichen Epithelzellen werden infiziert, jedoch werden nur marginal virale Proteine gebildet. Die Latenzphase endet, sobald der Hauptpromoter des LTR, P1, durch Glucocorticoide aktiviert wird.

Im Gegensatz zu anderen Retroviren, wurde bei MMTV ein zusätzliches Promoterelement innerhalb des LTR gefunden. Dieser P2 genannte Promoter wurde ursprünglich als ebenfalls hormonabhängig beschrieben. Allerdings konnte bis heute kein transkriptionelles Produkt dieses Promoters gefunden werden. Um die Regulation des Promoters P2 im Detail zu untersuchen, haben wir ein sensitives Reportersystem verwendet, welches auf der Expression des Luciferasegens basiert. Unerwarteterweise, haben wir dabei einen weiteren Promoter, P3, entdeckt, der sich in 3' Richtung in unmittelbarer Nachbarschaft zu P2 befindet. Bedingt durch ihre Nähe scheinen beide Promoterelemente einen Komplex im zentralen Teil des LTR von MMTV zu bilden.

Wir konnten zeigen, daß beide Promotoren durch einen in 3' Richtung befindlichen Enhancer reguliert werden. Der Promoter P2 wird zusätzlich von einem in 5' Richtung gelegenen Negativen Regulatorischen Element (NRE) beeinflusst. Weiters konnten wir zeigen, daß die konstitutive transkriptionelle Aktivität des P2/P3 Promoter Komplexes mit einer DNase hypersensitiven Stelle einhergeht, welche sich am zugehörigen Nucleosome feststellen ließ. Damit verbunden, konnten wir einen erhöhten Methylierungszustand der Lysine 4 und 36 des Histonproteins H3 (desselben Nucleosomes) detektieren. Beide Modifikationen werden hauptsächlich in Zusammenhang mit transkriptinell aktiven Region gefunden.

Unabhängig von hormoneller Stimulation, konnten wir eine doppelt gesplice mRNA identifizieren, welche am Promoterkomplex ihren Ausgang nimmt. Wir konnten weiters zeigen, daß die Menge an Transkript durch die Aktivität von Histone Deacetylasen beeinflusst wird, da die Verwendung des Histone Deacetylasen Inhibitors Trichostatin A (TSA) zu einer konzentrationsabhängigen Abnahme der neu entdeckten mRNA führte. Begleitet wurde diese Abnahme von einer gleichzeitigen Zunahme an Acetyl-Resten des entsprechenden Nucleosomes. Damit übereinstimmend, konnten wir eine hohe Konzentration des Enzymes Histone Deacetylase 1 an der P2/P3 Promoterregion feststellen.

Schließlich konnten wir zeigen, daß der Transkriptionsfaktor CCAAT displacement

protein (CDP) einen spezifischen Regulator des Promoterkomplexes darstellt. Durch Bindung an das NRE vermindert CDP die transkriptionelle Aktivität des P2 Promoters, wodurch das Auftreten des doppelt gesplitten Transkripts mit frühen Phasen der Zelldifferenzierung und damit mit der Latenzphase des Virus in Verbindung gebracht werden kann.

List of abbreviations

µg - microgram
µl - microlitre
µM - micromolar
AA - Aminoacid
AP - Alkaline Phosphatase
APC - Antigen presenting cell
APS - Ammonium persulfate
ATCC - American Type Culture Collection
ATP - Adenosine 5'-triphosphate
bp - basepair
°C - degrees centigrade
CD - Cluster of Differentiation
cDNA - complementary DNA
CDP - CCAAT displacement protein
ChIP - Chromatin immunoprecipitation
CIP - Calf Intestine Phosphatase
DC - Dendritic cell
DIG - digoxigenin
DMEM - Dulbecco's Modified Eagle Medium
DMSO - Dimethyl sulfoxide
DNA - Deoxyribonucleic acid
DNase - Deoxyribonuclease
dNTP - deoxynucleotide triphosphate
DTT - Dithiothreitol
ECL - Enhanced chemiluminescence
EDTA - Ethylenediaminetetraacetic acid
EGTA - Ethyleneglycoltetraacetic acid
FCS - Foetal Calf Serum
g - gram
GR - glucocorticoid receptor
HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRE - Hormone responsive element
HRP - Horseradish Peroxidase
Ig - Immunoglobulin

kb – kilobase
kDa – kilodalton
l – litre
LB – Luria Bertani
LTR – Long Terminal Repeat
M – Molar
mg – milligram
MHC – Major Histocompatibility Complex
ml – millilitre
mM – millimolar
MMTV – Mouse Mammary Tumor Virus
mRNA – messenger Ribonucleic Acid
ng – nanogram
nM – nanomolar
NRE – Negative regulatory element
PBS – Phosphate Buffered Saline
PCR – Polymerase Chain Reaction
pol – polymerase
RNA – Ribonucleic Acid
RNase – Ribonuclease
rpm – rotations per minute
RT – Room temperature
SSC – Saline sodium citrate buffer
SDS – Sodium Dodecyl Sulfate
SDS-PAGE – SDS-Poly-Acrylamide Gel Electrophoresis
TAP – tobacco acid pyrophosphatase
TBS – Tris buffered saline
TSA – Trichostatin A
TE – Trypsin/EDTA
TEB – Tris/EDTA Buffer
TEMED – Tetramethylethylenediamine
tRNA – transfer Ribonucleic Acid
U – Units
UV – Ultra Violet
V – Volt
v – volume
w – weight

Table of Contents

1.Introduction.....	15
1.1 Exogenous infection route.....	16
1.1.1 Dendritic cells.....	16
1.1.2 Lymphoid cells.....	17
1.1.3 Mammary cells.....	18
1.2 Endogenous infection route.....	18
1.3 Does MMTV infect human tissue?.....	19
1.4 The retroviral particle.....	20
1.4.1 The retroviral genome.....	21
1.4.1.1 MMTV LTR promoters.....	25
1.4.2 Viral proteins.....	26
1.4.2.1 GAG.....	26
1.4.2.2 POL.....	27
1.4.2.3 ENV.....	29
1.4.3 MMTV accessory proteins.....	29
1.4.3.1 Sag (Superantigen).....	29
1.4.3.2 Rem (Regulator of export of MMTV mRNA).....	30
1.4.3.3 DUTPase.....	31
1.4.4 Cellular APOBEC3	31
1.5 The retroviral replication cycle	32
1.6 The MMTV major P1 promoter.....	34
1.6.1 Stimulation by glucocorticoid hormones.....	34
1.6.2 Negative regulation.....	36
2.Aim of the project.....	39
3.Materials and methods.....	40
3.1 Tissue culture.....	40
3.1.1 Cell lines.....	40
3.1.2 Media and solutions.....	40
3.1.3 Cell culture.....	41
3.1.4 Cell counting.....	41
3.1.5 Freezing and thawing of cells.....	41
3.1.6 Synchronization of cells.....	42
3.1.7 Fluorescence activated cell sorting (FACS) analysis.....	42
3.1.7.1 Determination of eGFP expression.....	42
3.1.7.2 Probidium iodide staining.....	42
3.1.8 Transfection of eukaryotic cells.....	43
3.1.8.1 Calciumphosphate co-precipitation.....	43
3.1.8.2 Lipofectamine.....	43
3.2 Bacterial culture.....	44
3.2.1 Bacterial strains.....	44
3.2.2 Media and antibiotics.....	44
3.2.3 Transformation of bacteria.....	45
3.2.3.1 Heat shock.....	45
3.2.3.2 Electroporation.....	45
3.3 DNA methods.....	45
3.3.1 Preparation of plasmid DNA from bacteria.....	45
3.3.1.1 Small scale preparation - 'mini prep'.....	45
3.3.1.2 Large scale preparation of plasmid DNA - 'maxi prep'.....	46
3.3.2 Preparation of genomic DNA from cells.....	46
3.3.3 DNA separation on agarose gels.....	47
3.3.4 Plasmids	47
3.3.5 Plasmid construction.....	48
3.3.5.1 Cloning of luciferase expression constructs.....	48

3.3.5.2 Cloning of the CDP binding site mutations into the MMTV LTR.....	48
3.3.5.3 Cloning of a provirus containing the CDP binding site mutations.....	49
3.3.5.4 Cloning of CDP/p110 isoform.....	49
3.3.6 Homologous recombination.....	49
3.3.7 Oligonucleotide dimerization.....	50
3.3.8 Enzymatic reactions.....	50
3.3.8.1 Restriction digestions of DNA.....	50
3.3.8.2 Ligation of DNA fragments with T4 DNA ligase.....	50
3.3.8.3 Phosphorylation of DNA 5' ends.....	51
3.3.8.4 Dephosphorylation of DNA 5' ends.....	51
3.3.9 Purification of DNA solutions.....	51
3.3.9.1 Phenol/Chloroform extraction.....	51
3.3.9.2 Ethanol precipitation.....	51
3.3.10 Elution of DNA from an agarose gel.....	52
3.3.11 DNA quantification.....	52
3.3.12 Amplification of DNA fragments by PCR (polymerase chain reaction).....	52
3.3.12.1 DIG PCR.....	53
3.3.12.2 Real time PCR (qRT-PCR).....	53
3.3.13 Sequencing reactions.....	54
3.3.14 List of primers.....	54
3.3.14.1 Primers for CDP cloning and binding site mutations.....	54
3.3.14.2 Primers for sequencing.....	55
3.3.14.3 Primers for analysing the MMTV LTR.....	56
3.3.14.4 Primers used for selected methods.....	60
3.3.15 Southern blot hybridization.....	61
3.3.15.1 Separation and denaturation of DNA fragments.....	61
3.3.15.2 Southern blotting.....	61
3.3.15.3 Hybridization.....	61
3.3.15.4 Immunological detection.....	62
3.4 RNA methods.....	62
3.4.1 Extraction of total RNA.....	62
3.4.2 Extraction of mRNA.....	63
3.4.3 DNase digestion of RNA samples.....	63
3.4.4 Generation of first strand cDNA.....	64
3.4.5 RNA ligase-mediated rapid amplification of 5' and 3' ends (RLM-RACE).....	64
3.5 Protein methods.....	65
3.5.1 Extraction of cellular proteins.....	65
3.5.2 Extraction of nuclear and cytoplasmic protein fractions.....	66
3.5.3 Western blot analysis.....	66
3.5.3.1 SDS-PAGE.....	66
3.5.3.2 Western blotting.....	67
3.5.3.3 Coomassie staining of polyacrylamide gels.....	67
3.5.3.4 Immunological detection.....	67
3.5.4 Luciferase assay.....	68
3.5.5 Protein quantification.....	68
3.6 Chromatin methods.....	68
3.6.1 Isolation of chromatin.....	68
3.6.2 MNase digestion.....	69
3.6.3 Determination of nucleosomal borders.....	69
3.6.4 Chromatin immunoprecipitation (ChIP).....	70
4. Results.....	72
4.1 Two elements located within the MMTV LTR influence P2 activity.....	72
4.2 The P3 promoter is regulated by a downstream enhancer.....	77
4.3 A transcriptional initiation site is located downstream of the P3 promoter.....	78
4.4 In the absence of dexamethasone the P2/P3 promoter complex is more potent than the	

major P1 promoter	80
4.5 A doubly spliced transcript originates from the promoter complex.....	81
4.6 Determination of the nucleosome occupancy over the P2/P3 promoter complex.....	83
4.6.1 In silico analysis of the LTR allows multiple positions of nucleosome D.....	84
4.6.2 A nuclease hypersensitive site is detectable at the P2/P3 promoter region.....	87
4.7 Elevated methylation levels accompany transcription from the P2/P3 promoter complex	92
4.8 Post-translational acetylation inversely correlates with ORFe transcription.....	95
4.9 CDP is a repressor of the P2 promoter.....	99
4.10 Constitutive active P2 promoter gives rise to ORFe mRNA.....	103
4.11 The translated ORFe sequence harbours a transmembrane spanning domain and TRAF-2 binding motifs.....	106
5. Discussion.....	108
5.1 The P2/P3 promoter complex.....	108
5.2 Associated chromatin structure.....	109
5.3 The influence of acetylation.....	111
5.4 Model for the transcriptional regulation of ORFe.....	112
5.5 Putative ORFe function.....	114
6. References.....	117

1. Introduction

Retroviruses are RNA viruses which are characterised by their ability to integrate via a proviral DNA intermediate into the genome of the host cell. The transcription of all viral RNA species as well as the synthesis of the virally encoded proteins are mediated by the host cell machinery. Due to the integration process, tumour formation is frequently observed after retroviral infections. Indeed, this consequence often led to their discovery. In the 1930s, Bittner related particles present in the mother's milk of laboratory mouse strains to the high incidence of mammary tumours in those animals (Bittner, 1936).

Mouse mammary tumor virus (MMTV) belongs to the family *retroviridae*, which can be further divided into the subfamilies *orthoretrovirinae* and *spumavirinae* (table 1).

Genus	Type species
Subfamily Orthoretrovirinae	
<i>Betaretrovirus</i>	Mouse mammary tumor virus
<i>Gammaretrovirus</i>	Murine leukemia virus
<i>Deltaretrovirus</i>	Bovine leukemia virus
<i>Epsilonretrovirus</i>	Walleye dermal sarcoma virus
<i>Lentivirus</i>	Human immunodeficiency virus type 1
Subfamily Spumavirinae	
<i>Spumavirus</i>	Chimpanzee foamy virus
unclassified Retroviridae	
<i>Avian retroviruses</i>	Avian endogenous retrovirus EAV-HP

Table 1: the Family *Retroviridae* (according to the NCBI Taxonomy Browser)

MMTV is a non-acute transforming retrovirus, causing tumour formation in susceptible mice after a latency period of 7 to 10 month (Callahan and Smith, 2000; Ross, 2008). So far, no oncogene has been detected within the viral genome. Thus, MMTV is believed to induce tumours by insertional mutagenesis, a process which is further supported by the multiple cycles of mammary cell replication during pregnancy and lactation (Coffin *et al.*, 1997; Ross, 2008). Moreover, the integrated viral DNA is thought to influence cellular gene expression through the action of tissue specific enhancer sequences (Callahan and Smith, 2008).

Due to their activation in MMTV-associated tumours, several cellular proto-oncogenes have been identified including the *Wnt* and *Fibroblast growth factor (Fgf)* gene families and Notch-related genes (Morris *et al.*, 1990; Shackleford *et al.*, 1993; van Leeuwen and Nusse, 1995).

1.1 Exogenous infection route

1.1.1 Dendritic cells

Upon ingestion, the dendritic cells (DCs) of the intestinal tract are the first cells which are infected by the virus present in the mothermilk (figure 1.1) (Martin et al., 2002; Vacheron et al., 2002).

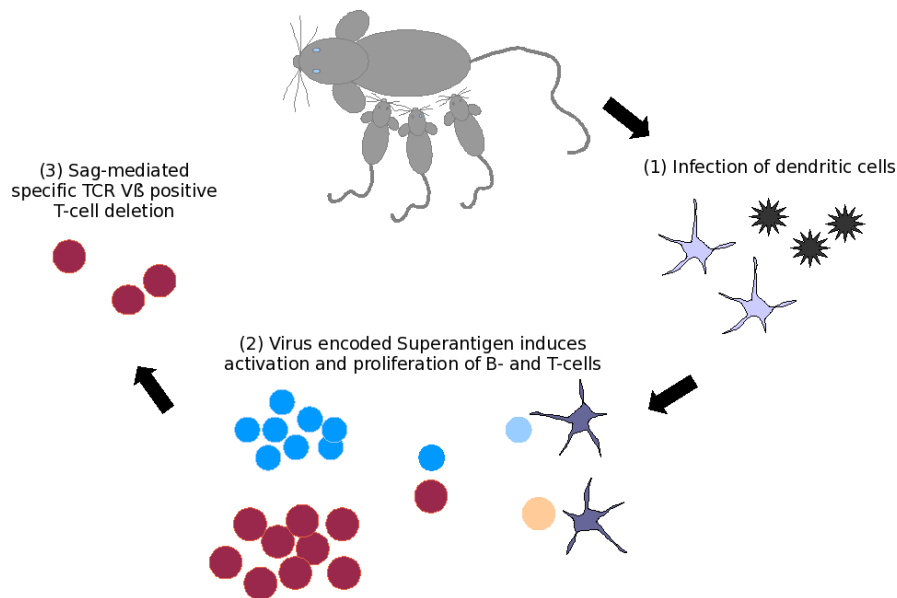


Figure 1.1 Exogenous infection route of MMTV. Viral particles present in the mother's milk are transmitted to the newborn pups. (1) First, the dendritic cells of the gut are infected by the virus. (2) Upon activation, these cells travel to the lymphatic tissue of the intestinal tract, where they express and present the viral encoded superantigen (Sag) to cognate B- and T-cells. In response, the lymphocytes are activated and start to proliferate clonally. (3) Later during infection, a Sag-mediated deletion of specific TCR V β positive T-cells is observed. (Figure modified from www.icmb.utexas.edu/images/upload/Directory_ID_1528/72_mmtvcycleLG.jpg)

A first interaction is mediated by the Toll-like receptors (TLRs) on the surface of the DCs, which are able to recognize invading organisms by pathogen-associated molecular patterns (PAMPs). Ligand binding induces signalling through the Toll/IL-1 receptor domain which ultimately leads to the initiation of inflammatory responses and the activation of the adaptive immunity (Underhill and Ozinsky, 2002; Takeda and Akira, 2007).

MMTV primarily interacts with TLR4, thereby inducing the differentiation of the DCs into major antigen-presenting cells (APCs) (Rassa *et al.*, 2002; Burzyn *et al.*, 2004). Among other things, this process causes the expression of higher levels of the recently identified viral entry receptor on the surface of the DCs, the transferrin receptor 1 (TfR1 or CD71) (Ross *et al.*, 2002). While most cells require iron and thus have the potential to express TfR1, the

highest levels are found in dividing cells and cells that are highly metabolically active. Thus, the tissue specificity of MMTV *in vivo*, which includes lymphoid tissue and mammary epithelial cells, can be explained in part by its receptor usage (Futran *et al.*, 1989; Schulman *et al.*, 1989; Brekelmans *et al.*, 1994). Two segments located on the surface of the TfR1 dimer, which are not involved in transferrin binding, have been shown to be essential for virus attachment (Wang *et al.*, 2006).

For delivering the viral genome into the cytoplasm, MMTV requires trafficking to a low pH compartment (\sim pH 5) (Wang *et al.*, 2008). This process is probably mediated by re-directing the internalized receptor-virus-complex to the late endosomes, although the trigger for the irreversible conformational change of the envelope protein is still unknown. No evidence for the involvement of enzymes located in the late endosome/lysosome compartment (such as cysteine proteases) could have been demonstrated so far (Wang *et al.*, 2008).

Upon activation, the DCs migrate to the lymphatic tissue of the gut, where they present viral peptides together with MHC-class II molecules to cognate T-cells, as described below (Abbas *et al.*, 2000; Vacheron *et al.*, 2002; Jude *et al.*, 2003). Moreover, they have been shown to be capable of producing infectious viral particles that can be transmitted to other cell types (Courreges *et al.*, 2007).

1.1.2 Lymphoid cells

For the establishment of infection in lymphoid cells, MMTV has evolved a strategy that takes advantage of the immune system of the host. The viral encoded superantigen (Sag), which is translated into a 37 kDa transmembrane glycoprotein, is expressed immediately after successful integration of the provirus (Choi *et al.*, 1991). T-cells expressing an appropriate variable region of the T cell receptor β chain (V β) show reactivity to Sag and start to expand clonally (Held *et al.*, 1993).

Thereupon, these activated T-lymphocytes provide B-cell help and produce cytokines that stimulate and recruit additional B- and T- cells, resulting in the establishment of a reservoir of infection-competent as well as infected cells (Held *et al.*, 1993). Later during infection, a gradual deletion or anergy of Sag-reactive T-cells is observed (Nakano *et al.*, 1993).

Besides T-cells, MMTV does infect B-cells that are located in the Peyer's patches of the gastrointestinal tract (Beutner *et al.*, 1994). A first unspecific interaction is mediated by the binding of viral envelope proteins to TLR4 molecules expressed on the surface of these APCs (Rassa *et al.*, 2002). However, early B-cell activation does require the presence of DCs (Courreges *et al.*, 2007). Upon infection, B-cells are stimulated and present Sag together with MHC class II molecules to cognate T-helper cells. In turn, these cells induce the proliferation of

infected B-cells, which is essential in order to achieve high infection levels (Beutner *et al.*, 1994; Karapetian *et al.*, 1994).

Several groups have demonstrated that no infection can be established in mice lacking B- or T-cells (Held *et al.*, 1993; Beutner *et al.*, 1994; Dzuris *et al.*, 1997). Moreover, both cell types are capable of shedding infectious viral particles to the mammary cells, making them a crucial factor for virus spread (Dzuris *et al.*, 1997; Golovkina *et al.*, 1998).

1.1.3 Mammary cells

The cells of the mammary gland epithelium represent the ultimate target of MMTV, as their successful infection ensures that viral particles are released into the mother's milk. Despite ongoing research, it is not known how the virus is actually delivered from the lymphoid cells to the mammary gland, although MMTV is highly specialized for replicating in these cells. In addition to the presence of a mammary specific enhancer element within the LTR (Mink *et al.*, 1990; Lefebvre *et al.*, 1991; Gouillieux *et al.*, 1991; Reuss and Coffin, 2000), the activity of its major P1 promoter is upregulated in the presence of glucocorticoids (section 1.6.1). Furthermore, the virus is thought to be able to increase the potential of mammary cells to differentiate into secretory gland cells, thereby optimizing the route of exogenous transmission (Squartini *et al.*, 1983).

Although several other tissues are infected as well (e.g. testis, salivary gland, kidney or lung), only the mammary gland epithelium is malignantly transformed upon MMTV infection and replication. It is believed that its regenerative activity represents a key factor for initiating tumour formation (Callahan and Smith, 2000).

1.2 Endogenous infection route

After successful integration into the DNA of the host cell, MMTV can additionally be transmitted through the germ line as an endogenous provirus. So far, more than 50 MMTV proviral loci (*Mtv*) have been identified (Hodes and Abe, 2001). Indeed, most laboratory mouse strains contain two to eight endogenous proviruses, although in general these do not encode functional viruses (Kozak *et al.*, 1987). In fact, only *Mtv-1* and *Mtv-2* are known to induce mammary tumours in mice (Moore *et al.*, 1979; Michalides *et al.*, 1981). However, almost all MMTV proviral loci show intact *Sag* coding regions, suggesting a positive selection for the retention of endogenous *Sags* (Ross, 2008). This idea is further supported by the finding that mice transgenic for the *sag* gene of the C3H strain are protected against exogenous MMTVs encoding the same *Sag* specificity (Golovkina *et al.*, 1992).

In contrast, endogenous *Mtvs* seem to be a prerequisite for the successful infection of viruses originating from different strains (Golovkina *et al.*, 1996; Bhadra *et al.*, 2006; Bhadra *et al.*, 2009). For instance, the expression of endogenous Sags was induced upon infection with the exogenous MMTV strain C3H (Xu *et al.*, 1996). Accordingly, only mice harbouring endogenous *Mtvs* (*Mtv*-6, -8 and -9) have been shown to be susceptible to tumorigenesis after infection with C3H MMTV as well as by the MMTV variant type B leukemogenic virus (TBLV) (Bhadra *et al.*, 2006). Hence, BALB/c-c oncogenic mice lacking all endogenous *Mtvs* are resistant to infection and tumorigenesis by exogenous C3H MMTV (Bhadra *et al.*, 2006). Interestingly, those mice could also not be infected with TBLV, a virus that lacks any functional *sag* gene. The mechanism for this resistance is not yet known (Bhadra *et al.*, 2006).

As a consequence of the negative selection during formation of the immune repertoire, Sags encoded by endogenous *Mtvs* promote the clonal deletion of reactive T-cells (Acha-Orbea and Palmer, 1991). Thus, mice harbouring inherited proviruses in their genome possess an altered T-cell repertoire. This could be demonstrated by the resistance of *Mtv*-null mice to the otherwise fatal Gram-negative bacterium, *Vibrio cholerae* (Scherer *et al.*, 1995; Barnett *et al.*, 1999; Bhadra *et al.*, 2006).

1.3 Does MMTV infect human tissue?

The involvement of MMTV in human pathogenesis was proposed a long time ago. Several groups have reported the association of a virus similar to MMTV with human breast cancer (Wang *et al.*, 1995; Etkind *et al.*, 2000; Wang *et al.*, 2001; Lawson *et al.*, 2004; Mok *et al.*, 2008). In these studies, MMTV-like *env* sequences have been detected by PCR amplification from isolated DNA or by immunological methods. Others, however, failed to reproduce those results (Witt *et al.*, 2003; Mant *et al.*, 2004; Bindra *et al.*, 2007). Moreover, a survey performed in almost 100 patients suffering from breast cancer did not reveal any immunological reactivity against MMTV proteins (Goedert *et al.*, 2006).

Recently, Melana and colleagues were able to isolate particles resembling *Betaretroviruses* from primary cultures of patients with metastatic breast cancer (Melana *et al.*, 2007). Additionally, a complete human mammary tumour virus (HTMV) proviral structure was identified in two tumours (Liu *et al.*, 2001). Hence, the finding that the human genome carries a number of distinct human endogenous retrovirus (HERV) sequences which exhibit sequence similarity to MMTV complicate a definite correlation (Mant and Cason, 2004).

On the other hand, studies of MMTV infecting or viruses adapting in order to infect human breast cancer cell lines have been reported (Lasfargues *et al.*, 1979; Howard and

Schlom, 1980; Indik *et al.*, 2005 & 2007). However, the mechanism of both events remains unclear, since human TfR1 does not function as an MMTV entry receptor (Ross *et al.*, 2002; Wang *et al.*, 2006). Experiments in human cells transfected with MMTV Env and hTfR1 (CD71) demonstrated that although both components are required, MMTV was unable to mediate the internalization of hTfR1 (Wang *et al.*, 2008). Alternatively, MMTV may use other receptors for entry into human cells as proposed by Indik and colleagues (Indik *et al.*, 2007).

Taken together, the effect of MMTV-like sequences on the aetiology of human breast cancer remains controversial.

1.4 The retroviral particle

Initially, MMTV was classified as simple retroviruses. Those viruses (e.g. Murine leukemia virus) are thought to encode only the minimal set of proteins and enzymes which are necessary for viral replication. In contrast, complex retroviruses (e.g. HIV-1) possess additional information for a variety of non-structural proteins that facilitate distinct steps of replication or counteract cellular and immunological anti-viral host responses. Since ongoing research in the field of MMTV revealed at least two additional proteins with a major impact on viral pathogenicity, MMTV is now considered as a complex retrovirus.

Retroviruses are enveloped viruses with an average diameter of 80 – 100 nm. Mature, infectious particles possess a core of electrondense material, which is located eccentrically in the subfamily *Betaretrovirus* (Coffin *et al.*, 1997). The protection of the genome, the interaction with the host as well as the delivery of the genome are mediated by the proteins of the viral particle, which is shown schematically in figure 1.2.

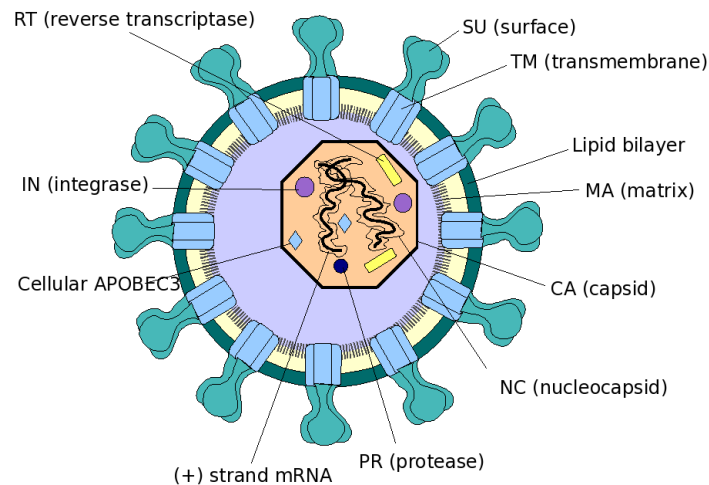


Figure 1.2 The retroviral particle. The composition of a *Betaretrovirus* is shown. The RNA genome is indicated by the two black lines, which are surrounded by the nucleocapsid proteins (marked by the thin black lines). This ribonucleoprotein complex is protected by a shell made of the capsid protein and the lipid bilayer, which is derived from the host cellular membrane. The knobs symbolize the surface part of the virally encoded Env protein. The cellular protein APOBEC3, which is incorporated into the retroviral particles is indicated by the blue diamond, all other geometrical shapes represent virally encoded proteins (modified from Flint *et al.*, 2009).

1.4.1 The retroviral genome

The retroviral genome consists of two identical linear, single-stranded RNA strands of positive polarity, ranging from 7 to 12 kb in size. (The genome of MMTV comprises 8.6 kb.) Only one integrated copy is usually detected after infection; therefore, the term pseudodiploid is used (Coffin *et al.*, 1997). In the mature virion, these two molecules are held together primarily by a sequence called the dimer linkage site, which is located in the 5' untranslated region of *gag* (Coffin *et al.*, 1997).

Since the genomic RNAs are generated by the RNA polymerase II of the host cell, they possess a 5' cap structure and are polyadenylated at their 3' end. A repeated (R) sequence is found at both ends of the viral RNA genome, which is followed by a unique sequence at the 5' end (U5). The untranslated region at the 3' end is referred to as U3 region (U3) and contains most of the transcriptional control elements (promoters and enhancers). The major transcriptional units of replication competent retroviruses from 5' to 3' direction are *gag*, *pol* and *env* (figure 1.3).

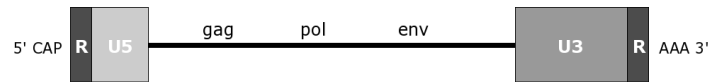


Figure 1.3 Retroviral RNA genome. Depicted is a linear, single stranded RNA genome. The repeat sequence (R) and the unique regions at the 5' end (U5) and the 3' end (U3), respectively, are symbolized by boxes. The transcriptional units (gag, pol, env) are indicated as a black line. The 5' cap-structure as well as the 3' polyA tail demonstrate the mRNA nature of the retroviral genome.

As a consequence of the reverse transcription step during infection, the genomic RNA is copied into a double stranded proviral DNA, which possesses identical structures on both ends of the coding region (figure 1.4). These are referred to as long terminal repeats (LTRs).

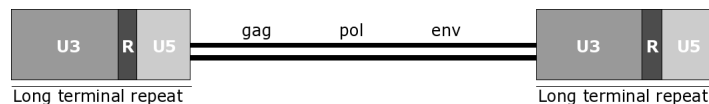


Figure 1.4 Proviral DNA. After reverse transcription the retroviral genome is referred to as provirus and comprises identical regions (long terminal repeats) on both ends of the double stranded DNA genome.

The viral DNA synthesis takes place within the cytoplasm in a large ribonucleoprotein structure called the reverse transcription complex (RTC), which is believed to result from structural rearrangements of the virion core (Mougel *et al.*, 2009). As the responsible enzymatic reaction represents a hallmark of retrovirus replication, it is described in detail below.

Similar to DNA-dependent polymerases, the viral encoded reverse transcriptase (RT) cannot initiate the DNA synthesis *de novo*. Therefore, it requires a specific primer. This task is accomplished by a cellular tRNA, which has been positioned at the primer binding site (pbs) of the RNA genome during virus assembly (figure 1.5).

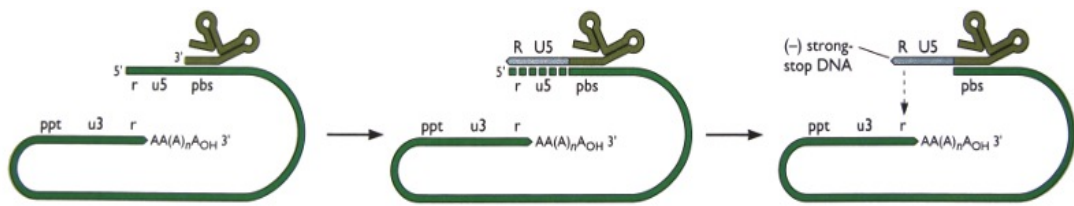


Figure 1.5 Retroviral reverse transcription (I). The retroviral RNA genome is shown in green. Its primer binding site (pbs) is occupied by the cellular tRNA, thereby allowing the initiation of (-) strand DNA synthesis. The initiation of the reverse transcription process until the formation of the (-) strong stop DNA product is shown (Flint *et al.*, 2009).

The reverse transcription starts, but quickly (after about 100 nucleotides) runs out of template, since it encounters the 5' end of the RNA genome. Immediately after being copied into the (-) strand DNA, the template is digested by the RNase H domain of the RT. The product including the tRNA primer accumulates and is called (-) strong-stop DNA.

In the following step, hydrogen bonding between the R sequence of the (-) strong-stop DNA and the complementary R sequence upstream of the poly(A) tail occurs, which allows the RT to continue copying. The substitution of one end of the RNA for another or *first template exchange* takes place prior to complete digestion of the copied 5' end of the viral RNA (figure 1.6).

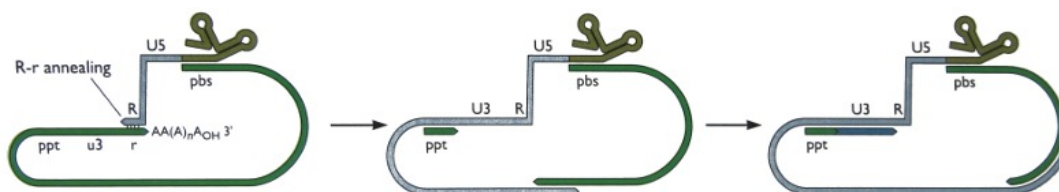


Figure 1.6. Retroviral reverse transcription (II). The first template exchange and the initiation of the (+) strand DNA synthesis are shown (Flint *et al.*, 2009).

One early digestion product of the genomic RNA is a fragment comprising the polypurine tract (ppt). Resourcefully, these few nucleotides serve as primer for the initiation of the (+) strand DNA synthesis, which starts prior to the completion of the (-) strand DNA. Again, the RT is forced to arrest after several nucleotides as it encounters a modified base which cannot be copied. The accumulating product is called the (+) strong-stop DNA (figure 1.7).

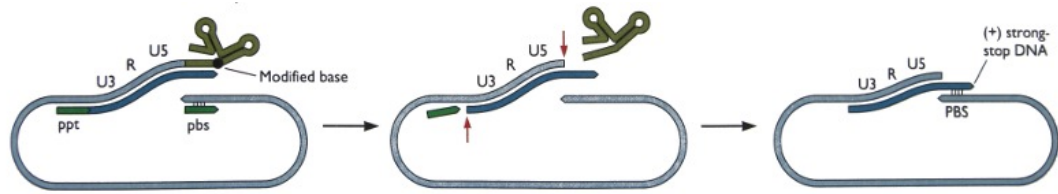


Figure 1.7. Retroviral reverse transcription (III). Generation of the (+) strong stop DNA (Flint *et al.*, 2009).

The removal of the tRNA primer by RNase H allows the single-strand 3' end to anneal to complementary sequences located at the corresponding 3' end of (-) strand DNA. This process is referred to as *second template exchange* and provides a circular DNA template for continued polymerization by the reverse transcriptase (Flint *et al.*, 2009).

Subsequently, (-) strand DNA synthesis continues until it reaches the end of viral DNA template. Both the production of (+) strong-stop DNA and the converging (-) strand DNA synthesis disengage the template ends during processing. The resulting product is the (-) strand of viral DNA equivalent of entire genome (in permuted order) annealed to (+) strong stop DNA (figure 1.8).

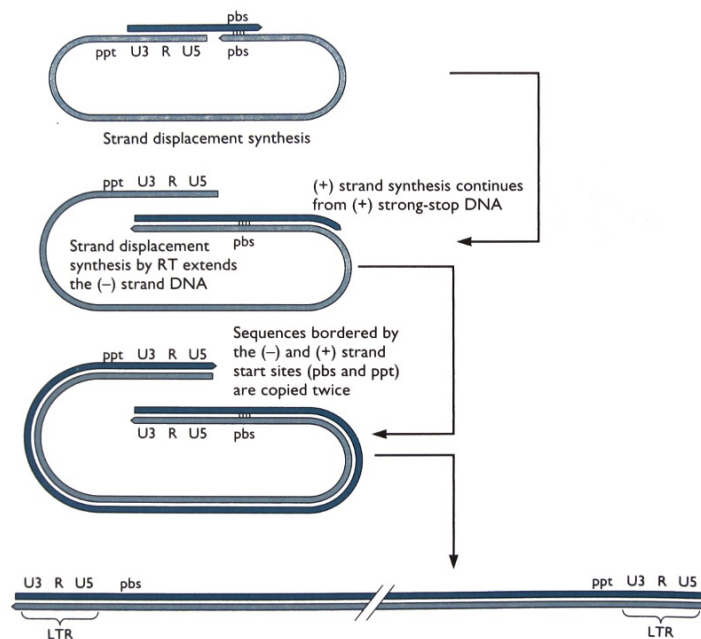


Figure 1.8. Retroviral reverse transcription (IV). The process from the second template exchange until the formation of the final linear DNA product is shown schematically (Figure modified from Flint *et al.*, 2009).

As a consequence of the whole process, the regulatory elements of the U3 region are now in the optimal position to direct the transcription of the viral genome.

1.4.1.1 MMTV LTR promoters

The current model for all retroviruses implies that a promoter located immediately upstream of the transcriptional initiation site (which is per definition the first base of the R region) gives rise to the full length RNA, which is subsequently either packaged into newly budding virions or used for translation of viral proteins. In MMTV, the major promoter is termed P1 promoter (Donehower *et al.*, 1981; Klemenz *et al.*, 1981; Kusk *et al.*, 1995). The mechanism of its activation will be described below (section 1.6)

In close vicinity to the P1 promoter, an initiator element (Inr) has been identified inside the R region (Pierce *et al.*, 1993; Fee *et al.*, 2002). Generally, the Inr motif encompasses the transcriptional start site and represents the probably most common occurring sequence motif in focused core promoters (Juven-Gershon *et al.*, 2008). These group of promoters possess either a single transcriptional start site or a distinct cluster within a short region. In vertebrates, only one-third or less are focused promoters, while the vast majority are dispersed core promoters, which have a number of transcriptional start sites distributed over a region of 50 to 100 nucleotides (Juven-Gershon *et al.*, 2008). Hence, a precise transcription initiation in TATA-box containing promoters generally requires both the TATA-box motif and an Inr-like element (Sandelin *et al.*, 2007).

Recently, it was successfully demonstrated that an MMTV provirus carrying a deletion of the P1 TATA-box is able to produce the viral envelope protein in an hormone-dependent way as efficiently as a non-deleted provirus (Rungaldier *et al.*, 2005). Thus, the MMTV Inr motif represents a classical initiator element that is able to direct specific transcription in the absence of the TATA-box in addition to its role in the start site selection process. However, the Inr element failed to substitute for the absence of the P1 promoter in producing full length genomic viral RNAs (Rungaldier *et al.*, 2005).

Originally, the existence of an additional promoter element located in the U3 region was reported by Günzburg and colleagues. They observed transcripts that include sequences upstream of the major P1 promoter, when they were characterizing messenger RNAs from either MMTV infected or transfected cells (Günzburg *et al.*, 1993). The start site of the novel transcriptionally active region, which has been termed P2 promoter, was assigned to position 696 (relative to the beginning of the LTR). The activity of the P2 promoter was shown to be enhanced in the presence of dexamethasone as well as in response to LPS-treatment (Günzburg *et al.*, 1993). Since its putative 1.2 kb transcripts were reported to be generated by splicing events occurring between the splice donor site of *gag* and the splicing acceptor site used for the ORF/Sag transcripts (Wheeler *et al.*, 1983; van Ooyen *et al.*, 1983), the authors concluded that the P2 promoter is responsible for Sag expression (Günzburg *et al.*, 1993).

1.4.2 Viral proteins

1.4.2.1 GAG

The *gag* gene is transcribed from full length, unspliced mRNA. Following translation, the MMTV polyprotein precursor (p77) is cleaved by the viral protease (PR) giving rise to the myristoylated matrix protein (p10), the capsid protein (p27) and the nucleocapsid protein (p14). Several other proteins are generated as well, yet their functions remain unknown (Hizi *et al.*, 1987).

The nucleocapsid proteins (NCs) are highly basic nucleic acid binding proteins that contain one or two conserved zinc finger domains. They have been shown to direct the selection, packaging and dimerization of the genomic RNA during virion assembly. Moreover, their association to the genomic RNA results in the congregation of other structural molecules, which in turn promotes the multimerization of the capsid proteins (CA) (Thomas and Gorelick, 2008). During early infection, the NC proteins contribute to the fidelity of the reverse transcription reaction by inhibiting self-initiation of the cDNA synthesis. Furthermore, they facilitate the annealing of complementary sequences as well as strand transfer and exchange reactions (Mougel *et al.*, 2009).

Upon retroviral infection, two ribonucleoprotein complexes are found within the cytoplasm of target cells, the Reverse-transcription-complex (RTC) and the Pre-integration-complex (PIC). Both complexes are encased by a protein shell built by the capsid (CA) protein, which represents an important scaffold protein. Accordingly, CA mutants have been shown to reduce viral infectivity by blocking either the process of reverse transcription or distinct steps during nuclear entry (most probably due to improper core disassembly events) (Thomas and Gorelick, 2008).

Located in between the envelope and the viral core is an additional protein layer, comprised of the membrane-associated matrix (MA) proteins. These polypeptides are co-translationally myristoylated. Although not sufficient, this modification is required for anchoring the Gag polypeptides to the plasma membrane of the host cell (Coffin *et al.*, 1997).

During the budding process, the Gag polyproteins are cleaved by the viral protease in a highly ordered cascade. Immature virions contain a roughly spherical shell of radially extended Gag molecules, whereas the N-terminal Gag (MA domain) is bound to inner viral membrane and the C-termini of the Gag molecules project into the centre of virus. A disruption of the regulated nature of these events is highly detrimental to viral maturation and infectivity (Adamson and Freed, 2008).

1.4.2.2 POL

The corresponding gene encodes for the essential enzymes protease (PR), reverse transcriptase (RT) and integrase (IN) (in 5' to 3' direction). In MMTV, these proteins are translated in different reading frames from full length, unspliced mRNA (Hizi *et al.*, 1987). Pseudoknot structures within the mRNA induce ribosomal frameshifting events that are required for the generation of the gag-pro and the gag-pro-pol precursors (Hizi *et al.*, 1987; Chamorro *et al.*, 1992; Chen *et al.*, 1995). As a consequence, all enzymatic proteins are fused to the Gag polyprotein, which provides a direct way to incorporate enzymes into the virion during assembly (Coffin *et al.*, 1997).

Pro encodes for the viral protease. The enzyme possesses a broad specificity with a maximum activity at pH from 4-6 and cleaves virally derived polyprotein precursors during the essential process of virion maturation (Menendez-Arias *et al.*, 1992). During these events, the liberated NC-RNA complex condenses at the centre of the core, the genome RNA dimer becomes more stable and the capsid is formed. In order to ensure that Gag proteins are not processed before they assemble, the protease is liberated at a very late stage of viral assembly from the Gag-Pol fusion polypeptide as has been recently demonstrated for the HIV protease (Ganser-pornillos *et al.*, 2008).

The translation products of *pol* are the viral enzymes reverse transcriptase and integrase. As has been described in section 1.4.1, the RT accounts for the viral DNA synthesis. The end product represents a direct substrate for the enzyme integrase, which specifically catalyses the insertion of the reverse transcribed viral genome into the host cell DNA. The actual reaction comprises three distinct steps (processing, joining and repair, figure 1.9), two of which are catalysed by the IN.

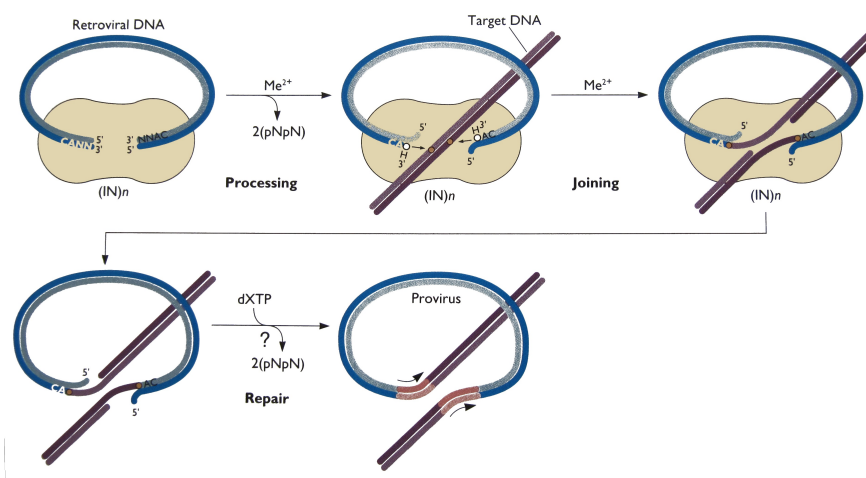


Figure 1.9 Overview of the enzymatic reactions necessary for retroviral integration. The proviral DNA represents the direct substrate for the viral encoded integrase. First, the enzyme removes

dinucleotides from the 3' ends of both LTRs. During the strand transfer reaction, the processed proviral DNA is joined covalently to staggered nicks at the target site that have been introduced by the integrase as well. Cellular enzymes are responsible for final repair reactions (Flint *et al.*, 2009).

In a first step, dinucleotides are removed from the 3' ends of both LTRs, thereby exposing a CA 5' overhang. In the subsequent event, the strand-transfer reaction, the 3' recessed ends are joined to 5' phosphates at the target site which have also been created by the integrase. Consequently, the proviral ends of all retroviruses are embedded by 5'-TG and CA-3'. (Occasionally, these dinucleotides extend into imperfect inverted repeat that can be as long as 20 bp for some viral genomes.) Finally, the exposed 5' dinucleotides of the LTR end are removed and the single-stranded region is repaired, resulting in direct repeat at each LTR-chromosomal DNA junction. This last step is mediated by cellular repair enzymes (Thomas and Gorelick, 2008). Figure 1.10 shows schematically the viral genome before and after integration.

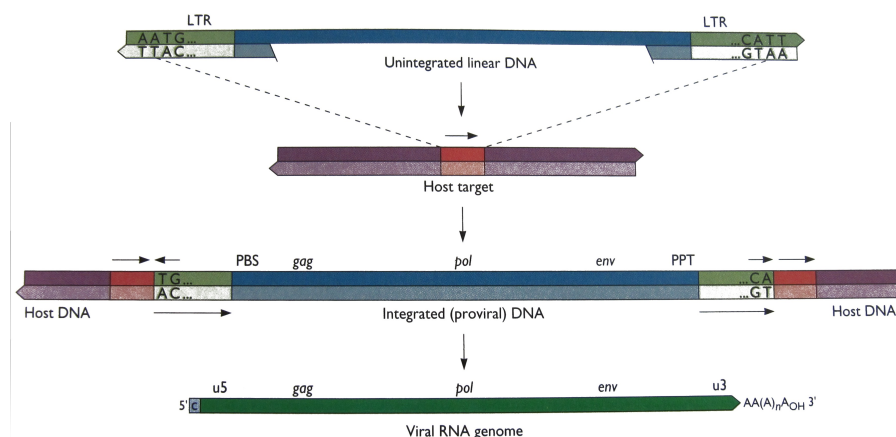


Figure 1.10 The retroviral genome before and after integration into the host cell DNA. Above, the unintegrated linear DNA is shown together with the host cell DNA. The respective target sequence is marked by an arrow. During integration, two base pairs on either side of the LTR are removed and a 6-bp duplication at the target site are generated. The overall enzymatic process results in an integrated proviral DNA that includes short, imperfect inverted repeats at its termini, which end with conserved 5'TG...CA-3' sequence (middle). The integrated provirus serves as template for the generation of new RNA viral genomes (Flint *et al.*, 2009).

Recently, MMTV integration sites have been analysed from acutely infected cells (Faschinger *et al.*, 2008). The experimental set-up was distinct from earlier studies, which mainly utilized tumour cells (i.e. clonally expanded cells) to determine common insertion sites and thereby most probably induced a bias (Gray *et al.*, 1986; Gallahan and Gallahan, 1987; Clausee *et al.*, 1993; Marchetti *et al.*, 1995; Theodorou *et al.*, 2007). This latest study revealed

that MMTV possesses the most random target sites selection among retroviruses at present investigated (Faschinger *et al.*, 2008).

The respective gene is encoded from the 3' end of *pol* region.

1.4.2.3 ENV

The outer lipid envelope of the retroviral particle is derived mainly from the host cell membrane, while the characteristic protrusions are built by integral viral glycoproteins. They are encoded by the *env* gene, which is translated from a single-spliced subgenomic RNA and proteolytically cleaved by host furin enzymes (Dudley and Varmus, 1981; Coffin *et al.*, 1997).

In MMTV, the surface polyprotein (SU or p52) mediates the binding to the cellular receptor TfR1, while the transmembrane protein (TM or p33) participates in the virion-cell membrane fusion process (Wang *et al.*, 2006; Ross, 2008). As mentioned earlier, the MMTV Env was additionally shown to interact with antigen presenting cells via the TLR4, thereby facilitating initial steps of infection (Rassa *et al.*, 2002; Burzyn *et al.*, 2004). Moreover, an involvement of the Env protein in tumour formation was demonstrated by ectopic expression of *env* in normal mammary epithelial cells, which resulted in a phenotypic transformation of the transfected cells. An immunoreceptor activation motif (ITAM) was identified as a critical, although not sufficient factor for the cellular transformation process (Katz *et al.*, 2005; Ross *et al.*, 2006).

Other oncogenic viruses do encode viral protein with ITAMS as well. For instance, the Env protein of the related *Betaretrovirus* Jagg siekte sheep retrovirus (JSRV) has been demonstrated to be required for the transformation of lung epithelial cells (Maeda *et al.*, 2001; Rai *et al.*, 2001; Wootton *et al.*, 2005).

1.4.3 MMTV accessory proteins

1.4.3.1 Sag (Superantigen)

As mentioned earlier, the superantigen is not displayed by the virus itself, but expressed from the infected antigen presenting cell. The resulting immune stimulation is essential for viral transmission. Sag is encoded from an ORF located within the LTR (Brandt-Carlson and Butel, 1991; Choi *et al.*, 1991 & 1992). Due to the low abundance of the protein and difficulties in the production of reactive antibodies, the regulation of Sag is largely unknown. Moreover, the promoter responsible for Sag expression could not be identified with certainty. Several candidates have been reported to account for *sag* mRNA expression (van Ooyen *et al.*, 1983; Günzburg *et al.*, 1993; Zhang *et al.*, 1996; Reuss and Coffin, 1998).

An MMTV variant, Type B leukemogenic virus (TBLV), induces T-cell lymphomas in infected animals after a very short latency period. The analysis of isolated viruses revealed major alterations within the LTR (Michalides *et al.*, 1986). Infected T-cells transcribe high levels of MMTV-specific RNAs due to the deletion of a negative regulatory element (Ball *et al.*, 1988; Hsu *et al.*, 1988). Besides, a substitution of 124 nucleotides is found in the U3 region of the virus that is thought to function as a T-cell specific enhancer (Mertz *et al.*, 2001). Both mutations alter the *Sag* coding region, resulting in the expression of a truncated protein. However, it seemed to be dispensable for the process of tumour formation, indicating that the loss of the NRE and putative other *cis*-acting elements are sufficient for altering the tissue specificity of the virus (Mustafa *et al.*, 2003).

To date, no B-cell malignancies induced by MMTV are known.

1.4.3.2 Rem (Regulator of export of MMTV mRNA)

The nuclear export of incompletely spliced RNAs is usually prevented in eukaryotic cells. Thus, viruses have evolved different strategies in order to circumvent these mechanisms. Simple retroviruses utilize *cis*-acting viral RNA segments like the constitutive transport element (CTE) of the *Betaretrovirus* Mason-Pfizer monkey virus (MPMV), which directly interacts with cellular proteins and thereby promote the export of the full length RNA (Bray *et al.*, 1994; Pasquinelli *et al.*, 1997).

In contrast, complex retroviruses encode auxiliary *trans*-acting proteins. The HIV encoded Rev protein binds to an RNA motif termed the Rev-responsive element (RRE) which is located in the *env* coding region (Felber *et al.*, 1989; Hadzopoulou-Cladaras *et al.*, 1989; Malim *et al.*, 1989). It acts as adaptor for the full length mRNA and contains a leucine-rich nuclear export signal (NES), which has been shown to interact with the nuclear export factor Crm1/Xpo1 (chromosome region maintenance-1/exportin-1) (Lin *et al.*, 2005). Rev and Rev-like proteins shuttle between nucleus and cytoplasm (Cullen, 2003).

Similar to HIV, MMTV encodes a double spliced transcript, termed *regulator of export of MMTV mRNA* (Rem) (Indik *et al.*, 2005; Mertz *et al.*, 2005). The *rem*-coding sequences are derived from the entire *env* signal-peptide coding region as well as parts of the SU and the TU and give rise to a 33 kDa protein (Indik *et al.*, 2005; Mertz *et al.*, 2005). Interestingly, all motifs necessary for RNA export are localized at the very N-terminus, making it similar in structure to the related HERV-K Rec protein (Mertz *et al.*, 2005). The binding site for Rem lies within a highly structured viral RNA segment at the *env*-U3 RNA junction and was termed Rem responsive element (RmRE) (Mertz *et al.*, 2005; Müllner *et al.*, 2008).

Recently, a similar protein was identified in the *Betaretrovirus* Jaagsiekte Sheep Retrovirus. Like Rec and Rem, Rej is a doubly spliced transcript encoded within the *env*

sequence. However, the mechanism of action may differ, since Rej does not appear to facilitate export of unspliced viral RNA. Nevertheless, it seems to be required for the synthesis and/or the translation of viral Gag polyproteins (Hofacre *et al.*, 2009).

1.4.3.3 DUTPase

Like herpesvirus, poxvirus and other retroviruses, MMTV encodes a dUTPase. This protein is generated by a ribosomal frameshift at the *gag-pro* junction (Bergmann *et al.*, 1994). Its function is the hydrolysis of dUTP, which leads to an increase in cellular dTTP levels. Thereby, the enzyme minimizes the misincorporation of uracil into the DNA during the process of reverse transcription. Since the expression levels of the cellular enzyme are low in differentiated cells (Chen *et al.*, 2002), MMTV relies on the activity of its own protein.

1.4.4 Cellular APOBEC3

APOBEC3 proteins (A3) are potent host restriction factors of viruses and retrotransposable endogenous mobile genetic elements (Harris and Liddament, 2004). Several mechanisms have been proposed for the antiviral activity of APOBEC3, among them a cytidine deaminase activity. During the process of reverse transcription, this property deaminates cytidine to uridine, thereby inducing hypermutations, which subsequently result in the degradation of the viral DNA and the generation of nonfunctional proviral genomes (Harris *et al.*, 2003).

Mice express a single APOBEC3 (mA3) protein (Cullen, 2006). Ross and co-workers have demonstrated that mA3 proteins are packaged into MMTV virions in an RNA-dependent fashion, thus inhibiting MMTV infection. Moreover, all tissues that are predominantly infected by MMTV express mA3 transcripts (Okeoma *et al.*, 2007). Accordingly, mA3^{-/-} mice have been shown to be more susceptible to MMTV infections, as viruses spread more rapidly and to an higher extent in these animals than in control groups (Okeoma *et al.*, 2007). Interestingly, upon viral infection, DCs increase their intracellular mA3 levels (Okeoma *et al.*, 2009a). These findings could explain why MMTV infects mature DCs less efficiently than immature DCs, although MMTV infection leads to an upregulation of the viral entry receptor TfR1 (Vacheron *et al.*, 2002).

Recently, it was demonstrated that mA3 inhibits infection by murine leukemia virus (MLV). The effect was not due to the enzymatic properties of the protein; instead, the restriction seemed to depend on the mA3 allele expressed (Takeda *et al.*, 2008). Similarly, mA3 could play a role in the genetics of susceptibility and resistance to MMTV infection (Okeoma *et al.*, 2009b). Since MMTV possesses a purine-rich genome, one could hypothesize

that MMTV do persist in mice because the virus became in evolutionary terms less susceptible to the inhibition by this protein (Okeoma *et al.*, 2007).

1.5 The retroviral replication cycle

The replication cycle within a single cell can be divided into early steps, which include attachment and binding to the target cell, viral entry as well as the processes of reverse transcription and integration and late steps, which comprise the transcription and translation of the viral proteins, their assembly and finally the release of the virion. A summary of these processes is shown in figure 1.11.

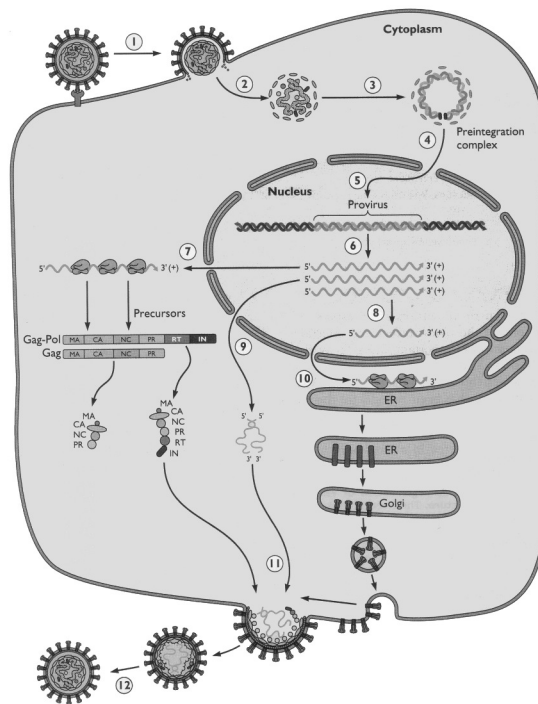


Figure 1.11 Retroviral life cycle. After binding and attachment to the cellular receptor (1), the viral envelope protein fuses with the cellular membranes and the viral core is released into the cytoplasm (2). The viral RNA genome is reverse transcribed into DNA (3) and transported via the preintegration complex to the nucleus (4). After integration (5), the proviral DNA is transcribed by the cellular RNA polymerase II (6). Unspliced (7) as well as spliced mRNAs (8) are translated and transported to the cell membrane (10). Some full length mRNAs are packaged into the nascent viral particles in order to serve as viral genomes (9). The assembly of the viral particle takes place at the cell membrane (11), from which the viral progeny buds off. Maturation of the viral particles is achieved by the condensation of the viral core (12). (Figure modified from Flint *et al.*, 2009)

Through interaction of the viral envelope protein with components of the cell surface, the envelope of the virus fuses with the plasma membrane of the host cell. Subsequently, the viral core undergoes a partial and progressive disassembly (uncoating), leading to the formation of the reverse-transcription complex. It is believed that the initiation of the reverse transcription is coupled to the onset of uncoating, since the process depends on a sufficient amounts of dNTPs. Nevertheless, limited linear DNA synthesis has been observed in virions prior to infection (Nisole and Saïb, 2004).

Besides the double-strand proviral copy of the viral genome, smaller quantities of two specific circular DNA products can be detected as well; however, they are considered as non-functional, dead-end products (Flint *et al.*, 2009). One is the LTR circle, which possibly occurs due to a failure of strand displacement during reverse transcription; the other is the two LTR circle. The latter is thought to be generated by the ligation of the ends of the linear viral DNA by the cellular enzyme DNA ligase. As the two LTR circle can be easily detected by PCR amplification, it is often used as a marker for the successful transport of the viral DNA to the nucleus.

The overall process of reverse transcription is very slow due to the periodic pausing of the enzyme, reaching only one tenth of the speed of DNA-dependent polymerases. Additionally, the fidelity rate of the enzyme is low, since it lacks any editing activity, thus contributing to the high mutation rates observed *in vivo*. Due to the utilization of the second RNA genome, genetic recombination events like misincorporations, rearrangements or deletions also occur with high probability. They are referred to as copy choice recombination (Flint *et al.*, 2009).

Four to eight hours after viral entry during infection, the preintegration complex (PIC), containing full-length viral DNA, localizes to the nucleus of the host cell. However, it is poorly understood how the structure passes through the nuclear pore complex, as it is too large for passive diffusion (Flint *et al.*, 2009).

Retroviruses depend on the host cell machinery for the transcription of their RNA genome as well as for the production of their proteins; therefore, several viruses including MMTV encode additional auxiliary proteins that modulate RNA Pol II transcription or regulate RNA splicing and export (see above).

For *Beta-*, *Deltaretroviruses* and *Spumaviruses*, the assembly of the viral particle takes place within the cytoplasm of the infected cells (Coffin *et al.*, 1997). Besides the RNA genome, which is recognized via the packaging signal, equimolar quantities of the reverse transcriptase and the integrase are incorporated as well as other relevant nonstructural proteins. Upon release of the viral particles, referred to as budding, the immature virions acquire a lipid envelope, which is derived from the cell membrane. Simultaneously, the maturation process starts that includes the proteolytic processing of virion Gag proteins. The

virions are now ready to infect a new cell.

1.6 The MMTV major P1 promoter

1.6.1 Stimulation by glucocorticoid hormones

The LTR of MMTV is widely used as a model system for studying hormone-dependent transcriptional regulation. The activity of the major P1 promoter is enhanced in the presence of steroid hormones which include glucocorticoids, progestins and androgens (Ringold *et al.*, 1975; Cato *et al.*, 1987). Upon binding to their receptors, these molecules are able to pass the nuclear membrane and adhere to sequences which are collectively termed hormone responsive element (HRE). Up to five consensus binding sites, which are represented by the hexamer TGTCT, are located between bps 994 and 1114 (relative to the beginning of the LTR) (Buetti and Diggelmann, 1983; Hynes *et al.*, 1983).

Additionally, two cellular transcription factors are required for proper hormone-induced transcription from the major P1 promoter (Mymryk *et al.*, 1995). Nuclear factor 1 (NF-1) binds at position 1114 to 1134 (Cordingley and Hager, 1988), while sequences recognized by the octamer binding factor (Oct-1) are found in the region from bps 1134 to 1156 (relative to the 5' end of the LTR) (Brüggemeier *et al.*, 1991). In the absence of hormones, the chromatin structure is thought to exclude both proteins from binding to their respective sites on the MMTV LTR (Hebbar and Archer, 2007).

Early studies on the chromatin organization of the P1 promoter revealed that an array of six phased nucleosomes covers the MMTV LTR (Richard-Foy and Hager, 1987). The mean centre positions of the regions between the nucleosomes were determined by MPE-Fe(II) cleavage as well as by micrococcal nuclease digestion of either uninduced or hormone-treated cells. (The centres of the cutting sites were determined by the indirect end-labelling method and are: 175, 368, 543, 750, 944, 1134, 1330 (relative to the beginning of the LTR) (Richard-Foy and Hager, 1987).)

The phasing was reproduced in chromatin reconstitution experiments comprising fragments of the MMTV LTR and core histone octamers (Pina *et al.*, 1990). In contrast, transiently introduced copies of the MMTV LTR do not acquire a specific chromatin structure, although they are hormone inducible (Lee and Archer, 1994).

The six positioned nucleosomes are referred to as nucleosomes A-F; nucleosome F is farthestmost distant from the transcriptional start site of the major P1 promoter (figure 1.12).

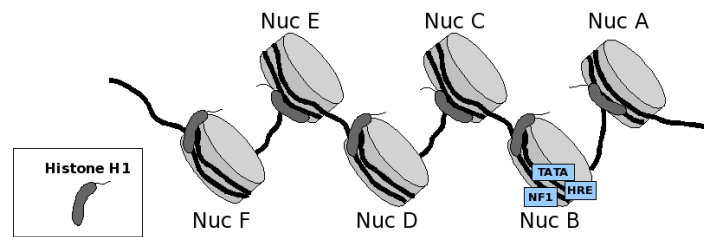


Figure 1.12 The chromatin organization of the MMTV LTR. Six nucleosomes are positioned over the MMTV LTR, which are referred to as nucleosome F-A (in 5' to 3' direction). The core histones are represented by the grey cylinders, while the linker histone H1 is shown in dark grey. The TATA-box of the major P1 promoter (TATA) as well as the hormone responsive element (HRE) and the binding site for transcription factors (the nuclear factor-1 binding site (NF-1) is shown exemplarily) are indicated as blue boxes. (Figure modified from Vicent *et al.*, 2008.)

In the presence of glucocorticoids, a hypersensitive site is detectable within nucleosome B that covers the HRE (Zaret and Yamamoto, 1984; Richard-Foy and Hager, 1987). The existence of nuclease hypersensitive site is either interpreted as a gap in the nucleosomal array (also referred to as nucleosome-free region) or is thought to result from a temporary removal of histone proteins (Rando and Ahmad, 2007). However, the presence of DNase hypersensitive sites do not automatically correlate with the rate of transcription from the corresponding promoters. Therefore, these sites are suggested to reflect rather the transcriptional competence than the actual promoter activity (Workman, 2006).

In MMTV, however, treatment with glucocorticoids results in the rapid induction of transcription driven by the P1 promoter and a concomitant remodelling of the chromatin in the presence of ATP (Fletcher *et al.*, 2000). Hormone-bound steroid receptors have been shown to enhance the association of chromatin remodelling complexes of the SWI/SNF family to the MMTV P1 promoter (Fletcher *et al.*, 2000; Nagaich *et al.*, 2004). During the final steps of promoter activation, the ATP-dependent catalytic subunit BRG1 (Brahma-related gene 1) is recruited together with RNA Pol II (Georgel *et al.*, 2003). Through its activity, the histone H2A and H2B subunits are displaced from nucleosome B, providing a more accessible DNA (Vicent *et al.*, 2004). Thus, steroid hormone binding facilitates the association of other relevant transcription factors (Fletcher *et al.*, 2000).

Furthermore, bound NF-1 has been proposed to stabilize GR/BRG1 interactions. In the absence of the transcription factor, the chromatin architecture of the promoter does not permit a strong association of the GR to the HRE, suggesting a cooperativity among these two DNA binding factors (Chavez and Beato, 1997; Hebbar and Archer, 2007).

This whole process is accompanied by the loss of the linker histone H1 from the promoter region (Bresnick *et al.*, 1992). Generally, the histone H1 stabilizes higher order

chromatin structures and acts as general transcriptional repressor (Lee and Archer, 1998). During GR-mediated chromatin remodelling, however, the histone H1 is phosphorylated and thereupon lost from the promoter (Bresnick *et al.*, 1992). Figure 1.13 summarises the essential steps during the MMTV P1 promoter activation.

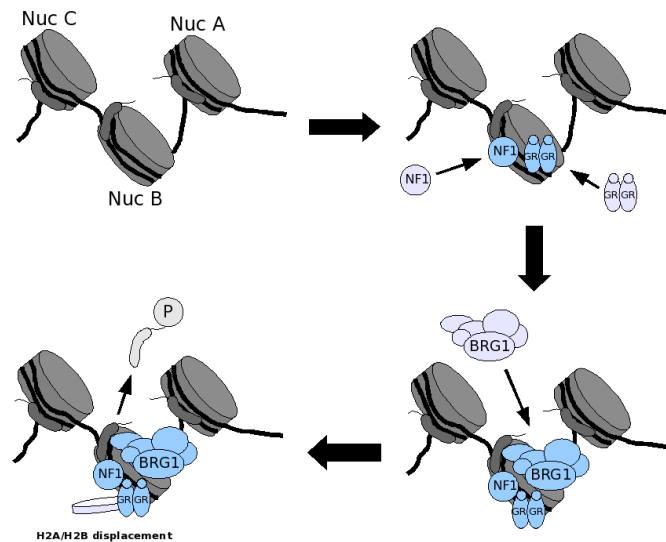


Figure 1.13 Nucleosome remodelling at the MMTV P1 promoter. The steps of the nucleosome B remodelling process are shown schematically. The binding of the glucocorticoid receptor (GR) and the transcription factor NF-1 promotes the association of the ATP-dependent chromatin remodelling complex BRG1, which induces the dislocation of the H2A/H2B dimer. These events are accompanied by the removal of the phosphorylated linker histone H1. (Figure modified from Vicent *et al.*, 2008).

Recently, the binding of the transcription factor forkhead box A1 (FoxA1) to the MMTV LTR was reported. It is speculated that FoxA1 replaces the linker histones from open chromatin and anchors nucleosomes to specific sites; therefore, the association of FoxA1 could alter the chromatin structure and affect transcription as well. Binding sites have been mapped to regions flanking the nucleosome B region (Holmqvist *et al.*, 2005).

1.6.2 Negative regulation

The activity of the major P1 promoter is additionally modulated by two effectors, which have been described to negatively influence transcription from the MMTV LTR.

The existence of a negative acting factor (Naf) was postulated by Salmons and colleagues during attempts to establish a packaging cell line containing full length MMTV proviral sequences under the control of the Rous sarcoma virus (RSV) promoter. Upon stable introduction of MMTV based retroviral vectors, they observed a downregulation in the

expression level of Gag proteins. They proposed that this effect was due to the *trans*-acting factor Naf (Salmons *et al.*, 1990). Nevertheless, neither the coding requirements nor the mechanism of its effector functions have yet been elucidated (Metzner *et al.*, 2006).

Several groups have described the presence of a negative regulatory element (NRE) within the LTR (Mink *et al.*, 1990; Lefebvre *et al.*, 1991; Bramblett *et al.*, 1995). Recently, it was demonstrated by Dudley and colleagues that the cellular transcription factor CCAAT displacement protein (CDP) binds to the NRE, thereby repressing MMTV transcription (Liu *et al.*, 1999; Zhu *et al.*, 2000).

CDP proteins are a family of transcription factors that contain a Cut homeodomain and one to three Cut repeats (Sansregret and Nepveu, 2008). (The term "Cut" originally derived from the "cut wing" phenotype observed in a mutant of *Drosophila melanogaster*. However, due to the variety of terms used for the same gene family in different species, the gene root of Cut-like was changed to CUX (*Cut*-homeobox). Accordingly, *CUX1* and *CUX2* refer to human genes, *cux1* and *cux2* to mouse genes and CUX1 and CUX2 to the respective proteins (Sansregret and Nepveu, 2008). However, for simplicity, the term CDP will be used throughout this thesis.)

CDP proteins are highly conserved and involved in the control of proliferation and differentiation processes (Nepveu, 2001). Several isoforms with distinct DNA binding and transcriptional properties are known that are generated by alternative transcriptional initiation, polyadenylation site selection and splicing as well as proteolytic processing (figure 1.14).

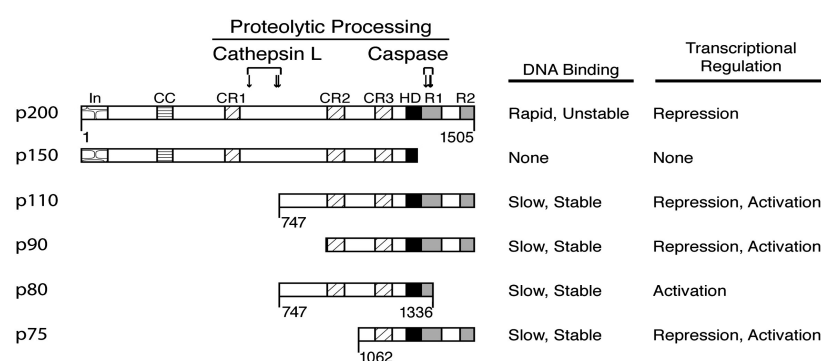


Figure 1.14 CDP isoforms. Several CDP isoforms have been described so far. With the exception of p75 that is generated by an transcriptional initiation site within intron 20, all proteins derive from the full length protein (p200). The protein domains are abbreviated as follows: Inhibitory domain (In), coiled coil domain (CC), Cut Repeat (CR), Homeodomain (HD), Repressing domain (R) (Sansregret and Nepveu, 2008).

Moreover, the DNA binding activity of CDP was shown to be regulated by numerous

post-translational modifications including phosphorylation and acetylation events, both of which inhibit the DNA binding and transcriptional activity of the protein (Sansregret and Nepveu, 2008).

CDP exerts a double influence on MMTV driven mammary carcinogenesis. On the one hand, the protein possesses a pivotal role in the differentiation process of the mammary gland. Studies revealed that CDP negatively regulates the transcription of multiple genes that are expressed in the differentiated mammary glands (Zhu *et al.*, 2004). Accordingly, its DNA binding activity has been shown to be high in the virgin mammary gland, whereas during development it continuously declined (Liu *et al.*, 1999; Zhu *et al.*, 2000).

On the other hand, CDP has been identified as repressor of MMTV transcription (Zhu *et al.*, 2000; Zhu and Dudley, 2002). Substitution mutations elevated both basal and glucocorticoid-induced reporter gene expression driven by the MMTV LTR. Furthermore, an additive effect was observed when combinations thereof were tested, suggesting that the binding sites contribute independently to the negative regulation (Zhu and Dudley, 2002). Additionally, mice infected with proviruses containing different CDP binding site mutations exhibited an increased expression of MMTV in the virgin mammary gland and suffered from an accelerated appearance of breast tumours (Zhu *et al.*, 2004).

2. Aim of the project

Almost twenty years ago, the CAAT- and the TATA-box as well as the transcriptional initiation site of the P2 promoter, which is located in the central part of the LTR, were determined by Günzburg and colleagues (Günzburg *et al.*, 1993). The activity of the promoter was shown to be enhanced upon glucocorticoid treatment and it was proposed to give rise to Sag mRNA in B-lymphocytes. However, difficulties in determining its putative transcriptional product complicated the attempts to obtain definitive proof.

The aim of the project was to investigate in more detail the regulation of the MMTV P2 promoter and ultimately link its activity to a transcriptional product.

Thus, in a first approach, we wanted to analyse possible regulatory elements within the LTR, since the influence of several regions on the major P1 promoter is controversial in the literature. We hypothesized that these sequences could possibly modulate the activity of the P2 promoter as well. Moreover, the transcriptional regulator CCAAT displacement protein (CDP) has been shown to bind to the negative regulatory element (NRE) of the LTR (Zhu and Dudley, 2002). Three binding sites are located in close vicinity to the P2 promoter elements. By mutating those CDP binding sites, we wanted to analyse whether the transcriptional activity of the P2 promoter is repressed by CDP.

In general, transcriptional active regions are free of nucleosomes to facilitate the binding of the transcriptional machinery (Segal *et al.*, 2006; Rando and Ahmad, 2007). The major P1 promoter is known to undergo a characteristic remodelling process upon hormonal stimulation (Richard-Foy and Hager, 1987; Bresnick *et al.*, 1992; Vicent *et al.*, 2008). Therefore, in a different approach, we wanted to investigate the influence of the nucleosome positioning on the activity of the P2 promoter, which has also been described as hormone inducible (Günzburg *et al.*, 1993). We asked whether the P2 promoter could be regulated similarly. Thus, we wanted to analyse the positioning of the nucleosome D that is covering the P2 promoter region both in the presence and in the absence of glucocorticoid hormones.

3. Materials and methods

3.1 Tissue culture

3.1.1 Cell lines

CrFK cells (CCL-94™), NMuMG cells (CRL-1636™), NIH3T3 cells (CRL-1658™) and A20 cells (TIB-208™) were obtained from the American Type Culture Collection (ATCC). HiB5 cells (rat neuronal progenitor cells; Renfranz *et al.*, 1991) and GR cells (mouse mammary carcinoma cells; Ringold *et al.*, 1975) were kindly provided by the cell bank of the Institute of Virology.

Two MMTV producer cell clones were established by stably transfecting CrFK cells and NMuMG cells with pMtv-2pAneoES (Rungaldier *et al.*, 2005), thereby generating C2AnE15 cells and N2AnE5 cells, respectively. The 5' LTR of MMTV was introduced stably into CrFK cells and NMuMG cells giving rise to populations of CK5'LTR cells and NM5'LTR cells, respectively.

3.1.2 Media and solutions

Dulbecco's Modified Eagle Medium (DMEM) containing 4,500 mg/l glucose and GlutaMAX™-I (Gibco BRL, Invitrogen) was supplemented with 10% (v/v) Foetal Calf Serum (FCS) (Gibco BRL, Invitrogen). The FCS was heat inactivated for 30 min at 56°C before use. This medium was used for monolayer cultures of CrFK, NIH3T3, GR and NMuMG cells. Stably transfected cell clones of CrFK cell and NMuMG cells were maintained in medium containing in addition 400 µg/ml and 800 µg/ml Geneticin (G418) (Gibco BRL, Invitrogen), respectively.

HiB5 cells were cultured in DMEM containing 4,500 mg/l glucose and GlutaMAX™-I (Gibco BRL, Invitrogen). The medium was supplemented with 0.11 mg/ml sodium pyruvate, 3.7 mg/ml sodium bicarbonat and 10 % (v/v) heat inactivated FCS.

RPMI 1640 medium containing 2 mM L-glutamine (Gibco BRL, Invitrogen) was supplemented with 1.5 g/l sodium bicarbonate, 4.5 g/l glucose, 10 mM HEPES, 1.0 mM sodium pyruvate, 0.05 mM 2-mercaptoethanol and 10 % (v/v) heat inactivated FCS prior to use. Suspension cultures of A20 cells were grown in this medium.

The buffer used to wash the cells was phosphate buffered saline (PBS). The 10x stock contains 14 mM KH₂PO₄, 27 mM Na₂HPO₄ and 1370 mM NaCl dissolved in H₂O_{dd}. The buffer was autoclaved before use.

The cells were detached from the culture flasks using a solution of 0.5% trypsin and 0.53 mM EDTA (TE) (Gibco BRL, Invitrogen).

The synthetic glucocorticoid dexamethasone (1 mM diluted in 96 % ethanol) was stored at -20°C and added to the cells in a final concentration of 1 µM.

3.1.3 Cell culture

The cells were cultured in flasks (Corning) with the corresponding medium supplemented with FCS and antibiotics at 37°C (33°C for HiB5 cells) in a humidified atmosphere containing 5 % CO₂. Before reaching complete confluence, the cells were washed with 1x PBS and incubated with a minimum amount of TE (see 3.1.2) at 37°C (33°C) until they detached from the bottom of the flask and separated from each other. By adding normal growth medium the enzymatic reaction was stopped, a defined volume of the cells was transferred into new flasks and supplemented with an appropriate volume of fresh medium.

A20 cells were passaged by transferring the suspension into falcon tubes, followed by a centrifugation step at 1000 rpm for 5 min. After removing the supernatant containing the used media, the cell pellet was resuspended with fresh medium and a defined volume was transferred to a new flask at a suitable dilution.

3.1.4 Cell counting

After the cells were trypsinized and detached from the flask surface, they were taken up in an appropriate volume of fresh medium. 100 µl of the cell suspension were loaded onto a haemocytometer (Bürker-Türk counting chamber). The cells were counted under a light microscope and the cell number per ml was calculated.

3.1.5 Freezing and thawing of cells

Exponentially growing cells were washed in 1x PBS, trypsinized, taken up in fresh medium and centrifuged at 1000 rpm for 5 minutes at 4°C. The supernatant was discarded and the cells were resuspended in freezing medium containing 75% DMEM, 20% FCS and 5% DMSO, transferred into cryo tubes (Nunc) and cooled down slowly in a freezing container (Nalgene). The cells were stored at -80°C or transferred to liquid nitrogen for long term storage. The cell density of the cryo-stocks was 5 x 10⁶ cells/ml.

The recovery of the frozen cells was done by quickly warming the cells to 37°C and transferring them to 10 ml of fresh medium. After a centrifugation step at 1000 rpm for 5 minutes, the pellet was resuspended in fresh medium and added to a new flask.

3.1.6 Synchronization of cells

In order to enrich the cells in the G₀/G₁ phase of the cell cycle, the cells were synchronized by serum starvation. CK5'LTR cells and NM5'LTR cells were plated in 175 cm² flasks in DMEM supplemented with 10 % FCS and antibiotics. After reaching 30 % to 40 % of cell density, the medium was replaced with DMEM containing 0.5 % FCS. Subsequently, the cells were incubated at 37°C for 48 hours to allow the entire population to respond to the serum withdrawal. By adding normal growth medium (DMEM supplemented with 10 % FCS) the cells were stimulated to reenter the cell cycle and to proceed simultaneously into the G₁ phase.

3.1.7 Fluorescence activated cell sorting (FACS) analysis

Fluorescence activated cell sorting allows the analysis and sorting of cells on the basis of size, surface attributes and expression of specific genes.

3.1.7.1 Determination of eGFP expression

In order to determine the transfection efficiency or the infectivity of different proviral constructs, FACS analysis was applied to analyse the enhanced green fluorescence protein marker gene (eGFP) expression. For this purpose, the cells were trypsinized and resuspended in growth medium. An aliquot of the cell suspension was transferred to a microcentrifuge tube and the cells pelleted by centrifugation at 1000 rpm for 5 minutes. The supernatant was discarded and the cells resuspended in 1 ml 4 % formaldehyde (in PBS) and mixed by vortexing. After incubation for 20 minutes, the cells were centrifuged as before, washed twice with 1 ml PBS and finally resuspended in 800 µl PBS. Just before FACS analysis, the cells were filtered through a nylon gauze to exclude any cell clumps. The fluorochromes were stimulated at 488 nm and 10,000 events (for transfections) or 100,000 events (for infections) per sample were analysed on a FACScalibur (Becton Dickinson) flow cytometer. The results were depicted using the Cell Quest Pro software (Becton Dickinson) giving the proportion of eGFP expressing cells and the extend of the eGFP expression in terms of mean fluorescence intensity (MFI).

3.1.7.2 Probidium iodide staining

The quality of synchronization was monitored by FACS analysis of cells stained with probidium iodide (PI). PI intercalates into the major groove of double-stranded DNA and produces upon excitation at 488 nm an highly fluorescent adduct which can be measured at 600 nm.

For the analysis, the trypsinized cells (see 3.1.3) were washed twice with ice-cold 1x

PBS. After the second centrifugation step, the pellet was resuspended in 950 µl of hypertonic lysis buffer (0.1 % sodium citrate, 0.2 % Triton-X 100, 20 µg/ml RNase A) and incubated at 37°C for 30 minutes. Then, 5 µl propidium iodide of a 1 mg/ml stock solution were added and the cells were incubated at RT for additional 30 minutes in the dark. Subsequently, the cells were filtered through a nylon gauze to exclude any cell clumps and analysed using a flow cytometer (FACScalibur, Becton Dickson).

A histogram plot of DNA content against cell numbers allowed the determination of the cell cycle phase. Cells in the G0/G1 phase of the cell cycle are thought to possess a diploid chromosome (DNA content: 2n), whereas cells in G2/M phase contain the doubled amount of DNA (4n). During the S-phase (phase of DNA synthesis), the DNA content ranges between 2n and 4n.

3.1.8 Transfection of eukaryotic cells

3.1.8.1 Calciumphosphate co-precipitation

24 hours prior to transfection, 2.5×10^5 cells were seeded into the wells of a 6-well plate and supplemented with 3 ml growth medium. On the day of transfection (70 % – 80 % cell density), per well 2 µg of the plasmid of interest were diluted in 120 µl H₂O_{dd} using polystyrol tubes (Greiner). (For the determination of the transfection efficiency, 0.5 µg of the plasmid pEGFPc1 (encoding the enhanced green fluorescence protein) were added in parallel.)

Subsequently, 120 µl of buffer A (0.1 M Hepes, pH 7.0; 0.5 M CaCl₂; sterile filtrated) were added. The solution was mixed by vortexing and incubated for 10 minutes at RT. Then, 240 µl of buffer B (0.05 M Hepes, pH 7.0; 0.28 M NaCl; 1.5 mM NaPO₄, pH to 7.0; sterile filtrated) were added dropwise while vortexing. The whole mixture was incubated for additional 15 minutes at RT. In the mean time, the medium of the cells was replaced with 2 ml fresh growth medium. Then, the transfection mix was equally distributed over the well. The plates were incubated at 37°C in a humidified atmosphere containing 5 % CO₂. Eight hours later, the cells were washed with 1x PBS to remove any remaining precipitates and covered with 3 ml fresh growth medium.

3.1.8.2 Lipofectamine

The day before transfection, the cells were plated in 6-well plates in a dilution that would allow a 90-95 % cell density 24 hours later. Per well, 4 µg DNA were diluted in 250 µl DMEM without FCS. Similarly, 10 µl of Lipofectamine™ 2000 (Invitrogen) were diluted in 250 µl DMEM without FCS, mixed gently and incubated for 5 minutes at RT. Then, the diluted DNA

was combined with the Lipofectamine™ 2000 mix, the solution was mixed gently and incubated for 20 minutes at RT to allow the formation of DNA-Lipofectamine™ 2000 complexes. Next, the medium of the cells was removed and replaced with 2 ml growth medium. 500 µl of the DNA-Lipofectamine™ complexes were added to each well and after gently mixing the solutions, the plates were incubated at 37°C in a CO₂ incubator. The medium was replaced with fresh growth medium after 6 to 8 hours.

3.2 Bacterial culture

3.2.1 Bacterial strains

The bacterial strain used for the amplification of plasmids was *E. coli* MAX Efficiency® DH5alpha™ (Invitrogen). For applying electroporations, we transformed ElectroMAX™ DH10B™ cells (Invitrogen). JC 5176 bacteria (kindly provided by Dr. Françoise Rouault) were used for homologous recombinations. OneShot® TOP10 chemically competent *E. coli* (Invitrogen) were transformed after TA-cloning.

3.2.2 Media and antibiotics

Luria-Bertani (LB) medium contained 10 g/l tryptone, 5 g/l yeast extract and 10 g/l NaCl. The pH was adjusted to 7.0. The medium was autoclaved before use and antibiotics were added.

LB low salt medium was made by combining 10 g/l tryptone, 5 g/l yeast extract and 5 g/l NaCl. The pH was adjusted to 7.5. LB low salt medium was used in combination with the antibiotic Zeocin.

SOC media (0.5 % yeast extract, 2 % select peptone, 10 mM NaCl, 25 mM KCl, 10 mM MgCl₂, 20 mM Glucose) was used for cell recovery upon transformation.

The solid bacteria medium was obtained by addition of 1.5% Bacto-agar to liquid media (LB) before autoclaving. After cooling down to 50°C, antibiotics were added to the final concentration and the solution was poured into Petri dishes.

The antibiotics used were Ampicillin sodium salt (Sigma), Kanamycin (Sigma), and Zeocin (Invitrogen). The stock solutions (50 mg/ml for Ampicillin and Kanamycin; 100 mg/ml for Zeocin) were kept at -20°C. The final concentrations of the antibiotics were 100 µg/ml, 50 µg/ml and 25 µg/ml, respectively.

3.2.3 Transformation of bacteria

3.2.3.1 Heat shock

An aliquot of competent cells (50 μ l) was thawed on ice for approximately 10 minutes. Then, the DNA was added (10 μ l for a ligation or 1 μ l for retransformation) to the bacteria. The suspension was incubated on ice for further 15 minutes. A heat shock was then applied to the cells at 42°C for 45 seconds. The cells were immediately chilled on ice for 5 minutes, upon which 400 μ l LB (or alternatively SOC medium) were added. The solution was incubated at 37°C for 30 to 60 minutes on an orbital shaker to allow the cells to recover and to induce the expression of the resistance gene. Aliquots thereof were plated in Petri-dishes containing the appropriate medium and antibiotics. After overnight incubation at 37°C, the plates were checked for colony formation.

3.2.3.2 Electroporation

Prior to the electroporation, the DNA (20 μ l of a ligation reaction) was precipitated once with butanol by adding 30 μ l of H_2O_{dd} and 500 μ l butanol and centrifuging the sample for 15 minutes at 13,000 rpm. The supernatant was decanted and the dried pellet resuspended in 30 μ l of sterile H_2O_{dd} . Thereof, 15 μ l were added to an aliquot (25 μ l) of electrocompetent bacteria. The suspension was transferred into a pre-chilled electroporation cuvette (1 mm electrode gap, Biozym). Immediately after applying the electric voltage (constant set at 1.8 kV, 200 Ω , 25 μ F; BioRad gene pulser II), the bacteria were taken up in 500 μ l of prewarmed SOC-media and incubated for one hour at 37°C while shaking. Aliquots were plated in Petri-dishes containing the appropriate selection medium and incubated overnight at 37°C to allow colony formation.

3.3 DNA methods

3.3.1 Preparation of plasmid DNA from bacteria

3.3.1.1 Small scale preparation - 'mini prep'

From a 3 ml bacterial culture grown overnight, 1.5 ml were harvested by centrifugation at 6,000 rpm for 5 min at room temperature. The cell pellet was resuspended in 300 μ l buffer P1 (50 mM glucose; 25 mM Tris-HCl, pH 8.0; 10 mM EDTA; 10% (v/v) RNase A). To lyse the cells, 300 μ l buffer P2 (0.2 N NaOH; 1 % (w/v) SDS) were added, the suspension was mixed gently and incubated for maximal 5 minutes at room temperature. The precipitation

of proteins and chromosomal DNA was done by adding 300 µl of buffer P3 (3 M potassium acetate; 5 M glacial acetic acid). The solution was gently mixed and the resulting precipitate separated by centrifugation at 13,000 rpm for 20 minutes at 4°C. The supernatant containing the plasmid DNA was transferred to a new tube and the DNA was precipitated by the addition of 650 µl isopropanol. After a centrifugation step, the pellet was washed in 1 ml of 70% ethanol. Finally, the DNA pellet was air-dried and resuspended in an appropriate volume H₂O_{dd}. The plasmid quality was suitable for restriction digests; for sequencing reactions, the resuspended DNA was precipitated with ethanol for a second time, followed by an additional washing step with 70% ethanol. Subsequently, the DNA was recovered in H₂O_{dd}.

3.3.1.2 Large scale preparation of plasmid DNA - 'maxi prep'

Larger amounts of pure plasmid DNA (up to 1.5 µg) were obtained by growing 200 ml of a bacterial culture overnight at 37°C. The bacterial cells were harvested by centrifugation at 6,000 rpm for 5 minutes at 4°C in an Avanti J-25 centrifuge (Beckman). After decanting the supernatant, the pellet was resuspended in 10 ml of buffer P1 (50 mM glucose; 25 mM Tris-HCl, pH 8.0; 10 mM EDTA; 10% (v/v) RNase A; Qiagen). Then, 10 ml of buffer P2 (0.2 N NaOH; 1 % (w/v) SDS; Qiagen) were added, the solution mixed vigorously by inverting the tube several times and incubated for maximal 5 minutes at RT. Subsequently, 10 ml of ice-cold buffer P3 (3 M potassium acetate; 5 M glacial acetic acid; Qiagen) were added. The samples were incubated on ice for 20 minutes after mixing thoroughly by inversion, followed by a 30 minutes centrifugation step at 10,000 rpm at 4°C. The supernatant was transferred into a new tube and centrifuged for an additional 30 minutes at 10,000 rpm at 4°C. In the mean time, a Qiagen tip 500 column per sample was equilibrated by applying 10 ml of equilibration buffer (QBT, Qiagen) onto the membrane. The supernatant was pipetted onto the column and allowed to pass through by gravitation. Subsequently, the column was washed twice with wash buffer (QC, Qiagen). The column was placed into a clean glass tube and the DNA was eluted by applying 15 ml elution buffer (QF, Qiagen) onto the membrane. 15 ml of isopropanol were immediately added to the flow-through and the solution was centrifuged at 16,000 rpm for 15 min at 4°C. The DNA was washed once with 70 % ethanol and after an additional centrifugation step, the pellet was dried in a vacuum dryer (Oligo prep OP120 concentrator, Savant) and resuspended in 400 µl TE-buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA).

3.3.2 Preparation of genomic DNA from cells

Generally, the Qiagen DNeasy Tissue kit involving a silica-gel membrane for DNA binding was used according to the manufacturer's guidelines.

Approximately 5×10^6 cells were harvested by trypsinization and washed once with PBS. The supernatant was decanted and the pellet resuspended in 200 μ l PBS. Then, 20 μ l proteinase K (Qiagen) and 200 μ l of buffer AL (Qiagen) were added to each sample. The solution was vortexed vigorously and incubated at 70°C for 10 minutes. Afterwards, the solution was mixed with 200 μ l ethanol and pipetted into a DNeasy Mini spin column placed in a 2 ml collection tube. After a centrifugation at 8,000 rpm for one minute, the flow-through was discarded. To wash the DNeasy membrane, 500 μ l AW1 buffer (Qiagen) were applied to the column, followed by a centrifugation at 8,000 rpm for one minute. As before, the flow-through was discarded before 500 μ l of buffer AW2 (Qiagen) were loaded onto the membrane. Each sample was centrifuged at 14,000 rpm for 3 minutes. After discarding the flow-through, the membranes were dried by an additional centrifugation step at 8,000 rpm for one minute. Subsequently, the spin column was placed onto a clean 1.5 ml Eppendorf tube and 200 μ l AE buffer (Qiagen) were pipetted directly onto the membrane. The samples were incubated for one minute at RT. Finally, the DNA was eluted by centrifugating the columns at 8,000 rpm for one minute.

3.3.3 DNA separation on agarose gels

The DNA diluted in H_2O_{dd} was mixed with 1/5 of the sample volume 5 x loading buffer (30 % glycerol; 0,25 % bromphenolblue; 0.25 % xylene cyanol; 0.25 % Orange G; 10 mM EDTA) and separated according to size via agarose gel electrophoresis. Depending on the expected fragment size of the DNA of interest, the percentage of agarose dissolved in 1x TAE buffer (40 mM Tris-base; 40 mM acetic acid; 1 mM EDTA) was between 0.7 and 2.0 % (w/v). Gels were run at 80 to 120 Volt (power pac, Biorad) in 1x TAE buffer in tanks. The DNA was visualized after ethidium bromide staining (5 μ g/ml in 1 x TAE buffer) under UV (ultraviolet) light.

For gels with an agarose content higher than 1.2 %, the ethidium bromide was added to the dissolved agarose before pouring the solution into the gel chamber. These gels were analysed immediately after electrophoresis under the UV light.

The intensity of the DNA band was compared either to a 100 bp molecular size marker (Fermentas) or a 1 kb DNA ladder (Invitrogen) in order to estimate the DNA size and quantity.

3.3.4 Plasmids

The plasmids used in this work were the pCR®4-TOPO (Invitrogen), which was supplied linearised with single 3' thymidine overhangs and a topoisomerase covalently bound

to the vector. Due to the template-independent terminal transferase activity of the Taq polymerase, PCR inserts are efficiently ligated into this vector. For longer PCR amplification products, the vector pCR®-XL-TOPO® (Invitrogen) was used.

The plasmid pGR102 (Salmons *et al.*, 1985) was used to amplify Mtv-2 and Mtv-8 LTR sequences.

We utilized the plasmid pcDNA3.1/Zeo (Invitrogen) for the cloning and expression of the different CDP isoforms.

The plasmid pLTRpAneo (Rungaldier *et al.*, 2005) was used to study the positioning of the nucleosome D.

For reporter gene assays, we used the promoterless plasmid pZluc (Brasier *et al.*, 1989; Maxwell *et al.*, 1989).

3.3.5 Plasmid construction

3.3.5.1 Cloning of luciferase expression constructs

To analyse the regulation of the LTR promoters, different fragments of the MMTV LTR were generated by PCRs using MMTV-specific oligonucleotides carrying a BamHI (or HindIII) restriction site at their respective 5' end. After digestion with the relevant enzymes, the fragments were gel purified and inserted upstream of the luciferase gene in the promoterless expression plasmid pZluc. The plasmid pGR102Cla (that was obtained after ClaI-digestion and self-ligation of the plasmid pGR102) was used as template for Mtv-2 LTR sequences. The plasmid pGR102AflII, which have been generated by religation of the previously digested pGR102EGFPSac (Indik *et al.*, 2005) with AflII, was used to amplify Mtv-8 LTR sequences.

3.3.5.2 Cloning of the CDP binding site mutations into the MMTV LTR

Two PCRs were performed with primer pairs m596f - m693r and m693f - HindIIIrev in order to introduce the CDP binding site mutations and the mutation of the P3 TATA-box into the LTR of MMTV. The plasmid pZ1303 containing the Mtv-2 LTR upstream of the luciferase gene was used as template. Subsequently, the two fragments were used as starting material of a third PCR using m596f together with HindIIIrev to obtain a single fragment spanning the whole region. The fragment was subcloned into the TOPO-XL-Vector, and after amplification, digested with StuI and HindIII. In the following step, the plasmid pZ1303 was digested with the same enzymes, dephosphorylated and ligated with the above fragment, giving rise to pZP2⁺P1⁺dCDP⁻.

To be able to analyse the P2 promoter activity in the absence of the major P1

promoter, the LTR harbouring the CDP binding site mutations, yet lacking the promoter elements of P1, was recovered by a PCR using the oligonucleotides Z1 – X1156mtv2. The amplified fragment was digested with BamHI and XhoI and ligated with the vector pZgzP2+delR harbouring mutations of the P3 promoter and a short deletion encompassing the P1 promoter and the Inr element. The vector was digested with the same enzymes and dephosphorylated before use. As a result, the plasmid pZgzP2⁺delR_dCDP⁻ was obtained.

3.3.5.3 Cloning of a provirus containing the CDP binding site mutations

The plasmid pZP2⁺P1⁺dCDP⁻ (see 3.3.5.2) was double-digested with AarI and SacI. The fragment of interest (567 bps) was gel-purified and ligated with vector pLTRpAneo (Rungaldier *et al.*, 2005) that had been digested with the same enzymes and dephosphorylated before use. The resulting plasmid was digested with BsaI giving rise to a 7184 bp fragment, which was subsequently introduced into the linearized and dephosphorylated pGR102EGFPSac (Indik *et al.*, 2005) by homologous recombination. The resulting plasmid was named pMtv2pAneoESdCDP_m.

3.3.5.4 Cloning of CDP/p110 isoform

The transcriptionally active isoform of CDP was obtained by a long-template PCR using the Expand high fidelity PCR system (Roche) in combination with the primer pair mCDP1815for – mCDP3987rev. As template reverse transcribed mRNA from NIH3T3 cells was used. The obtained fragment (~2.200 bps) was subcloned into the plasmid pCR-XL-TOPO (Invitrogen) and amplified. To introduce restriction sites on either end of the fragment, a PCR was performed using the primer pair p110exfor – p110NotIrev. The resulting fragment was again subcloned into the plasmid pCR-XL-TOPO (Invitrogen) and amplified. After double-digestion with NheI and NotI, the fragment was inserted into the plasmid pcDNA3.1/zeo (Invitrogen) that has been digested with the same enzymes and dephosphorylated before use. The final construct was named pCMV110.

3.3.6 Homologous recombination

Under physiological conditions, homologous recombination occurs between two identical (homologous) DNA molecules during cell division. The same mechanism can be used for efficiently linking vector and fragment DNA *in vitro*, if an homologous sequence of at least 100 bp is present on either end of each molecule.

For the reaction 100 ng of the vector in combination with a 5 times excess in

fragment DNA were used. Taking into account the length of both sequences, the following calculation was applied: $100 \text{ ng} \times 5 / (\text{bps of the vector} / \text{bps of the fragment})$ gives the amount of the fragment (in ng). The appropriate amounts of both, vector and fragment were combined in an Eppendorf tube together with 4 μl of a 5 x ligase buffer containing polyethylenglycol (Roche). $\text{H}_2\text{O}_{\text{dd}}$ was added to a total volume of 20 μl . The reaction was mixed gently and incubated two hours at RT. Then, the mixture was used to transform JC 5176 bacteria (Capaldo-Kimball and Barbour, 1971).

The recombination event was examined by restriction digest of corresponding mini preps (see 3.3.1.1 and 3.3.8.1). Since JC 5176 bacteria possess only a very low potential to amplify non-bacterial DNA, an aliquot of the mini prep containing the correct recombinated plasmid was retransformed into DH5alpha bacteria in order to achieve a sufficient amount of DNA.

3.3.7 Oligonucleotide dimerization

Oligonucleotides complementary to each other were hybridized according to the following protocol: In a total volume of 50 μl , 20 μM of each oligonucleotide (diluted in 250 mM Tris-HCl, pH 7.7) were heated for 5 minutes at 95°C using an PCR machine (Eppendorf master cycler). Then, the temperature was lowered to 70°C, followed by a controlled gradually cool down to room temperature. After one hour incubation at RT, the solution was gradually cooled to 4°C. The mixture containing the hybridized product was aliquoted and stored at 4°C until further use.

3.3.8 Enzymatic reactions

3.3.8.1 Restriction digestions of DNA

Restriction endonucleases and 10 x reaction buffer were used according the manufacturer's instructions. Analytical digestions were performed in a total of 15 μl with 0.5 ng to 1.0 μg DNA; preparative digestions in a total volume of 100 μl using 2 μg to 5 μg of DNA.

3.3.8.2 Ligation of DNA fragments with T4 DNA ligase

Generally, a 20 μl ligation reaction mix contained 100 ng of digested vector DNA and a defined amount of insert DNA in 1 mM ATP, 1x T4-DNA Ligase buffer (30 mM Tris-HCl, pH 7.8; 10 mM MgCl_2 ; 10 mM DTT) and 0.5 μl T4-DNA ligase (Roche, 10 U/ μl). The amount of

insert (in ng) was calculated in relation to the length of both, vector and fragment using the following formula: $100 \text{ ng vector} \times 10 / (\text{size of vector (bps)} / \text{size of insert (bps)})$.

The reactions were incubated overnight at 16°C, before the enzyme was inactivated by incubating the reaction mix at 65°C for 10 minutes.

3.3.8.3 Phosphorylation of DNA 5' ends

Phosphorylation of 5' ends was performed using T4 polynucleotide kinase (10 U/μg DNA, Roche) in the buffer supplied with the enzyme (70 mM Tris-HCl, pH 7.6; 10 mM MgCl₂; 5 mM DTT) and 1 mM ATP. The reaction was incubated for 30 min at 37°C prior to phenol/chloroform extraction (see 3.3.9.1) and ethanol precipitation (see 3.3.9.2).

3.3.8.4 Dephosphorylation of DNA 5' ends

In order to avoid self-ligation, cleaved vectors were incubated with 1 U/pmol DNA ends of calf intestine phosphatase (CIP) (New England Biolabs) in 20 mM Tris-HCl, pH 8.0 for 45 minutes at 37°C. The enzyme was inactivated by incubation at 65°C for 45 minutes. Afterwards, the reaction mixture was purified by elution from an agarose gel (see 3.3.10).

3.3.9 Purification of DNA solutions

3.3.9.1 Phenol/Chloroform extraction

To purify DNA from enzymes or other proteins and to remove other contaminants, one volume of saturated phenol was added to the DNA. The solution was vortexed at the highest setting for 5 to 10 seconds. Subsequently, the phases were separated by centrifugation at 13,000 rpm for 5 minutes. Then, the supernatant was transferred into a new tube and reextracted with one volume of phenol/chloroform/isoamyl-alcohol (Invitrogen). Following an additional centrifugation step, the upper aqueous phase was transferred into a new tube and precipitated with ethanol (see 3.3.9.2).

3.3.9.2 Ethanol precipitation

To purify or concentrate DNA, 10% (v/v) 3 M sodium acetate, pH 5.6 and 2 volumes of absolute ethanol were added to the DNA solution. The mixture was incubated at -80°C for at least 20 minutes and subsequently centrifuged at 14,000 rpm for 15 minutes at 4°C. After discarding the supernatant, 500 μl of 70 % ethanol were added to the pellet. The sample was

incubated for at least 5 minutes at -20°C. After a centrifugation at 14,000 rpm for 5 minutes at 4°C the pellet was dried in a speed vac (Oligo prep OP120, Savant) and taken up in an appropriate volume of H₂O_{dd} or TE-buffer.

3.3.10 Elution of DNA from an agarose gel

Elution of DNA from agarose gels was performed with the Wizard® SV Gel and PCR Clean-Up system kit (Promega). Briefly, the excised DNA band was placed into an 1.5 ml Eppendorf tube and 1 µl membrane binding solution (Promega) per mg of gel slice was added. The reaction mix was vortexed and incubated at 50°C to 65°C until the gel slice was dissolved completely. The solution was transferred to a SV minicolumn (Promega) placed in a 2 ml collection tube and incubated at RT for one minute. After centrifugation at 13,000 rpm for one minute the flow-through was discarded. Then, the column was washed once with 700 µl membrane wash solution (Promega), followed by a second washing step with 500 µl of membrane wash solution. Each time, the samples were centrifugated at 13,000 rpm for one minute and the flow-through was discarded. Before eluting the DNA, any residual ethanol was removed by an additional centrifugation at 13,000 rpm for 3 minutes. The DNA was recovered by applying 50 µl H₂O_{dd} onto the membrane and collected after a short incubation by centrifugation at 13,000 rpm for one minute.

3.3.11 DNA quantification

The concentration of DNA solutions was determined by UV spectroscopy at 260 nm in a spectrophotometer (Gene Quant Pro, Amersham). OD₂₆₀ of 1 corresponds to 50 µg/ml of DNA. Dilutions of 1/40 of the DNA solutions were prepared in H₂O_{dd} and measured against a blank containing H₂O_{dd}.

Alternatively, dilutions of vector or fragment DNA were loaded onto an agarose gel and the DNA concentrations was estimated by comparison with the ladder bands 1636 and 506/517 of the 1 kb DNA marker (Invitrogen, Cat.-N° 1561-016) corresponding to 50 ng and 30 ng DNA, respectively.

3.3.12 Amplification of DNA fragments by PCR (polymerase chain reaction)

A standard PCR of 50 µl contained 10 ng of template DNA, 0.2 µg of each primer, 1x Taq-polymerase buffer (supplied with the enzyme), 0.2 mM dNTPs (Invitrogen) and 0.5 U Taq-polymerase (Invitrogen). The reaction was prepared on ice and transferred to a Perkin Elmer GeneAmp PCR System 2400. Initially, the DNA was denaturated at 94°C and amplified in 25 to

35 cycles of denaturing at 94°C for 30 seconds, 30 to 90 seconds of annealing temperature (based on the melting temperature of the primers) and 30 to 90 seconds of extension temperature at 72°C. A final step of extension temperature was performed for 7 minutes at 72°C. Afterwards, the reactions were cooled down to and stored at 4°C until further analysis.

Alternatively, 1.0 U Platinum Taq-polymerase (Invitrogen) was used together with 1 x Platinum Taq PCR buffer (Invitrogen), 1.5 mM MgCl₂ (Invitrogen), 0.2 mM dNTPs, 0.2 µg of each primer and 2.0 µl template in a total volume of 50 µl. Initially, the enzyme was activated by an incubation at 94°C for 2 minutes. Then, the DNA was denatured at 94°C for 30 seconds, followed by an annealing step at 55°C for 30 seconds and an elongation phase at 72°C for 30 seconds. The last three steps were repeated 25 times, followed by final extension step at 72°C for 7 minutes. The reactions were cooled down and stored at 4°C.

For long template PCRs (up to 5 kb) the Expand high fidelity PCR system (Roche) was used according to the manufacturer's instructions. Briefly, mix 1 (containing 200 µM of each dNTP, 300 nM of each primer and a variable amount of template DNA in a final volume of 25 µl) was combined with mix 2 (containing 19.25 µl of H₂O_{dd}, 1x Expand high fidelity buffer and 2.6 U Expand high fidelity enzyme mix in a final volume of 25 µl). After an initial hot start at 94°C for 2 minutes to activate the polymerase enzyme, 35 cycles were carried out, each containing a denaturation step at 94°C for 15 seconds, followed by an annealing step at 55°C for 30 seconds and a subsequent elongation phase at 68°C for 5 minutes. After 15 cycles, 5 additional seconds of elongation time were successively added to each remaining cycle. The final elongation step was performed at 68°C for 7 minutes. Subsequently, the samples were cooled down to 4°C.

3.3.12.1 DIG PCR

The PCR DIG probe synthesis kit (Roche) was used for the generation of digoxigenin (DIG)-dUTP labelled probes. The reaction was performed according to the manufacturer's instructions, except for the PCR conditions which were adjusted to the platinum Taq-polymerase (see 3.3.12).

The synthesis of the labelled probe was checked by running an aliquot of the PCR on an agarose gel. The DIG-labelled probe was expected to migrate slower than the unlabelled control PCR product.

3.3.12.2 Real time PCR (qRT-PCR)

The quantitative real-time PCRs were performed with the Power SYBR® Green PCR mastermix (Applied Biosystems) according to the manufacturer's instructions.

Each reaction within a 96-well plate contained 2x Power Sybr Green mastermix (Applied Biosystems), 150 nM of each primer and 5 µl template in a total volume of 25 µl. The polymerase was activated during a 10 minutes incubation period at 95°C, followed by 40 cycles of a denaturation phase at 95°C for 15 seconds and a subsequent 1 minute annealing/extension phase at 60°C.

The amplification steps as well as the data analysis of the generated dissociation curves were performed on a 7500 Real-Time PCR System (Applied Biosystems).

3.3.13 Sequencing reactions

The DNA samples were purified by ethanol precipitation (see 3.3.9.2), the DNA concentration was determined and 10 µl aliquots were prepared containing 0.8 µg plasmid DNA. If other primers than the standard M13 primer was used, 1 pmol of the selected primer was added to the sequencing tube. For the analysis, the samples were sent to the Microsynth laboratory (Microsynth AG, Balgach, Suisse).

3.3.14 List of primers

3.3.14.1 Primers for CDP cloning and binding site mutations

Label	Sequence (orientation 5' to 3')
m596f	CCTCAGGCCTAGAAGTAAAAAAGGGAAAAAAGAGTGTGTTTGTCGACGTC GGAGACAGGTGGTGGCAACC Sense primer for introducing point mutations of the distal CDP binding sites into the MMTV LTR together with StuI and AvrII restriction sites
m693f	CCATATACAGGAAGATATGACCTAGGTTGGGATAGGTGGGTCACTATCAAC GGCTAGCAAGTGTCTACAGATCCC Sense primer for introducing point mutation of the distal CDP binding sites into the MMTV LTR together with a NcoI restriction site
m693r	GGGATCTGTAGAACAATTGCTAGCCGTTGATAGTGACCCACCTATCCCAAC CTAGGTCATATCTTCCTGTATATGG Antisense primer for introducing point mutations of distal CDP binding sites together with a NcoI restriction
Hindrev	GGAGGAAGCTTGACCGGATCTGCGG Antisense primer for introducing a HindIII restriction site 130bp downstream of the MMTV P1 promoter
X1156Mtv-2rev	TTTTTCTCGAGTTTACATAAGCATTACATAAGACTTGG Antisense primer for introducing a XhoI restriction site upstream of the P1 promoter

mCDP13for	CATGTTGTGCGTAGCCGGAGCCAAGT Sense primer for cloning of CDP (mouse origin)
mCDP77for	AACAGGCAAGATGAGAGCGAACAGT Sense primer for cloning of CDP (mouse origin)
mCDP231for	AAGCAGAGGCAGCCTTCTTGACTGTG Sense primer for cloning of CDP (mouse origin)
mCDP553for	GAGAAGGAGAGAAAGCTGCAAGAGAC Sense primer for cloning of CDP (mouse origin)
mCDP778for	ACAGCTCTCATCGGCCAACCACTCTC Sense primer for cloning of CDP (mouse origin)
AY1192for	ACGCTGCGCATCTCCAACAGTGACCT Sense primer for cloning of CDP (mouse origin)
AY1922rev	AGTCTGTTCAGGCTCTGGCCTGGATT Antisense primer for cloning of CDP (mouse origin)
mCDP1815for	TCCGTGGCAAGGAGCCATTCCACAAG Sense primer for cloning the C-terminal part of CDP (mouse origin)
mCDP2620for	CTCAGCATCCGCAGAGTACTGGAA Sense primer for cloning the C-terminal part of CDP (mouse origin)
mCDP3987rev	CGGCTACTTCGAGCTGAAGGTGAGTC Antisense primer for cloning the C-terminal part of CDP (mouse origin)
mCDP4522rev	CAGGCGGTGGATGATGCTGTTCAAGT Antisense primer for cloning the C-terminal part of CDP (mouse origin)
p110exfor	AAGCTAGCACCATGGTCCGTGGCAAGGAG Sense primer for recovering the p110/CDP isoform from the TOPO® vector and introducing a Nhe I restriction site
p110notIrev	AAAAACGCCGGCGAACTAGCTACTTCGAGC Antisense primer for recovering the p110/CDP isoform from the TOPO® vector and to introduce a Not I restriction site

Table 2: List of primers used for CDP cloning and binding site mutation

All primers were designed according to the sequence of the *Mus musculus* CCAAT displacement protein CDP (Cutl1) mRNA (GenBank Accession number AY037807.1).

3.3.14.2 Primers for sequencing

Target	Label	Sequence (orientation 5' to 3')
CDP	838zeof	CCACTGCTTACTGGCTTATC Sense oligo for analysing CDP/p110 isoform cloned into pcDNA3.1/zeo
	1045zeor	GGCAACTAGAAGGCACAGTC Antisense oligo for analysing CDP/p110 isoform

	321cdp110r	cloned into pcDNA3.1/zeo AGCGGATGGCATCGTCAGAG Antisense oligo for analysing CDP/p110 isoform
	1843cdp110f	cloned into pcDNA3.1/zeo GTGACAGTCAGCCTTGTGAG Sense oligo for analysing CDP/p110 isoform
	702cdp110f	cloned into pcDNA3.1/zeo CTCACTTCTCCGAAGAGAC Sense oligo for analysing CDP/p110 isoform
	1616cdp110r	cloned into pcDNA3.1/zeo GTTGTCTGTCAGCACCTCTT Antisense oligo for analysing CDP/p110 isoform
TOPO®Vector	M13 forward	cloned into pcDNA3.1/zeo GTAAAACGACGGCCAG Sense oligo for analysing inserts of pCR®4-TOPO and pCR®-XL-TOPO®
	M13 reverse	CAGGAAACAGCTATGAC Antisense oligo for analysing inserts of pCR®4- TOPO and pCR®-XL-TOPO®

Table 3: List of primers used for sequencing**3.3.14.3 Primers for analysing the MMTV LTR**

Label	proviralGR102*	<i>Mtv-2 LTR</i> **	Sequence (orientation 5' to 3')
33for	30/49 and 8603/8622		CCGAGAGTGTCTACTTA Sense oligo for the generation of the DIG- labelled probe
510for	9080/9100	510/530	GAGACGCTCAACCTCAATTGA Sense primer for nucleosome E region
538rev	9108/9087	538/517	CCTGCTCTTCAATTGAGTTGA Antisense primer for nucleosome E region
548rev	9118/9099	548/529	AGTCCTTGACCTGCTCTTC Antisense oligo for the generation of the DIG- labelled probe
552for	546/565 and 9122/9141	552/571	GGCCTCAGGCCTAGAAGTAA Sense primer for nucleosome E/D region
568rev	9138/9119	568/549	CTTCTAGGCCTGAGGCCAAT Antisense primer for nucleosome E/D region
LM602for	599/619 and 9175/9195	605/625	GAGACAGGTGGTGGCAACCAG Sense primer for nucleosome D

622rev	616/595 and 9192/9171	622/601	GTTGCCACCACCTGTCTCTCTAT Antisense primer for nucleosome D region
638for	9215/9242	645/672	CATCTACAGACCAACAGATGCCCCCTTA Sense primer for nucleosome D region
646for	9216/9235	646/665	ATCTACAGACCAACAGATGC Sense primer for nucleosome D
666rev	9236/9214	666/644	GGCATCTGTTGGTCTGTAGATGT Antisense primer for nucleosome D region
672for	9243/9263	673/693	CCATATACAGGAAGATATGAC Sense primer for nucleosome D region
691for	9261/9278	691/708	GACTTAAATTGGGATAGG Sense primer for nucleosome D region
+702	9274/9293	704/723	ATAGGTGGGTCACAATCAAC Sense primer binding immediately downstream of the P2 promoter initiation site
710rev	9280/9261	710/691	CACCTATCCCAATTTAAGTC Antisense primer for nucleosome D region
723for	9293/9313	723/743	CGGCTATAAAGTGTATACAG Sense primer for nucleosome D
LM723rev	9296/9273	726/703	GCCGTTGATTGTGACCCACCTATC Antisense primer for nucleosome D
739for	733/752 and 9309/9328	739/758	TACAGATCCCTCCCCTTTTCG Sense primer for nucleosome D/C region
762for	756/775 and 9332/9351	762/781	AAGACTCGCCAGAGCTAGAC Sense primer for nucleosome C
P2+P3mtv-2	9403/9384	833/814	ACCTGTTGTTTCATGTCGTC Antisense oligo binding downstream of the P3 promoter initiation site
LM772for	9342/9366	772/796	AGAGCTAGACCTCCTTGGTGTATGC Sense primer for nucleosome C
771mtv2	9343/9365	773/795	GAGCTAGACCTCCTTGGTGTATG Sense primer complementary to the P3 promoter initiation site
796rev	9366/9347	796/777	GCATACACCAAGGAGGTCTA Antisense primer for nucleosome C region
876rev	9437/9419	867/849	GTGCATTCCTGTTCTCTAGA Antisense primer for nucleosome C region
LM922rev	9492/9465	922/895	GTTCTATCAGTCCAGCCACTGCCCCCTTC Antisense primer for nucleosome C/B
980for	9553/9575	983/1005	GCGGTTCCCAGGGTTTAAATAAG Sense primer for nucleosome B region

3. Materials and methods

1073rev	9662/9643	1092/1073	AGCTCAGATCAGAACATTTG Antisense primer for nucleosome B region
LM1099rev	9672/9651	1102/1081	GAACACTAAGAGCTCAGATCAG Antisense primer for nucleosome B
P1STmtv2	9767/9789	1197/1219	GCAACAGTCCTAACATTCTTCTC Sense primer complementary to the initiator region of the P1 promoter
-1295mtv-2	9869/9851	1299/1281	GGATCTGCGGGGGACCCT Antisense primer complementary to the region downstream of the splice acceptor site of the MMTV LTR (SA ₁₂₇₅)
-1412	1431/1409		CTTGTGATGATAGCCAGACAAGA Antisense primer binding to the <i>gag</i> untranslated region
-1442	1443/1424		GTCCGTCCGCTCTTGTGAT Antisense primer binding to the <i>gag</i> untranslated region
-2754	2755/2730		CCTGCTAATACAGTCTTCTTGACTCC Antisense primer binding within the <i>gag</i> gene
-3466	3466/3446		GGTGGGTACTAGCGATATTCC Antisense primer binding within the <i>gag</i> gene
-5394	5394/5373		GATAATGGCAGTTTGTAGGGTA Antisense primer binding within the <i>pol</i> gene
-5535	5511/5485		CTCTTTAATTATAGGCTCCCTTCCTTG Antisense primer binding within the <i>pol</i> gene
-6652	6626/6605		GGCGTTGCTTCTTCCGCTCAG Antisense primer binding downstream of the splice acceptor of the <i>env</i> gene (SA ₆₅₅₄)
6890R	6763/6746		CTCCTCCGCTTCGGAGAT Antisense primer binding at the 3' end of the <i>pol</i> polyprotein (within the leader sequence of the <i>env</i> gene)
-7803	7803/7781		GCTGCATAGTCGTAGGCAGAAGA Antisense primer binding within the <i>env</i> gene
-7933	7933/7911		AGCGACGAATCGCTTAGCTCGAA Antisense primer binding within the <i>env</i> gene
-8464	8464/8442		GACTAAAAGCAGGGCTCCAACAC Antisense primer binding within the <i>env</i> gene
delS	9140/9116	570/546	TACTTCTAGGCCTGAGGCCAATAGTCC Antisense primer binding upstream of the CAAT-box of the P2 promoter within the U3

		region
TBP2_Z	9211/9254	GATCCTTACATCTACAGACCAACAGATGCCCCCT TACCATATACAGGA Sense oligo for cloning the minimal P2 promoter containing BamHI and HindIII restriction sites
TBP2_Zc	9254/9211	AGCTTCCTGTATATGGTAAGGGGGCATCTGTTGG TCTGTAGATGTAAG Antisense oligo for cloning the minimal P2 promoter containing BamHI and HindIII restriction sites
TBP3_Z	9281/9310	GATCCGGTCACAATCAACGGCTATAAAGTGTTAT ACAGATCCCTCA Sense oligo for cloning the minimal P3 promoter containing BamHI and HindIII restriction sites
TBP3_Zc	9310/9281	AGCTTGAGGGATCTGTATAACACTTTATAGCCGT TGATGTGACCG Antisense oligo for cloning the minimal P3 promoter containing BamHI and HindIII restriction sites
Z1	3/17 and 8576/8590	AAAAAAGGATCCGCGCCTGCAGCAGAA Sense oligo introducing a BamhI restriction site
Z2	303/319 and 8879/8895	AAAAAAGGATCCACAGCCAACCTTCCTCTT Sense oligo introducing a BamhI restriction site
Z3	9105/9125	AAAAAAGGATCCAGGTGCAAGGACTATTGGC Sense oligo introducing a BamhI restriction site
Z3.mtv8	529/548	AAAAAAGGATCCAGGTGCAAGGACTAAGGGC Sense oligo introducing a BamhI restriction site
Z.P2c	9251/9231	AAAAAAAGCTTCCTGTATATGGTAAGGGGGCATC Antisense oligo introducing a HindIII restriction site
HenhA2	742/767 and 9318/9333	AAAAAAAGCTTCTCCCCTTTCGTGAAAGACTCGC CAG Sense oligo introducing a BamhI restriction site
HenhA2rev	9566/9543	AAAAAAAGCTTACCCTGGGAACCGCAAGGTTG GGC Antisense oligo introducing a HindIII restriction site
d.B.mtv8	536/553	AAAAAGGATCCAAGGACTAAGGGCCTCAG Sense oligo introducing a BamhI restriction

P2mtv8.C.H	677/661	site AAAAAGCTTCCTGTATATGGTAGCGG Antisense oligo introducing a HindIII restriction site
------------	---------	---

Table 4: List of primers used for analysing the MMTV LTR

*The proviralGR102 sequence corresponds to the MMTV proviral sequences of plasmid pGR102 (Salmons *et al.*, 1985) that harbours a 5' LTR (bp 1-1327) originating from Mtv-8 and a 3' LTR (8574-9900) derived from Mtv-2 flanking *gag*, *pol* and *env* coding regions.

** The numbers shown in *italic* indicate the respective sequence in relation to the first base of the Mtv-2 LTR.

3.3.14.4 Primers used for selected methods

Methode	Label	Sequence (orientation 5' to 3')
LM-PCR	Linker PCR1	GTAATACGACTCACTATAGGGC Sense primer for LM-PCR analysis
	Linker PCR2	AGGGCTCCGCTTAAGGGAC Sense primer (nested) for LM-PCR analysis
	LinkerNuc+	GTAATACGACTCACTATAGGGCTCCGCTTAA GGGAC Sense Oligonucleotide for generation of the linker-hybrid
	LinkerNuc-	PO ₄ -GTCCCTTAAGCGGAG-NH ₂ Antisense oligonucleotide for generating the linker-hybrid
RLM-Race	GeneRacer™ 5' Primer	CGACTGGAGCACGAGGACACTGA Sense primer for the reverse transcribed RNA linker region
	GeneRacer™ 5' Nested Primer	GGACTGACATGGACTGAAGGAGTA Sense primer (nested) for the reverse transcribed RNA linker region

Table 5: List of primers used for selected methods

3.3.15 Southern blot hybridization

3.3.15.1 Separation and denaturation of DNA fragments

The integration of plasmid DNA into the genome of stably transfected cells was analysed by Southern blotting and subsequent hybridization analysis using a digoxigenin-labelled probe. Briefly, genomic DNA was isolated from the cells of interest (see 3.3.2) and 20 µg thereof were digested overnight with selective restriction enzymes. The fragmented DNA was loaded onto a 1.0 % agarose gel and separated by electrophoresis (see 3.3.3). The separation of the DNA fragmentation was monitored by ethidium bromide staining. Next, the dye was removed by washing the gel for 15 minutes in H₂O_{dd}. To achieve single-stranded molecules suitable for the subsequent hybridization, the DNA was denatured by incubating the gel two times in 10 volumes of denaturation solution (1.5 M NaCl; 0.5 M NaOH) for 15 minutes each. The pH of the gel was lowered by washing the gel twice (each time for 15 minutes) in 10 volumes of neutralization solution (1.5 M NaCl; 0.5 M Tris-HCl, pH 7.0) to ensure proper binding of the DNA to the membrane. Before transferring the DNA to the PVDF membrane (Hybond-P membrane, Amersham), the gel was equilibrated in 20x SSC buffer (3M NaCl; 300 mM sodium citrate, pH 7.0) for 10 minutes.

3.3.15.2 Southern blotting

The transfer was set up by placing one sheet of Whatman filter paper soaked with blotting buffer (20x SSC) into a plastic tray (electrophoresis chamber), followed by the gel (downside up), the membrane, two additional sheets of Whatman filter paper and several sheets of paper towels. The blot was covered by a plastic plate and a ~ 0.5 kg weight. On the following day, the blot was decomposed and the membrane UV-crosslinked in an UV transilluminator at 254 nm wavelength to immobilize the DNA. Then, the membrane was rinsed shortly with H₂O_{dd}, air dried and stored at 4°C until further use.

3.3.15.3 Hybridization

For the hybridization, the DIG High Prime DNA labelling and detection starter kit II (Roche) was used according to the manufacturer's instructions. For the detection, the membrane was soaked with 2x SSC buffer, followed by a 30 minutes incubation in an appropriate volume (10 ml/100 cm² filter) of prewarmed prehybridization solution (DIG Easy Hyb buffer, Roche) while constantly rotating (Biometra Compact Line OV4 hybridization oven). The optimal hybridization temperature was calculated according to the GC content and the percent homology of the probe to the target applying the following equation: $T_{\text{optimal}} = T_m -$

20 to 25°C, whereas $T_m = 49.82 + 0.41 (\% G + C) - (600/\text{length of hybrid in base pairs})$. Meanwhile, 10 µl of the DIG-labelled probe (see 3.3.12.1) were diluted 1:5 in H₂O_{dd} and denatured at 97°C for 5 minutes. Immediately thereafter, the probe was placed on ice. The prehybridization solution was decanted and replaced with prewarmed hybridization solution containing the denatured probe diluted 1:10,000 in an appropriate volume of DIG Easy Hyb buffer (3.5 ml/100 cm² membrane). The membrane was incubated at the optimal hybridization temperature (T_{optimal}) overnight during constant rotation.

3.3.15.4 Immunological detection

On the following day, the hybridization solution was removed and the membrane rinsed briefly with H₂O_{dd}, washed twice with 50 ml of low stringency wash buffer (2x SSC; 0.1% SDS) at RT for a total of 10 minutes, followed by two washing steps with 50 ml prewarmed high stringency buffer (0.5xSSC, 0.1% SDS) at 68°C. Then, the membrane was rinsed briefly in wash buffer (0.1 M maleic acid; 0.15 M NaCl, pH 7.5; 0.3 % (v/v) Tween 20).

For the subsequent immunological detection, an anti-digoxigenin antibody conjugated to alkaline phosphatase (AP) was used. For this purpose, the membrane was incubated for 30 minutes with 1x blocking solution (Roche; 10x stock diluted in maleic acid buffer (0.1 M maleic acid; 0.15 M NaCl, pH 7.5)). The blocking solution was removed and replaced with an appropriate volume of an anti-digoxigenin-AP antibody solution (diluted 1:10,000 in 1x blocking solution). After incubation at RT for 30 minutes, the membrane was washed twice for a total of 30 minutes with 100 ml of wash buffer. The membrane was equilibrated in 20 ml detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5) for 5 minutes at RT. Then, 1 ml of CSPD ready-to-use (Roche) containing the substrate for the chemiluminescent reaction was applied onto a plastic slide and covered immediately by the DNA side of the membrane. The membrane was incubated for 5 minutes at RT, followed by 10 minutes at 37°C in order to enhance the luminescent reaction. Finally, the membrane was exposed to an autoradiography film (Hyperfilm MP high performance, Amersham).

3.4 RNA methods

3.4.1 Extraction of total RNA

For extraction of total RNA, the RNeasy mini kit from Qiagen was used according the manufacturer's instructions. Briefly, 5×10^6 cells were harvested by trypsinization, washed once with 1x PBS and pelleted by centrifugation prior to lysis. The cells were resuspended in 600 µl RLT buffer (Qiagen). Subsequently, each sample was pipetted directly onto a

QIAshredder spin column (Qiagen) in order to homogenize the lysate and centrifuged at 13,000 rpm for 2 minutes. 600 μ l of 70 % ethanol were added, the solutions were mixed by pipetting and transferred to an RNeasy mini column (Qiagen). After a short centrifugation at 10,000 rpm for 15 seconds, the flow-through was discarded. The membrane containing spin column was washed once with 700 μ l of RW1 buffer (Qiagen), followed by two washing steps with 500 μ l of RPE buffer (Qiagen). After discarding the last flow-through of RPE buffer, the spin column was centrifugated for an additional minute to remove any traces of the buffer. The RNA was recovered by adding 50 μ l H₂O_{dd} onto the membrane and by centrifugation at 10,000 rpm for one minute. Finally, the concentration of the RNA was determined photometrically.

3.4.2 Extraction of mRNA

5 x 10⁶ cells were harvested by trypsinization and washed once in ice-cold PBS. The cells were transferred into a 1.5 ml Eppendorf tube, pelleted by centrifugation and placed immediately on ice. In the following step, each sample was resuspended in 400 μ l extraction buffer (Quickprep micro mRNA purification kit, Amersham Biosciences). Then, 800 μ l elution buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA) were added and the solutions mixed thoroughly by vortexing.

In order to prepare a cleared cellular homogenate the cell lysate was centrifuged at 13,000 rpm for one minute at RT. Similarly, for each sample, 1 ml aliquots of Oligo(dT)-cellulose (Quickprep micro mRNA purification kit, Amersham Biosciences) were centrifuged. The buffer of the Oligo(dT)-cellulose was removed and replaced by 1 ml of the cell lysates. The tube was mixed gently for 3 minutes and then centrifugated at 13,000 rpm for 10 seconds. The supernatant was discarded and the mRNA bound to the oligo(dT)-cellulose was resuspended in 1 ml high salt buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 0.5 M NaCl) and centrifuged for 10 seconds at maximum speed. This washing step was repeated four times. Then, the pellet was washed twice with 1 ml of low salt buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 0.1 M NaCl). Subsequently, the pellet was resuspended in 300 μ l low salt buffer, transferred to a MicroSpin column (Quickprep micro mRNA purification kit, Amersham Biosciences) and centrifuged for 5 seconds at maximum speed. The column was washed three times with 500 μ l low salt buffer. After the last centrifugation, the mRNA was recovered in 100 μ l of prewarmed (65°C) elution buffer and eluted by centrifugating each sample at top speed for 5 seconds. The concentration of the mRNA was determined photometrically.

3.4.3 DNase digestion of RNA samples

DNA contamination of RNA samples was eliminated by incubating 200 μ l (or a

multiple) of the RNA sample with 20 µl of 10x DNase buffer (Roche) and 1 µl of DNase I (10 U/µl; Roche). The digest was incubated for 30 minutes at RT, followed by additional 30 minutes at 37°C. Subsequently, the enzyme was inactivated by incubating the solution at 75°C for 5 minutes. The samples were precipitated with 96 % ethanol (see 3.3.9.2), washed with 70 % ethanol and resuspended in an appropriate amount of DEPC H₂O_{dd}.

Alternatively, the samples were treated with 1 µl TURBO DNase (2 U/µl; Ambion) together with 1x TURBO DNase buffer (Ambion) in a 50 µl reaction. After 30 minutes at 37°C, 0.1 volumes of the DNase inactivating reagent (Ambion) were added in order to remove the DNase enzyme as well as divalent cations, which would otherwise catalyse RNA degradation. The solution was incubated for two minutes at RT and mixed occasionally. Subsequently, the tubes were centrifugated at 10.000 rpm for 90 seconds. The supernatant containing the DNA-free RNA was transferred into a new tube and used immediately for downstream analysis or stored at -80°C.

3.4.4 Generation of first strand cDNA

For the reverse transcription reaction, 200 ng of DNase treated mRNA (see 3.4.2 and 3.4.3) or 2 µg of total RNA (see 3.4.1) were combined with 4 µl dNTP (10 mM, Applied Biosystems), 200 pMol primer and H₂O_{dd} in a final volume of 15 µl. (If mRNA was used as starting material, either a poly-(dT) primer composed of 50 % dTTTV and 50 % dTTTG oligonucleotides or a gene-specific primer was used.)

The reaction was incubated for 5 minutes at 65°C and placed on ice for at least one minute. Then, 4 µl DTT (0.1 M Stock), 6 µl of a 5 x buffer (reverse transcriptase buffer, Invitrogen) and 80 U of RNaseOUT™ ribonuclease inhibitor (Invitrogen) were added, the reaction was mixed gently and incubated for two minutes at 42°C. Afterwards, 1 µl of Superscript™ II reverse transcriptase (Invitrogen) was added to the solution. The reaction was mixed gently and incubated at 42°C for additional 50 minutes.

Finally, the enzyme was inactivated at 70°C for 15 minutes and 2 µl RNase H (1U/µl) (Invitrogen) were added. The solution was incubated for 20 minutes at 37°C to allow the removal of any RNAs complementary to the cDNA. The first strand cDNA was immediately placed on ice and stored at -20°C until further use.

3.4.5 RNA ligase-mediated rapid amplification of 5' and 3' ends (RLM-RACE)

This method ensures the exclusive amplification of full-length transcripts which are then, in the second step, used for the RACE (rapid amplification of 5' and 3' cDNA ends) analysis. Truncated mRNAs and non-mRNAs are eliminated by a preliminary dephosphorylating

reaction which is followed by the removal of the 5' cap structure from full-length mRNAs. As a consequence, the 5' phosphate that is required for the subsequent ligation to the GeneRacer™ RNA-oligonucleotide ('linker') is exclusively present on previously capped mRNA. After reverse transcription, either the 5' end or the 3' end of the generated first strand DNA can be analysed.

Per sample, 200 µg of mRNA were gently mixed with 1x CIP buffer (GeneRacer™ kit, Invitrogen), 40 U of RNase Out (Invitrogen), 10 U CIP (GeneRacer™ kit, Invitrogen) and a variable volume of DEPC water for a total volume of 10 µl. The solution was incubated at 50°C for one hour. Then, the RNA was purified by a phenol/chloroform extraction and ethanol precipitated. After a washing step with 70 % ethanol, the pellet was resuspended in 7 µl DEPC water. The 5' cap structure was removed by adding 1x tobacco acid pyrophosphatase (TAP) buffer (GeneRacer™ kit, Invitrogen), 40 U of RNaseOUT™ and 0.5 U of TAP (GeneRacer™ kit, Invitrogen). The reaction was incubated for one hour at 37°C. Again, the RNA was precipitated as above and recovered in 7 µl DEPC water. For the ligation to the linker, the dephosphorylated, decapped RNA was added to one aliquot of lyophilized GeneRacer™ RNA Oligo (GeneRacer™ kit, Invitrogen) and incubated for 5 minutes at 65°C to denature RNA secondary structures. Then, the reaction was placed on ice and 1x ligase buffer (GeneRacer™ kit, Invitrogen), 1 mM ATP, 40 U of RNaseOUT™ (Invitrogen) and 5 U of T4 RNA ligase (Invitrogen) were added. After incubation for one hour at 37°C, the RNA was precipitated as above, resuspended in 10 µl DEPC water and reverse transcribed (see 3.4.4) using a gene-specific primer.

The generated first strand cDNA was then used for the subsequent amplification step. The PCR products were purified by agarose gel electrophoresis (see 3.3.3 and 3.3.10) and subcloned into the pCR®-TOPO vector. Upon amplification in bacteria and purification of the plasmid DNA, the sequence of the inserts was analysed.

3.5 Protein methods

3.5.1 Extraction of cellular proteins

Whole cell lysates were generated 48 hours after transfection of cells grown in a 6-well plate. Briefly, the plates were put on ice, the medium was removed and the cells washed once with ice-cold PBS. Upon addition of 200 µl lysis buffer (25 mM KPO₄, pH 7.8; 8 mM MgCl₂, 1 mM DTT, 1 mM EDTA, 1 % (v/v) Triton X-100, 15 % (v/v) glycerol), the cells were scraped from the bottom of the well, transferred into a pre-chilled Eppendorf tube and placed on ice. During the following incubation step (15 minutes on ice), the samples were vortexed vigorously several times. Thereafter, the tubes were centrifuged at 14,000 rpm for 5 minutes at 4°C and the supernatant containing the cellular proteins transferred into a new tube and

either analysed immediately or stored at -80°C.

3.5.2 Extraction of nuclear and cytoplasmic protein fractions

Approximately 5×10^6 cells were harvested by trypsinization, washed once with ice-cold PBS and centrifuged at 4°C for 5 minutes at 1,000 rpm. The supernatant was removed and 200 µl of ice-cold CER I (Cytoplasmic extraction reagent I, Pierce) supplemented with 5 µl/ml of proteinase inhibitor cocktail (Sigma) was added to the cells. After resuspending the pellet by vigorously vortexing for 15 seconds, the solution was transferred into a 1.5 ml Eppendorf tube and incubated on ice for 10 minutes. Subsequently, 11 µl of ice-cold CER II (Cytoplasmic extraction reagent II, Pierce) were added and the mixture was again vortexed for 5 seconds. After one minute incubation on ice, the solution was again vortexed and centrifuged at 13,000 rpm for 5 minutes at 4°C in a microcentrifuge. The supernatant containing the cytoplasmic extract was immediately transferred into a new, pre-chilled tube and stored at -80°C.

The remaining pellet (containing the nuclei) was resuspended in 100 µl of ice-cold NER (Nuclear extraction reagent, Pierce) supplemented with 5 µl/ml of proteinase inhibitor cocktail (Sigma). The solution was intensively vortexed on the highest setting for 15 seconds and immediately placed on ice for 10 minutes. The last two steps were repeated five times for a total period of 50 minutes. Then, the mixture was centrifuged at 13,000 rpm for 10 minutes at 4°C. The supernatant representing the nuclear extract was immediately transferred into a new pre-chilled tube and placed on ice. The nuclear protein fraction was stored at -80°C until further use.

3.5.3 Western blot analysis

3.5.3.1 SDS-PAGE

SDS-Polyacrylamide gel electrophoresis (PAGE) was performed in a SE 260 Mini-vertical gel electrophoresis unit (Hoefer), assembled according to the instructions of the manufacturer. The separation gel (1.5 M Tris-base, pH 8.8; 10 % (v/v) Acrylamide, 0.1 % SDS, 0.05 % APS, 1 µl/ml gel Temed) was poured in a dual gel caster (Hoefer), overlaid with isopropanol and rinsed with water after polymerization. Then, the stacking gel (0.5 M Tris-HCl, pH 6.8; 5 % (v/v) Acrylamide, 0.4 % SDS, 0.05 % APS, 1 µl/ml gel Temed) was prepared. After polymerization, the gel unit was assembled and the gel submerged with in 1x Laemmli-buffer (25 mM Tris-base, 0.19 M glycine, 0.1 % (w/v) SDS; pH 8.3). Protein samples mixed with 2x protein sample buffer (100 mM Tris- base, pH 6.8; 5 % (w/v) SDS, 20 % glycerol, 0.02 % bromphenolblue, 10 % (v/v) β-mercaptoethanol) were denaturated for 5 minutes at 95°C

and loaded onto the gel. The separation of protein samples including a protein marker (PageRuler™ Prestained protein Ladder, Fermentas) was performed at 95 volt (Power Pac 300 electrophoresis power supply, Biorad).

3.5.3.2 Western blotting

For Western blot analysis, protein samples were separated by SDS-PAGE (see 3.5.3.1). The proteins were then transferred to a PVDF membrane (Hybond-P membrane, Amersham). For this purpose, the membrane was presoaked with methanol and placed together with the gel and 3 pieces of Whatman™ 3MM paper in a semi-dry electric blotter device (Hoefer). The transfer was performed in transfer buffer (25 mM Tris-base, pH 8.3; 0.19 M glycine; 20 % (v/v) methanol) at 1.1 mA/cm² for 1-2 hours depending on the expected size of the protein to be analysed. The transfer efficiency was verified by Coomassie staining of the polyacrylamid gels (see 3.5.3.3).

3.5.3.3 Coomassie staining of polyacrylamide gels

After the blotting was completed the SDS-PAGE were removed from the transfer unit and stained for 1 hour in Coomassie light staining buffer (50 ml isopropanol, 20 ml acetic acid and 0.1 g Coomassie Brilliant blue R250 in 200 ml H₂O_{dd}, filtered before use). Then, the gel was rinsed shortly in water and destained by in 10 % acetic acid for approximately two hours on a shaker. The intensity of the protein bands should indicate equal loading of the samples and the running quality of the gel.

3.5.3.4 Immunological detection

To saturate the non-specific binding sites the membrane was incubated in blocking buffer (5 % milk powder solution in 1x TTBS buffer (20 mM Tris, pH 8.0; 0.9 % NaCl; 0.1 % Tween 20) at constant agitation overnight at 4°C. On the following day, the membrane was incubated with the primary antibody diluted in 1x TTBS buffer for at least one hour at room temperature. After two washing steps with 1 x TTBS at RT for a total of 30 minutes, the membrane was incubated with an horseradish peroxidase (HRP) labelled secondary antibody (DAKO; diluted 1:10,000 in 1x TTBS) for one hour at RT. To minimize unspecific binding events, the membrane was washed three times (10 minutes each) with 1x TTBS buffer, followed by one washing step with 1x TBS buffer (20 mM Tris, pH 8.0; 0.9 % NaCl).

ECL plus Western blotting detection reagents (Amersham) were used to visualise the HRP-labelled secondary AB. 2 ml solution "A" and 50 µl solution "B" were mixed together and

immediately applied to the membrane for a maximum of 5 minutes. The blot was placed into a plastic foil and exposed to an autoradiography film (Hyperfilm ECL, Amersham).

3.5.4 Luciferase assay

The luciferase assays were performed 48 hours post transfection. After extraction of the cellular proteins (see 3.5.1), the protein concentration of each sample was determined using a colorimetric assay (see 3.5.5). 20 µg of the protein lysate was pipetted into a plastic tube and transferred into the luminometer (AutoLumat LB 953). By automatic injection a defined amount of reaction mix (100 mM KPO₄, pH 7.8; 125 µM luciferin, 0.4 mM Mg ATP) containing the luciferase substrate was added to each sample and the amount of emitted light was measured. The quantity of emitted light is directly proportional to the activity of the regulatory elements that were responsible for the enzyme production.

3.5.5 Protein quantification

The protein content of samples was determined using a colorimetric assay according to the Lowry method. For the generation of the standard curve, 5-6 dilution of bovine serum albumine (BSA; New England Biolabs) ranging from 0.2 mg/ml to 2.0 mg/ml protein were prepared. Then, 5 µl of each standard and each sample were pipetted into a 96-well plate bottom microtiter plate in duplicate. Subsequently, 25 µl of working reagent A' (20 µl reagent S per 1000 µl reagent A, BioRad DC protein assay kit II) and 200 µl of reagent B (BioRad DC protein assay kit II) were added. After the plate was incubated for 15 minutes at RT in the dark, the absorbance was read at 630 nm by a microplate reader (Elx 800, Bio Tek). The protein concentration of the samples was determined based on the standard curve.

3.6 Chromatin methods

3.6.1 Isolation of chromatin

CK5'LTR cells and NM5'LTR cells, respectively, were seeded in 6x 175 cm² culture flasks and synchronized (see 3.1.6). Two hours after reentering the cell cycle simultaneously in the G₁ phase, the medium was replaced with 5 ml of 1 % formaldehyde (in 1x PBS) to fix the cells. After incubating the cells for 5 minutes at RT, the formaldehyde was removed. The cells were washed twice with ice-cold 1x PBS, scraped, collected in prechilled 15 ml falcon tubes and placed on ice. After a brief centrifugation at 1000 rpm for 5 minutes at 4°C, the supernatant was removed. The pellet was resuspended in 5 ml homogenization buffer (10 mM Tris-HCl, pH

7.4; 15 mM NaCl; 60 mM KCl; 1 mM EDTA; 0.1 mM EGTA; 0.1 % (v/v) NP-40; 5 % (w/v) sucrose; 0.15 mM spermine; 0.5 mM spermidine), transferred to a Dounce homogenizer and incubated for 5 minutes on ice. In order to disrupt the cell membrane, 10 to 15 strokes with the Dounce pestles were performed. The lysis was controlled by observing an aliquot of the cells suspension under a light microscope. Then, the mixture was transferred carefully onto a layer of wash buffer containing 10 % sucrose (10 mM Tris-HCl, pH 7.4; 15 mM NaCl; 60 mM KCl; 10 % (w/v) sucrose; 0.15 mM spermine; 0.5 mM spermidine) and centrifugated at 1000 rpm for 20 minutes at 4°C. The supernatant was removed and the pellet gently resuspended in 5 ml wash buffer (10 mM Tris-HCl, pH 7.4, 15 mM NaCl, 60 mM KCl, 0.15 mM spermine, 0.5 mM spermidine). The samples were centrifugated at 800 rpm for 5 minutes at 4°C. The supernatant was removed and the pellet containing the isolated nuclei was kept on ice for further downstream manipulations.

3.6.2 MNase digestion

When nuclei (or permeabilized cells) are exposed to micrococcal nuclease (MNase) in the presence of a divalent cation, the enzyme introduces double-strand cuts within the inter-nucleosomal region. This property was used to achieve mononucleosomal particles for either the determination of nucleosomal positions or for the chromatin immunoprecipitation (ChIP) experiments.

Briefly, the pelleted nuclei (see 3.6.1) were gently resuspended in either 375 µl (for determining the nucleosomal borders) or 250 µl (for ChIP analysis) MNase buffer (25 mM KCl; 4 mM MgCl₂; 1 mM CaCl₂; 12.5 % (v/v) glycerol; 50 mM Tris-HCl, pH 7.4; 0.1 µM Aprotinin, 10 µM Leupeptin; 0.2 mM PMSF) and transferred to an 1.5 ml Eppendorf tube. Micrococcal nuclease (Sigma, diluted in 0.1 M NaPO₄ buffer, pH 7.0; 2.5 µM CaCl₂) was added at a final concentration of 2 U/ml (for determination of the nucleosome position) or 40 U/ml (for ChIP experiments). The samples were incubated at 37°C for 5 minutes while shaking slowly, before the enzymatic reaction was stopped by adding 0.5 M EDTA to a final concentration of 50 mM.

3.6.3 Determination of nucleosomal borders

In order to remove any associated proteins from the MNase-digested chromatin, the solution was diluted (1:2) in proteinase K buffer (20 mM EDTA, 200 mM NaCl, 1 % (w/v) SDS, 0.25 mg/ml glycogen) and 1/10 of the sample volume proteinase K (2.5 mg/ml; Sigma) was added. The reaction was incubated at 37°C for up to two hours. In order to reverse the protein-DNA crosslinks, the whole digest was then incubated for 6 hours (alternatively overnight) at 60°C while gently shaking slowly. Subsequent to a phenol-chloroform extraction

(see 3.3.8.1), the DNA was ethanol-precipitated (see 3.3.9.2) and resuspended in 100 μ l H_2O_{dd} . To avoid any RNA contamination, the samples were treated with 20 μ g RNase A (10 mg/ml; Sigma) at 37°C for 10 minutes. Again, the DNA was purified via phenol-chloroform extraction, precipitated by ethanol and recovered in 100 μ l H_2O_{dd} .

For the determination of the nucleosomal borders, the fragmented DNA originating from MNase digested chromatin (see 3.6.2) was loaded onto a 2 % agarose gel and separated via electrophoresis. The band corresponding to \sim 147 bp (i.e the DNA protected by a single nucleosome) was extracted from the gel (see 3.3.10) and recovered in 40 μ l H_2O_{dd} . In the following step, the DNA fragments were treated with T4 polynucleotide kinase (see 3.3.8.3) in order to phosphorylate the cleaved double-stranded substrate. After an ethanol-precipitation, 0.2 μ g of each sample were ligated to (\sim 60 pmol) the linker-hybrid (see 3.3.7 and 3.3.8.2). The ligated fragments were precipitated by ethanol, resuspended in 20 μ l H_2O_{dd} and stored at 4°C. Ligation-mediated PCRs were performed with an oligonucleotide corresponding to the linker-hybrid and a gene specific primer.

3.6.4 Chromatin immunoprecipitation (ChIP)

This method is used to determine the association of proteins or protein modifications to specific genomic regions. Briefly, fragments of cross-linked chromatin are immunoprecipitated with an antibody against a desired protein or protein modification. Thus, DNA sequences that directly or indirectly cross-link with the protein of interest are selectively enriched in the immunoprecipitated sample. The subsequent reversal of the cross-links by heating permits the recovery and quantitative analysis of the immunoprecipitated DNA. The amounts of specific genomic regions in control and immunoprecipitated samples are determined individually by quantitative PCR.

In order to examine the MMTV LTR region, ChIP was performed on single nucleosomes that were generated by enzymatic digestion using micrococcal nuclease (see 3.6.2). Immediately after terminating the enzymatic reaction with EDTA, the chromatin was diluted (1:1) in 250 μ l FA lysis buffer (50 mM HEPES, pH7.5 with KOH, 150 mM NaCl, 1 mM EDTA, 1 % (v/v) Triton X-100, 0.1 % (w/v) sodium deoxycholate, 0.1 % SDS (v/v); 0.2 mM PMSF; 10 μ l/ml protease inhibitor cocktail, Sigma). Simultaneously, an aliquot was subjected to proteinase K treatment and analysed via electrophoresis to determine the quality of the MNase digestion.

In order to reduce any non-specific associations of any material to the protein G beads that were used for the subsequent immunoprecipitation, a preclearing step was performed by adding 40 μ l ChIP-grade protein G agarose beads (Cell Signaling) to the digested chromatin. The solution was incubated for two hours at 4°C during constant rotation.

Subsequent to a centrifugation at 14,000 rpm for 10 minutes, the supernatant was transferred into a new eppendorf tube. Additionally, an aliquot (50 µl) labelled as 'Input-DNA' was removed from the isolated, MNase digested chromatin and transferred into a new pre-chilled tube. All samples including the 'Input-DNA' were stored at -80°C until further use.

For each immunoprecipitation, 100 µl of the precleared chromatin were thawed on ice and diluted in 400 µl FA lysis buffer. 10 µl of the antibody of interest were added and the samples were incubated overnight at 4°C with constant rotation. Normal rabbit IgG (Cell Signaling) was used as negative control. On the following day, 40 µl protein G beads were added, followed by an additional incubation step of two to four hours at 4°C during constant rotation. Then, the samples were centrifuged at 6,000 rpm for 90 seconds at 4°C and the supernatant discarded.

Next, the beads were washed twice with 1 ml of FA lysis buffer and twice with 1 ml of FA lysis buffer containing 0.5 M NaCl, followed by an additional washing step with ChIP wash buffer (10 mM Tris-HCl, pH 8.0; 0.25 M LiCl; 1 mM EDTA; 0.5 % (v/v) Nonidet P-40; 0.5 % (w/v) sodium deoxycholate.) Each washing step included 10 minute of constant rotation in the cold, followed by a centrifugation at 6,000 rpm for 90 seconds at 4°C. Then, 150 µl ChIP elution buffer (50 mM Tris-HCl, pH 7.5; 10 mM EDTA; 1 % (v/v) SDS) were added to each sample. At this point the 'Input-DNA' sample was thawed on ice and diluted in 250 µl ChIP elution buffer. The chromatin-antibody complexes were dissociated from the beads by incubating the samples for 30 minutes at 65°C while gentle rotating. The protein G agarose beads were pelleted by a centrifugation at 6,000 rpm for one minute and the supernatant containing the immunoprecipitated chromatin was carefully transferred to a new tube. Then, 5 M NaCl was added to a final concentration of 20 mM. Finally, the samples were treated with 2 µl proteinase K and incubated at 65°C for two hours (or alternatively overnight) to remove the protein-DNA crosslinks. Then, the DNA was extracted by phenol-chloroform and precipitated with ethanol. After resuspending the samples in an appropriate volume of H₂O_{dd}, they were subjected to qRT-PCR analysis (3.3.12.2).

4. Results

The aim of the present study was to characterise in more detail the central LTR of MMTV, which harbours the P2 promoter element. Moreover, we wanted to investigate the conditions under which transcription initiation is allowed to occur upstream of the major P1 promoter. The Mtv-2 and Mtv-8 sequences used in this work are derived from the plasmid pGR102 (Salmons *et al.*, 1985).

4.1 Two elements located within the MMTV LTR influence P2 activity

The P2 promoter is located in the central part of the proviral MMTV LTR (Günzburg *et al.*, 1993). Within this region, a number of regulatory elements have been identified that were shown to modulate the activity of the major P1 promoter. We wanted to investigate whether these regulatory sequences exert a similar influence on the P2 promoter. Preliminary experiments revealed that the transcriptional potential of the P2 promoter is extremely low (data not shown). Therefore, we decided to use the promoterless luciferase expression plasmid pZluc for the analysis (Brasier *et al.*, 1989; Maxwell *et al.*, 1989). This vector contains three polyadenylation signals (one for each reading frame) upstream of the multiple cloning site that impede any readthrough from the plasmid backbone. We started by testing Mtv-2 LTR sequences, because this provirus is known to be responsible for the high incidence of mammary tumors in GR mice (Golovkina *et al.*, 1996).

The activity of a 40 bp fragment containing the minimal P2 core promoter including the TATA-box (pTBP2, construct 1) was compared with the luciferase expression of other fragments that additionally possess defined regions located immediately upstream of the minimal P2 promoter. Either the complete U3 region (pZ678, construct 2), the region from position 301 (relative to the beginning of the LTR) up to the minimal promoter (pZ370, construct 3) or the region ranging from position 557 until the P2 core promoter (pZ151, construct 4) were linked individually to luciferase coding regions. Several independent cell lines from different species (feline kidney CrFK cells, mouse mammary tumor GR cells, mouse fibroblast NIH 3T3 cells, mouse lymphocyte A20 cells, and rat hippocampal HiB5 stem cells) were transfected and transient luciferase activity was measured 48 hours later.

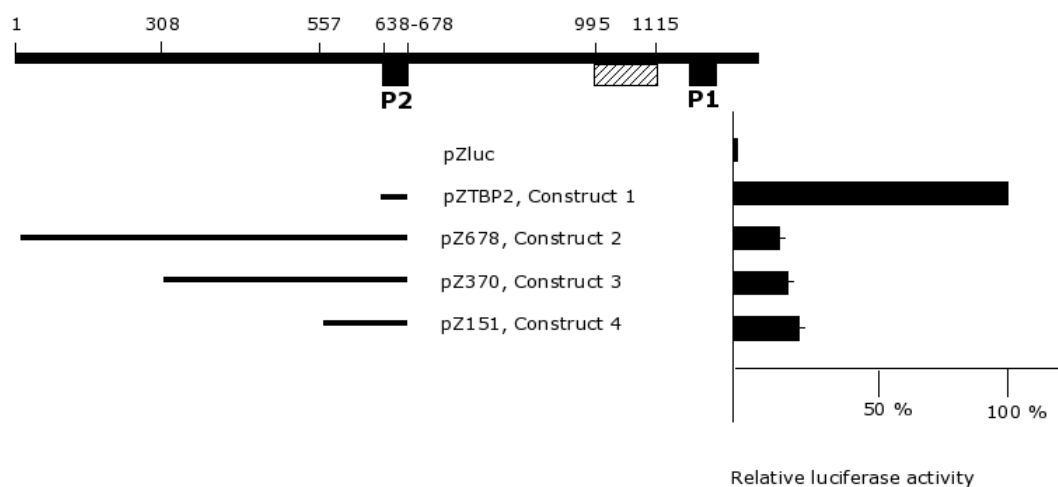


Figure 4.1 Influence of upstream located U3 sequences on P2 promoter activity. The U3 region is shown schematically on the top. Numbers indicate the position relative to the beginning of the LTR. The P2 and P1 core promoter elements are represented by filled boxes. The hormone-responsive element (HRE) is symbolized by the hatched box. On the left, the sequences used in the four different constructs are indicated. On the right, the corresponding luciferase expression levels are shown. The expression of the minimal P2 promoter was taken as 100%. All transfections were performed in the absence of dexamethasone. Each bar reflects the mean value of at least three independent experiments with standard errors of less than 5% of the mean values.

Figure 4.1 shows the result of one representative analysis which was performed in CrFK cells, although similar luciferase expression patterns were obtained in all cell lines tested (data not shown). Here, the transcriptional activity of the minimal P2 promoter was set to 100%. A decrease in P2 promoter activity in all three constructs tested can be observed, indicating the existence of an upstream located negative regulatory element (NRE). As the smallest fragment (pZ151, construct 4) is sufficient to attenuate P2 activity, the location of the NRE can be narrowed down to the region from position 557 to 638, immediately adjacent to the P2 promoter. We observed the inhibitory effect in all cell lines tested, indicating a tissue unspecific mechanism of repression.

Next, we analysed the effect of immediately downstream located LTR sequences on the basal activity of the P2 core promoter. As for the previous set of constructs, we inserted various downstream located fragments into the promoterless luciferase reporter plasmids, all of which include the region from bp 744 to bp 995 (constructs 5 to 8, figure 4.2). The influence of this region on the transcriptional activity of the major P1 promoter is controversial in the literature, because the presence of a possible enhancer (Gouilleux *et al.*, 1991; Reuss and Coffin, 2000) as well as several NREs (Mink *et al.*, 1990; Lefebvre *et al.*, 1991; Bramblett *et*

al., 1995) have been described.

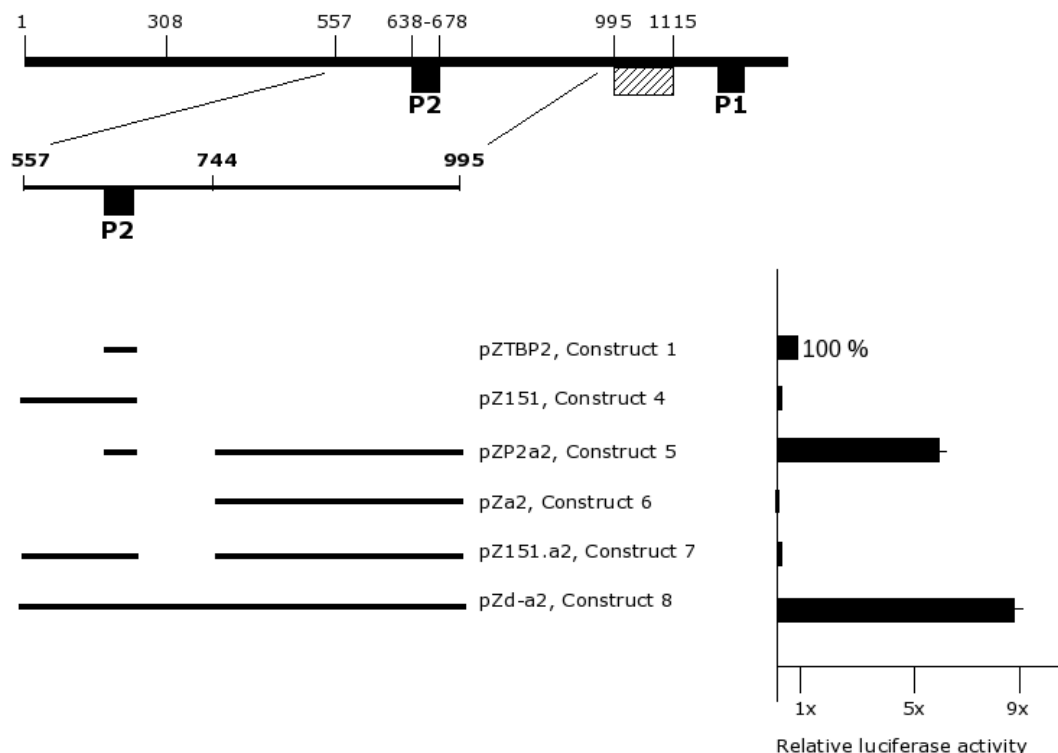


Figure 4.2 Influence of downstream located U3 sequences on P2 promoter activity. The U3 region is shown schematically above. Numbers indicate the position relative to the beginning of the LTR. The P2 and P1 core promoter elements are represented by filled boxes, while the HRE is symbolized by the hatched box. The lines on the left indicate the regions used and analysed in luciferase reporter gene assays. The bars represent the relative reporter gene expression levels of the tested constructs after transient transfection of CrFK cells in the absence of dexamethasone. Each bar reflects the mean value of at least three independent experiments with standard errors of less than 5% of the mean values.

As can be seen in figure 4.2, a 6-fold induction of luciferase expression was observed when the region from bp 744 to bp 993 was linked to the P2 promoter core, (pZP2a2, construct 5, figure 4.2). This positive effect is caused by a true enhancer element, as the region on its own does not possess any transcriptional potential (pZa2, construct 6). When both the upstream located NRE and the enhancer were present in the same construct, the activity of the promoter was again downregulated, indicating that the NRE abrogates the influence of the downstream enhancer element (pZ151.a2, construct 7). Interestingly, when the entire region from the beginning of the NRE up to the end of the enhancer (bp 557 to bp 995, pZd-a2, construct 8) was analysed, an even higher luciferase expression level was observed than that previously obtained from the P2 promoter under the influence of its enhancer (pZP2a2, construct 5). We conclude that the region between bps 678 and 744 does contain an element with a strong positive effect on transcription.

To find out more about the properties of this region, we analysed the corresponding sequence in different MMTV variants and found two potential TATA-boxes and one CAAT-like motif (underlined sequences, figure 4.3).

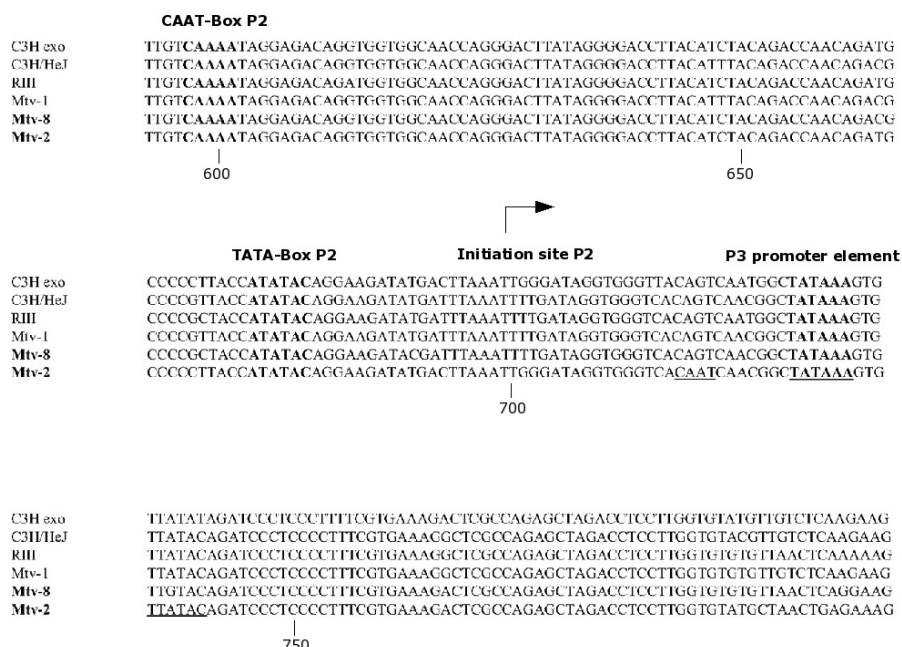


Figure 4.3 Alignment of the central LTR promoter complex. The sequences of Mtv-8 and Mtv-2 correspond to the 5' LTR and 3' LTR of the plasmid pGR102, respectively (Salmons *et al.*, 1985). The TATA-box of P2 is shown in boldface type, while potential CAAT- and TATA-boxes for the P3 promoter are underlined. The P2 transcriptional initiation site is indicated by an arrow. The GeneBank accession numbers are as follows: AF228552 for C3H exo, AF228551 for CeH/HeJ, AF136900 for RIII and AF228550 for Mtv-1.

Only one of these motifs is conserved within the Mtv-8 provirus; therefore, this element was tested first for its transcriptional activity. To this end, different fragments of the central LTR (encompassing the region from bp 557 to bp 995) were amplified from an Mtv-8 background and analysed in transient transfection assays. The results are summarized in figure 4.4.

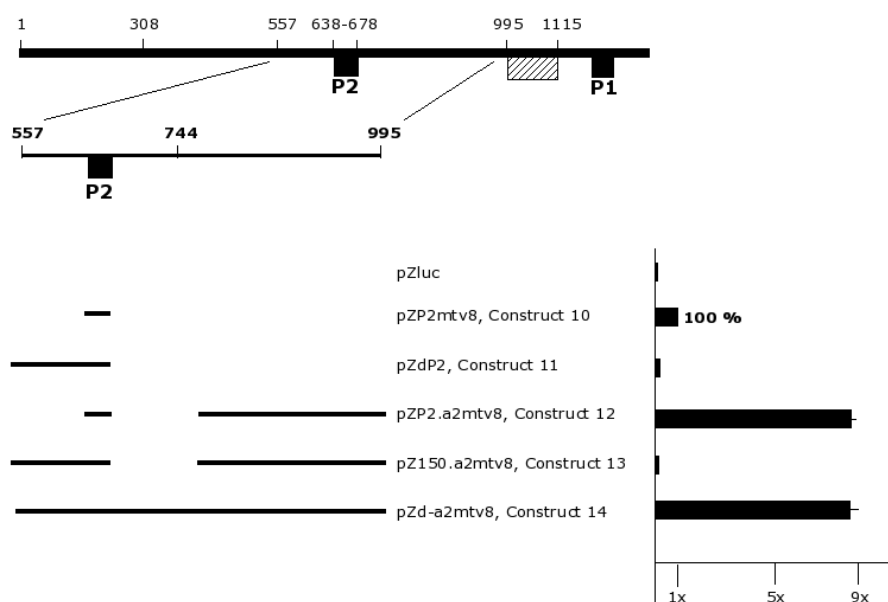


Figure 4.4 The central LTR Promoter region is functional in Mtv-8. A scheme of the MMTV U3 region is presented above. The core promoter elements are symbolized by filled boxes, while the hatched box represents the HRE. The lines on the left side indicate the regions used and analysed in luciferase reporter gene assays, whereas the bars on the right represent the relative reporter gene expression levels of the tested constructs. The analysis was performed in CrFK cells in the absence of dexamethasone. The transcriptional activity of the minimal P2 promoter was set to 100%. Each bar reflects the mean value of at least three independent experiments with standard errors of less than 5% of the mean values.

The activity of the P2 promoter originating from the Mtv-8 isolates followed the same pattern of regulation as the previously tested Mtv-2 sequences. The upstream located regulatory element abrogated its activity (pZdP2mtv8, construct 11 versus minimal Mtv-8 P2 promoter, construct 10), irrespective of the presence of the downstream enhancer (pZ150.a2mtv8, construct 13). However, in the absence of the NRE, the same enhancer increased P2 activity up to 10-fold (pZP2.a2, construct 12). The level of reporter gene expression remained elevated when the complete region from the beginning of the NRE up to the end of the enhancer was analysed, reflecting the transcriptional activity of the short region between the P2 promoter core (bp 638-678) and position 744 (pZd-a2mtv8, construct 14). We conclude that the TATA-motif present in Mtv-8 represents the promoter core of the transcriptional active element. In accordance with previous MMTV nomenclature, the novel promoter was termed P3.

4.2 The P3 promoter is regulated by a downstream enhancer

To further characterise the novel promoter, we tested its transcriptional potential and possible influencing sequences of the LTR using the same transient reporter gene assays as described above. The results of the analysis are summarized in figure 4.5.

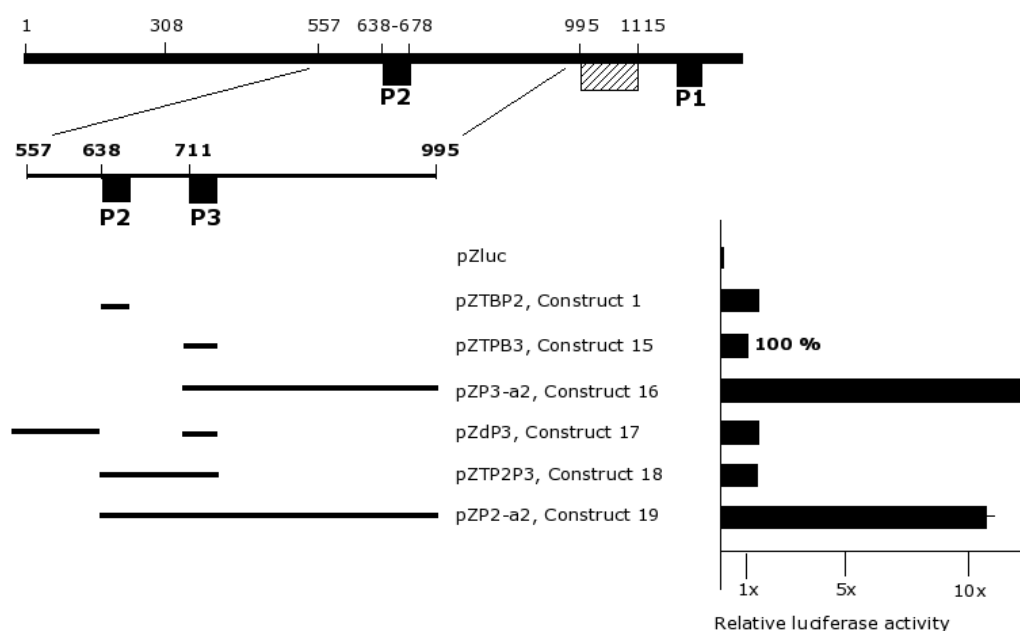


Figure 4.5 Influence of immediately upstream and downstream located regions on the activity of the P3 core promoter. A scheme of the MMTV U3 region is presented on the top. The core promoter elements are symbolized by filled boxes, while the HRE is indicated by the hatched box. The sequences used and analysed are indicated by the thin lines. The bars represent the relative luciferase expression levels of the tested constructs after transient transfection of CrFK cells in the absence of dexamethasone. The transcriptional activity of the minimal P3 promoter was set to 100%. Each bar reflects the mean value of at least three independent experiments with standard errors of less than 5% of the mean values.

This time, the activity of the minimal P3 promoter was set to 100% and compared to the relative activities of the other constructs tested. Both minimal core promoters show comparable transcriptional potential (pZTBP2, construct 1, compared with pZTBP3, construct 15, figure 4.5). The presence of the downstream regulatory element increased the relative luciferase expression (pZP3-a2, construct 16, compared to construct 15). However, the addition of the NRE did not affect P3 activity (pZdP3, construct 17, compared to construct 15). The two promoters do not work in synergy, as no increase in reporter gene expression is observed in the fragment encompassing both promoters (pZTBP2P3, construct 18, compared to construct 1 and construct 15). As expected, high levels of luciferase expression were

observed when the complete region from bp 638 up to bp 995 was analysed (pZP2-a2, construct 19 compared to construct 15 and 16).

Summarizing these transient transfection experiments, two transcriptional active promotor elements are located within the central LTR of MMTV, whose TATA-boxes are only 47 bp apart from each other. Both are regulated by the same enhancer element, while the activity of the more upstream situated P2 promoter is additionally influenced by an NRE.

4.3 A transcriptional initiation site is located downstream of the P3 promoter

For the characterization of the novel P3 promoter, we decided to utilize the plasmid pGR102 which carries the complete coding information of an integrated provirus (Salmons *et al.*, 1985). Originally, the 5'LTR and the beginning of *gag* was derived from Mtv-8, while the remaining sequences were comprised of Mtv-2 based elements. In order to ensure that both LTRs carry the same sequence in their central part, we replaced the former 5' LTR with Mtv-2 based sequences to generate an Mtv-2/Mtv-8/Mtv-2 hybrid. Additionally, a CMV-driven EGFP expression cassette was introduced into the 3' LTR to facilitate the detection of integrated copies. NMuMG cells, which are derived from the mammary glands of an adult Namru mouse (Owens *et al.*, 1972) and thus represent a natural reservoir of MMTV, and CrFK cells, which are permissive to MMTV yet lack any endogenous murine retroviral sequences, were transfected and stable clones, referred to as N2AnE5 and C2AnE15, respectively, were produced. The ability to produce infectious viral particles was monitored periodically by FACS analysis (data not shown).

In order to determine the respective transcriptional initiation site of transcripts originating from the P3 promoter, we performed an RLM-RACE (RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends). This method allows the specific amplification of capped transcripts and, in consequence, the determination of their respective 5' or 3' ends. In brief, this method starts by dephosphorylating the pool of isolated RNAs. Thereby, truncated mRNA and non-mRNA are eliminated from the subsequent ligation step while full-length, capped mRNA would remain unaffected. In the following step, the cap-structure is removed leaving a 5' phosphate that is required for the subsequent ligation to the provided GeneRacer™ RNA oligonucleotide. Then, the ligated RNAs are reverse transcribed. In order to obtain the 5' ends, the first strand cDNAs are amplified with the GeneRacer™ 5' Primer and a sequence specific reverse primer.

For the analysis, we isolated messenger RNAs from GR cells, a productively infected mouse mammary cell line, as well as from clone N2AnE5, both of which have been grown in

the absence of dexamethasone. After dephosphorylation, decapping and ligation, the RNAs were reverse transcribed with a reverse primer corresponding to the *gag* 5' untranslated region (1443/1424) (figure 4.6A). Subsequently, RT-PCRs were performed with the GeneRacer™ 5' Primer and a nested reverse primer (1431/1409).

In both MMTV producer cell lines, we obtained a PCR product of approximately 240 bp in length (figure 4.6B, lane 1 and 2). The size of the amplicon indicates that the corresponding transcript represents a spliced message. (In contrast, the expected length of an unspliced product would be ~700 bps.) A splice donor (SD₈₂₅) and splice acceptor site (SA₁₂₇₅) within the LTR have been described previously (Xu *et al.*, 1997).

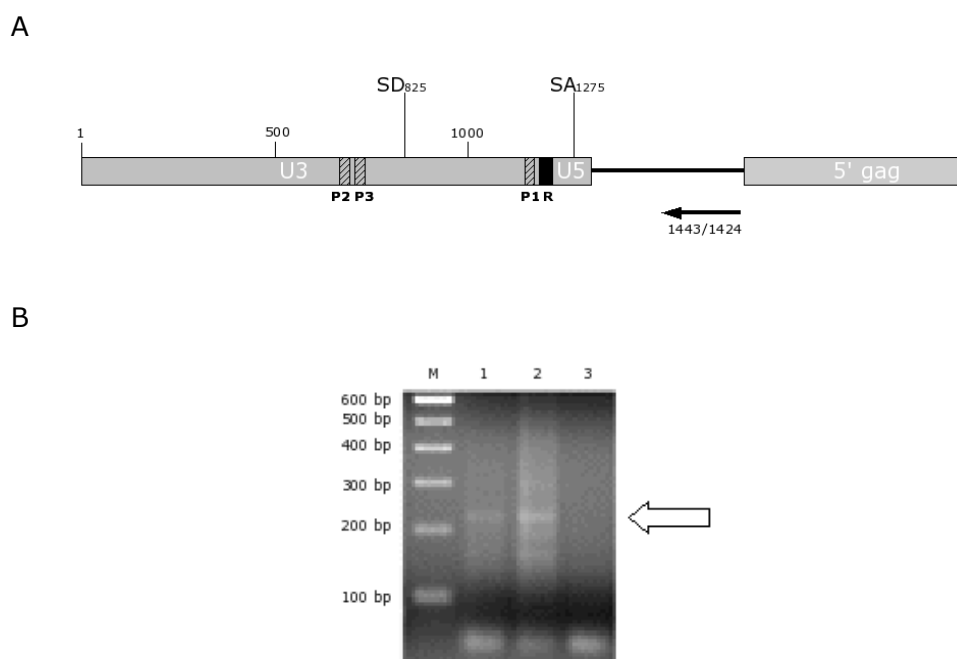


Figure 4.6 Determination of the transcriptional initiation site. Messenger RNA was isolated from GR cells and N2AnE5 cells, decapped and ligated to the GeneRacer™ RNA oligonucleotide. After reverse transcription, PCRs were performed with the GeneRacer™ 5' Primer and a gene-specific reverse primer (1431/1409). (A) The 5' LTR and the beginning of the *gag* gene of MMTV are shown. The known splice donor and splice acceptor sites are indicated as well as the sequence specific primer binding site used for the generation of first strand cDNAs. (B) The amplification products obtained from GR (lane 1) and N2AnE5 (lane 2) first strand cDNA are shown. PCRs performed in the absence of any template are shown in lane 3. The arrow points to the MMTV-specific 5' RACE PCR product.

The sequence analysis of the gel purified and subcloned GeneRacer PCR products revealed a corresponding transcriptional initiation site at position 760 relative to the beginning of the LTR, which is therefore 61 nucleotides downstream of the initiation site of the P2

promoter that has been determined by Günzburg and colleagues (Günzburg *et al.*, 1993) (figure 4.7).

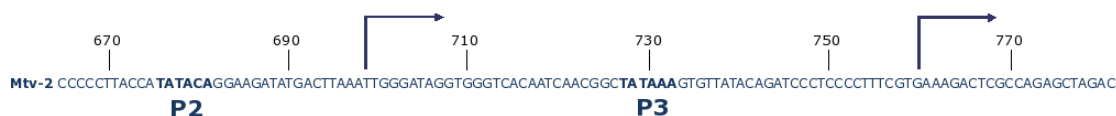


Figure 4.7 Transcriptional initiation sites in the central LTR of MMTV. The P2/P3 promoter region derived from the Mtv-2 proviral sequence is shown. The TATA-boxes of both promoters are displayed in bold face. The arrows indicate the transcriptional start sites of the P2 promoter (Günzburg *et al.*, 1993) and the P3 promoter, respectively. The numbers refer to the beginning of the LTR.

Thus, the central LTR of MMTV harbours two core promoters elements and two defined transcriptional start sites in close vicinity. Moreover, the RLM-RACE revealed that messages initiating downstream of the P3 promoter utilize the splice donor and acceptor sites of the LTR at position 825 and 1275, respectively.

4.4 In the absence of dexamethasone the P2/P3 promoter complex is more potent than the major P1 promoter

To determine the strength of the MMTV LTR Promoter complex, the activity of the LTR promoter complex was compared to the basal transcription initiation rate of the major P1 promoter. To this end, mRNA was isolated from the MMTV-producer cell clone N2AnE5 and reverse transcribed with the oligonucleotide complementary to the *gag* 5' untranslated region (1443/1424). The use of this primer allows the detection of transcripts initiating from both promoter regions (P2/P3 promoter and P1 promoter, respectively). Subsequently, the first strand cDNAs were subjected to semi-quantitative RT-PCR analysis. Two primer pairs were used simultaneously, which either amplify a short region downstream of the P3 promoter (expected length of 62 bps) or the very 5' end of transcripts originating from the major P1 promoter (expected length of 76 bps). (The primer binding sites are indicated in figure 4.8A). Aliquots were withdrawn from the reaction tube after 15, 20, 25, 30, 35 and 40 cycles. Figure 4.8B shows the amplification products separated on a 2.5% agarose gel.

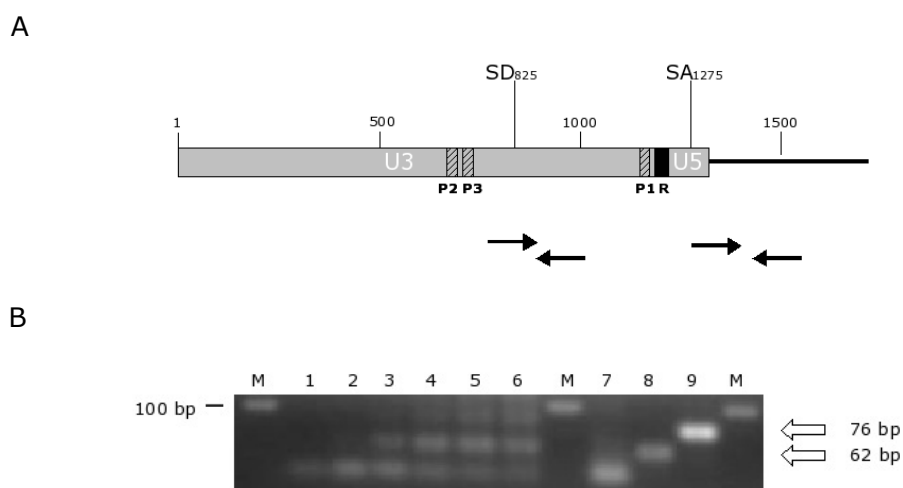


Figure 4.8 Semi-quantitative analysis of transcriptional initiation from the LTR promoters. (A) The 5' LTR and the *gag* 5' untranslated region is shown schematically. The primer binding sites used for determining the initiation rate from both promoter regions of the MMTV LTR are indicated as arrows. (B) First strand cDNA from isolated mRNA of N2AnE5 cells that had been grown in the absence of dexamethasone were used for semi-quantitative PCR analysis (lane 1 to 6). Two primer pairs, specific either for transcripts originating from the P2/P3 promoter complex or from the major P1 promoter were used simultaneously. 10 μ l aliquots were taken at the end of the following cycles: 15 (lane 1), 20 (lane 2), 25 (lane 3), 30 (lane 4), 35 (lane 5) and 40 (lane 6-9). The PCR performed in the absence of template cDNA is shown in lane 7 (negative control). As positive control, a plasmid template was used for the amplification of the P2/P3 transcripts (lane 8) and for the P1 transcripts (lane 9).

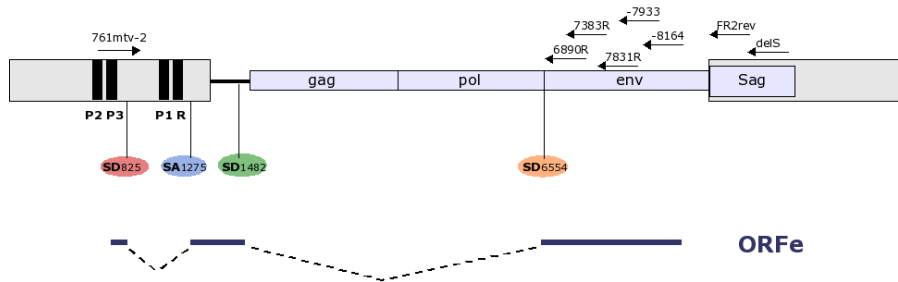
While the amplification products of the promoter complex are visible after cycle 25 (lane 3, figure 4.8B), those of the P1 promoter do not emerge until cycle 30 (lane 4, figure 4.8B). Thus, in the absence of glucocorticoid hormones relatively more mRNA is transcribed from the P2/P3 promoter region than from the major P1 promoter.

4.5 A doubly spliced transcript originates from the promoter complex

In order to facilitate the detection of the putative full length transcript of the central LTR promoter complex, we analysed CrFK cell clones that have been stably transfected with a full-length provirus harbouring a short deletion of the P1 promoter in its 5'LTR. Total RNA was extracted from cells that have been grown in the absence of dexamethasone and reverse transcribed with poly(dT)-primers. Subsequently, first strand cDNA PCRs were performed with a forward primer located immediately downstream of the identified transcription initiation site at position 760 and a collection of reverse primers homologous to the proviral sequence. While

no product was obtained with primers complementary to *gag*- or *pol* coding regions, a strong signal emerged with primers located downstream of the *env* splice acceptor site (SA₆₅₅₄). All amplification products were subcloned and revealed a double spliced message after sequence analysis (figure 4.9).

A



B

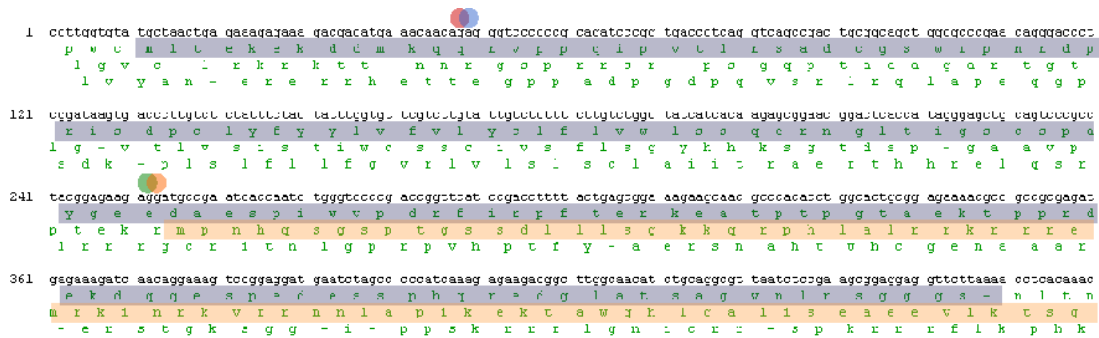


Figure 4.9 A double spliced message originates from the P2/P3 promoter complex. (A) The newly identified transcript is shown in relation to the genome of the MMTV provirus. The active splice donor and acceptor sites are indicated by coloured dots. Exons are shown in blue, while introns are represented by dashed lines. (B) The ORFe reading frame (highlighted in blue) differs from the coding information of the envelope gene (marked in light red). Utilized splicing sites of the ORFe transcript are indicated by coloured dots.

As already shown, the first intron of the amplified transcript is placed within the LTR. The second intron starts at position 1482 (SD₁₄₈₂) and extends until position 6554 (SA_{env}). The complete message was easily recovered using the forward primer in combination with reverse primer (7255/7238). Moreover, we could demonstrate that the message does not depend on the deletion of the major P1 promoter, since the same product was detected in cDNA from GR cells as well as from the MMTV producer clones N2AnE5 and C2AnE15 (data not shown). The sequence carries an ORF of 459 nucleotides, which is translated into a 20 kDa protein of yet unknown function (Rouault *et al.*, 2007). Since the third exon is located within the *env*-coding region, the putative transcript was labelled ORFe.

In summary, the LTR of MMTV harbours two transcriptional active regions, both of which are comprised of two promoter elements: the major promoter P1, which is accompanied by an downstream positioned initiator element (Inr), and the central promoter complex, harbouring the P2 and the P3 promoter. Therefore, we asked whether the activities of the central LTR promoters are regulated in a similar way to that of the major P1 promoter and the nearby located initiator element.

4.6 Determination of the nucleosome occupancy over the P2/P3 promoter complex

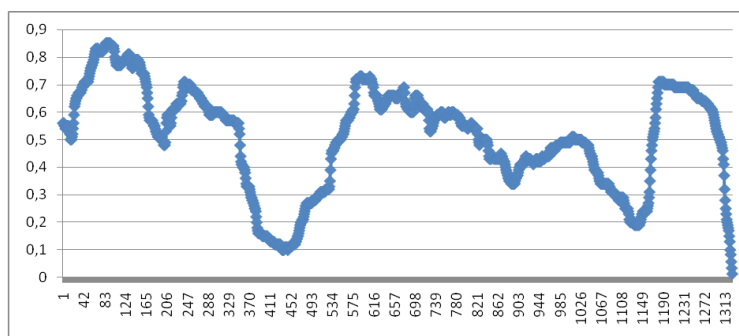
It is well known that upon hormone stimulation, a characteristic nucleosome remodelling process takes place at nucleosome B, which covers the hormone responsive element and the TATA-box of the major P1 promoter. Furthermore, studies have revealed that the neighbouring nucleosomes C and A are involved in an extended chromatin rearrangement as well (Fletcher *et al.*, 2000; Vicent *et al.*, 2004). However, further upstream regions were not investigated during these analyses. We hypothesized that the activity of the central LTR promoters (P2 and P3, respectively) could be regulated by alternating positions of the corresponding nucleosome. Depending on the presence of glucocorticoids, a nucleosome free site could be generated for either of the two promoters to allow transcription to take place. Therefore, we decided to study the position of the nucleosome D, which covers the P2/P3 promoter region on a base pair level of resolution in the presence and in the absence of dexamethasone.

In contrast to regulatory proteins, nucleosomes do not recognize a specific sequencing motif. They must be able to associate with the entire genome in order to facilitate the packaging of the DNA. Hence, this unspecific binding makes it difficult to predict regions occupied by nucleosomes. Nevertheless, genome wide studies revealed two main properties of the DNA that influences nucleosome binding. One characteristic feature is given by a 10 bp periodicity of A/T nucleotides, which facilitates the bending of the double helix around the nucleosome core particle (Satchwell *et al.*, 1986; Widom, 2001). In addition, these studies demonstrated that there are a number of short sequences on the DNA itself that act as nucleosome disfavoured motifs (Segal *et al.*, 2006; Field *et al.*, 2008). Current computer prediction programs take advantage of both features.

4.6.1 *In silico* analysis of the LTR allows multiple positions of nucleosome D

We started by subjecting the LTR of Mtv-2 and Mtv-8, respectively, to the nucleosome prediction model developed by the group of Eran Segal (Segal *et al.*, 2006; Kaplan *et al.*, 2009). Mtv-2 was chosen as a representative of an integrated virus able to produce infectious viral particles, while Mtv-8 served as example of a defective provirus. Figure 4.10A shows the predicted nucleosome occupancy over the Mtv-2 LTR.

A



B

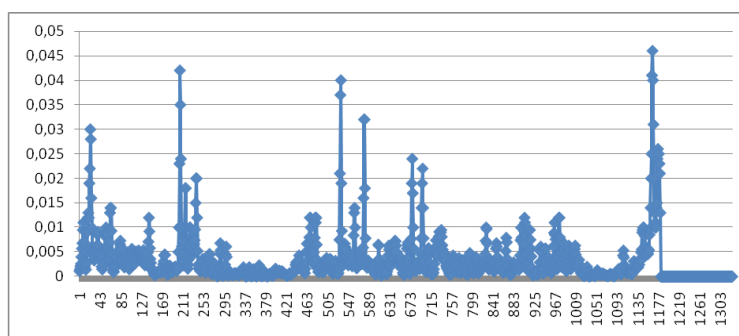


Figure 4.10 *In silico* analysis of the Mtv-2 LTR. (A) The probability of the nucleosomal occupancy over the Mtv-2 LTR is shown. The numbers on the x-axis refer to the base pair position of the LTR, while the corresponding probability is indicated on the y-axis. (B) The graph displays the probability of a specific Mtv-2 LTR sequence to serve as transcriptional initiation site.

As mentioned in section 1.6.1, the MMTV LTR is covered by an array of six positioned nucleosomes, termed F-A (in 5' to 3' direction). The hormone-responsive element and the TATA-box of the major promoter are covered by nucleosome B, while the nucleosome D is occupying the central LTR including the P2/P3 promoter complex.

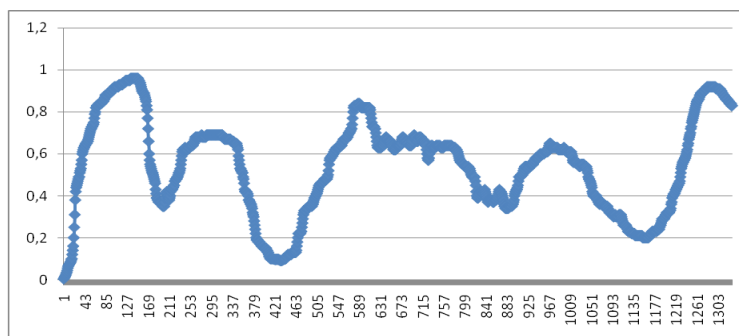
The calculated nucleosomal occupancy over the Mtv-2 LTR revealed distinct positions for the first two nucleosomes in respect to the beginning of the LTR (nucleosomes F and E, according to the literature), since relatively low probabilities are found around position 1-15

and 180-200. These are followed by a strong nucleosome disfavouring sequence of more than 100 bp (positions 370-470). The probability increases unequivocally around bp 534, reaching an internal high at position 600. This is followed by a constant decrease until position 880, which complicate a prediction of the border between the corresponding nucleosomes D and C. Two local low points at bp 890 and at bp 1120, both interrupted by internal highs, indicate the respective positions of the nucleosomes B and A.

Besides the nucleosome occupancy, the model calculates the probability for transcriptional initiation sites of the sequence of interest. The results in regard of the Mtv-2 LTR are shown in figure 4.10B. The highest probability for an initiation site is found at position 1171, which correlates to the known transcriptional start of the major promoter P1. In addition, a cluster of potential start sites is indicated over the central part of the U3 region (position 680-700), where the promoters P2 and P3 have been identified. Interestingly, transcriptional initiation sites are predicted even further upstream: at position 25, around position 210-240 and within the region ranging from position 530 to 580. It is tempting to speculate that the very N-terminal start site indicates initiation from the open reading frame that encodes for Sag. However, no transcriptional activity has as yet been reported from the other two indicated regions.

The predicted nucleosome occupancy over the Mtv-8 LTR, which is shown in figure 4.11A, resembles that calculated from the Mtv-2 sequence (figure 4.10A).

A



B

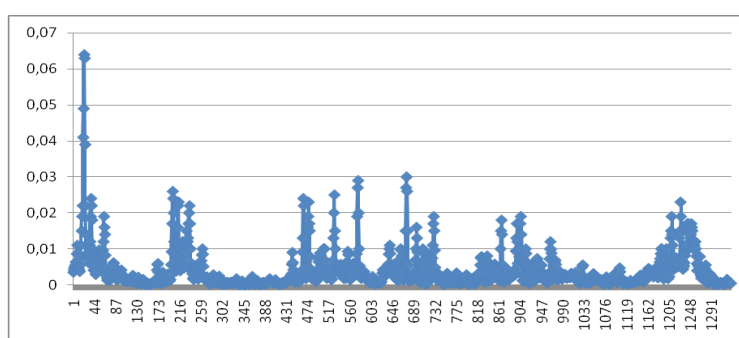


Figure 4.11 Analysis of the Mtv-8 LTR. (A) The probability of nucleosomal occupancy over the Mtv-8 LTR is shown. The numbers on the x-axis indicate the base pair positions relative to the beginning of the LTR. (B) The graph represents the computed transcriptional initiation site distribution over the Mtv-8 LTR.

The positions of the first two nucleosomes relative to the beginning of the LTR are predicted with high probability, followed by a short, but strong disfavouring region of about 100 bp around position 420. Then an almost continuous increase reaches its top level at bp 610. The probability drops about 20 %, but stays at a constant level until position 750, before it decreases further. Again, it is not possible to locate the border between nucleosome D and nucleosome C. In contrast to Mtv-2, there seems to be a second, although less strong, disfavouring region of about 100 bps from position 820 to 920. A broad, lower peak together with a high, but smaller peak allow the localization of the positions of nucleosome B and A, respectively.

In accordance to the nucleosome positions, the predicted transcriptional initiation sites of Mtv-8 resemble those of Mtv-2 (compare figure 4.11B with figure 4.10B). However, a prominent exception is illustrated by the predicted transcriptional start site at the very N-terminus, which appears to be twice as high as the other ones. The viral encoded superantigen is known to be expressed from almost all endogenous MMTV sequences (Scherer *et al.*, 1995). Therefore, this result supports our previous assumption that this peak corresponds to the initiation site of Sag.

In conclusion, it is possible to predict the putative 5' border of nucleosome D that is covering the P2/P3 promoter region, because it is strongly affected by an upstream located nucleosome disfavoured region. However, the nucleosome occupancy calculated merely from the DNA sequence does not allow the determination of the 3' border of nucleosome D.

4.6.2 A nuclease hypersensitive site is detectable at the P2/P3 promoter region

For the *in vivo* study, we stably transfected cells with a plasmid containing only the 5' LTR of Mtv-2 in order to compensate for possible differences in the distribution of the nucleosomes over both LTR regions. Stable populations were produced in the mammary epithelial cell line NMuMG as well in the MMTV permissive feline cat kidney cells CrFK and are referred to as NM5'LTR cells and CK5'LTR cells, respectively.

The successful integration of the plasmid DNA into the host cells was demonstrated by Southern blot analysis. Genomic DNA was isolated from both populations, blotted and subsequently hybridized to digoxigenin (DIG)-labelled probe corresponding to the 5' end of the LTR. Figure 4.12 shows that the Mtv-2 LTR is integrated in transfected CrFK cells (lane 2, figure 4.12) as well as in transfected NMuMG cells (lane 4, figure 4.12); the original plasmid was used as positive control (figure 4.12, lane 5). In contrast to the CrFK cells, which are free of any endogenous MMTV proviruses (lane 1, figure 4.12), a number of endogenous *Mtvs* could be detected in the parental NMuMG cell line (lane 3, figure 4.12). However, none of them is identical to the transfected LTR of Mtv-2 (compare lane 3 with lane 4, figure 4.12).

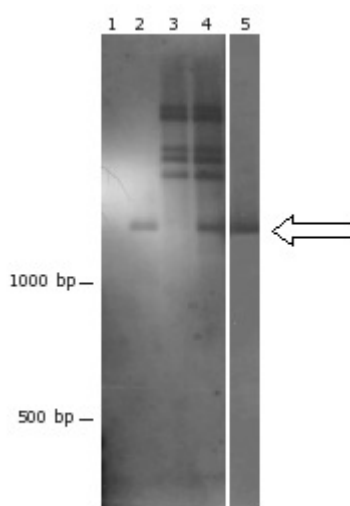


Figure 4.12 Stable integration of the 5'LTR into CrFK and NMuMG cells. Genomic DNA was digested to completion with restriction enzymes BamHI, EcoRI and HindIII, separated on a 1% agarose gel and transferred onto a nylon membrane. Hybridization was performed with a DIG-labelled probe spanning a 500 bp fragment corresponding to the 5' end of the MMTV LTR. The arrow points to the Mtv-2

based LTR of the transfected plasmid. The hybridization results obtained from non-transfected CrFK cells (lane 1), CK5'LTR cells (lane 2), non-transfected NMuMG cells (lane 3) and NM5'LTR cells (lane 4) are shown. Plasmid DNA was used as positive control (lane 5).

In order to ensure the functionality of the integrated single LTR, we determined the level of transcripts initiating from both promoter regions by semi-quantitative RT-PCR analysis. To compensate for differences in the nucleosome occupancy during transcription and mitosis, the cells were synchronized in their cell cycle. This was achieved by arresting the cells using serum starvation for 48 hours. At the end of the starvation period, the cells were allowed to reenter the cell cycle by the addition of normal growth medium, with half of the cells being stimulated with dexamethasone. Total RNA was extracted after 30 min and 120 min and reverse transcribed with an oligonucleotide corresponding to the *gag* 5' untranslated region (1443/1424); subsequently, RT-PCRs were performed. The forward primers were located downstream of either the transcriptional initiation site of the major P1 promoter (1197/1219) or the transcriptional start site of the P3 promoter (773/795).

Figure 4.13A shows the amplification products originating from the P1 promoter. An increase of its transcriptional activity is found in response to hormonal stimulation after 30 min (lane 1) as well as after 120 min (lane 3) compared to the basal levels detected in untreated cells (lane 2 and 4, respectively). In contrast, no differences in the amount of transcripts initiating at the central promoter complex are detectable in response to glucocorticoid induction. Moreover, the amount of transcripts stayed constant over the whole time period investigated (lane 1 to 4, figure 4.13B). Results with NM5'LTR cells are shown, although similar pictures were also obtained with CK5'LTR cells.

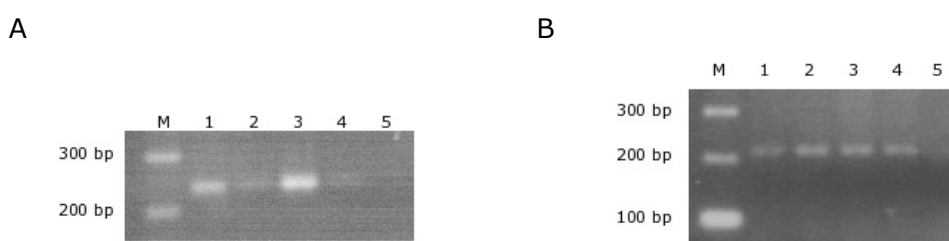


Figure 4.13 Transcriptional initiation from the LTR promoters in NM5'LTR cells. NM5'LTR cells were growth arrested by serum starvation for 48 hours. Total RNA was extracted after the cells were allowed to reenter the cell cycle for 30 min and 120 min, respectively. Half of the cells of each time point were stimulated for the same period with dexamethasone. After reverse transcription, RT-PCRs were performed with primers located downstream of the initiation site of the P1 promoter (A) or the central promoter complex (B). The amplification products of synchronised NM5'LTR cells treated with dexamethasone (lane 1) or without any hormones (lane 2) for 30 min and those of NM5'LTR cells treated with dexamethasone (lane 3) or without any hormones for 120 min (lane 4) are shown. Lane 5

represents RT-PCRs performed in the absence of a cDNA template.

We decided to focus on CK5'LTR cells in order to prevent any influence originating from endogenous *Mtv* sequences as this could not be excluded in a NMuMG background. Moreover, the strong difference between untreated and dexamethasone induced cells regarding the transcriptional activity of the major promoter after 120 min prompted us to choose this time point for further analysis.

The strategy for the determination of the nucleosomal borders was adopted from Soutoglou and Taliandis (Soutoglou and Taliandis, 2002) and is shown schematically in figure 4.14.

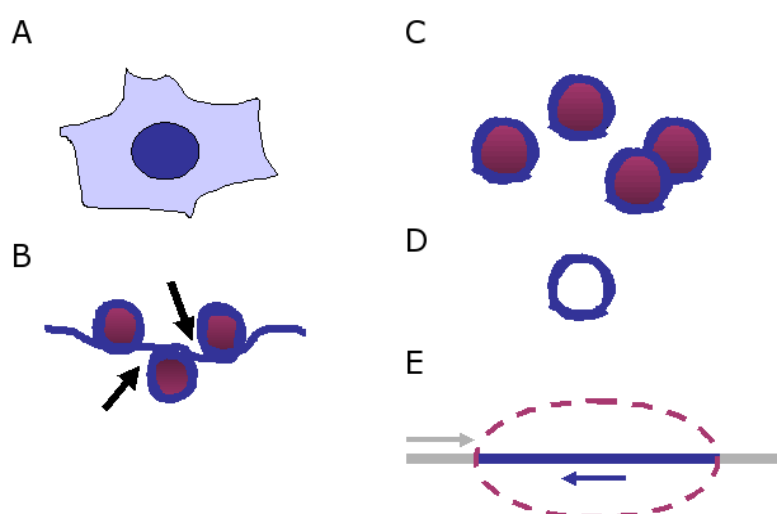


Figure 4.14 Determination of the nucleosomal borders via LM-PCR. The steps for the analysis of the position of a specific nucleosome are shown schematically. First, the chromatin is isolated from the cells of interest (A) and treated with micrococcal nuclease (B) to achieve mononucleosomal particles (C). The histone proteins are digested by proteinase K treatment and the remaining DNA fragments are electrophoretically separated (D). Finally, those fragments that have been protected by a single nucleosome are extracted and ligated to the double stranded linker-oligonucleotide. The subsequent PCR analysis is performed with a linker-specific primer (grey arrow) and a sequence-specific primer (blue arrow) (E).

The first steps included the isolation of the chromatin (figure 4.14A), which was subsequently digested with micrococcal nuclease (MNase), an enzyme that cuts preferentially inter-nucleosomal DNA (figure 4.14B). Then, the generated mono- and multinucleosomal particles (figure 4.14C) were digested with proteinase K to remove histone as well as associated non-histone proteins. Afterwards, the fragmented DNA (figure 4.14D) was

separated electrophoretically and those fragments that had been protected by a single nucleosome (appropriate length of 147 bps) were extracted and purified. Before ligating the asymmetric linker oligonucleotide to the MNase-cleaved substrate, it was necessary to phosphorylate the 5' hydroxyl-groups that were generated by the enzyme. The ligation-mediated (LM-) PCRs were performed with a primer pair comprising of a linker- and a sequence-specific primer (figure 4.14E).

Prior to the analysis, the determination of the optimal concentration of the enzyme was required to prevent overdigestion and thus the generation of subnucleosomal particles, which would otherwise distort the interpretation of the results. Figure 4.15 shows one representative agarose gel of chromatin from synchronized CK5'LTR cells treated with suitable amounts of micrococcal nuclease.

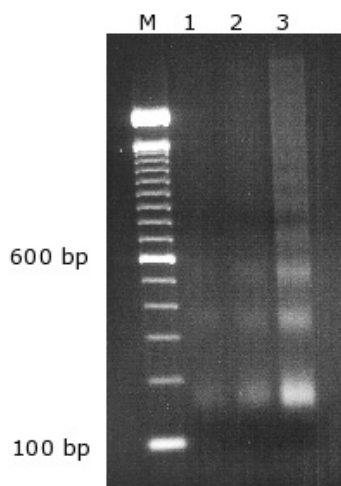


Figure 4.15 Agarose gel electrophoresis of purified DNA fragments after micrococcal nuclease digestion of the respective chromatin. Chromatin was extracted from synchronized CK5'LTR cells and digested with micrococcal nuclease. The respective DNA fragments were purified and separated by agarose gel electrophoresis. 2 μ l (lane 1), 5 μ l (lane 2) and 10 μ l (lane 3) of the sample were applied to the gel.

A characteristic laddering of the DNA is visible, comprising of the mononucleosomal fragments together with their multiples. These are generated when one or more consecutive inter-nucleosomal regions escaped digestion. In addition, this analysis allows the determination of the proper quality of the nucleosomal preparation, which is guaranteed if six or more distinct bands are visible.

Independent preparations of mononucleosomes made from synchronized CK5'LTR cells that had been either stimulated with glucocorticoids or left untreated were subjected to LM-PCR analysis. The use of an MMTV-specific forward (645/672) and reverse primer

(726/703), each in combination with an oligonucleotide corresponding to the reverse transcribed linker sequence, allowed the determination of the position of the nucleosome D that is covering the P2/P3 promoter region.

Figure 4.16 shows the amplification products of the *in vivo* analysis of the position of nucleosome D. Irrespective whether the cells have been stimulated with dexamethasone or left untreated, two prominent bands appeared when the 5' border of the nucleosome was analysed (figure 4.16A, left picture). In contrast, a single product was amplified using a MMTV-specific primer for determining the 3' end of the nucleosome (figure 4.14A, right picture).

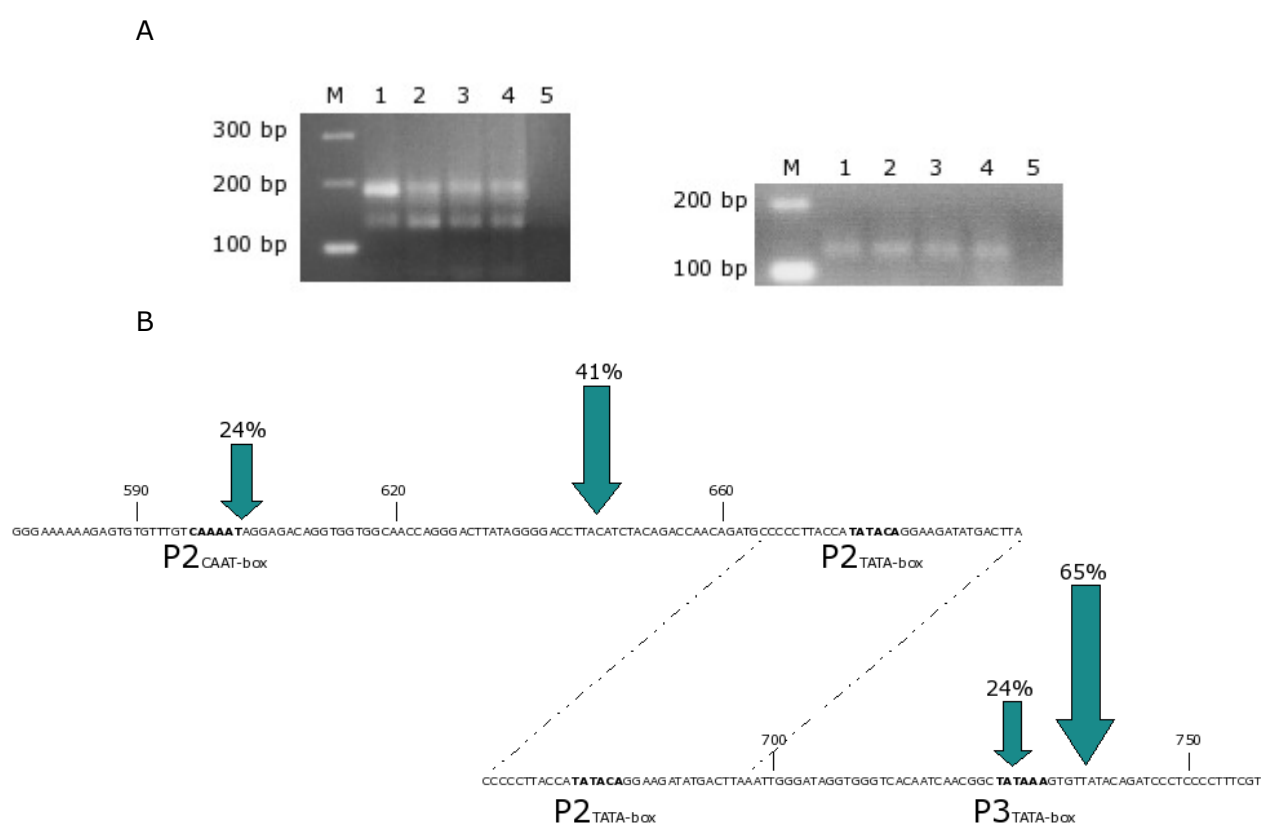


Figure 4.16 Determination of the nucleosome D positioning in CK5'LTR cells. (A) The LM-PCRs results for determining the 5' border (left) and the 3' border (right) of the nucleosome D (covering the P2/P3 promoter region) after electrophoretic separation are shown. Lane 1: CK5'LTR cells stimulated with dexamethasone, preparation I; lane 2: CK5'LTR cells left untreated, preparation I, lane 3: CK5'LTR cells stimulated with dexamethasone, preparation II; lane 4: CK5'LTR cells left untreated, preparation II. Lane 5 shows the LM-PCR performed in the absence of any DNA template. (B) The sequence of the P2/P3 promoter region is shown. The arrows indicate the MNase cleavage sites obtained after sequence analysis of at least 30 mini-preps that were made after subcloning the LM-PCR products. Only digestion sites that were hit at least three times were included in the interpretation of the analysis.

The bands were extracted from the gel, purified and sequenced after an additional

subcloning step. The analysis is presented in figure 4.16B. The total amount of samples (each representing an individual clone) that contained both primer sequences and aligned to the MMTV LTR were set at 100 %. Only those digestion sites that occurred at least three times were included in the analysis of the nucleosomal borders.

Two clusters of nuclease digestion sites corresponding to the two amplified fragments were detected at the 5' end of the nucleosome. The upstream located site overlaps with the CAAT-box of the P2 promoter, while the centre of the downstream located site is found in between the CAAT-box and the TATA-box of the P2 promoter. At the 3' border of the nucleosome, the nuclease digestion sites were distributed around the TATA-box of the P3 promoter.

In summary, the more upstream located 5' nuclease digestion site corresponds together with the detected 3' border approximately to the theoretical size of a nucleosomal DNA fragment (147 bps). Moreover, both sites lie in close vicinity of the nucleosome D positions that were originally identified by Richard-Foy and Hager (Richard-Foy and Hager, 1987). Therefore, we conclude that the nuclease digestion sites around position 644 represent an additional hypersensitive site that indicates an open chromatin formation at nucleosome D. Interestingly, both, the CAAT-box of the P2 promoter as well as the TATA-box of the P3 promoter are located at the very end of the nucleosome, a property which increases the accessibility of these sequences to be bound by putative transcription factors.

However, we were unable to detect any difference of the nucleosomal position in dependence of glucocorticoid treatment.

4.7 Elevated methylation levels accompany transcription from the P2/P3 promoter complex

Besides the positioning of the nucleosomes, post-translational modifications of histone proteins have a major impact on the transcriptional activity of the underlying DNA. These modifications influence the chromatin structure either directly by altering DNA-histone interactions or indirectly by providing binding sites for non-histone proteins (Thorne *et al.*, 2009). The constant rate of transcription initiation from the P2/P3 promoter complex as well as the results from the nucleosome positioning analysis prompted us to investigate associated histone modifications by chromatin immunoprecipitation (ChIP).

Briefly, the chromatin was isolated from the cells of interest and digested with a suitable enzyme into parts of appropriate length. After a preclearing step to minimize unspecific binding events, antibodies against the modifications of interest were allowed to bind

to their corresponding epitopes. Then, protein G beads, which adhere to the immunoglobulins due to their high affinity to the Fc (Fragment crystallizable) region, were added to the solution. After several washing steps, the bound fraction (comprising of the antibody associated to specific subgenomic regions of the chromatin) was eluted. Subsequently, the corresponding DNA-fragments were isolated and purified. The last step included a qRT-PCR analysis in order to allocate the investigated modifications to their corresponding regions on the DNA and to allow a comparison of their relative amounts.

For the isolation of the chromatin and the generation of subgenomic units, the same protocol was used as for the determination of the nucleosomal borders with the exception that this time we wanted to achieve a complete digestion of the chromatin (i.e. mononucleosomal fragments). Synchronized CK5'LTR cells grown in the absence of hormones were used as starting material for the ChIP analysis. Figure 4.17 shows the digested chromatin after extraction of the DNA and separation on a 2% agarose gel.

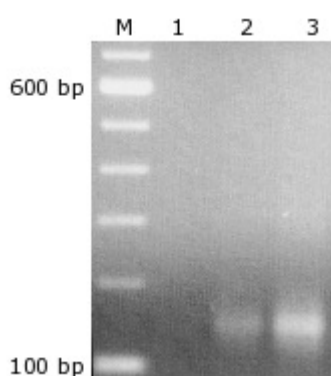


Figure 4.17 Mononucleosomal DNA fragments obtained after complete digestion with micrococcal nuclease. Agarose gel electrophoresis of DNA fragments which were purified after nuclease digestion of isolated chromatin of CK5'LTR cells. 2 µl (lane 1), 5 µl (lane 2) and 10 µl (lane 3) of the sample have been loaded onto the gel.

Only mononucleosomal fragments (147 bp) are present in the preparation; therefore, a direct comparison of the relative amounts of the respective modifications between nucleosome D, which covers the P2/P3 promoter complex, and nucleosome B, which occupies the hormone-responsive sites of the P1 promoter, would be possible.

A marker of choice for active chromatin represents the trimethylation of lysine 4 on histone H3 (H3K4m3) (Ng *et al.*, 2003; Pokholok *et al.*, 2005). Although lysine methylation itself has no inherent effect on the histone charge, its predominant role lies in the recruitment of additional factors (Fuchs *et al.*, 2009). Independently performed experiments revealed an equal distribution of this specific modification over the nucleosomes B, C and D. The results are represented in figure 4.18 as percent of the input-sample.

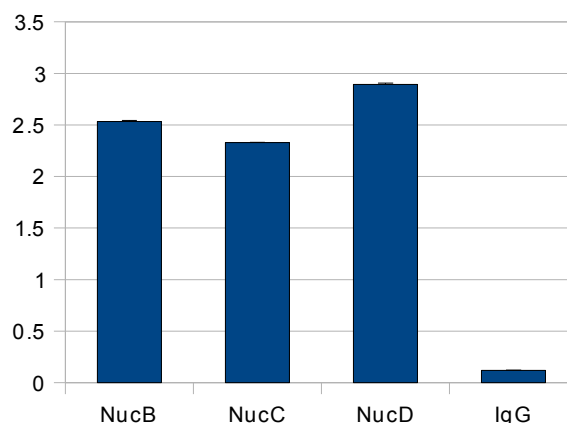


Figure 4.18 Active chromatin formation is found equally distributed over the central nucleosomes of the MMTV LTR. ChIP analyses were conducted in synchronized CK5'LTR cells with an antibody targeting methyl-groups on histone H3 (H3K4m3 antibody, Cell Signaling Technology). The IgG antibody was used as control. The relative amounts of the respective DNA were determined by qRT-PCR with primers specific for each nucleosome. The results are represented as percentage of the input-sample which was set to 100%. The mean of at least three independent experiments is shown.

Although H3K4m3 is known to promote transcriptional elongation and regulate RNA processing events (Santos-Rosa *et al.*, 2002; Sims *et al.*, 2007), the same modification has been shown to be able to recruit repressing protein complexes as well (Berger, 2007).

Therefore, we additionally investigated the distribution of trimethylation at lysine 36 of histone 3 (H3K36m3). This modification is catalysed by the histone methyltransferase Set2, which specifically binds to the RNA polymerase II. Thereby, H3K36m3 is directly linked to the passage of RNA polymerase II and thus acts as potent sign of transcriptional elongation (Berger, 2007; Fuchs *et al.*, 2009).

The results of the ChIP analysis performed with an antibody against H3K36m3 are shown in figure 4.19. In the absence of glucocorticoid stimulation, the highest levels of these modifications are found at nucleosome D, which occupies the P2/P3 promoter region.

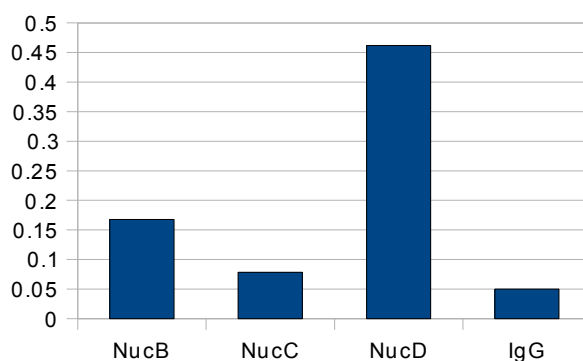


Figure 4.19 A marker for transcriptional elongation is found predominantly at nucleosome D.

Chromatin immunoprecipitation of CK5'LTR cells was performed with an antibody targeting H3K36m3 (Cell Signaling Technology). The IgG antibody was used as control. The relative amount of the respective DNA was determined by qRT-PCR with primers specific for the nucleosome B, C or D regions. All results are depicted as percentage of the input-sample which was set to 100%. The mean of at least three independent experiments is shown.

The results correspond to the findings obtained by semi-quantitative RT-PCR (figure 4.8), which indicated that without dexamethasone induction, relatively more RNA is transcribed from the promoter complex than from the major P1 promoter.

4.8 Post-translational acetylation inversely correlates with ORFe transcription

Subsequent to the passage of the DNA-polymerase, the initiation signals have to be suppressed to prevent an otherwise constitutive activation of the respective gene. This task is mediated by the recruitment of enzymes that possess repressive deacetylase activities (Joshi and Struhl, 2005). Thus, we investigated in the following step the impact of the acetylation status of the corresponding histone proteins on the transcriptional activity of the central promoter complex.

According to the commonly accepted model, the acetylation of histones neutralizes their positive charge. Thereby, the tight association with the negatively charged DNA backbone is outweighed and gene activation is allowed to take place. However, a few exceptions (*c-myc*, cyclin D1) have been described (Van Lint *et al.*, 1996; Lallemand *et al.*, 1996). In regard of the MMTV LTR, controversial results have been reported. Several groups detected an activation dependent increase in histone acetylation over the MMTV LTR (Lambert and Nordeen, 2003; Li *et al.*, 2003; Astrand *et al.*, 2009) while others notified deacetylation events upon induction

(Sheldon *et al.*, 2001; Mulholland *et al.*, 2003; Wilson *et al.*, 2002; Aoyagi and Archer, 2007). As no one has separately analysed the effect of acetylation on the major promoter and the central promoter complex, we thought that this approach might help to explain previous observations.

In a preliminary experiment, synchronized CK5'LTR cells were treated with increasing concentrations of the deacetylase inhibitor trichostatin A (TSA). Total RNA was isolated, reverse transcribed and transcription initiation was analysed either at the P2/P3 promoter complex or at the P1 promoter/Initiator region. The results of the RT-PCRs are shown in figure 4.20.

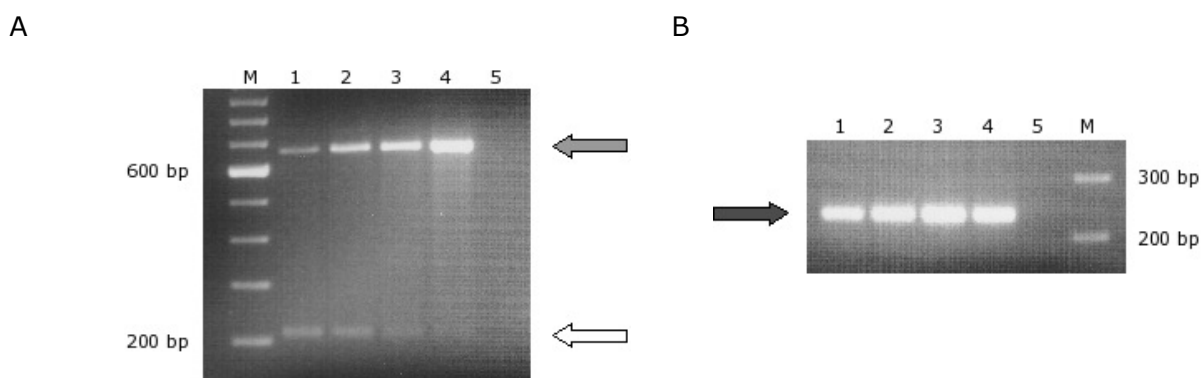


Figure 4.20 Effect of trichostatin A on the transcriptional initiation rate of the LTR promoters.

Synchronized CK5' cells were treated for 18 hours with increasing concentrations of TSA. Subsequently, total RNA was reverse transcribed with poly(dT) primers and transcriptional initiation was analysed by RT-PCR with oligonucleotides specific for the P2/P3 promoter region (A) and the P1 promoter (B), respectively. The amplification products of untreated cells (lane 1), cells incubated with 5 ng/ml TSA (lane 2), cells incubated with 50 ng/ml TSA (lane 3) and cells incubated with 100 ng/ml TSA (lane 4) are shown. RT-PCRs performed without any cDNA template are shown in lane 5.

While no effect was observed after 2 hours of incubation (data not shown), overnight treatment lead to a shift in the nature of transcriptions initiating at the P2/P3 promoter complex. The amount of unspliced transcripts accumulated upon increasing TSA concentrations (grey arrow, figure 4.20A), while at the same time the level of spliced messages decreased (white arrow, figure 4.20A). In contrast, only a slight change was detected in the basal transcriptional activity of the P1 promoter (figure 4.20B). We concluded that transcription from the central promoter complex is regulated by acetylation and at least one level of regulation does involve the process of RNA splicing.

In the following step, we asked whether we could identify the mechanism of repression by investigating the acetylation level of the corresponding nucleosomes. Chromatin immunoprecipitations of synchronized CK5'LTR cells were performed, which were either left

untreated or incubated with 100 ng/ml TSA for 18 hours, the highest tolerated repressing dose. To prevent any deacetylation events during the procedure, all buffers and reagents contained 5 mM sodium butyrate. Antibodies against the acetylated histone H3 as well as antibodies recognizing acetyl-groups on either lysine 8 or lysine 12 of histone H4 were used. The amount of immunoprecipitated DNA was determined by qRT-PCR using primers for nucleosome B and D, respectively. Figure 4.21 shows the results obtained for nucleosome B (panel A) and nucleosome D (panel B).

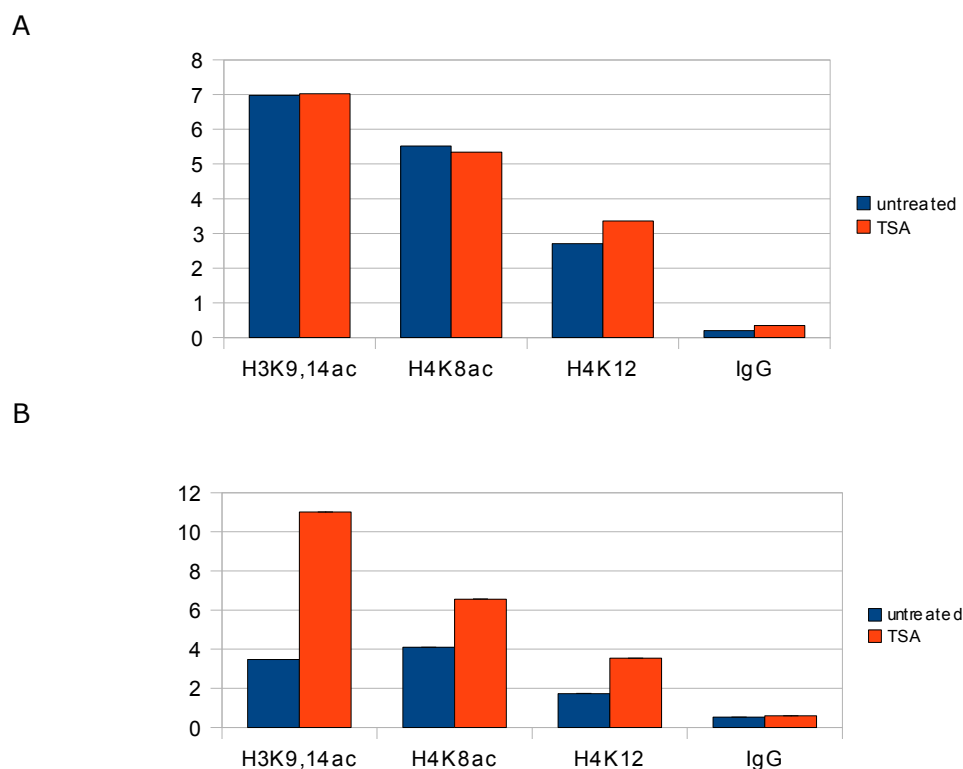


Figure 4.21 Effect of TSA treatment on the acetylation level of nucleosome B and nucleosome D. Chromatin immunoprecipitation of CK5'LTR cells either left untreated (blue bars) or incubated with 100 ng/ml TSA (red bars) were performed against acetyl-groups of histone H3 (H3K9,4Ac antibody, Cell Signaling Technology) and histone H4 (H4K8Ac antibody and H4K12Ac antibody, both from Cell Signaling Technology). The IgG antibody was used as control. The relative amount of the respective DNA was determined by qRT-PCR with primers specific for nucleosome B (A) and nucleosome D (B), respectively. All results are depicted as percentage of the input-sample which was set to 100%.

In the absence of hormones, we detected no change of the nucleosome B acetylation status in response to the TSA treatment (figure 4.21A). This finding corresponds to the constant rate of transcription initiation from the major P1 promoter under these conditions (figure 4.20B). In contrast, an overall increase in acetyl-groups of histone H3 as well as of histone H4 following the incubation with TSA was detected at nucleosome D, which occupies

the P2/P3 promoter region (figure 4.21B). Thus, the downregulation of ORFe specific transcripts upon TSA is accompanied by increasing levels of histone acetylation.

Consequently, their upregulation would require the presence of a histone deacetylase complex. The mammalian HDAC superfamily consists of 11 proteins which share a highly conserved deacetylase domain. According to their homology with yeast proteins, they can be classified into four families, which differ in structure, subcellular localization and expression patterns (Haberland *et al.*, 2009). Class I HDACs (HDAC1, 2, 3 and 8) are localized predominantly in the nucleus and exert a high enzymatic activity toward histone substrates. HDAC1 and HDAC2 are nearly identical and generally found together in repressive complexes. HDAC3 is recruited to distinct complexes, while HDAC8 possesses no specificity at all (Haberland *et al.*, 2009). In contrast, class IIa (HDAC 4, 5, 7 and 9) specifically bind to the transcription factor myocyte enhancer factor 2 (MEF2) and to the chaperone protein 14-3-3. The enzymes of class IIb (HDAC 6,10) are the main cytoplasmic deacetylases found in mammalian cells. Little is known about HDAC11, the only member of class IV HDACs.

Given their known association to histone proteins, we concentrated on determining the level of class I HDACs. We performed a chromatin immunoprecipitation with an antibody against the most prominent member HDAC1. The results are represented in figure 4.22.

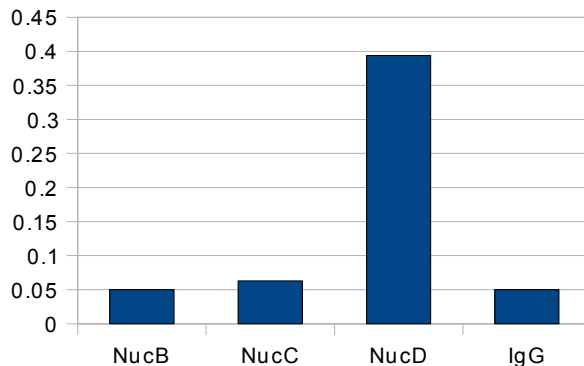


Figure 4.22 HDAC1 is predominantly associated with nucleosome D. Chromatin immunoprecipitation of CK5'LTR cells was performed with an anti-HDAC1 antibody (Cell Signaling Technology). The IgG antibody was used as control. The relative amount of the respective DNA was determined by qRT-PCR with primers specific for the nucleosome B, C or D regions. All results are depicted as percentage of the input-sample which was set to 100%.

Indeed, eight times more HDAC1 complexes were found associated with nucleosome D compared to nucleosome B or nucleosome C regions, suggesting a tight association of a potential transcriptional regulator to the P2/P3 promoter region. However, HDACs lack any DNA-binding domain, depending on another protein in order to exert their activities in a certain

genomic context.

4.9 CDP is a repressor of the P2 promoter

Within the MMTV LTR, a putative binding partner of these enzymes has been identified by the group of J. Dudley (Bramblett *et al.*, 1995). When they were characterizing proteins that bind to the NRE, they identified two major complexes. These included either the transcription factor CCAAT displacement protein (CDP) or the special AT-rich sequence binding protein 1 (SATB1) (Liu *et al.*, 1997). The latter exhibits a tissue specific expression pattern and has been suggested to be responsible for the suppression of MMTV expression in T-cells (Liu *et al.*, 1997 & 1999); in contrast, CDP is expressed ubiquitously (Sansregret and Nepveu, 2008). It functions either by competing with transcriptional activators for the common binding site, referred to as active repression, or by recruiting the histone deacetylase HDAC1 enzyme, a phenomenon known as passive repression (Nepveu, 2001).

At least seven CDP binding sites have been mapped throughout the MMTV LTR (Zhu *et al.*, 2000; Zhu and Dudley, 2002). Indeed, three of those are located in close vicinity of the P2 promoter elements. Two strong binding sites overlap with the CAAT-box and the transcriptional initiation site of P2, respectively, whereas a third weaker motif is found near the P2 TATA-box (figure 4.23).

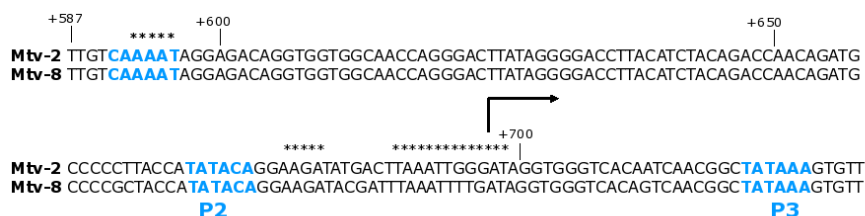


Figure 4.23 CDP binding site distribution over the P2/P3 promoter region. The sequences of the P2/P3 promoter complex of the Mtv-2 provirus (top line) and of the Mtv-8 provirus (bottom line) are shown. The CAAT-box and the TATA-box of the P2 promoter as well as the TATA-box of the P3 are highlighted in blue. The P2 transcriptional start sites is represented by an arrow. CDP binding sites, determined by Dudley and colleagues (Zhu and Dudley, 2002) are marked with asterisks. The base pairs are depicted relative to the beginning of the LTR.

From the above discussion, we assumed that CDP not only acts as global repressor of MMTV as shown by others (Zhu *et al.*, 2000 & 2004; Zhu and Dudley, 2002), but additionally as a specific negative regulator of the P2 promoter. To investigate this notion, the influence of CDP was prevented by introducing point mutations into the two strong binding

sites of P2 (abbreviated in figure 4.25 as m596 and m693, respectively). Furthermore, the P3 TATA-box was mutated to prevent possible interfering transcription from the neighbouring promoter (figure 4.25).

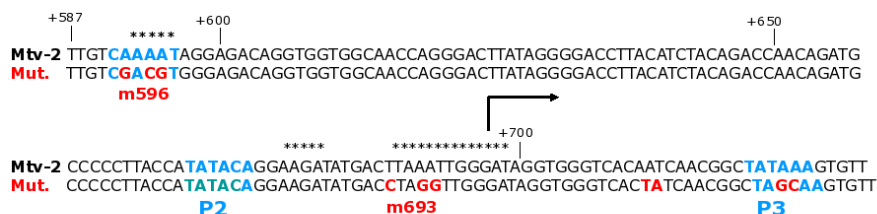


Figure 4.25 Introduced point mutations of the CDP binding site within the central U3 region.

The sequences of the P2/P3 promoter complex of the Mtv-2 provirus before (top line) and after mutation (bottom line) are shown. The CAAT-box and the TATA-box of the P2 promoter as well as the TATA-box of the P3 promoter are highlighted in blue, while the introduced mutations are shown in red. Putative CDP binding sites are represented by asterisks.

In addition, transcription from the P1 promoter was abolished by introducing a short deletion encompassing the P1 TATA-motif and the R initiator element. The so mutated LTR was inserted upstream of the luciferase gene in the pZluc vector (pZgzP2⁺delR_dCDP-). For comparison, we used a plasmid containing the same promoter mutations, yet the CDP binding sites were left unmodified (pZgzP2⁺delR). The transcriptional activity of each construct was analysed in CrFK and NMuMG cells, respectively. After transient transfection, half of the cells were stimulated with dexamethasone for 24 hours, as the activity of P2 has been described to be enhanced by glucocorticoids (Günzburg *et al.*, 1993). After a further 24 hours of incubation, the luciferase activity was measured. Figure 4.26 shows the outcome of one representative analysis.

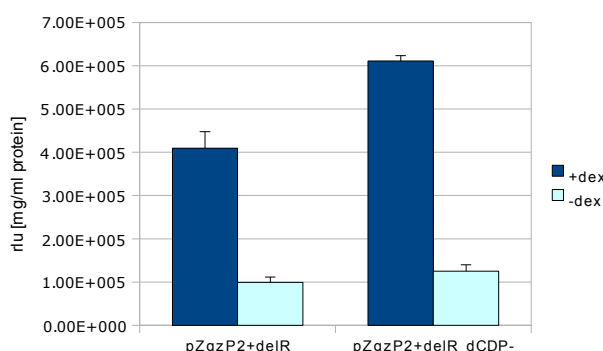


Figure 4.26 Effect of the mutation of the CDP binding sites on P2 promoter activity. NMuMG cells were transiently transfected with the indicated plasmids. The transcriptional potential of the MMTV LTR

containing only a functional P2 promoter (pZgzP2⁺delR) was compared to the activity of the P2 promoter in the presence of CDP binding site mutations (pZgzP2⁺delR_dCDP-). Half of the cells were incubated with dexamethasone for 24 hours (blue bars) while the others were left untreated (turquoise bars). The luciferase expression was measured 48 hours post transfection. Error bars represent the standard deviation of three independent experiments.

In the presence of dexamethasone, the luciferase expression driven by the P2 promoter increased when the CDP binding sites were inactivated by mutation (compare construct pZgzP2⁺delR with construct pZgzP2⁺delRdCDP-, blue bars, figure 4.26). The same effect occurred without hormonal stimulation, albeit to a much lower extent (turquoise bars, figure 4.26). One explanation for the low expression levels that were observed in the absence of dexamethasone could be that they actually reflect the transcription levels of the uninduced promoter. Besides, CDP has been described to lose its DNA-binding activity depending on the differentiation status of the cells (Zhu *et al.*, 2000). Therefore, differences in the amount of the biological active form of CDP between the hormone stimulated (i.e. differentiated cells) and the uninduced (i.e. undifferentiated) state of the respective cell could further distort the effect of the introduced binding site-mutations.

In order to test the latter hypothesis, we decided to determine the endogenous levels of the transcription factor by immunodetection. We repeatedly tried to quantify the amount of CDP in the cell lines used; nevertheless, we did not succeed due to the extremely low sensitivity level of the antibodies. Moreover, the cell lines we used for studying the regulation of the P2/P3 promoter complex have been selected due to their potential for expressing high amounts of viral proteins. Consequently, they should possess low endogenous CDP levels.

Hence, to ensure equal amounts of the transcription factor, we decided to express the protein exogenously. Previous studies have shown that the activity of CDP is elevated in proliferating cells and decreases during terminal differentiation (Nepveu, 2001). We started with the isolation and reverse transcription of messenger RNA from different cell lines. Numerous RT-PCRs were performed, but due to the complexity of the gene (CDP possesses more than 33 exons and is organized over a region of over 340 kb (Zeng *et al.*, 2000)), we had to overcome major difficulties.

Nevertheless, we were able to amplify the 110 kD isoform (p110) from NIH3T3 cells (mouse embryonic fibroblasts). Under physiological conditions, this N-terminal truncated protein is generated in a cell cycle dependent manner via proteolytic cleavage of the full length CDP polypeptide (p200) (Moon *et al.*, 2001; Truscott *et al.*, 2004; Goulet *et al.*, 2004). Hence, the removal of the auto-inhibitory domain and Cut repeat (CR) 1 changes the DNA binding specificity and kinetics of the protein. CDP isoforms containing the CR1 domain bind rapidly but

only transiently to DNA. They prefer a distinct nucleotide repeat sequence than isoforms lacking CR1 and carry the CCAAT-displacement activity (Truscott *et al.*, 2004). In contrast, p75, p90 and p110 exhibit much slower DNA binding kinetics and, depending on the context, either repress or activate transcription (Sansregret and Nepveu, 2008). Moreover, the C-terminal domain has been shown to recruit the HDAC1 deacetylase (Li *et al.*, 1999).

Figure 4.27 shows a western blot detecting CDP/p110 in transfected NMuMG cells, demonstrating that CDP/p110 can be expressed successfully in eukaryotic cells. According to its function as transcription factor, it is found exclusively in the nuclear fraction.

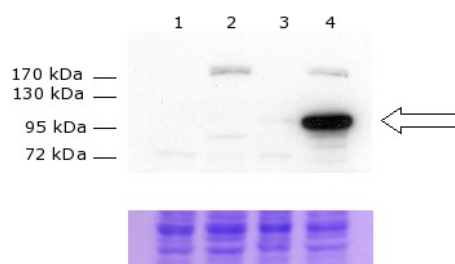


Figure 4.27 Immunoblot showing the expression of CDP/p110 in nuclear extracts of NMuMG cells. NMuMG cells were transfected with the pCMV110 expression vector. Nuclear and cytoplasmic fractions (see methods) were separated on a 10% SDS-PAGE. The protein was detected by immunoblotting using an anti-CDP antibody (M-222, Santa Cruz Biotechnology). Lane 1: cytoplasmic extract of non-transfected NMuMG cells, lane 2: nuclear extract of non-transfected NMuMG cells, lane 3: cytoplasmic extract of NMuMG cells transfected with pCMV110, lane 4: nuclear extract of NMuMG cells transfected with pCMV110. The coomassie-staining of the gel is shown below.

In the following experiment, the reporter gene driven by the P2 promoter in the context of the full-length LTR (pZgzP2⁺delR) or in combination with the NRE (pZ150) was co-transfected with equal amounts of the CDP/p110 expression vector. The luciferase expression was measured 48 hours later. This time, the cells were not stimulated with dexamethasone to avoid any influence of the hormonal treatment on the activity of CDP. Figure 4.28 shows a decrease of P2 activity in the presence of CDP/p110 (pCMV110). A downregulation was observed in the context of the full length LTR (figure 4.28A) as well as in the situation when the P2 promoter elements are solely under the control of the NRE (figure 4.28B).

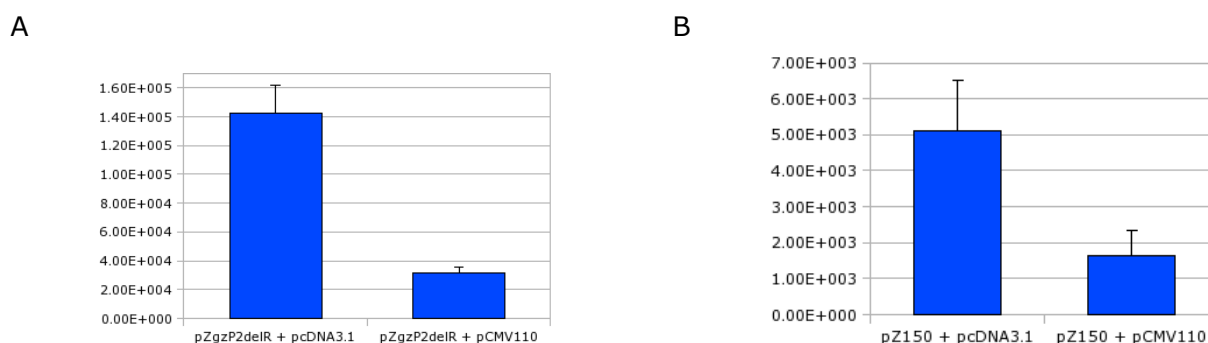


Figure 4.28 Influence of CDP/p110 on P2 promoter activity. The relative luciferase activities driven by the P2 promoter in the presence of the expression vector (pcDNA3.1) and under the influence of exogenously expressed CDP/p110 (pCMV110), respectively, are shown. (A) Analysis of the luciferase activity in the context of the full length MMTV LTR. Mutations of the P3 promoter and a deletion comprising the P1 promoter and the R initiator element have been introduced. (B) Analysis of the luciferase activity of the LTR fragment comprising the minimal P2 promoter and the NRE.

In conclusion, this experiment shows that the activity of the P2 promoter is downregulated if CDP binds to the cognate sites in the vicinity of the promoter. Consequently, the introduction of these binding site mutations (m596 and m693) into the proviral context should allow the identification of the corresponding transcriptional product of P2.

4.10 Constitutive active P2 promoter gives rise to ORFe mRNA

We demonstrated earlier that, independent of hormonal stimulation, a double-spliced transcript (ORFe) initiates from the P2/P3 promoter complex (Rouault *et al.*, 2007). Furthermore, we showed that the P3 promoter functions independently of the NRE, while under the same conditions, the activity of the P2 promoter decreases. Therefore, we concluded that ORFe is transcribed from the constitutive active P3 promoter.

In order to identify the transcriptional product of the active P2 promoter, we inserted the 5' LTR containing the CDP binding site mutations (m596 and m693, respectively) and the mutated P3 TATA-box into the full length proviral Mtv-2/Mtv-8 hybrid containing an EGFP expressing cassette within the 3' LTR. The P1 promoter as well as the initiator element were left unmodified to ensure proper transcription of all other viral related RNA species. Two 100 % EGFP-expressing clones that were produced in NMuMG cells, were chosen at random and their total RNA was isolated and reverse transcribed with poly(dT) primers. First, the functionality of the P2 promoter in the proviral context was tested by performing RT-PCRs with a forward primer binding immediately downstream of the P2/P3 promoter complex and a reverse primer

corresponding to the *gag* 5' untranslated region (figure 4.29A). As can be seen in figure 4.29B, the nature of transcripts differed in the clones analysed. While spliced messages prevailed in clone A, unspliced transcripts were found predominantly in clone B.

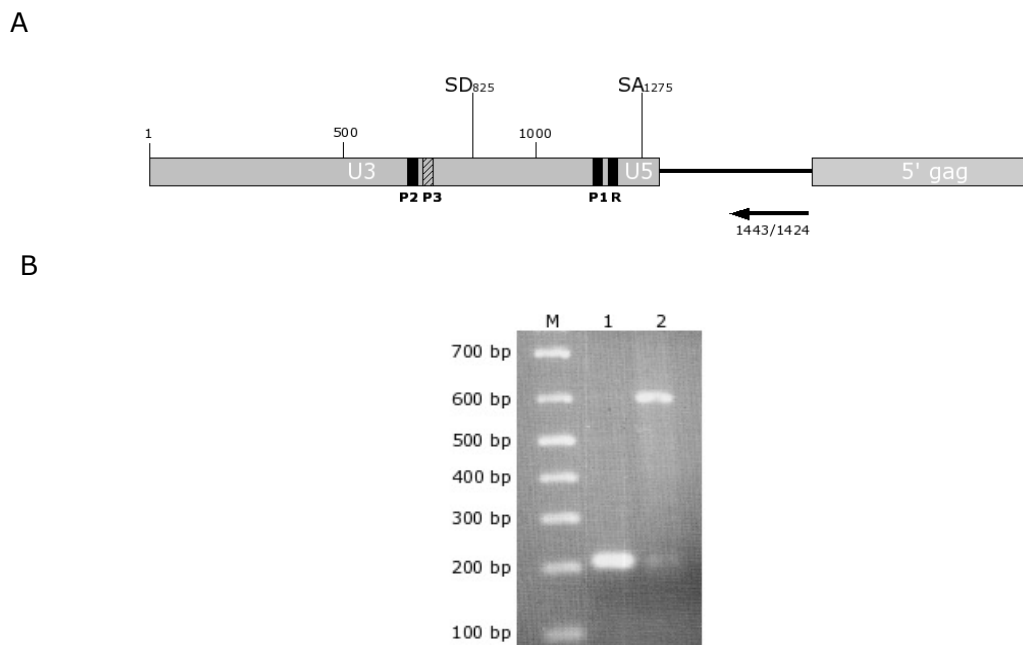


Figure 4.29 Transcripts initiating downstream of the P2/P3 promoter complex after mutation of the P3 TATA-box and the CDP binding sites. (A) The 5'LTR and the beginning of *gag* are shown. The functional LTR promoters as well as the initiator element are symbolized by black boxes, while the mutated P3 promoter is depicted as hatched box. The numbers indicate the position relative to the beginning of the LTR. The primer utilized for reverse transcription is represented by the black arrow. (B) Total RNA was extracted and reverse transcribed with poly(dT) primers. The transcriptional activity of the mutated promoter complex was analysed by RT-PCRs with oligonucleotides (773/795) and (1431/1409). The amplification products of clone A (lane 1) and clone B (lane 2) are shown.

For the identification of the full length P2 product, another set of RT-PCRs was conducted utilizing the same forward primer (773/795) in combination with different reverse primers corresponding to *gag*, *pol* and *env* proviral sequences. While no specific product could be amplified with primers corresponding to *gag* or *pol* regions (data not shown), a specific PCR product was detected with an oligonucleotide (6627/6601) that binds immediately downstream of the *env* splice donor (SD₆₅₅₄) (figure 4.30A). Interestingly, the signal was detected solely in the clone A, but could not be amplified from clone B (figure 4.30B).

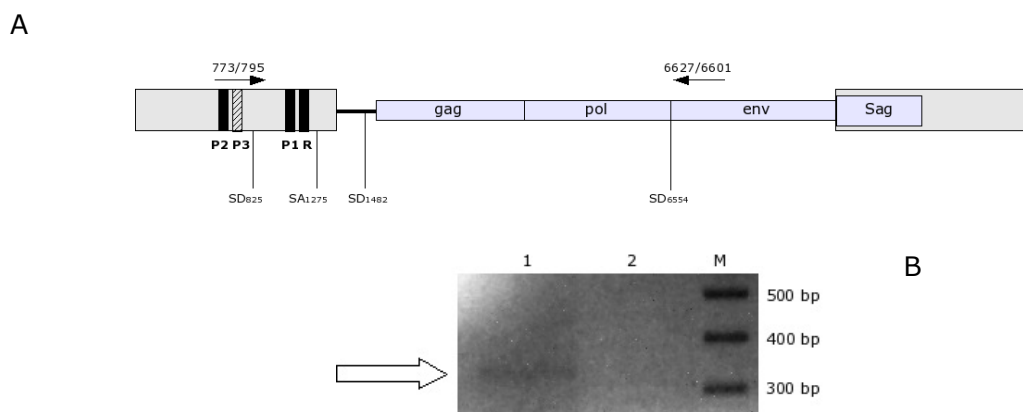


Figure 4.30 A double spliced message originates from the P2 promoter. (A) Primer binding sites for the identification of the putative P2 promoter transcript. The proviral DNA of MMTV is shown. The unmodified promoter elements of the LTR are indicated by black boxes, while the mutated P3 promoter is symbolized by the hatched box. The splice donor (SD) and splice acceptor (SA) sites of the LTR as well as those utilized for the single-spliced *env* message are indicated. Forward and reverse primer binding sites are represented as arrows. (B) Total RNA was extracted from two cell clones that were stably transfected with the proviral construct and reverse transcribed with poly(dT) primers. Subsequent RT-PCRs were performed with a forward primer located downstream of the transcriptional initiation site of the P2 promoter in combination with reverse primer corresponding to the *env* coding region (6627/6601). The results obtained with cDNA from clone A (lane 1) and clone B (lane 2) are shown.

The sequencing analysis of the amplification product revealed the double spliced ORFe message. Thus, ORFe mRNAs are transcribed from the intact P2/P3 promoter complex as well as from the active P2 promoter. In the latter situation, however, the amount of the double spliced transcript was almost below the detection limit, suggesting that the activity of the P2 promoter alone is not sufficient for constitutive ORFe transcription levels.

According to previous studies, the DNA-binding activity of CDP inversely correlates with the differentiation status of the mammary gland epithelial (Zhu *et al.*, 2000). Therefore, in a related experiment CDP binding was prevented by a long-term dexamethasone treatment which should mimic the hormonal status during pregnancy and lactation. NMuMG cells that had been stably transfected with a Mtv-2 based provirus harbouring a mutation of the P3 TATA-box within the 5' LTR sequence were grown in the presence of dexamethasone for 96 hours. (The CDP binding sites were wild-type in the applied construct.) The doubly spliced ORFe transcript was successfully amplified from all tested samples (data not shown), indicating that, in the absence of the P3 promoter, the detection of ORFe mRNA depends on the inhibition of CDP binding.

4.11 The translated ORFe sequence harbours a transmembrane spanning domain and TRAF-2 binding motifs

In order to obtain a first impression about the putative function of ORFe, we performed a detailed *in silico* analysis of its *in vitro* translated aminoacid sequence (152 AA). However, no results were obtained using *BLAST* (basic local alignment tool) search, even though different protein databases (UniProt, Protein database Europe (PDBe), NCBI/BLAST) were used. The only exception was a correlation with the gag-pol polyproteins of human and mouse *betaretroviruses*, which was not surprising, since the start of *env* is known to overlap with the end of *pol*.

When we looked for secondary structure elements, we were however able to identify a transmembrane spanning domain (AA 44-61) by using either the prediction program *proteinpredict* or *interproScan*. The latter application utilizes the TMHMM server which is a specialized prediction tool for the identification of transmembrane helices. Figure 4.31 depicts the out-put graph obtained with TMHMM (www.cbs.dtu.dk/services/TMHMM/).

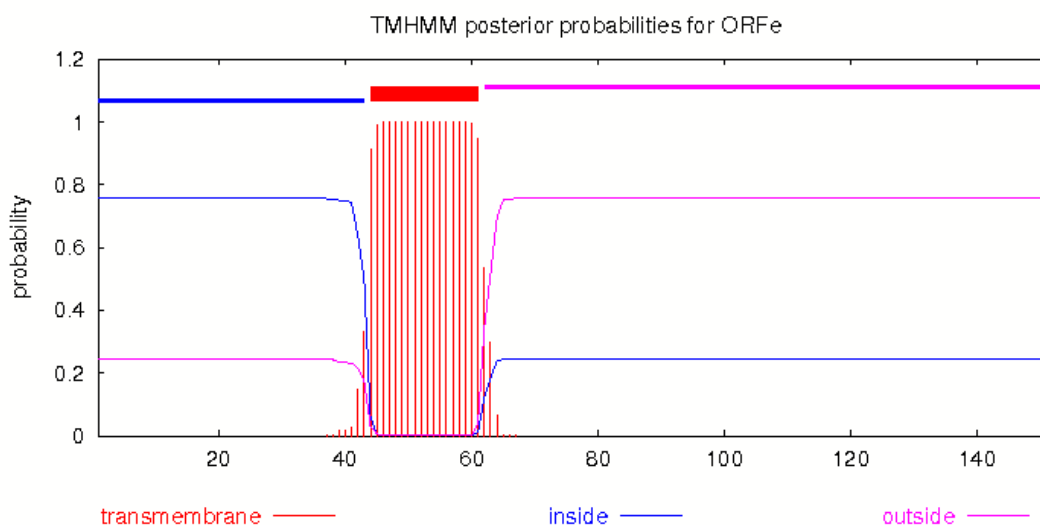


Figure 4.31 Probabilities of transmembrane helices within ORFe. The output graph of the applied prediction program TMHMM (version 2.0) is shown. The number of the ORFe AA are depicted on the x-axis, while the probability to form a transmembrane domain is shown on the y-axis. The localization of the inside and outside segments of the sequence can not be calculated by the algorithm, therefore each of the two possibilities is indicated (compare the thin blue and pink lanes).

The program does not include any algorithm capable of distinguishing outer and inner compartments. Thus, in terms of the protein, those domains are interchangeable.

Furthermore, by applying *ELM* (Eukaryotic Linear Motif resource for functional sites in proteins), we were able to identify one major TRAF2-binding consensus motif (AA 62-65) as well as one minor TRAF2-binding motif (AA 131-136). Tumor necrosis factor receptor associated factors (TRAFs) belong to a family of proteins that bind to the cytoplasmic domain of the tumor necrosis factor receptor (TNFR) family members (Tsitsikov *et al.*, 1997). To date, six members of the TRAF family have been identified, whereby TRAF2 is found associated with the intracellular domain of the type II TNF receptor (TNF-RII) and CD40. CD40 as well as the TNF-R family members are type I membrane proteins with conserved extracellular cysteine-rich domains.

TNF-RI and TNF-RII are present on almost all cell types. Upon ligand binding, they stimulate the recruitment of neutrophils and macrophages to the site of infection and activate these cells in order to eradicate invading microbes. Most of these physiological tasks, however, are mediated by TNF-RI (Abbas *et al.*, 2000).

In contrast, CD40 is predominantly expressed on B-cells, macrophages, dendritic cells and endothelial cells. Particularly B-lymphocytes and dendritic cells are inevitable for the establishment of MMTV infection. Moreover, CD40 mediates the activation of the transcription factor NF- κ B leading to the (T cell-dependent) B-cell activation as well as to the activation of macrophages, dendritic cells and endothelial cells (Abbas *et al.*, 2000). The stimulation and subsequent proliferation of T- and B-cells represents a hallmark of the MMTV infection.

Taken together, this finding allows the prediction of the subcellular orientation of the protein (outside: AA 1-43; transmembrane: AA 44-61; inside: AA 62-152) and additionally suggests a major role of ORFe in the activation of cellular signal transduction pathways. Further analyses of the protein are needed to verify these predictions.

5. Discussion

Mouse mammary tumor virus accounts for the majority of mammary tumours found in common laboratory mouse strains as well as in wild animals. During the course of infection, a virally encoded superantigen excessively stimulates the immune system of the host, leading to an amplification of a sub-set of both B- and T-cells. The virus infects and replicates in these cells until it finally reaches the mammary gland epithelial tissue. There, it remains latent until the hormonal activation during pregnancy and lactation leads to the production of high amounts of viral particles. These are finally passed on via the mother's milk to the newborn pups. In addition to this exogenous route of infection, MMTV can be transmitted endogenously, if the virus infects and integrates into the cells of the germ line. As a consequence, these proviruses are inherited in a Mendelian fashion.

According to the current model of retroviral replication, a single promoter located at the boundary between the U3 and the R region is responsible for the generation of all mRNA species, both full-length and spliced. In MMTV, this major promoter, P1, has been studied extensively because its transcriptional activity is highly enhanced in the presence of glucocorticoid hormones.

However, unique among retroviruses, an additional promoter, P2, was identified in MMTV upstream of the P1 promoter within the central part of the U3 region (Günzburg *et al.*, 1993). Initially, we focused on the analysis of the regulation of the P2 promoter. However, at the very beginning of our studies, we detected another promoter in the close vicinity, which was termed the P3 promoter. The project was therefore extended to investigate the regulation of these two promoters and to identify possible transcripts.

5.1 The P2/P3 promoter complex

Using an improved luciferase reporter construct, we could demonstrate that the P2 promoter is influenced by an upstream negative regulatory element and a downstream enhancer. In contrast, the activity of the adjacent P3 promoter is exclusively modulated by the downstream enhancer element. Originally, this enhancer was thought to influence transcription from the MMTV P1 promoter in cooperation with the hormone responsive element (Gouilleux *et al.*, 1991). However, we did not observe any hormone dependency of the P2 or of the P3 promoter activity in connection with this region (data not shown). Neither did we detect any alterations in the luciferase expression levels between the different cell lines tested. Both the

P2 and the P3 promoter showed the same relative activities in mammary cells, lymphocytes and fibroblasts of murine origin as well as in MMTV permissive feline kidney cells and rat hippocampal stem cells. These data suggest that the two promoters are not regulated by tissue specific factors.

The direct comparison between Mtv-2 and Mtv-8 based sequences allowed us to determine the P3 promoter core, which is represented by the consensus sequence TATAAA. In contrast, the P2 TATA-motif does contain a cytosine at its fifth position (TATACA). This is uncommon and, according to the literature, indicative of a less frequent promoter usage (Bjornsdottir and Myers, 2008; Juven-Gershon *et al.*, 2008). Nevertheless, the central part of the U3 region including both promoters appears to be highly conserved among different MMTV isolates. Moreover, by analysing the 5' ends of transcripts initiating from the central promoter region, we detected an additional initiation site at position 760 (relative to the beginning of the LTR). This is in close vicinity (61 bp) to the previously described initiation site of the P2 promoter (Günzburg *et al.*, 1993). Both promoter elements showed transcriptional activity in reporter gene assays; we therefore conclude that the two promoters together with their respective transcriptional initiation sites form a promoter complex in the central part of the LTR. According to our knowledge, this is the first time that a promoter complex has been described within the LTR of a retrovirus.

Furthermore, we were able to show that in the absence of hormones, the complex is even more potent in initiating transcription than the major P1 promoter. Nevertheless, in our experimental system, the global amount of RNAs initiating from the central promoter complex did not change upon glucocorticoid stimulation. In contrast, Günzburg and colleagues reported an increase of the transcriptional activity of the P2 promoter in response to hormones (Günzburg *et al.*, 1993). Thus, the observed constitutive level of transcripts can be explained by the inability of the central LTR promoters to respond to the synthetic glucocorticoid dexamethasone. On the other hand, each of the two promoters may be inversely regulated, so that any changes would cancel themselves out.

5.2 Associated chromatin structure

Previous analyses have shown that the MMTV LTR is covered by an array of six positioned nucleosomes termed F-A (in the 5' to 3' direction). The TATA-box of the major P1 promoter, the hormone responsive element and the binding sites of relevant transcription factors are covered by a single nucleosome (nucleosome B) that undergoes a characteristic remodelling process upon hormonal induction (Richard-Foy and Hager, 1987; Truss *et al.*, 1995; Nagaich *et al.*, 2004; Vicent *et al.*, 2004 & 2008). Accordingly, a single nucleosome (nucleosome D) has been shown to occupy the P2/P3 promoter region (Richard-Foy and Hager,

1987).

In order to decipher the hormone dependency of the central promoters, we hypothesised that a slight translational movement of the nucleosome D may allow either one or the other promoter to be uncovered from these basic protein complexes and thus may render the relevant promoter accessible to the transcriptional machinery. Therefore, we determined the respective borders of the corresponding nucleosome (nucleosome D) at a base-pair level of resolution in the presence and absence of hormones. Our results regarding the 5' and 3' end of the nucleosome D correlate with previous low-resolution mappings (Richard-Foy and Hager, 1987). However, we did not detect any difference in the positioning of the nucleosome upon hormone induction. This finding seems to argue against a translational movement (sliding) as a putative regulatory mechanism.

However, independent of dexamethasone stimulation, we nevertheless observed a nuclease hypersensitive site within nucleosome D indicating either a nucleosome free region (i.e. the exchange of nucleosomes in and out of the chromatin) or the partial removal of histone proteins (Rando and Ahmad, 2007). Each possibility would allow the underlying sequence to be more easily accessible to DNA-binding proteins, thus reflecting the transcriptional competence of the investigated region. Interestingly, this additional site has been described previously using *in vitro* assembled chromatin on the MMTV LTR (Belikov *et al.*, 2000). However, due to the applied experimental system, the nucleosome positioning over the LTR as well as the nuclease hypersensitive sites within nucleosome B (indicating the hormone response of the P1 promoter) and nucleosome D (occupying the P2/P3 promoter region) were exclusively observed in the presence of glucocorticoids.

By investigating the corresponding post-translational histone modifications in the absence of dexamethasone, we were able to relate the 'open chromatin' formation at the nucleosome D with signs of active transcription from the P2/P3 promoter region. The nucleosomes D, C, and B (covering the P2/P3 promoter complex, the region in between and the P1 promoter, respectively) exhibit elevated methylation levels at lysine 4 of histone H3 (H3K4me3). This peculiar methylation increases when genes become active and is known to specifically localize to 5' ends of ORFs (Heintzman *et al.*, 2007). However, it has been shown that this modification recruits not only activating proteins, but also transcriptionally repressing factors (Berger, 2007).

Furthermore, we detected (in absence of dexamethasone) the highest levels of methylation at lysine 36 on histone H3 (H3K36me3) at nucleosome D compared to the nucleosome C and B regions. This modification is known to be intimately linked to the elongation phase of the RNA polymerase II. The CTD-phosphorylated (elongating) enzyme harbours intrinsic histone acetyltransferase activities to facilitate its passage through the nucleosomes. However, by recruiting the Set2 histone methyltransferase (that specifically

modifies H3K36), it promotes the association of the Rpd3S histone-deacetylase complex. In turn, the transcription elongation-induced acetyl-groups are removed (Workman, 2006). Taken together, the elevated methylation levels at nucleosome D provide further evidence for the transcriptional activity of the P2/P3 promoter complex.

During our work on the regulation of these promoters, a colleague in our laboratory was independently able to identify (in the proviral context) a double-spliced message carrying a 459 nucleotide long ORF ("ORFe") that initiates from the promoter complex. The transcript comprises sequence elements from the LTR, the 5' untranslated region and parts of *env* (Rouault *et al.*, 2007). Hence, the close vicinity of the two promoters complicate a direct correlation between the activity of either promoter with the emergence of the double spliced transcript.

5.3 The influence of acetylation

Besides methylation, active transcription has additionally been associated with increased acetylation levels at histone proteins (Strahl and Allis, 2000). However, in terms of the MMTV LTR, this modification has been controversial. The accumulation of intracellular acetylation levels has been associated with an increase in the promoter activity as well as with transcriptional repression (Bartsch *et al.*, 1996; Lambert and Nordeen, 1998; Wilson *et al.*, 2002; Aoyagi and Archer, 2007).

By investigating the P2/P3 promoter region and the region of the hormone dependent P1 promoter (i.e. the corresponding nucleosomes) separately, we detected differences in the response to the histone deacetylase (HDAC) inhibitor trichostatin A. While in our hands, the transcriptional activity of the major P1 promoter (in the absence of hormones) remained almost unaffected, the amount of ORFe specific transcripts declined upon TSA treatment, indicating that the generation of the double spliced ORFe mRNA depends on the activity of histone deacetylating enzymes. Furthermore, we could demonstrate that the effect of TSA is accompanied by an increase of acetyl-groups at the corresponding nucleosome D, suggesting that this kind of regulation specifically modulates the activity of the P2/P3 promoter complex. In contrast, the acetylation levels of nucleosome B that is covering the P1 promoter did not change upon TSA treatment.

In accordance with our observation that the generation of the double spliced ORFe message depends on active histone deacetylases, we could show that the highest levels of HDAC1, the most prominent representative of these enzymes, are found at nucleosome D compared to nucleosome C and B regions. This finding indicates that these regulatory enzymes are intimately linked to the region containing the P2/P3 promoter complex. Interestingly, TSA

treatment did not change the overall rate of transcripts initiating from the P2/P3 promoter complex. Instead, it increased the emergence of unspliced RNAs (in regard to the first intron). Although it has recently been demonstrated that TSA is able to inhibit mRNA splicing *in vitro* (Kuhn *et al.*, 2009), no direct relationship between histone acetylation and mRNA splicing has as yet been reported, indicating the involvement of other proteins.

Hence, the idea that this particular modification targets a non-histone protein is further supported by the fact that the response to TSA was observed after an long-term (overnight) incubation period. In contrast, a shorter treatment (2 hours) failed to induce any detectable changes.

5.4 Model for the transcriptional regulation of ORFe

HDACs themselves lack any DNA-binding capacity (Haberland *et al.*, 2009). Thus, they have to associate with other proteins in order to exert their function in a certain genomic context. Binding sites for the transcriptional regulator CCAAT displacement protein (CDP) have been recently mapped to the NRE of the MMTV LTR (Zhu *et al.*, 2000; Zhu and Dudley 2002). This ubiquitously expressed transcription factor is known to prevent the transcription of numerous genes during early developmental stages (Nepveu, 2001; Sansregret and Nepveu, 2008). Moreover, it has been demonstrated that CDP is able to recruit HDAC1 in order to establish a repressing chromatin architecture (Li *et al.*, 1999).

Using LTR deletion constructs, we could demonstrate in luciferase reporter gene assays that the activity of the P2 promoter is indeed modulated by the NRE. Moreover, the co-transfection of the exogenously expressed CDP p110 isoform resulted in a decrease of the transcriptional activity of the P2 promoter. This downregulation was observed in the context of the full length LTR as well as in situations when the P2 promoter was exclusively under the control of the NRE. Furthermore, by mutating the CDP binding sites as well as the TATA-box of the P3 promoter, we were able to detect a specific transcript of the P2 promoter in the context of the full-length provirus. The sequence of this transcript corresponded to the double-spliced ORFe message.

However, we noticed major discrepancies in the amount of produced mRNAs. If the transcriptional elements of both promoters (P2 and P3) are functional, the level of double spliced ORFe mRNA is much higher than in the situation in which only the TATA-box of the P2 promoter is left unmodified. This results in only marginal amounts of the transcript. We conclude therefore that the P3 promoter is responsible for the generation of constitutive levels of ORFe mRNA. Furthermore, we propose that, by transcriptional interference, the activity of the upstream located P2 promoter negatively influences the neighbouring P3 promoter.

Thus, the overall genomic organization of the central promoter complex is highly reminiscent of a special mode of transcriptional regulation that was originally described for the *his3* promoter of *S. cerevisiae*. Here, two TATA-boxes, in most cases one canonical TATA sequence (TATAAA) and one noncanonical sequence, are arranged in close vicinity to each other. Both promoter elements are able to direct transcription, either of them from distinct start sites. Usually, the downstream promoter contains the canonical TATA motif and is able to increase its transcriptional activity upon induction, whilst the upstream one is used for low levels of gene expression and additionally can be repressed by specific factors (Butler and Kadonaga, 2002). According to this example and taking in account all our findings, we propose the following model for the regulation of the MMTV central promoter complex:

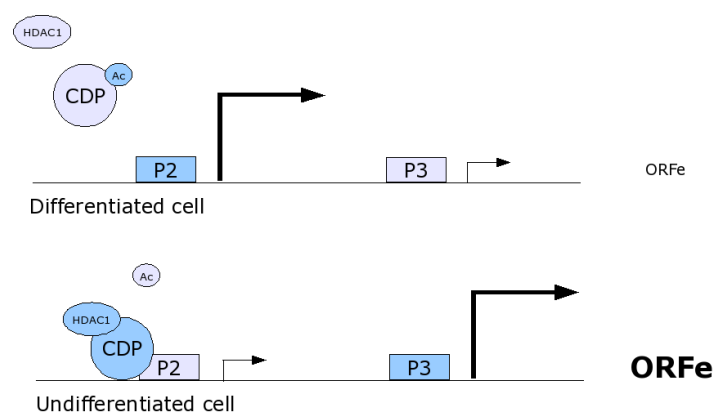


Figure 5.1 Regulation of the P2/P3 promoter complex. The TATA-boxes of the P2 and the P3 promoter, respectively, are indicated as boxes, while putative regulators are depicted as circles/ovals. The transcriptional initiation sites are marked by arrows. The size of the transcript 'ORFe' correlates with the observed levels of double-spliced mRNA. (A) In differentiated cells, CDP (CCAAT displacement protein) is post-translationally modified by acetylation. Consequently, its DNA-binding activity is prevented, which allows the P2 promoter to be fully active. Simultaneously, the activity of the adjacent P3 promoter is downregulated by transcriptional interference. (B) In undifferentiated cells, CDP is able to bind to its cognate sites on the DNA. Furthermore, it recruits the HDAC1 (histone deacetylase 1) enzyme that removes acetyl-groups from histones as well as other histone-associated proteins. Consequently, the P3 promoter becomes active and transcribes constitutive levels of double spliced ORFe mRNA.

The level of double spliced ORFe mRNA is regulated by the mutually exclusive activity of two adjacent promoters. The upstream located P2 promoter is active when the binding of the transcriptional repressor CDP is prevented. Studies have shown that the ability of CDP to bind DNA is lost upon post-translational acetylation of the protein (Li *et al.*, 2000). Under these conditions, transcripts are assumed to initiate predominantly from the P2 promoter. As a consequence, the activity of the P3 promoter is suppressed by transcriptional

interference.

In contrast, an active P3 promoter implies that the P2 promoter has to be repressed. CDP is known to recruit HDAC1 upon DNA binding, thus allowing the removal of acetyl-moieties from histone as well as non-histone proteins. Moreover, CDP has been shown to be highly active during early developmental stages. Consequently, the activity of the P2 promoter should be suppressed in undifferentiated tissues, allowing proper transcription from the downstream located P3 promoter.

Taken together, the generation of constitutive levels of ORFe mRNA is dependent on the CDP-mediated downregulation of the P2 promoter. Hence, the amount of ORFe transcripts is inversely correlated with the differentiation status of the infected cell. According to our model, the previously reported increase in P2 transcriptional activity upon glucocorticoid stimulation (Günzburg *et al.*, 1993) rather represents a consequence of the diminishing DNA-binding activity of CDP during cellular differentiation than a direct influence of the hormone itself.

5.5 Putative ORFe function

In order to further characterize the ORFe protein, we performed an *in silico* analysis of the translated ORFe sequence (152 AA). With this approach, we uncovered two major characteristics of the putative protein that allowed us to deduce its possible function during MMTV infection.

First, we identified a transmembrane spanning domain ranging from AA 43 to AA 61. Interestingly, diverse viral species which do infect lymphocytes harbour the coding information for similar proteins within their genome. For instance, the related retrovirus HTLV-1 (Human T-lymphotropic virus type-1) encodes for the protein p12 which contains two transmembrane regions. Upon expression, this protein accumulates in the ER and *cis*-Golgi apparatus, where it binds key regulatory proteins that are involved in calcium-mediated cell signalling pathways. It is assumed that p12 plays a role during the establishment of HTLV-1 infection through the activation of host cells (Michael *et al.*, 2004). Besides, several members of the *gamma-herpesvirus* subfamily express so-called terminal membrane proteins (TMPs), which are named according to their localization within the viral genome (Brinkmann and Schulz, 2006). The most prominent member, the latent infection membrane protein-1 (LMP1) is expressed by the Epstein-Barr virus (EBV) during the latency phase of infection. This protein lacks any intrinsic enzymatic activity. Instead, the six transmembrane spanning domains induce the spontaneous aggregation of the protein within the plasma membrane (Brinkmann and Schulz, 2006). As a consequence, LMP1 mediates the association of cellular proteins of the tumor necrosis factor

receptor (TNFR) signaling pathway, resulting in the nuclear translocation of the ubiquitous transcription factor NF- κ B (McFarland *et al.*, 1999). These events finally stimulate differentiation and proliferation processes and additionally protect virus-infected cells from apoptosis. In summary, the signalling pathways induced by LMP1 are remarkably similar to those mediated by CD40, a type I membrane glycoprotein of the TNFR-superfamily (Abbas *et al.*, 2000). Under physiological conditions, CD40 is predominantly expressed by APCs which include B-lymphocytes, dendritic cells, macrophages and endothelial cells. Upon ligand binding, it provides key activation signals in T-cell dependent B-cell activation such as the upregulation of surface molecules, the stimulation of antibody production and isotype switching as well as the secretion of cytokines and protection from apoptosis (van Kooten and Banchereau, 2000; Bishop *et al.*, 2002).

Interestingly, the same cellular pathway is also exploited by HIV-1. Martin and colleagues demonstrated recently that HIV-1 incorporates CD40L (a surface molecule that is expressed by activated helper T-cells) into emerging viral particles (Martin *et al.*, 2007). Thereby, the virus triggers an abnormal activation of B-cell lymphocytes that is supposed to result in a more efficient spreading of the virus to CD4⁺ T-cells (Martin *et al.*, 2007).

Furthermore, the direct interaction of CD40 with CD40L has already been shown to represent a critical step during early phases of MMTV infection. Mice which were defective of CD40L failed to induce any deletion of cognate T-cells (a hallmark of MMTV infection) after being exposed to infectious viruses. Furthermore, the response of T-cells in CD40L-minus mice was impaired *in vitro*, but could be reconstituted by adding B-cells that had been activated by recombinant CD40L (Chervonsky *et al.*, 1995).

Second, we were able to identify one major and one minor TRAF-2 (TNFR-associated factor-2) binding motif in the C-terminal part of the protein (AA 62 to AA 65 and AA 131 to 136). TRAFs are common signal transducers that associate intracellularly to ligand-bound receptors of the TNFR superfamily. To date, six TRAFs are known, the majority of which promote cell survival via activation of protein kinase cascades (NF- κ B, AP-1), leading to the proliferation and differentiation of the respective cell (Lee and Lee, 2002).

Taken together, we suggest that ORFe mimics CD40 activation and therefore is supposed to be expressed during the establishment of MMTV infection. Furthermore, we propose that upon TRAF-2 binding, ORFe induces the activation and subsequent proliferation of B-lymphocytes, thereby contributing to the amplification of virus infected cells. This hypothesis implies that ORFe is predominantly expressed from exogenous viruses. We have provided (preliminary) evidence for this assumption by demonstrating that ORFe mRNA is detectable in GR cells (productively infected with a Mtv-2 provirus) as well as from NMuMG cells that have been stably transfected with a Mtv-2 based proviral sequence (N2AnE5). In contrast, no signal was obtained from parental NMuMG cells which harbour three, yet different endogenous

retroviruses (Rouault *et al.*, 2007). Further experiments are needed to test these hypotheses. Indeed, a detailed characterization of ORFe might uncover additional mechanisms during the establishment of MMTV infection.

In conclusion, we were able to demonstrate by using an epigenetic approach that the P2/P3 promoter complex represents a transcriptional active region within the U3 region of MMTV. In contrast to the major P1 promoter, both promoters of the complex seem to work independently of hormonal stimulation. Hence, the tight transcriptional regulation of their product suggests a regulatory function of the corresponding ORFe protein. This idea is further supported by the presence of two TRAF-2 binding motifs which were identified in the C-terminal part of the protein. Furthermore, we were able to establish a model showing that the amount of ORFe mRNA is dependent on the activity of the transcriptional regulator CDP. Accordingly, ORFe expression is supposed to occur in undifferentiated cells and therefore during the latency phase of MMTV infection.

Thus, the LTR of MMTV does contain two elements that are inevitable for the early stages of infection; the coding information for Sag and the P2/P3 promoter complex. We propose that similar to Sag, which upon expression accounts for the differentiation and proliferation of T-cells, the transcriptional product of the P2/P3 promoter complex (ORFe) induces the activation of infected B-cells.

6. References

- Abbas, A. K., A. H. Lichtman, J. S. Pober (2000) *Cellular and Molecular Immunology*, 4th Edition, W.B. Saunders Company
- Acha-Orbea, H., E. Palmer (1991) MIs: A retrovirus exploits the immune system. *Immunol. Today*, 12, 356-361
- Adamson, C. S., E. O. Freed (2008) Recent progress in antiretrovirals – lessons from resistance. *Drug Discov. Today*, 13, 424-432
- Aoyagi, S. and T. K. Archer (2007) Dynamic histone acetylation/deacetylation with progesterone receptor-mediated transcription. *Mol. Endocrinology*, 21 (4), 843-856
- Astrand, C., S. Belikov, O. Wrange (2009) Histone acetylation characterizes chromatin presetting by NF1 and Oct1 and enhances glucocorticoid receptor binding to the MMTV promoter. *Exp. Cell Res.*, 315 (15), 2604-2615
- Ball, J. K., H. Diggelmann, G. A. Dekaban, G. F. Grossi, R. Semmler, P. A. Waight, and R. F. Fletcher (1988) Alterations in the U3 region of the long terminal repeat of an infectious thymotropic type B retrovirus. *J. Virol.*, 62, 2985-2993
- Banan, M., I. C. Rojas, W. H. Lee, H. L. King, J. V. Harriss, R. Kobayashi, C. F. Webb, P. D. Gottlieb (1997) Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. *J. Biol. Chem.*, 272 (29), 18440-18452
- Barnett, A., F. Mustafa, T. J. Wrona, M. Lozano, J. P. Dudley (1999) Expression of mouse mammary tumor virus superantigen mRNA in the thymus correlates with kinetics of self-reactive T-cell loss. *J. Virol.*, 73, 6634-6645
- Bartsch, J., M. Truss, J. Bode and M. Beato (1996) Moderate increase in histone acetylation activates the mouse mammary tumor virus promoter and remodels its nucleosome structure. *PNAS*, 93, 10741-10746
- Belikov S., P. H. Holmqvist, C. Astrand, O. Wrange (2004) Nuclear factor 1 and octamer transcription factor 1 binding preset the chromatin structure of the mouse mammary tumor virus promoter for hormone induction. *J. Biol. Chem.*, 279 (48), 49857-49867
- Belikov, S., B. Gelius, G. Almouzni and O. Wrange (2000) Hormone activation induces nucleosome positioning in vivo. *EMBO J.*, 19, 1023-1033
- Berger, S. (2007) The complex language of chromatin regulation during transcription. *Nature*, 447, 407-412
- Bergman, A. C., O. Björnberg, J. Nord, P. O. Nyman, A. M. Rosengren (1994) The protein p30, encoded at the gag-pro junction of mouse mammary tumor virus, is a dUTPase fused with a nucleocapsid protein. *Virology*, 204 (1), 420-424
- Beutner, U., E. Kraus, D. Kitamura, K. Rajewsky, and B. T. Huber (1994) B cells are essential for murine mammary tumor virus transmission, but not for presentation of

- endogenous superantigens. *J. Exp. Med.*, 179, 1457-1466
- Bhadra, S., M. M. Lozano, and J. P. Dudley (2009) BALB/*Mtv*-Null mice responding to strong mouse mammary tumor virus superantigens restrict mammary tumorigenesis. *J. Virol.*, 83 (1), 484-488
- Bhadra, S., M. M. Lozano, S. M. Payne, J. P. Dudley (2006) Endogenous MMTV proviruses induce susceptibility to both viral and bacterial pathogens. *PloS Pathogens*, 2 (12), 1134-1143
- Bhattacharjee, R. N., T. K. Archer (2006) Transcriptional silencing of the mouse mammary tumor virus promoter through chromatin remodeling is concomitant with histone H1 phosphorylation and histone H3 hyperphosphorylation at M phase. *Virology*, 346 (1), 1-6
- Bindra, A., S. Muradrasoli, R. Kisekka, H. Nordgren, F. Wärnberg, J. Blomberg (2007) Search for DNA of exogenous mouse mammary tumor virus-related virus in human breast cancer samples. *J. Gen. Virol.*, 88 (P6), 1806-1809
- Bishop, G. A., B. S. Hostager, and K. D. Brown (2002) Mechanisms of TNF receptor-associated factor regulation in B lymphocytes. *J. Leukoc. Biol.*, 72, 19-23
- Bittner, John J. (1936) Some possible effects of nursing on the mammary gland tumor incidence in mice. *Science*, 84 (2172), 162
- Bjornsdottir, G., and L. C. Myers (2008) Minimal components of the RNA polymerase II transcription apparatus determine the consensus TATA box. *Nucl. Acids Res.*, 36, 2906-2916
- Bramblett, D., C. L. Hsu, M. Lozano, K. Earnest, C. Fabritius, J. Dudley (1995) A redundant nuclear protein binding site contributes to negative regulation of the mouse mammary tumor virus long terminal repeat. *J. Virol.*, 69 (12), 7868-7876
- Brandt-Carlson, C., and J. S. Butel (1991) Detection and characterization of a glycoprotein encoded by the mouse mammary tumor virus long terminal repeat gene. *J. Virol.*, 65, 6051-6060
- Brasier, A. R., J. E. Tate, and J. F. Habener (1989) Optimized use of the firefly luciferase assay as a reporter gene in mammalian cell lines. *Bio. Techniques*, 7, 1116-1122
- Bray, M., S. Prasad, J. W. Dubay, E. Hunter, K. T. Jeang, D. Rekosh, M. L. Hammariskjöld (1994) A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *PNAS*, 91 (4), 1256-1260
- Brekelmans P., P. van Soest, J. Boerman, P. P. Platenburg, P. J. Lienen, and W. van Ewijk (1994) Transferrin receptor expression as a marker of immature cycling thymocytes in the mouse. *Cell. Immunol.* 159, 331-339
- Bresnick, E. H., M. Bustin, V. Marsaud, H. Richard-Foy, G. L. Hager (1992) The transcriptionally-active MMTV promoter is depleted of histone H1. *Nucl. Acids Res.*, 20 (2), 273-278
- Brinkmann, M. M., and T. F. Schulz (2006) Regulation of intracellular signalling by the terminal membrane proteins of members of the *Gammaherpesvirinae*. *J. Gen. Virol.*,

87, 1047-1074

- Brüggemeier, U., M. Kalff, S. Franke, C. Scheidereit, M. Beato (1991) Ubiquitous transcription factor OTF-1 mediates induction of the MMTV promoter through synergistic interaction with hormone receptors. *Cell*, 64 (3), 565-572
- Buetti, E., H. Diggelmann (1983) Glucocorticoid regulation of mouse mammary tumor virus: identification of a short essential DNA region. *EMBO J.*, 2 (8), 1432-1429
- Burzyn, D., J. C. Rassa, D. Kim, I. Nepomnaschy, S. R. Ross, and I. Piazzon (2004) Toll-like receptor 4-dependent activation of dendritic cells by a retrovirus. *J. Virol.*, 78 (2), 576-584
- Butler, J. E. F., and J. T. Kadonaga (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.*, 16, 2583-2592
- Callahan, R., and G. H. Smith (2000) MMTV-induced mammary tumorigenesis: gene discovery, progression to malignancy and cellular pathways. *Oncogene*, 19, 992-1001
- Capaldo-Kimball, F., and S. D. Barbour (1971) Involvement of recombination genes in growth and viability of *Escherichia coli* K-12. *J. Bacteriology*, 106, 204-212
- Cato, A. C., D. Henderson, H. Ponta (1987) The hormone responsive element of the mouse mammary tumour virus DNA mediates the progestin and androgen induction of transcription in the proviral long terminal repeat region. *EMBO J.*, 6 (2), 363-368
- Chamorro M., N. Parkin, H. E. Varmus (1992) An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *PNAS*, 89 (2), 713-717
- Chen, R., H. Wang, L. M. Mansky (2002) Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J. Gen. Virol.*, 83, 2339-2345
- Chen, X., M. Chamorro, S. I. Lee, L. X. Shen, J. V. Hines, I. Tinoco Jr., H. E. Varmus (1995) Structural and functional studies of retroviral RNA pseudoknots involved in ribosomal frameshifting: nucleotides at the junction of the two stems are important for efficient ribosomal frameshifting. *EMBO J.*, 14 (4), 842-852
- Chervonsky, A. V., J. Xu, A. K. Barlow, M. Khery, R. A. Flavell, C. A. Janeway Jr. (1995) Direct physical interaction involving CD40 ligand on T cells and CD40 on B cells is required to propagate MMTV. *Immunity*, 3, 139-146
- Choi, Y., J. W. Kappler, and P. Marrack (1991) A superantigen encoded in the open reading frame of the 3' long terminal repeat of the mouse mammary tumor virus. *Nature*, 350, 203-207
- Choi, Y., P. Marrack and J. W. Kappler (1992) Structural analysis of a mouse mammary tumor virus superantigen. *J. Exp. Med.*, 175, 847-852
- Clausse, N., R. Smith, C. M. Calberg-Bacq, G. Peters, C. Dickson (1993) Mouse mammary-tumor virus activates Fgf-3/Int-2 less frequently in tumors from virgin than from parous mice. *Int. J. Cancer*, 55 (1), 157-163
- Coffin, J. M., S. H. Hughes and H. E. Varmus (1997) Retroviruses. *Cold Spring Harbor*

Laboratory Press, Cold Spring Harbor, NY

- Cordingley, M. G., G. L. Hager (1988) Binding of multiple factors to the MMTV promoter in crude and fractionated nuclear extracts. *Nucl. Acids Res.*, 16 (2), 609-628
- Courreges, M. C., D. Burzyn, I. Nepomnachy, I. Piazzon, and S. R. Ross (2007) Critical role of dendritic cells in Mouse Mammary Tumor Virus in vivo infection. *J. Virol.*, 81 (8), 3769-3777
- Cullen, B. R. (2003) Nuclear mRNA export: insights from virology. *Trends biochem. Sci.*, 28 (8), 419-424
- Cullen, B. R. (2006) Role and Mechanisms of action of the APOBEC3 family of antiretroviral resistance factors. *J. Virol.*, 80 (3), 1067-1076
- Denis, F., N. H. Shoukry, M. Delcourt, J. Thibodeau, N. Labrecque, H. McGrath, J. S. Munzer, N. G. Seidah, and R.-P. Sekaly (2000) Alternative proteolytic processing of mouse mammary tumor virus superantigens. *J. Virol.*, 74, 3067-3073
- Donehower, L. A., A. L. Huang, and G. L. Hager (1981) Regulatory and coding potential of the Mouse Mammary Tumor Virus long terminal redundancy. *J. Virol.*, 37 (1) 226-238
- Dudley, J. P., H. E. Varmus (1981) Purification and translation of murine mammary tumor virus mRNA's. *J. Virol.*, 39 (1), 207-18
- Dzuris J. L., T. V. Golovkina, S. R. Ross (1997) Both T and B cells shed infectious mouse mammary tumor virus. *J. Virol.*, 71 (8), 6044-6048
- Etkind, P., J. Du, A. Khan, J. Pilitteri, P. H. Wiernik (2000) Mouse mammary tumor virus-like ENV gene sequences in human breast tumors and in a lymphoma of a breast cancer patient. *Clin. Cancer Res.*, 6 (4), 1273-1278
- Faschinger, A., F. Rouault, J. Sollner, A. Lukas, B. Salmons, W. H. Günzburg, S. Indik (2007) Mouse mammary tumor virus integration site selection in human and mouse genomes. *J. Virol.*, 82 (3), 1360-1367
- Fee, B. E., J. W. Steinke, J. Pierce, D. O. Peterson (2002) Initiation site binding protein and the initiator-like promoter element of mouse mammary tumor virus. *Virology*, 302, 185-194
- Felber, B. K., M. Hadzopoulou-Cladaras, C. Cladaras, T. Copeland, and G. N. Pavlakis (1989) Rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *PNAS*, 86, 1495-1499
- Field, Y., N. Kaplan, Y. Fondufe-Mittendorf, I. K. Moore, E. Sharon, Y. Lubling, J. Widom, E. Segal (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Computational Biology*, 4 (11), e1000216
- Fletcher, T. M., B. W. Ryu, C. T. Baumann, B. S. Warren, G. Frago, S. John, G. L. Hager (2000) Structure and dynamic properties of a glucocorticoid receptor-induced chromatin transition. *Mol. Cell Biol.*, 20 (17), 6466-6475
- Flint, S. J., L. W. Enquist, V. R. Racaniello, A. M. Skala (2009) Principles of Virology, 3th Edition, *AMS Press, American Society for Microbiology, Washington*

- Fragoso, G., S. John, M. S. Roberts, G. L. Hager (1995) Nucleosome positioning on the MMTV LTR results from the frequency-biased occupancy of multiple frames. *Genes Dev.*, 9 (15), 1933-1947
- Fuchs, S. M., R. N. Laribee, B. D. Strahl (2009) Protein modifications in transcription elongation. *Bioch. Biophys. Acta*, 1789, 26-36
- Futran, J., J. D. Kemp, E. H. Field, A. Vora, and R. F. Ashman (1989) Transferrin receptor synthesis is an early event in B cell activation. *J. Immunol.* 143, 787-792
- Gallahan, D., and R. Callahan (1987) Mammary tumorigenesis in feral mice: identification of a new *int* locus in mouse mammary tumor virus (Czech II)-induced mammary tumors. *Virology*, 61, 66-74
- Ganser-Pornillos, B. K., M. Yeager, W. I. Sundquist (2008) The structural biology of HIV assembly. *Curr. Opin. Struct. Biol.*, 18 (2), 203-217
- Goedert, J. J., C. S. Rabkin, S. R. Ross (2006) Prevalence of serologic reactivity against four strains of mouse mammary tumour virus among US women with breast cancer. *Br. J. Cancer*, 94 (4), 548-551
- Golovkina T. V., J. P. Dudley, S. R. Ross (1998) B and T cells are required for mouse mammary tumor virus spread within the mammary gland. *J. Immunol.*, 161 (5), 2375-2382
- Golovkina, T. V., A. V. Chervonsky, J. P. Dudley, and S. R. Ross (1992) Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell*, 69, 637-645
- Golovkina, T. V., O. Prakash, S. R. Ross (1996) Endogenous mouse mammary tumor virus Mtv-17 is involved in Mtv-2-induced tumorigenesis in GR mice. *Virology*, 218 (1), 14-22
- Gouilleux, F., B. Sola, B. Couette, and H. Richard-Foy (1991) Cooperation between structural elements in hormon-regulated transcription from the mouse mammary tumor virus promoter. *Nuc. Acids Res.*, 19, 1563-1569
- Goulet, B., A. Baruch, N. S. Moon, M. Poirier, L. L. Sansregret, A. Erickson, M. Bogyo, A. Nepveu (2004) A cathepsin L isoform that is devoid of a signal peptide localizes to the nucleus in S phase and processes the CDP/Cux transcription factor. *Mol. Cell*, 14, 207-219
- Gray, D. A., C. M. McGrath, R. F. Jones, and V. L. Morris (1986) A common mouse mammary tumor virus integration site in chemically induced precancerous mammary hyperplasia. *Virology*, 148, 360-368
- Günzburg, W. H., F. Heinemann, S. Wintersperger, T. Miethke, H. Wagner, V. Erfle and B. Salmons (1993) Endogenous superantigen expression controlled by a novel promoter in the MMTV long terminal repeat. *Nature*, 364, 154-158
- Haberland, M., R. L. Montgomery, E. N. Olson (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat. Rev. Genet.*, 10 (1), 32-42
- Hadzopoulou-Cladaras, M., B. K. Felber, C. Cladaras, A. Athanassopoulos, A. Tse and G. N. Pavlakis (1989) The rev (trs/art) protein of human immunodeficiency virus type 1 affects viral mRNA and protein expression via a cis-acting sequence in the env

- region. *J. Virol.*, 63, 1265-1274
- Harris, R. S., A. M. Sheehy, H. M. Craig, M. H. Malim and M. S. Neuberger (2003) DNA deamination: not just a trigger for antibody diversification but also a mechanism for defense against retroviruses. *Nat. Immunol.*, 4, 641-643
- Harris, R. S., and M. T. Liddament (2004) Retroviral restriction by APOBEC proteins. *Nat. Rev. Immunol.*, 4, 868-877
- Hebbar, P. B., T. K. Archer (2007) Chromatin dependent cooperativity between site-specific transcription factors in vivo. *J. Biol. Chem.*, 282 (11), 8284-8291
- Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford and B. Ren (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics*, 39, 311-318
- Held, W., G. A. Waanders, A. N. Shakhov, L. Scarpellino, H. Acha-Orbea, and H. R. MacDonald (1993) Superantigen-induced immune stimulation amplifies mouse mammary tumor virus infection and allows virus transmission. *Cell*, 74, 529-540
- Hizi, A., L. E. Henderson, T. D. Copeland, R. C. Sowder, H. C. Krutzsch, S. Oroszlan (1987) Analysis of gag proteins from mouse mammary tumor virus. *J. Virol.*, 63 (6), 2543-2549
- Hodes R. J., R. Abe (2001) Mouse Endogenous Superantigens: Mls and Mls-like determinants encoded by mouse retroviruses. *Curr. Prot. Immunol.*, Appendix 1: Appendix 1F
- Hofacre, A., T. Nitta, and H. Fan (2009) Jaagsiekte Sheep Retrovirus encodes a regulatory factor, Rej, required for synthesis of Gag Protein. *J. Virol.*, 83 (23), 12483-12498
- Holmqvist, P. H., S. Belikov, K. S. Zaret, O. Wrangé (2005) FoxA1 binding to the MMTV LTR modulates chromatin structure and transcription. *Exp. Cell Res.*, 302 (2), 593-603
- Howard, D. K., J. Schlom (1980) Isolation of a series of novel variants of murine mammary tumor viruses with broadened host ranges. *Int. J. Cancer*, 25 (5), 647-654
- Hsu, C.-L. L., C. Fabritius, and J. Dudley (1988) Mouse mammary tumor virus provirus in T-cell lymphomas lack a negative regulatory element in the long terminal repeat region. *J. Virol.*, 62, 4644-4652
- Hynes, N. A., J. Van Ooyen, N. Kennedy, P. Herrlich, H. Ponta, B. Groner (1983) Subfragments of the large terminal repeat cause glucocorticoid-responsive expression of mouse mammary tumor virus and of an adjacent gene. *PNAS*, 80 (12), 3637-3641
- Indik, S., W. H. Günzburg, B. Salmons, F. Rouault (2005) Mouse mammary tumor virus infects human cells. *Cancer Res.*, 65 (15), 6651-6659
- Indik, S., W. H. Günzburg, P. Kulich, B. Salmons, F. Rouault (2007) Rapid spread of mouse mammary tumor virus in cultured human breast cells. *Retrovirology*, 4, 73
- Johnson, T. A., C. Elbi, B. S. Parekh, G. L. Hager, S. John (2008) Chromatin remodeling complexes interact dynamically with a glucocorticoid receptor-regulated promoter. *Mol. Biol. Cell*, 19 (8), 3308-3322

- Joshi, A. A., K. Struhl (2005) Baf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol. Cell*, 20 (6), 971-978
- Jude, B. A., Y. Pobezinskaya, J. Bishop, S. Parke, R. M. Medzhitov, A. V. Chervonsky, and T. V. Golovkina (2003) Subversion of the innate immune system by a retrovirus. *Nat. Immunol.*, 4, 573-578
- Juven-Gershon, T., J. Y. Hsu, J. W. Theisen, J. T. Kadonaga (2008) The RNA polymerase II core promoter – the gateway to transcription. *Curr. Opin. Cell Biol.*, 20, 253-259
- Kaplan, N., I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, D. Tillo, Y. Field, E. M. LeProust, T. R. Hughes, J. D. Lieb, J. Widom and E. Segal (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, 458 (7236), 362-366
- Karapetian, O., A. N. Shakhov, J. P. Kraehenbuhl, and H. Acha-Orbea (1994) Retroviral infection of neonatal Peyer's patch lymphocytes: the mouse mammary tumor virus model. *J. Exp. Med.*, 180, 1511-1516
- Katz, E., M. H. Lareef, J. C. Rassa, S. M. Grande, L. B. King, J. Russo, S. R. Ross, J. G. Monroe (2005) MMTV Env encodes an ITAM responsible for transformation of mammary epithelial cells in three-dimensional culture. *J. Exp. Med.*, 201 (3), 431-439
- Klemenz, R., M. Reinhardt, H. Diggelmann (1981) Sequence determination of the 3' end of the mouse mammary tumor virus RNA. *Mol. Biol. Rep.*, 7 (1-3), 123-126
- Kozak, C., G. Peters, R. Pauley, V. Morris, R. Michalides, J. Dudley, M. Green, M. Davisson, O. Prakash, A. Vaikdya, J. Hilgers, A. Verstraeten, N. Hynes, H. Diggelmann, D. Peterson, J. C. Cohen, C. Dickson, N. Sarkar, R. Nusse, H. Varmus and R. Callahan (1987) A standardized Nomenclature for endogenous Mouse Mammary Tumor Viruses. *J. Virol.*, 61, 1651-1654
- Kuhn, A. N., M. A. van Santen, A. Schwienhorst, H. Urlaub, and R. Lührmann (2009) Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. *RNA*, 15, 153-175
- Kusk, P., K. E. Carlson, B. S. Warren, G. L. Hager (1995) Role of the TATA box in transcription of the mouse mammary tumor virus RNA. *Mol. Endocrinol.*, 9, 1180-1192
- Lallemand, F., D. Courilleau, M. Sabbah, G. Redeuilh, J. Mester (1996) Direct inhibition of the expression of cyclin D1 gene by sodium butyrate. *Biochem. Biophys. Res. Commun.*, 229 (1), 163-169
- Lambert, J. R. and S. K. Nordeen (2003) CBP recruitment and histone acetylation in differential gene induction by glucocorticoids and progestins. *Mol. Endocrinology*, 17 (6), 1085-1094
- Lambert, J. R., S. K. Nordeen (1998) Steroid-selective initiation of chromatin remodeling and transcriptional activation of the Mouse Mammary Tumor Virus promoter is controlled by the site of promoter integration. *J. Biol. Chem.*, 273, 32708-32714
- Lasfargues, E. Y., W. G. Coutinho, A. S. Dion (1979) A human breast tumor cell line (BT-474) that supports mouse mammary tumor virus replication. *In Vitro*, 15 (9), 723-729
- Lawson, J. S., D. D. Tran, C. Ford, W. D. Rawlinson (2004) Elevated expression of the tumor suppressing protein p53 is associated with the presence of mouse mammary tumor-

- like env gene sequences (MMTV-like) in human breast cancer. *Breast Cancer Res. Treat.*, 87 (1), 13-17
- Lee, H., J. Guo, M. Li, J. K. Choi, M. DeMaria, M. Rosenzweig, J. U. Jung (1998) Identification of an immunoreceptor tyrosine-based activation motif of K1 transforming protein of Kaposi's sarcoma-associated herpesvirus. *Mol. Cell Biol.*, 18, 5219-5228
- Lee, H. L., T. K. Archer (1994) Nucleosome-mediated disruption of transcription factor-chromatin initiation complexes at the mouse mammary tumor virus long terminal repeat in vivo. *Mol. Cell Biol.*, 14 (1), 32-41
- Lee, H. L., T. K. Archer (1998) Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. *EMBO J.*, 17 (5), 1454-1466
- Lee, N. K., and S. Y. Lee (2002) Modulation of life and death by the Tumor Necrosis Factor Receptor-Associated Factors (TRAFs). *J. Biochem. Mol. Biol.*, 35, 61-6654
- Lefebvre, P., D. S. Berard, M. G. Cordingley, and G. L. Hager (1991) Two regions of the mouse mammary tumor virus long terminal repeat regulate the activity of its promoter in mammary cell lines. *Mol. Cell. Biol.*, 11, 2529-2537
- Li, S., B. Aufiero, R. L. Schiltz, and M. J. Walsh (2000) Regulation of the homeodomain CCAAT displacement/*cut* protein function by histone acetyltransferases p300 / CREB-binding protein (CBP)-associated factor and CBP. *PNAS*, 97, 7166-7171
- Li, S., L., Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, M. J. Walsh (1999) Transcriptional repression of the cystic fibrosis transmembrane conductance regulator gene, mediated by CCAAT displacement protein/*cut* homologue, is associated with histone deacetylation. *J. Biol. Chem.*, 274 (12), 7803-7815
- Li, X., J. Wong, S. Y. Tsai, M. J. Tsai, and B. W. O'Malley (2003) Progesterone and glucocorticoid receptors recruit distinct coactivator complexes and promote distinct patterns of local chromatin modification. *Mol. Cell. Biol.*, 23 (11), 3763-3773
- Lievens, P. M. J., J. J. Donady, C. Tufarelli and E. J. Neufeld (1995) Repressor activity of CCAAT displacement protein in HL-60 myeloid leukemia cells. *J. Biol. Chem.*, 270 (21), 12745-12750
- Lin, L. I., L. I. Hai Shan, C. D. Pauza, M. Burkrinsky, R. Y. Zhao (2005) Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions. *Cell Research*, 15, 923-934
- Liu, B., Y. Wang, S. M. Melana, I. Pelisson, B. Najfeld, J. F. Holland, B. G. Pogo (2001) Identification of a proviral structure in human breast cancer. *Cancer Res.*, 61 (4), 1754-1759
- Liu, J., A. Barnett, E. J. Neufeld, J. P. Dudley (1999) Homeoproteins CDP and SATB1 interact: potential for tissue-specific regulation. *Mol. Cell Biol.*, 19(7), 4918-4926
- Liu, J., D. Bramblett, Q. Zhu, M. Lozano, R. Kobayashi, S. R. Ross and J. P. Dudley (1997) The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. *Mol. Cell Biol.*, 17, 5275-5287
- Lu, J., W.-H. Lin, S.-Y. Chen, R. Longnecker, S.-C. Tsai, C.-L. Chen, C.-H. Tsai (2006) Syk tyrosine kinase mediates Epstein-Barr Virus latent membrane protein 2A-induced cell

- migration in epithelial cells. *J. Biol. Chem.*, 281, 8806-8814
- Maeda, N., M. Palmarini, C. Murgia, H. Fan (2001) Direct transformation of rodent fibroblast by jaagsiekte sheep retrovirus DNA. *PNAS*, 98, 4449-4454
- Malim, M. H., J. Hauber, S. Y. Le, J. V. Maizel, and B. R. Cullen (1989) The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*, 338, 254-257
- Mant, C., D. Gillett, C. D'Arrigo, J. Cason (2004) Human murine mammary tumour virus-like agents are genetically distinct from endogenous retroviruses and are not detectable in breast cancer cell lines or biopsies. *Virology*, 318 (1), 393-404
- Mant, C., J. Cason (2004) A human murine mammary tumour virus-like agent is an unconvincing aetiological agent for human breast cancer. *Rev. Med. Virol.*, 14 (3), 169-177
- Marchetti, A., F. Buttitta, S. Miyazaki, D. Gallahan, G. H. Smith, and R. Callahan (1995) *Int-6*, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. *J. Virol.*, 69, 1932-1938
- Marim, M., K. M. Rose, S. L. Kozak, D. Kabat (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat. Med.*, 9 (11), 1398-1403
- Martin, G., J. Roy, C. Barat, M. Quellet, C. Gilbert, and M. J. Tremblay (2007) Human Immunodeficiency Virus Type 1-associated CD40 Ligand transactivates B lymphocytes and promotes infection of CD4⁺ T cells. *J. Virol.*, 81, 5872-5881
- Martin, P., S. R. Ruiz, G. Martinez del Hoyo, F. Anjuere, H. H. Vargas, M. Lopez-Bravo, and C. Ardavin (2002) Dramatic increase in lymph node dendritic cell numbers during infection by the mouse mammary tumor virus occurs by a CD62L-dependent blood-borne DC recruitment. *Blood*, 99, 1282-1288
- Maxwell, I. H., G. S. Harrison, W. M. Wood, and F. Maxwell (1989) A DNA cassette containing a trimerized SV40 polyadenylation signal which efficiently blocks spurious plasmid-initiated transcription. *Bio. Techniques*, 7, 276-280
- McFarland, E. D. C., K. M. Izumi and G. Mosialos (1999) Epstein-Barr virus transformation: involvement of latent membrane protein 1-mediated activation of NF- κ B. *Oncogene*, 18, 6959-6964
- Melana, S. M., I. Nepomnaschy, M. Sakalian, A. Abbott, J. Hasa, J. F. Holland, B. G. Pogo (2007) Characterization of viral particles isolated from primary cultures of human breast cancer cells. *Cancer Res.*, 67 (18), 8960-8965
- Menendez-Arias, L., M. Young, S. Oroszlan (1992) Purification and characterization of the mouse mammary tumor virus protease expressed in *Escherichia coli*. *J. Biol. Chem.*, 267 (33), 24134-24139
- Mertz, J. A., F. Mustafa, S. Meyers, and J. P. Dudley (2001) Type B leukemogenic virus has a T-cell-specific enhancer that binds AML-1. *J. Virol.*, 75, 2174-2184
- Mertz, J. A., M. S. Simper, M. M. Lozano, S. M. Payne, J. P. Dudley (2005) Mouse mammary tumor virus encodes a self-regulatory RNA export protein and is a complex retrovirus. *J. Virol.*, 79 (23), 14737-14747

- Metzner, C., B. Salmons, W. H. Günzburg, M. Gmeiner, I. Miller, B. Gesslbauer, A. Kungl, J. A. Dangerfield (2006) MMTV accessory factor Naf affects cellular gene expression. *Virology*, 346 (1), 139-150
- Michael, B., A. Nair, and M. D. Lairmore (2004) Role of accessory proteins of HTLV-1 in viral replication, T cell activation, and cellular gene expression. *Front. Biosci.*, 9, 2556-2576
- Michalides R., R. van Nie, R. Nusse, N. E. Hynes, B. Groner (1981) Mammary tumor induction loci in GR and DBAf mice contain one provirus of the mouse mammary tumor virus. *Cell*, 23 (1), 165-173
- Michalides, R., E. Wagenaar (1986) Site-specific rearrangements in the long terminal repeat of extra mouse mammary tumor proviruses in murine T-cell leukemias. *Virology*, 154, 76-84
- Michalides, R., E. Wagenaar, J. Hilkens, J. Hilgers, B. Groner, N. E. Hynes (1982) Acquisition of proviral DNA of mouse mammary tumor virus in thymic leukemia cells from GR mice. *J. Virol.*, 43 (3), 819-829
- Michalides, R., E. Wagenaar, P. Weijers (1985) Rearrangements in the long terminal repeat of extra mouse mammary tumor proviruses in T-cell leukemias of mouse strain GR result in a novel enhancer-like structure. *Mol. Cell Biol.*, 5 (4), 823-830
- Mink, S., H. Ponta, A. C. Cato (1990) The long terminal repeat region of the mouse mammary tumour virus contains multiple regulatory elements. *Nuc. Acids Res.*, 18 (8), 2017-2024
- Mok, M. T., J. S. Lawson, B. J. Iacopetta, N. J. Whitaker (2008) Mouse mammary tumor virus-like env sequences in human breast cancer. *Int. J. Cancer*, 122 (12), 2864-2870
- Moon, N. S., P. Premdas, M. Truscott, L. Leduy, G. Berube, and A. Nepveu (2001) S phase-specific proteolytic cleavage is required to activate stable DNA binding by the CDP/Cut homeodomain protein. *Mol. Cell. Bio.*, 21 (18), 6332-6345
- Moore, D. H., C. A. Long, A. B. Vaidya, J. B. Sheffield, A. S. Dion, and E. Y. Lasfargues (1979) Mammary tumor viruses. *Adv. Cancer Res.*, 29, 347-418
- Morris, D. W., P. A. Barry, H. D. Bradshaw Jr, R. D. Cardiff (1990) Insertion mutation of the int-1 and int-2 loci by mouse mammary tumor virus in premalignant and malignant neoplasms from the GR mouse strain. *J. Virol.*, 64 (4), 1794-1802
- Morrison, J. A., N. Raab-Traub (2005) Roles of the ITAM and PY motifs of Epstein-Barr Virus latent membrane protein 2A in the inhibition of epithelial cell differentiation and activation of β -catenin signaling. *J. Virol.*, 79, 2375-2382
- Mougel, M., L. Houzet and J.-L. Darlix (2009) When is it time for reverse transcription to start and go? *Retrovirology*, 6, 24-32
- Mulholland, N. M, E. Soeth and C. L. Smith (2003) Inhibition of MMTV transcription by HDAC inhibitors occurs independent of changes in chromatin remodeling and increased histone acetylation. *Oncogene*, 22, 4807-4818
- Müllner, M., B. Salmons, W. H. Günzburg and S. Indik (2008) Identification of the Rem-responsive element of mouse mammary tumor virus. *Nuc. Acids Res.*, 36 (19),

6284-6294

- Mustafa, F., S. Bhadra, D. Johnston, M. Lozano, J. P. Dudley (2003) The type B leukemogenic virus truncated superantigen is dispensable for T-cell lymphomagenesis. *J. Virol.*, 77, 3866-3870
- Mymryk, J. S., D. Berard, G. L. Hager, T. K. Archer (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 (1), 26-34
- Nagaich, A. K., D. A. Walker, R. Wolford and G. L. Hager (2004) Rapid periodic binding and displacement of the glucocorticoid receptor during chromatin remodeling. *Mol. Cell*, 14, 163-174
- Nakano, H., T. Yoshimoto, T. Kakiuchi, A. Matsuzawa (1993) Nonspecific augmentation of the lymph node T cells and I-E-dependent selective deletion of V beta 14+ T cells by Mtv-2 in the DDD mouse. *Eur. J. Immunol.*, 23 (10), 2434-2339
- Nepveu, A. (2001) Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. *Gene*, 270 (1-2), 1-15
- Ng, H. H., F. Robert, R. A. Young, K. Struhl (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell*, 11 (3), 709-19
- Nishio, H., M. J. Walsh (2004) CCAAT displacement protein/cut homolog recruits G9a histone lysine methyltransferase to repress transcription. *PNAS*, 101 (31), 11257-11262
- Nisole, S., A. Saïb (2004) Early steps of retrovirus replicative cycle. *Retrovirology*, 14, 1-9
- Okeoma, C. M., A. Low, W. Bailis, H. Y. Fan, B. Matija Peterlin, and S. R. Ross (2009a) Induction of APOBEC3 *in vivo* causes increased restriction of retrovirus infection. *J. Virol.*, 83 (8), 3486-3495
- Okeoma, C. M., J. Petersen, and S. R. Ross (2009b) Expression of murine APOBEC3 alleles in different mouse strains and their effect on mouse mammary tumor virus infection. *J. Virol.*, 83 (7), 3029-3038
- Okeoma, C. M., N. Lovsin, B. Matija Peterlin and S. R. Ross (2007) APOBEC3 inhibits mouse mammary tumour virus replication *in vivo*. *Nature*, 445, 927-930
- Owens, R. B., A. J. Hackett (1972) Tissue culture studies of mouse mammary tumor cells and associated viruses. *J. Natl. Cancer Inst.*, 49 (5), 1321-1332
- Pasquinelli, A. E., R. K. Ernst, E. Lund, C. Grimm, M. L. Zapp, D. Rekosh, M. L. Hammar-skjöld, J. E. Dahlberg (1997) The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J.*, 16 (24), 7500-7510
- Pierce, J., B. E. Fee, M. G. Toohey, D. O. Peterson (1993) A mouse mammary tumor virus promoter element near the transcription initiation site. *J. Virol.*, 67, 415-424
- Pina, B., U. Brüggermeier, M. Beato (1990) Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter. *Cell*, 60 (5), 719-

- Pokholok, D. K., C. T. Harbison, S. Levine, M. Cole, N. M. Hannett, T. I. Lee, G. W. Bell, K. Walker, P. A. Rolfe, E. Herbolsheimer, J. Zeitlinger, F. Lewitter, D. K. Gifford, R. A. Young (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, 122 (4), 517-527
- Rai, S. K., F. M. Duh, V. Vigdorovich, A. Danilkovitch-Miagkova, M. I. Lerman, A. D. Miller (2001) Candidate tumor suppressor HYAL2 is a glycosylphosphatidylinositol (GPI)-anchored cell-surface receptor for jaagsiekte sheep retrovirus, the envelope protein which mediates oncogenic transformation. *PNAS*, 98, 4443- 4448
- Rando, O. J., K. Ahmad (2007) Rules and regulation in the primary structure of chromatin. *Curr. Opin. Cell Biol.*, 19 (3), 250-256
- Rassa, J. C., J. L. Meyers, Y. Zhang, R. Kudaravali, and S. R. Ross (2002) Murine retroviruses activate B cells via interaction with toll-like receptor 4. *PNAS*, 99, 2281-2286
- Renfranz, P. J., M. G. Cunningham and R. D. G. McKay (1991) Region-specific differentiation of the hippocampal stem cell line HiB5 upon implantation into the developing mammalian brain. *Cell*, 66 (4), 713-729
- Reuss, F. U., and J. M. Coffin (1995) Stimulation of mouse mammary tumor virus superantigen expression by an intragenic enhancer. *PNAS*, 92 (20), 9293-9297
- Reuss, F. U., and J. M. Coffin (2000) The mouse mammary tumor virus transcription enhancers for hematopoietic progenitor and mammary gland cells share functional elements. *J. Virol.*, 74, 8183-8187
- Reuss, F. U., J. M. Coffin (1998) Mouse mammary tumor virus superantigen expression in B cells is regulated by a central enhancer within the pol gene. *J. Virol.*, 72 (7), 6073-6082
- Richard-Foy, H., G. L. Hager (1987) Sequence-specific positioning of nucleosomes over the steroid-inducible MMTV promoter. *EMBO J.*, 6 (8), 2321-2328
- Ringold, G. M., K. R. Yamamoto, G. M. Tomkins, M. Bishop, H. E. Varmus (1975) Dexamethasone-mediated induction of mouse mammary tumor virus RNA: a system for studying glucocorticoid action. *Cell*, 6 (3), 299-305
- Rong Zeng, W., E. Soucie, N. Sung Moon, N. Marin-Soudant, G. Berube, L. Leduy, A. Nepveu (2000) Exon/intron structure and alternative transcripts of the CUTL1 gene. *Gene*, 241 (1), 75-85
- Ross, S. R., J. J. Schofield, C. J. Farr, and M. Bucan (2002) Mouse transferrin receptor 1 is the cell entry receptor for mouse mammary tumor virus. *PNAS*, 99, 12386-12390
- Ross, S. R., J. W. Schmidt, E. Katz, L. Cappelli, S. Hultine, P. Gimotty, J. G. Monroe (2006) An immunoreceptor tyrosine activation motif in mouse mammary tumor virus envelope protein plays a role in virus-induced mammary tumors. *J. Virol.*, 80 (18), 9000-9008
- Ross, S. (2008) MMTV infectious cycle and the contribution of virus-encoded proteins to transformation of mammary tissue. *J. Mammary Gland Biol. Neoplasia*, 13 (3), 299-307

- Rouault, F., S. Badihi Nejad Asl, S. Rungaldier, E. Fuchs, B. Salmons, W. H. Günzburg (2007) Promoter complex in the central part of the mouse mammary tumor virus long terminal repeat. *J. Virol.*, 81 (22), 12575-12581
- Rungaldier, S., S. Badihi Nejad Asl, W. H. Günzburg, B. Salmons, F. Rouault (2005) Abundant authentic MMTV-Env production from a recombinant provirus lacking the major LTR promoter. *Virology*, 342, 201-214
- Salmons, B., B. Groner, C. M. Calberg-Bacq, H. Ponta (1985) Production of mouse mammary tumor virus upon transfection of a recombinant proviral DNA into cultured cells. *Virology*, 144, 101-114
- Salmons, B., V. Erfle, G. Brem, W. H. Günzburg (1990) Naf, a trans-regulating negative-acting factor encoded within the mouse mammary tumor virus open reading frame region. *J. Virol.*, 64 (12), 6355-6359
- Sandelin, A., P. Carninci, B. Lenhard, J. Ponjavic, Y. Hayashizaki, D. A. Hume (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat. Rev. Genet.*, 6, 424-436
- Sansregret, L., A. Nepveu (2008) The multiple roles of CUX1: insights from mouse models and cell-based assays. *Gene*, 412 (1-2), 84-94
- Santos-Rosa, H., R. Schneider, A. J. Bannister, J. Sherriff, B. E. Bernstein, N. C. Emre, S. L. Schreiber, J. Mellor, T. Kouzarides (2002) Active genes are tri-methylated at K4 of histone H3. *Nature*, 419 (6905), 407-411
- Satchwell, S. C., H. R. Drew, A. A. Travers (1986) Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.*, 191, 659-675
- Scherer, M. T., L. Ignaowicz, A. Pullen, J. Kappler, P. Marrack (1995) The use of mammary tumor virus (*Mtv*)-negative and single-*Mtv* mice to evaluate the effects of endogenous viral superantigens on the T-cell repertoire. *J. Exp. Med.*, 182, 1493-1504
- Schulman, H. M., P. Ponka, A. Wilczynska, Y. Gauthier, G. Shyamala (1989) Transferrin receptor and ferritin levels during murine mammary gland development. *Biochim. Biophys. Acta*. 1010 (1), 1-6
- Segal, E., Y. Fondufe-Mittendorf, L. Chen, A. Thaström, Y. Field, I. K. Moore, J. P. Wang, J. Widom (2006) A genomic code for nucleosome positioning. *Nature*, 442 (7104), 772-778
- Shackleford G. M., C. A. MacArthur, H. C. Kwan, H. E. Varmus (1993) Mouse mammary tumor virus infection accelerates mammary carcinogenesis in Wnt-1 transgenic mice by insertional activation of int-2/Fgf-3 and hst/Fgf-4. *PNAS*, 90 (2), 740-744
- Sheldon, L. A., M. Becker and C. L. Smith (2001) Steroid hormone receptor-mediated histone deacetylation and transcription at the mouse mammary tumor virus promoter. *J. Biol. Chem.*, 276 (35), 32423-32426
- Sims, R. J. 3Rd, S. Milhouse, C. F. Chen, B. A. Lewis, H. Erdjument-Bromage, P. Tempst, J. L. Manley, D. Reinberg (2007) Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell*, 28 (4), 665-676

- Soutoglou, E., I. Talianidis (2002) Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science*, 295 (5561), 1901-1904
- Squartini, F., F. Basolo, M. Bistocchi (1983) Lobuloalveolar differentiation and tumorigenesis: two separate activities of mouse mammary tumor virus. *Cancer Research*, 43, 5879-5882
- Strahl, B. D., C. D. Allis (2000) The language of covalent histone modifications. *Nature*, 403 (6765), 41-45
- Takeda, E., S. Tsuji-Kuwahara, M. Sakamoto, M. A. Langlois, M. S. Neuberger, C. Rada, M. Miyazawa (2008) Mouse APOBEC3 restricts friend leukemia virus infection and pathogenesis in vivo. *J. Virol.*, 82 (22), 10998-11008
- Takeda, K., S. Akira (2007) Toll-like receptors. *Curr. Protoc. Immunol.*, Chapter 14, Unit 14.12
- Theodorou, V., M. A. Kimm, M. Boer, L. Wessels, W. Theelen, J. Jonkers, and J. Hilken (2007) MMTV insertional mutagenesis identifies genes, gene families and pathways involved in mammary cancer. *Nat. Genet.*, 39, 759-769
- Thomas, J. A. and R. J. Gorelick (2008) Nucleocapsid protein function in early infection processes. *Virus Res.*, 134, 39-63
- Thorne, J. L., M. J. Campbell, B. M. Turner (2009) Transcription factors, chromatin and cancer. *Int. J. Biochem. Cell Biol.*, 41 (1), 164-175
- Truscott, M., L. Raynal, Y. Wang, G. Bérubé, L. Leduy and A. Nepveu (2004) The N-terminal region of the CCAAT displacement protein (CDP)/Cux transcription factor functions as an autoinhibitory domain that modulates DNA binding. *J. Biol. Chem.*, 279 (48), 49787-49794
- Truss, M., J. Bartsch, A. Schelbert, R. J. G. Hache, and M. Beato (1995) Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter in vivo. *EMBO J.*, 14, 1737-1751
- Tsitsikov, E. N., D. A. Wright, R. S. Geha (1997) CD30 induction of human immunodeficiency virus gene transcription is mediated by TRAF2. *PNAS*, 94 (4), 1390-1395
- Underhill, D. M., and A. Ozinsky (2002) Toll-like receptors: key mediators of microbe detection. *Curr. Opin. Immunol.*, 14, 103-110
- Vacheron, S., S. A. Luther, H. Acha-Orbea (2002) Preferential infection of immature dendritic cells and B cells by mouse mammary tumor virus. *J. Immunol.*, 168 (7), 3470-3476
- Van Kooten, C., and J. Banchereau (2000) CD40-CD40 ligand. *J. Leukoc. Biol.*, 67, 2-17
- Van Leeuwen F., R. Nusse (1995) Oncogene activation and oncogene cooperation in MMTV-induced mouse mammary cancer. *Semin. Cancer Biol.*, 6 (3), 127-33
- Van Lint, C., S. Emiliani, E. Verdin (1996) The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. *Gene Expr.*, 5 (4-5), 245-253
- Van Ooyen, A. J., R. J. Michalides, R. Nusse (1983) Structural analysis of a 1.7-kilobase mouse mammary tumor virus-specific RNA. *J. Virol.*, 46 (2), 362-370

- Vicent, G.P., A. S. Nacht, C. L. Smith, C. L. Peterson, S. Dimitrov, M. Beato (2004) DNA instructed displacement of histones H2A and H2B at an inducible promoter. *Mol. Cell*, 16 (3), 439-452
- Vicent, G. P., C. Ballare, A. S. Nacht, J. Clausell, A. Subtil-Rodriguez, I. Quiles, A. Jordan, M. Beato (2008) Convergence on chromatin of non-genomic and genomic pathways of hormone signaling. *J. Steroid Biochem. Mol. Biol.*, 109 (3-5), 344-349
- Wang, E., L. Albritton, and S. R. Ross (2006) Identification of the Segments of the Mouse Transferrin Receptor 1 Required for Mouse Mammary Tumor Virus Infection. *J. Biol. Chem.*, 281 (15), 10243-10249
- Wang, E., O.-A. Nyamekye, Q. Ying, L. Meertens, T. Dragic, R. A. Davey, S. R. Ross (2008) Mouse mammary tumor virus uses mouse but not human transferrin receptor 1 to reach a low pH compartment and infect cells. *Virology*, 381, 230-240
- Wang, Y., I. Pelisson, S. M. Melana, J. F. Holland, B. G. Pogo (2001) Detection of MMTV-like LTR and LTR-env gene sequences in human breast cancer. *Int. J. Oncol.*, 18 (5), 1041-1044
- Wang, Y., J. F. Holland, I. J. Bleiweiss, S. Melana, X. Liu, I. Pelisson, A. Cantarella, K. Stellrecht, S. Mani, B. G. Pogo (1995) Detection of mammary tumor virus env-gene like sequences in human breast cancer. *Cancer Res.*, 55 (22), 5173-5179
- Wheeler, D. A., J. S. Butel, D. Medina, R. D. Cardiff, G. L. Hager (1983) Transcription of mouse mammary tumor virus: identification of a candidate mRNA for the long terminal repeat gene product. *J. Virol.*, 46 (1), 42-49
- Widom, J. (2001) Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.*, 34, 269-324
- Wilson, M. A., A. R. Ricci, B. J. Deroo, and T. K. Archer (2002) The histone deacetylase inhibitor trichostatin A blocks progesterone receptor-mediated transactivation of the mouse mammary tumor virus promoter *in vivo*. *J. Biol. Chem.*, 277 (17), 15171-15181
- Witt, A., B. Hartmann, E. Marton, R. Zeilinger, M. Schreiber, E. Kubista (2003) The mouse mammary tumor virus-like env sequence is not detectable in breast cancer tissue of Austrian patients. *Oncol. Rep.*, 10 (4), 1025-1029
- Wootton, S. K., C. L. Halbert, A. D. Miller (2005) Sheep retrovirus structural protein induces lung tumours. *Nature*, 434, 904-907
- Workman J. L. (2006) Nucleosome displacement in transcription. *Genes Dev.*, 20, 2009-2017
- Xu, L., T. J. Wrona, and J. P. Dudley (1996) Exogenous mouse mammary tumor virus (MMTV) infection induces endogenous MMTV *sag* expression. *Virology*, 215, 113-123
- Xu, L., T. J. Wrona, and J. P. Dudley (1997) Strain-specific expression of spliced MMTV RNAs containing the *superantigen* gene. *Virology*, 236, 54-65
- Zaret, K. S., K. R. Yamamoto (1984) Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. *Cell*, 38 (1), 29-38

- Zhang, D. J., V. K. Tsiagbe, C. Huang, and G. J. Thorbecke (1996) Control of endogenous mouse mammary tumor virus superantigen expression in SJL lymphomas by a promoter within the env region. *J. Immunol.*, 157, 3510-3517
- Zhang, Y., J. C. Rassa, M. E. DeObaldia, L. M. Albritton, S. R. Ross (2003) Identification of the receptor binding domain of the mouse mammary tumor virus envelope protein. *J. Virol.*, 77 (19), 10468-10478
- Zhu, Q., J. P. Dudley (2002) CDP binding to multiple sites in the mouse mammary tumor virus long terminal repeat suppresses basal and glucocorticoid-induced transcription. *J. Virol.*, 76 (5), 2168-2179
- Zhu, Q., K. Gregg, M. Lozano, J. Liu, J. P. Dudley (2000) CDP is a repressor of mouse mammary tumor virus expression in the mammary gland. *J. Virol.*, 74 (14), 6348-6357
- Zhu, Q., U. Maitra, D. Johnston, M. Lozano, J. P. Dudley (2004) The homeodomain protein CDP regulates mammary-specific gene transcription and tumorigenesis. *Mol. Cell Biol.*, 24 (11), 4810-4823

Curriculum vitae

Eva Fuchs

Personal Details

Date of birth	January 20, 1975
Nationality	Austrian
Address	Eckmüllnergasse 1/2/10 1160 Vienna

Education

University of Vienna and University of Veterinary Medicine, Austria	Doctoral studies in Molecular Biology PhD thesis: 'Transcriptional Regulation of the Mouse Mammary Tumor Virus LTR' Supervisor: Prof. Dr. Timothy Skern and Dr. Francoise Rouault	10/2007 -present
University of Vienna, Austria	Studies of Molecular Biology Diploma thesis: 'The Role of PDGF Signaling in late Stage Tumor Progression' Supervisor: Prof. Dr. Wolfgang Mikulits	02/2001-06/2006
University of Applied Sciences, Vienna, Austria	Studies of Biomedical Laboratory Science Degree: 'Dipl. Med.-techn. Analytikerin' Diploma Thesis: 'Staphylococcus aureus Infektionen unter besonderer Berücksichtigung des MRSA' Supervisor: OA. Dr. Dietmar Wolf	10/1994-10/1997

Occupational activities

University of Veterinary Medicine, Vienna, Austria	Early stage Researcher / Virology	04/2006-10/2009
Medical University, Vienna, Austria	Research Assistant / Cancer Research	09/2002-01/2005
	Research Assistant / Pharmacology	06/2001-08/2002
	Research Assistant / Oncology	12/1999-05/2001
University of Veterinary Medicine, Vienna, Austria	Research Assistant / Virology	07/1998-11/1999

Publications

Francoise Rouault, Shiva Badihi Nejad Asl, Stefanie Rungaldier, **Eva Fuchs**, Brian Salmons, and Walter H. Günzburg (2007) Promoter complex in the central part of the mouse mammary tumor virus long terminal repeat. *Journal of Virology*, 81(22), 12572-12581

Alexandra N.M. Fischer*, **Eva Fuchs***, Mario Mikula*, Heidemarie Huber, Hartmut Beug, and Wolfgang Mikulits (2007) PDGF essentially links TGF- β signaling to nuclear beta-catenin accumulation in hepatocellular carcinoma progression. *Oncogene*, 26(23), 3395-3405

*These authors contributed equally to this work.

Verena Proell, Mario Mikula, **Eva Fuchs** and Wolfgang Mikulits (2005) The plasticity of p19 ARF null hepatic stellate cells and the dynamics of activation. *Biochimica et Biophysica Acta*, 1744(1), 76-87

Alexandra N.M. Fischer, Blanca Herrera, Mario Mikula, Verena Proell, **Eva Fuchs**, Josef Gotzmann, Rolf Schulte-Hermann, Hartmut Beug and Wolfgang Mikulits (2005) Integration of Ras subeffector signaling in TGF- β mediated late stage hepatocarcinogenesis. *Carcinogenesis*, 26(5), 931-942

Mario Mikula, **Eva Fuchs**, Heidemarie Huber, Hartmut Beug, Rolf Schulte-Hermann, and Wolfgang Mikulits (2004) Immortalized p19ARF null hepatocytes restore liver injury and generate hepatic progenitors after transplantation. *Hepatology*, 39(3), 628-634