Gene Ther Mol Biol Vol 6, 169-181, 2001

A retroviral model for tissue-specific transcription: lessons for gene therapy

Review Article

Quan Zhu and Jaquelin P. Dudley*

Section of Molecular Genetics and Microbiology and Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX

*__Correspondence__: Jaquelin P. Dudley Address: Section of Molecular Genetics and Microbiology, The University of Texas at Austin, 100 W. 24th Street, Austin, TX 78705; Telephone: (512) 471-8415; Fax: (512) 471-7088; E-mail: jdudley@uts.cc.utexas.edu

Key words: MMTV, T-cell lymphoma, steroid reseptor

Abbreviations: β-chain, (Vβ); envelope proteins, (Env); GA-binding protein, (GABP); glucocorticoid receptor, (GR); homeodomain, (HD); hormone responsive elements, (HREs); long terminal repeats, (LTRs); major histocompatibility complex, (MHC); Mouse mammary tumor virus, (MMTV); negative regulatory elements, (NREs); nuclear factor 1, (NF1); Nucleosome A, (Nuc-A); octamer, (OCT); splice acceptor, (SA); Splice donor, (SD); T-cell receptor, (TCR)

Received: 31 October 2001; accepted: 12 November 2001; electronically published: February 2004

Summary

Mouse mammary tumor virus (MMTV) induces breast cancer in mice by transmission of virus from infected mothers to susceptible offspring through milk. During milk-borne MMTV transmission, virus must be transferred to B and T cells in the gut-associated lymphoid tissue, and these lymphocytes carry the infection to the mammary gland. Wild-type MMTV strains have been selected for optimal virus expression in lactating mammary gland cells, while minimizing gene expression and integration in other cell types. In particular, the MMTV transcriptional control region contains binding sites for both transcriptional repressor proteins, e.g., SATB1 and CDP, and positive factors, e.g., glucocorticoid receptor. Studies of MMTV transcriptional regulation may provide important lessons for the design of effective and safe gene therapy vectors.

I. Introduction

Mouse mammary tumor virus (MMTV) is a betaretrovirus that was shown in the 1930s to induce breast cancer in mice (Bentvelzen et al, 1972; Nandi and McGrath, 1973; Dudley, 1999). Female mice derived from high-mammary-cancer incidence strains transmit the virus to their offspring through the milk (**Figure 1**). More recent experiments have shown that MMTV traverses the gastrointestinal tract until the specialized M cells of the small intestine take up the virus (Golovkina et al, 1999). MMTV then infects B cells in the gut-associated lymphoid tissue (Karapetian et al, 1994). These infected B cells express a virally-encoded protein, called superantigen or Sag, at the surface in conjunction with major histocompatibility complex (MHC) class II antigen (Acha-Orbea et al, 1991; Choi et al, 1991; Janeway, Jr., 1991). Sag is a type II transmembrane glycoprotein (Korman et al, 1992) that, upon recognition by the variable region of the β-chain (Vβ) of the T-cell receptor (TCR), causes the release of cytokines and proliferation of bystander B and T cells (Acha-Orbea, 1992). It is widely thought that this proliferation allows MMTV infection of additional lymphoid cells that provide a means for viral trafficking to epithelial cells in the mammary gland (Golovkina et al, 1992; Held et al, 1993; Beutner et al, 1994). Viral replication in mammary gland epithelial cells of female mice results in release of high levels of MMTV particles into the milk (Nandi and McGrath, 1973). Genetically engineered mice that lack B cells or Sag-reactive T cells cannot be infected by milk-borne viruses (Golovkina et al, 1992; Held et al, 1993), and MMTV proviruses carrying a frameshift mutation within the *sag* gene are not infectious by the milk-borne route (Golovkina et al, 1995). These results are consistent with a role for B and T cells in Sag-mediated amplification of MMTV-infected lymphoid cells. However, experiments also have shown that both B cells and Sag-reactive T cells are required for MMTV dissemination within the mammary gland

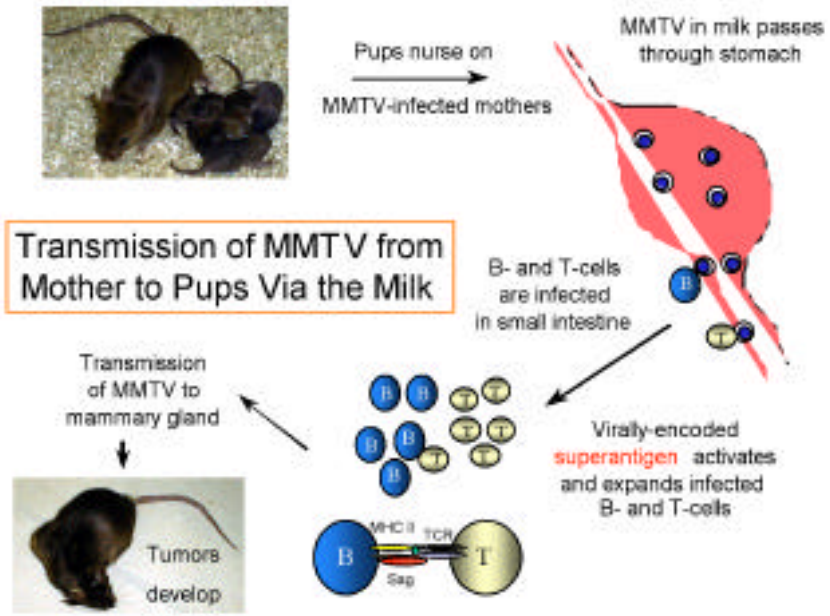

Figure 1: Life cycle of MMTV. MMTV is produced in the mammary glands of infected female mice and is transmitted to newborn pups through the mother's milk. The ingested virus infects B and T cells in the gut-associated lymphoid tissues. The infected B cells express a superantigen (Sag) that activates subsets of T cells and provide more targets for viral infection. In mammary gland, hormonal stimulation during pregnancy and lactation dramatically increases MMTV replication and allows insertional mutagenesis of proto-oncogenes and the development of mammary tumors.

(Golovkina et al, 1998), although the mechanism for transmission is not clear. Therefore, the MMTV life cycle requires virus replication in B cells, T cells, and mammary gland cells.At the cellular level, MMTV particles bind to a cell surface receptor that is believed to mediate viral fusion with the plasma membrane. Two such receptors (MTVR1 and MTVR2) have been described, but little is known about the exact nature of the receptors or the process of viral internalization (Hilkens et al, 1983; Golovkina et al, 1998). The incoming viral genome, consisting of two identical copies of single-stranded, positive-sense RNA, is replicated using a virally-encoded RNA-dependent DNA polymerase or reverse transcriptase (Telesnitsky and Goff, 1997). The process of reverse transcription in the cytoplasm generates a double-stranded DNA or provirus that has characteristic long terminal repeats (LTRs) that are not present in genomic RNA (**Figure 2**). The provirus then enters the nucleus through an unknown mechanism and integrates into the host cell chromosomes at sites that are believed to be relatively random (Pryciak et al, 1992; Withers-Ward et al, 1994; Weidhaas et al, 2000). Entry into the nucleus is probably dependent, or at least accelerated by nuclear envelope breakdown during mitosis, and this may explain why Sag-induced proliferation of lymphoid cells is necessary for efficient viral dissemination during milk-borne infection (**Figure 1**). A preintegration complex consisting of some virion proteins, including the integrase protein or IN, mediates integration. IN catalyzes cleavage of the ends of the linear double-stranded DNA as well as a staggered break in cellular DNA (Brown, 1997). Following integration, the proviral 5' LTR is recognized by host RNA polymerase II and is transcribed into a full-length RNA that is structurally identical to virion RNA. Thus, complete retroviral replication requires both the retrovirally-encoded reverse transcriptase and host enzymes (Rabson and Graves, 1997).

The full-length MMTV RNA may take one of several different pathways in infected cells. The RNA may be spliced to give a sub-genomic RNA that encodes the viral Env, yet a fraction of the genome-length RNA always is exported directly to the cytoplasm. In the cytoplasm, genomic RNA is translated into virion structural proteins as well as the reverse transcriptase and IN. Genomic RNA also is used directly for packaging by the virion or Gag proteins, and for betaretroviruses, the initial assembly into particles occurs in the cytoplasm (Dickson and Peters, 1983). Envelope proteins are translated on membrane-bound ribosomes, while processing and glycosylation of Env proteins occurs in the endoplasmic reticulum and Golgi prior to budding of assembled particles (Swanstrom and Wills, 1997).

Synthesis of the Sag protein reportedly occurs using viral RNAs originating from four different promoters, including two within the LTR and two within the *env* gene (Wheeler et al, 1983; Elliott et al, 1988; Miller et al, 1992; Jarvis et al, 1994; Reuss and Coffin, 1995; Arroyo et al, 1997) (**Figure 2**). However, mutagenesis of a unique

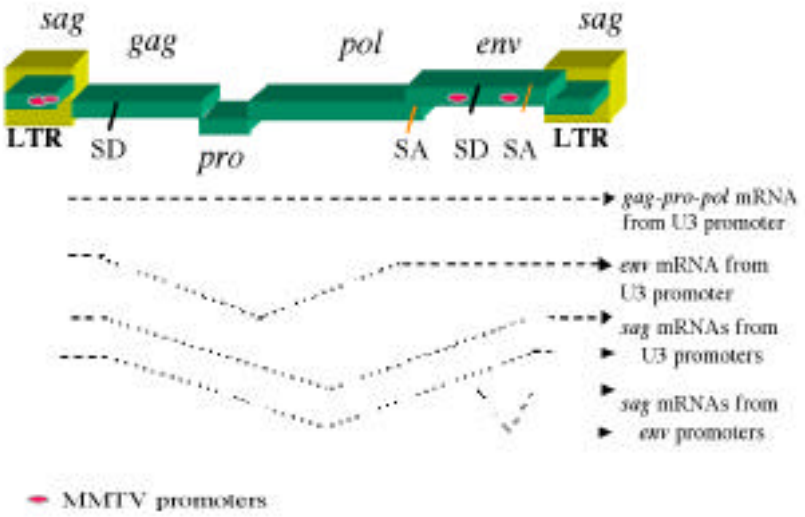

Figure 2: MMTV provirus structure and transcripts
Yellow boxes represent the proviral long terminal repeats (LTRs) and arrows represent viral transcripts. Portions of arrows that are interrupted by V-shaped structures indicate regions removed by splicing. Reported MMTV promoters are shown by ovals. The sag open reading frame in the 5' LTR lacks a promoter. The thinner gray boxes show viral open reading frames. Splice donor sites (SD) and splice acceptor sites (SA) are denoted by slashes. (Adapted from Mustafa et al, 2000.)

splice donor site within the viral *env* gene of an infectious MMTV provirus is sufficient to eliminate production of the Sag (Mustafa et al, 2000). These results indicate that only the spliced mRNA originating within the *env* region is required for Sag production. Since the promoter controlling virion protein synthesis is located within the LTR and an intragenic *env* promoter regulates Sag levels (Mustafa et al, 2000), MMTV may dissociate these two functions in some cell types (Wrona et al, 1998).

II. Factors that influence tissue-specific MMTV transcription and disease

A. MMTV-induced mammary cancers

MMTV is known to induce mammary cancers in mice, and the tumor incidence increases with increasing numbers of pregnancies (Nandi and McGrath, 1973). The relationship between tumor frequency and pregnancies appears to be related to the higher numbers of proviral insertions that occur with longer times for active MMTV replication. If each single insertion has a relatively low probability of affecting a given oncogene, then additional insertions will improve the chance of oncogene activation. Analysis of MMTV-induced mammary tumors has shown that proviral insertions are frequently detected near a subset of oncogenes, including *Wnt-1/int-1*, *fgf-3/int-2*, *notch-like/int-3*, *fgf-4/hst*, *aromatase/intH/int-5*, *int-6*, *Wnt-10b*, and *fgf-8* (Morris et al, 1991; Kwan et al, 1992; Lee et al, 1995; MacArthur et al, 1995; Marchetti et al, 1995). Since the probability of MMTV insertion near any one gene is relatively small, the presence of proviral insertions near two different oncogenes in a clonal tumor population suggests that there is cooperativity between such genes for tumor growth (Jonkers and Berns, 1996). Such co-operativity between the *Wnt-1* and *fgf-3* genes has been confirmed by the detection of dual insertions in MMTV-induced breast cancers (Clausse et al, 1993) and by experiments with transgenic mice (Shackleford et al, 1993; van Leeuwen and Nusse, 1995). Mice carrying the *Wnt-1* transgene display mammary gland hyperplasia, with approximately 50% of the animals developing mammary cancer (Brown et al, 1986; Rijsewijk et al, 1987). MMTV infection of *Wnt-1* transgenic mice increases the frequency of mammary tumors, and these tumors often have proviral insertions near *fgf-3* and *hst* (Shackleford et al, 1993), confirming their ability to co-operate with *Wnt-1*. Mice carrying both *fgf-3* and *Wnt-1* transgenes also have an increased incidence of mammary tumors compared to mice carrying either transgene alone (Li et al, 2000).

B. MMTV-induced T-cell lymphomas

A number of MMTV variants induce T-cell lymphomas in mice rather than mammary cancer (Michalides et al, 1982; Dudley and Risser, 1984; Ball et al, 1985; Lee et al, 1987). These MMTVs invariably have a 350 to 500-bp deletion within the U3 region of the LTR that overlaps with the transcriptional control region that regulates the synthesis of virion structural genes as well as the coding region for superantigen (Michalides et al, 1985; Lee et al, 1987; Hsu et al, 1988; Ball et al, 1988) (**Figure 3**). Substitution of the U3 region converts a mammotropic MMTV into a lymphomagenic virus, indicating that this region is necessary and sufficient to alter the type of tumor produced (Yanagawa et al, 1993). What then is the molecular basis for this change in disease specificity?

Many experiments indicate that the MMTV U3 region has negative regulatory elements (NREs) that

suppress viral transcription from the standard LTR promoter (Morley et al, 1987; Hsu et al, 1988; Lee et al, 1991; Bramblett et al, 1995). At least two types of NREs have been described that map within the region deleted in lymphomagenic MMTVs (Lee et al, 1991; Bramblett et al, 1995). The first type of NRE has been mapped to several elements localized between –364 and –427 and between the proximal and distal hormone response elements (HRE) (ca. –160 to –140) relative to the start of genomic RNA (Langer and Ostrowski, 1988; Mink et al, 1990; Lee et al, 1991). Deletion of these elements has been shown to elevate MMTV LTR-reporter gene expression in fibroblast cells, but its effect in lymphoid cells is unknown (Lee et al, 1991). The NRE region between –433 and –418 has been reported to bind to a nuclear protein of approximately 100 kDa purified from HeLa cells (Kang and Peterson, 1999).

A second type of NRE also was identified using deletion analysis of an MMTV LTR-reporter construct (Bramblett et al, 1995). In transient transfection experiments in mink lung cells, sequential deletions revealed that there were at least two NREs (called promoter-proximal and promoter-distal) located between –655 and –165 relative to the start of the genomic RNA, a region encompassing the deletions found in lymphomagenic MMTVs (Bramblett et al, 1995). Transgenic animals expressing MMTV LTR-reporter constructs have been shown to recapitulate the tissue-specific expression of endogenous MMTV proviruses resident within most mouse genomes (Ross et al, 1990). A subset of LTR deletion constructs was used in transgenic mouse experiments to confirm the presence of two NREs. Deletion of either region was sufficient to allow MMTV transcription in tissues where the wild-type virus was not expressed (i.e., brain, heart, skeletal muscle, and liver) (Henrard and Ross, 1988). Both wild-type and NRE-deletion mutants were highly expressed in the most permissive tissue, lactating mammary gland, and at lower levels in a variety of semi-permissive tissues, including lymphoid and reproductive tissues (Ross et al, 1990). These experiments confirmed that deletion of specific LTR sequences could dramatically alter MMTV transcriptional specificity. Such results also suggested that relief of MMTV transcriptional suppression in specific tissues, e.g., lymphoid cells, could lead to increased mutagenic insertions that result in leukemias (Bramblett et al, 1995; Liu et al, 1997).

C. Molecular basis for disease specificity

Studies of MMTV disease variants indicated that a key region within the LTR was responsible for tissue-specific transcription and disease specificity (Michalides and Wagenaar, 1986; Bramblett et al, 1995; Mertz et al, 2001). What is the molecular basis for tissue-specific transcription and how can this affect the type of tumor induced? Experiments in a variety of genetic systems suggest that the binding of proteins to DNA regulatory elements controls transcription (Beato, 1996). Since retroviruses, including MMTV, have relatively simple genomes that encode few genes, cellular protein factors must mediate the majority of transcriptional events (Rabson and Graves, 1997).

1. Steroid receptors

The MMTV LTR has served as a model transcriptional element for many years, and early studies indicated that viral transcription is inducible by glucocorticoids and several other steroid hormones (Parks et al, 1974; Payvar et al, 1981). Addition of glucocorticoid hormones to MMTV-infected cells typically gives 10- to 50-fold increases in the level of viral RNA (Ringold et al, 1977). Subsequently, the HRE has been mapped upstream of the transcription initiation site for genomic RNA (**Figure 3**) (Groner et al, 1982; Majors and Varmus, 1983). Linkage of the HRE to heterologous promoters is sufficient to confer hormone responsiveness (Hynes et al, 1983; Chandler et al, 1983).

The HRE consists of several independent receptor-binding sites that have a similar consensus sequence, TGTTCT (Buetti and Kuhnel, 1986; Kuhnel et al, 1986). Ligand binding to the receptor (e.g., glucocorticoid receptor or GR) allows entry into the nucleus and binding to the HRE. Binding of hormone receptor to the MMTV HRE results in nucleosomal changes near the promoter that then lead to binding of nuclear factor 1 (NF1) (Beato, 1996) and recruitment of the basal transcription machinery. Nucleosome A (Nuc-A) has been mapped over the TATA box and the transcription start site, whereas the octamer (OCT) motifs (see below) are located between Nuc-A and Nuc-B (**Figure 3**). Upon addition of the synthetic glucocorticoid, dexamethasone, transcription from the MMTV LTR is increased and the DNA encompassed by Nuc-B becomes hyper-sensitive to many reagents, such as restriction enzymes, nucleases or chemical probes (Zaret and Yamamoto, 1984; Archer et al, 1992), suggesting that Nuc-B acquires a more “open” configuration. The hormone-induced MMTV promoter also shows increased binding by NF1, Oct-1 (OTF-1), and TBP (Lee and Archer, 1994), whereas in the absence of hormone the positioning of Nuc-A and -B excludes ubiquitous transcription factors from the promoter. Hormone-activated GR modifies Nuc-B to allow binding of NF1 and other factors, leading to formation of the transcription initiation complex (Hager et al, 1993; Archer, 1993). NF1 binds strongly to sites on free DNA but is unable to bind in a nucleosomal context (Archer et al, 1991). Conversely, GR has a lower affinity for sites in free DNA than the same sites bound to nucleosomes. However, in vivo footprinting experiments have not been able to reproducibly detect GR bound to the HREs in the presence of hormone (Lee and Archer, 1994). These results suggest that GR interacts with its cognate sites in a “hit and run” manner (Lee and Archer, 1994).

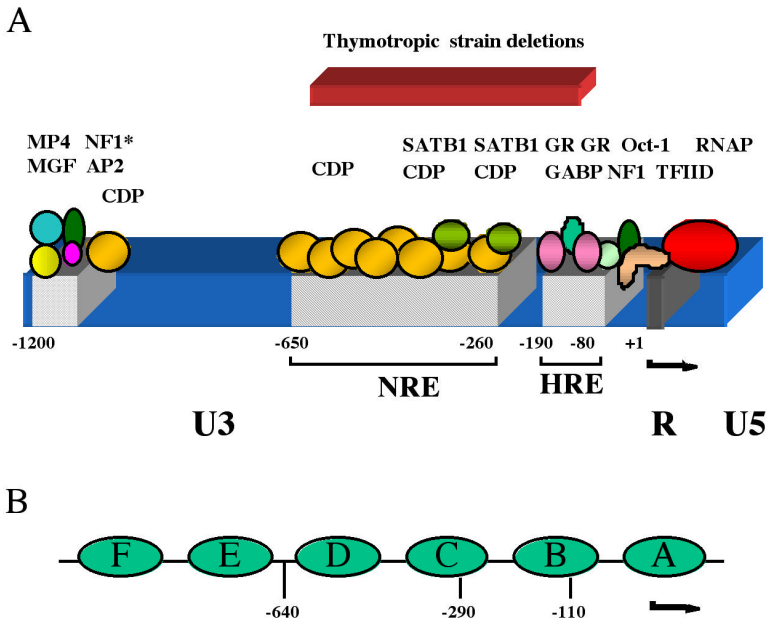

Figure 3: Transcription factors that bind the MMTV LTR
A. Ovals and circles represent different transcription factors that bind to different regions in the MMTV LTR. The maximal deletion observed in acquired MMTV proviruses from T-cell tumors is shown above the LTR. Numbers below the LTR refer to the distance (in bp) from the transcriptional start site of genomic RNA (+1) (Brandt-Carlson et al, 1993). Circles do not indicate the size of the proteins or their interactions with each other. Abbreviations: GR (glucocorticoid receptor), TFIID (transcription factor IID), RNAP (RNA polymerase II), Oct-1 (octamer-binding protein 1), NF1 (nuclear factor 1), NF1*, a member of the NF1 family, CDP (CCAAT displacement protein), SATB1 (special AT-rich binding protein 1), GABP (GA-binding protein), hormone responsive elements (HREs), and NREs. There appear to be at least eight CDP-binding sites, two SATB1-binding sites, and six GR binding sites in the LTR U3 region. **B**. Relative positions of nucleosomes in the MMTV LTR. Nucleosomes A to F are indicated as ovals. (Adapted from Fletcher et al, 2000.)

Hormone receptor binding is thought to contribute to the high levels of MMTV RNA and virions produced in the lactating mammary gland during milk-borne transmission (Dudley, 1999). However, since functional glucocorticoid receptors are found in many tissues, including liver where MMTV RNA is not expressed (Henrard and Ross, 1988; Ross et al, 1990), these receptors do not explain the tissue-specific nature of MMTV expression.

2. Special AT-rich binding protein 1

The LTR region deleted from leukemogenic MMTVs has served as a starting point for the isolation of tissue-specific transcription factors. Using probes derived from the LTR NREs, gel shift experiments revealed the presence of two DNA-binding complexes that were referred to as NBP and UBP (Bramblett et al, 1995). The NBP complex was present in both T-cell and lung-cell extracts, but was absent in mammary cell extracts, whereas UBP was present in all extracts tested. Purification of the NBP complex showed that it was composed of a previously identified protein, special AT-rich binding protein 1 or SATB1, that was enriched in thymocytes and T-cells (Dickinson et al, 1992; Liu et al, 1997).

SATB1 originally was isolated as a protein that binds to nuclear matrix- or scaffold-associated regions (MARs or SARs) localized in the immunoglobulin heavy chain intronic enhancer (Dickinson et al, 1992); this factor appears to bind to the MARs associated with the bases of chromatin loops (de Belle et al, 1998). MARs are AT-rich stretches of DNA that have been associated with binding to the nuclear framework in the nucleus and serve to regulate cellular processes such as transcription and DNA replication (Boulikas, 1995). In addition to the MMTV LTR, SATB1 binds to the regulatory elements of the CD8, TCRβ, gp-*phox*, and immunoglobulin heavy chain genes

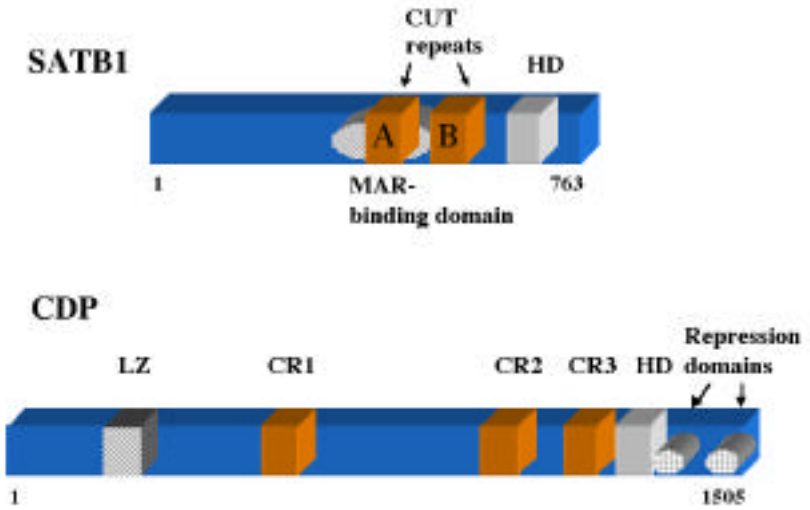

Figure 4: CDP and SATB1 protein structure The Cut-like repeats A and B, atypical homeodomain (HD), and a MAR-binding domain are shown in SATB1. The three Cut repeats (CR1, CR2, and CR3), a HD, and a leucine-zipper region are indicated within the human CDP protein. The ovals show the C-terminal repression domains.

(Dickinson et al, 1992; Banan et al, 1997; Chattopadhyay et al, 1998; Hawkins et al, 2001) and has been described variously as a transcriptional repressor (Bramblett et al, 1997; Kohwi-Shigematsu et al, 1997; Liu et al, 1997) or activator (Banan et al, 1997), depending on the regulatory element analyzed.

SATB1 contains three DNA-binding domains, consisting of two Cut-like repeats (A and B) and an atypical homeodomain (Dickinson et al, 1997) (**Figure 4**). However, the major MAR-binding domain appears to be localized in the region of Cut-like repeat A (Nakagomi et al, 1994). SATB1 has been shown to bind to the proximal MMTV NRE, and mutation of the binding site at +924 (-271 from the start of genomic RNA) upregulates MMTV transcription from the standard LTR promoter (Liu et al, 1997). MMTV LTR-reporter genes that contain the SATB1-binding mutation at +924 have been used for the construction of transgenic mice that show the highest levels of expression in lymphoid tissues, in contrast to wild-type MMTV LTRs that have optimal expression in lactating mammary gland (Ross et al, 1990; Liu et al, 1997). SATB1-null mice have stunted growth, small thymi and spleens, and thymocyte development is blocked at the CD4+CD8+ stage of differentiation (Alvarez et al, 2000). Thus, SATB1 binding to the MMTV LTR appears to be a major determinant of tissue-specific expression in lymphoid cells.

3. CCAAT-displacement protein

The second major DNA-binding activity localized to the MMTV NRE, initially called UBP (Liu et al, 1997), was identified as the murine equivalent of the human CCAAT-displacement protein or CDP (Neufeld et al, 1992). CDP is a 180 to 190 kDa protein that contains a leucine-zipper region near the N-terminus, four DNA-binding domains [three Cut repeat domains (CR1, 2, and 3) and a homeodomain (HD)], and two C-terminal repression domains (**Figure 4**). Expression of individual binding domains as GST-fusion proteins indicated that each region bound to a slightly different AT-rich sequence (Aufiero et al, 1994). More recent experiments suggest that the DNA-binding domains function in pairs and that CR1 and CR2 interact to induce the originally described displacement activity for the CCAAT-binding factor (CBF) on the sperm histone H2B-1 gene in sea urchin embryos (Moon et al, 2000). Such results emphasize the diversity of sites that may be recognized by CDP.

The *CUTL1* gene (encoding CDP) is also known as *cut* in *Drosophila*, *Clox* in dogs, and *Cux-1* in mice (Blochlinger et al, 1988; Andres et al, 1992; Valarche et al, 1993), and murine and avian cells contain a second gene referred to as *Cux-2* (Quaggin et al, 1996). The *Drosophila* Cut protein is expressed in a variety of embryonic and adult tissues, including the peripheral and central nervous systems, Malpighian tubules, ovarian follicle cells, cells of the wing margin, adepithelial cells of the wing and leg discs, muscle cells, and cone cells of the eye (Bodmer et al, 1987). Both lethal and viable *cut* mutations have been identified, and the best characterized of these mutations are the embryonic lethal type that allow the transformation of external sensory organs into internal chordotonal organs (Bodmer et al, 1987). Such mutations indicate that Cut is a major determinant of cell-type specification in *Drosophila* (Bodmer et al, 1987). Similarly, a role for CDP in cell-type specification has been suggested by experiments in mammalian cells, and human CDP or murine Cux-1 can at least partially rescue some of the *cut* mutations in flies (Ludlow et al, 1996).

The role of Cux/CDP in mammals has been investigated by germ line manipulations of the gene in mice. Initial attempts to knockout the gene yielded an exon skipping mutant that produced a truncated form of

the *CUTL1* gene that lacked 246 amino acids in the CR1 DNA-binding domain (ΔCR1), but was capable of binding to DNA (Tufarelli et al, 1998). Mice homozygous for this mutation had curly whiskers and wavy hair and exhibited a failure to thrive among pups born to mutant females. Although the exact nature of the defect was not determined, preliminary experiments indicated a defect in maternal milk that was associated with a decrease in ε-casein expression (Tufarelli et al, 1998). More recently, the CR3 and HD of the murine *Cux* gene have been replaced by an in-frame *lacZ* gene to give homozygous mutant mice that lack nuclear Cux expression and the ability to repress a target reporter gene in transient assays (Ellis et al, 2001). Mice homozygous for the CR3-HD mutation died after birth due to defects in maturation of the lung epithelium. The mutant phenotype was more severe in an inbred background, and homozygous mutant mice on an outbred background showed growth retardation and defects of the hair follicles (Ellis et al, 2001). Our experiments using the ΔCR1 mice also suggest that breeding of this mutation onto the BALB/c background exacerbates the lethality of the CDP-mutant phenotype (Zhu, Lozano, and Dudley, unpublished results). Such data confirm the essential role of Cux/CDP in the normal developmental program of several tissues, including the lungs, hair follicles, and mammary glands.

CDP is a transcriptional repressor of multiple cellular genes, including gp91-*phox*, TCRβ, CD8, immunoglobulin heavy chain, and c-*myc*, as well as several viral genomes, including MMTV and human papilloma viruses (Skalnik et al, 1991; Dufort and Nepveu, 1994; Banan et al, 1997; Pattison et al, 1997; Chattopadhyay et al, 1998; Ai et al, 1999; Wang et al, 1999; Zhu et al, 2000). Several reports indicate that CDP is expressed at high levels in undifferentiated cells, and that upon terminal differentiation, CDP-mediated repression is lost (Nepveu, 2001). Since MMTV is expressed at high levels in differentiated cells of the lactating mammary gland and at much lower levels in virgin glands, we examined CDP binding to the NRE using nuclear extracts derived from several different developmental stages. These results showed that CDP binding to the MMTV LTR was highest in virgin mammary extracts, but was undetectable in extracts from the lactating mammary glands (Liu et al, 1997; Zhu et al, 2000). Sp1 binding to a consensus sequence actually increased during murine mammary development, indicating that nuclear extracts from lactating mammary gland were not degraded (Zhu et al, 2000). These results suggested that there was an inverse relationship between the presence of CDP and the transcriptional activity of the MMTV LTR.

Transient transfection assays in mouse mammary cells showed that the activity of an MMTV LTR-reporter construct was diminished in a dose-dependent manner, depending on the amount of CDP-expression vector present (Zhu et al, 2000). Further experiments also revealed that CDP could repress both basal and glucocorticoid-induced levels of LTR-reporter expression (Zhu and Dudley, in press). These data suggested that MMTV RNA levels are highest in the lactating mammary gland due to the presence of active steroid hormone receptors to promote nucleosomal rearrangements near the viral promoter and the absence of CDP that may interfere with the function of steroid receptors and other positively-acting factors.

Further experiments were performed to determine if CDP-mediated repression of MMTV expression was a direct result of DNA binding to the viral negative regulatory elements. Multiple CDP-binding sites have been mapped on the MMTV LTR by DNase I footprinting and direct DNA-binding experiments (Zhu et al, 2000; Zhu and Dudley, in press) (**Figure 3**). Mutation of two independent binding sites, one in the proximal and one in the distal NRE, were shown to greatly diminish CDP binding to the MMTV LTR, and such mutations were sufficient to elevate MMTV LTR-reporter gene activity in both transient and stable transfection assays (Zhu et al, 2000; Zhu and Dudley, in press). If CDP is a transcriptional repressor in undifferentiated mammary gland, then CDP-binding site mutations should increase virus transcription and replication in early stages of breast development, thus decreasing the latency of tumor development. Two of these mutations, one at +692 (-503) and another at +838 (-357), have been transferred into the LTR of an infectious MMTV provirus (Shackleford and Varmus, 1988) and used for the inoculation of susceptible BALB/c mice. Preliminary results indicate that the latency of mammary tumors induced by CDP-mutant viruses is reduced compared to tumors induced by the wild-type virus. In addition, the growth rate and number of tumors induced by the CDP mutants is accelerated relative to that observed for the wild-type virus (Zhu, Lozano and Dudley, in preparation).

These results confirm that CDP is a developmentally regulated transcription factor that suppresses MMTV expression in early stages of mammary gland development. CDP-binding sites in the LTR presumably are retained to minimize the number of mutagenic integration events in the life of the mouse, thus allowing transmission to increased numbers of offspring.

4. Other factors affecting tissue-specific MMTV expression

Cell-type specific and ubiquitous factors also have been shown to control MMTV expression in mammary cells. Stewart et al first reported the ability of the MMTV LTR to direct mammary-specific transcription (Stewart et al, 1984). One enhancer-like element that localized to the 5' end of the LTR (-1072 to -1052) bound to a mammary-specific factor called MP-4. Deletion of this region decreased both glucocorticoid-induced and basal transcription from the MMTV promoter (Haraguchi et al, 1997). Transgenic mouse experiments defined a region

between –1166 and –987 that directed MMTV LTR-transgene expression in mammary and salivary gland tissues (Mok et al, 1992). This enhancer functioned in both lactating and non-lactating mammary glands. At least six functional cis-acting elements have been mapped to this enhancer, including mp4 and mp5 (Lefebvre et al, 1991; Mellentin-Michelotti et al, 1994) and F2, F3, F11, and F12 (Mink et al, 1992). Some of the transcription factors that bind to these elements appear to be related to AP-2 and NF1/CTF (Mellentin-Michelotti et al, 1994; Kusk et al, 1996). The 5' end of the LTR also contains a sequence motif TTCGGAGAA that potentially binds to mammary gland factor (MGF) (Gouilleux et al, 1994). MGF (otherwise known as Stat5a) is a transcription factor regulated by prolactin through phosphorylation by the JAK family of tyrosine kinases (Wakao et al, 1994). Stat5 or a related protein may also bind to an MMTV LTR sequence near +520 (-675 relative to the start of MMTV genomic RNA) (Qin et al, 1999).

Several transcription factors have been shown to participate in T-cell or lymphoid-specific MMTV transcription. Some MMTV variants that induce T-cell lymphomas, e.g., type B leukemogenic virus or TBLV, have both the characteristic LTR deletion that removes the negative regulatory elements as well as a triplication of sequences flanking the deletion (Ball et al, 1988). The structure of the triplicated region is reminiscent of many retroviral enhancer elements, and transient transfection experiments have shown that the triplicated region in the TBLV LTR allows greatly enhanced expression in T-cell lines, but not other cell types tested (Mertz et al, 2001). Linker scanning mutations within the LTR triplication revealed a critical region for T-cell specific expression in transient assays, and this region was used for the identification of cellular DNA-binding factors. At least three DNA-binding activities were identified within this region, including two unknown factors called NF-A and NF-B, and AML-1/Runx1 (Mertz et al, 2001). Overexpression of the transcriptionally active form of AML-1 in mammary cells increased the activity of TBLV LTR-reporter constructs, suggesting that AML-1 contributes to the T-cell specific nature of the TBLV enhancer.

Other regulatory elements that affect tissue-specific MMTV transcription also have been described. Several cellular binding activities have been mapped upstream of the distal HRE and downstream of the known CDP-binding sites; these activities have been referred to as DRa and DRc (Cavin and Buetti, 1995). DRa was present in tissues that were permissive for MMTV transcription, whereas DRc was ubiquitously expressed (Cavin and Buetti, 1995). One factor that binds to two areas located between –139 and –164 appears to be the heterodimeric Ets factor GA-binding protein (GABP). GABP was shown to increase the hormone-responsiveness of an MMTV LTR-reporter gene in transient assays performed in B cells (Aurrekoetxea-Hernandez and Buetti, 2000). The MMTV LTR also contains three overlapping sites related to the consensus octamer sequence ATGCAAAT (Bruggemeier et al, 1991; Huang et al, 1993), and it has been suggested that the transcription factor Oct-2 participates in B-cell specific MMTV transcription (Kim and Peterson, 1995).

III. Implications for gene therapy

One of the many important issues of gene therapy is tissue-specific expression of therapeutic genes. The MMTV LTR is a well-characterized retroviral transcriptional unit, and it is suitable for further manipulations in both tissue culture and in mouse model systems. In particular, the MMTV LTR clearly has been subject to selection so that virus expression is optimal in lactating mammary cells, but not in other cell types or undifferentiated mammary cells. Although MMTV must be transmitted through B and T lymphocytes to developing epithelial cells in the mammary glands of offspring, expression is suppressed in these cell types to minimize potential mutagenic events that will shorten the life of the infected mother. Furthermore, the LTR of an MMTV variant (TBLV) shows high levels of expression in lymphocytes. Such naturally occurring variants selected *in vivo* can serve as potential tools for cell-type specific gene delivery systems.

Recent data have indicated that the expression pattern of the MMTV promoter during cellular differentiation appears to result from overlapping and sophisticated positive and negative elements in the LTR. The dissection of viral *cis*-elements and identification of cellular *trans*-factors, such as SATB1 and CDP that are in turn developmentally regulated, make the MMTV LTR amenable to further manipulation for specific purposes. For example, inclusion or enhancement of MMTV negative regulatory elements should minimize viral expression in the majority of tissues, other than mammary tissue. Conversely, disruption of negative regulation may be required before the MMTV LTR can be used in directing therapeutic gene expression in mammary tumor cells that are relatively undifferentiated. In conclusion, further studies of fundamental gene expression should allow the development of additional strategies for the design of effective and safe gene therapy vectors.

Acknowledgements

We acknowledge the helpful comments by members of the Dudley lab. This work was supported by grants CA34780, CA52646, and CA7760 from the National Institutes of Health.

References

Acha-Orbea H (**1992**) Retroviral superantigens. **Chem Immunol** 55, 65-86.

Acha-Orbea H, Shakhov AN, Scarpellino L, Kolb E, Muller V, Vessaz-Shaw A, Fuchs R, Blochlinger K, Rollini P, Billotte J, Sarafidou M, MacDonald HR and Diggelmann H (**1991**)

Clonal deletion of V β 14-bearing T cells in mice transgenic for mammary tumour virus. **Nature** 350, 207-211.

Ai W, Toussaint E and Roman A (**1999**) CCAAT displacement protein binds to and negatively regulates human papillomavirus type 6 E6, E7, and E1 promoters. **J Virol** 73, 4220-4229.

Alvarez JD, Yasui DH, Niida H, Joh T, Loh DY and Kohwi-Shigematsu T (**2000**) The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. **Genes Dev** 14, 521-535.

Andres V, Nadal-Ginard B and Mahdavi V (**1992**) Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding regulatory proteins differentially expressed during development. **Development** 116, 321-334.

Archer TK (**1993**) Nucleosomes modulate access of transcription factor to the MMTV promoter in vivo and in vitro. **Ann N Y Acad Sci** 684, 196-198.

Archer TK, Cordingley MG, Wolford RG and Hager GL (**1991**) Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. **Mol Cell Biol** 11, 688-698.

Archer TK, Lefebvre P, Wolford RG and Hager GL (**1992**) Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation [published erratum appears in Science 1992 Apr 10;256(5054):161]. **Science** 255, 1573-1576.

Arroyo J, Winchester E, McLellan BS and Huber BT (**1997**) Shared promoter elements between a viral superantigen and the major histocompatibility complex class II-associated invariant chain. **J Virol** 71, 1237-1245.

Aufiero B, Neufeld EJ and Orkin SH (**1994**) Sequence-specific DNA binding of individual cut repeats of the human CCAAT displacement/cut homeodomain protein. **Proc Natl Acad Sci U S A** 91, 7757-7761.

Aurrekoetxea-Hernandez K and Buetti E (**2000**) Synergistic action of GA-binding protein and glucocorticoid receptor in transcription from the mouse mammary tumor virus promoter. **J Virol** 74, 4988-4998.

Ball JK, Arthur LO and Dekaban GA (**1985**) The involvement of a type-B retrovirus in the induction of thymic lymphomas. **Virology** 140, 159-172.

Ball JK, Diggelmann H, Dekaban GA, Grossi GF, Semmler R, Waight PA and Fletcher RF (**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, Rojas IC, Lee WH, King HL, Harriss JV, Kobayashi R, Webb CF and Gottlieb PD (**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, 18440-18452.

Beato M (**1996**) Chromatin structure and the regulation of gene expression: remodeling at the MMTV promoter. **J Mol Med** 74, 711-724.

Bentvelzen P, Daams JH, Hageman P, Calafat J and Timmermans A (**1972**) Interactions between viral and genetic factors in the origin of mammary tumors in mice. **J Natl Cancer Inst** 48, 1089-1094.

Beutner U, Kraus E, Kitamura D, Rajewsky K and Huber BT (**1994**) B cells are essential for murine mammary tumor virus transmission, but not for presentation of endogenous superantigens. **J Exp Med** 179, 1457-1466.

Blochlinger K, Bodmer R, Jack J, Jan LY and Jan YN (**1988**) Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in *Drosophila*. **Nature** 333, 629-635.

Bodmer R, Barbel S, Sheperd S, Jack JW, Jan LY and Jan YN (**1987**) Transformation of sensory organs by mutations of the *cut* locus of *D. melanogaster*. **Cell** 51, 293-307.

Boulikas T (**1995**) Chromatin domains and prediction of MAR sequences. **Int Rev Cytol** 162A, 279-388.

Bramblett D, Hsu CL, Lozano M, Earnest K, Fabritius C and Dudley J (**1995**) A redundant nuclear protein binding site contributes to negative regulation of the mouse mammary tumor virus long terminal repeat. **J Virol** 69, 7868-7876.

Bramblett D, Liu J, Lozano M, Xu L, Wrona T, Rajan L, Ross SR and Dudley J (**1997**) Mouse mammary tumor virus: a virus that exploits the immune system. **Leukemia** 11 Suppl 3, 183-186.

Brandt-Carlson C, Butel JS and Wheeler D (**1993**) Phylogenetic and structural analyses of MMTV LTR ORF sequences of exogenous and endogenous origins. **Virology** 193, 171-185.

Brown AM, Wildin RS, Prendergast TJ and Varmus HE (**1986**) A retrovirus vector expressing the putative mammary oncogene *int-1* causes partial transformation of a mammary epithelial cell line. **Cell** 46, 1001-1009.

Brown PO (**1997**) Integration. In **Retroviruses**. J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press), 161-204.

Bruggemeier U, Kalff M, Franke S, Scheidereit C and Beato M (**1991**) Ubiquitous transcription factor OTF-1 mediates induction of the MMTV promoter through synergistic interaction with hormone receptors. **Cell** 64, 565-572.

Buetti E and Kuhnel B (**1986**) Distinct sequence elements involved in the glucocorticoid regulation of the mouse mammary tumor virus promoter identified by linker scanning mutagenesis. **J Mol Biol** 190, 379-389.

Cavin C and Buetti E (**1995**) Tissue-specific and ubiquitous factors binding next to the glucocorticoid receptor modulate transcription from the mouse mammary tumor virus promoter. **J Virol** 69, 3759-3770.

Chandler VL, Maler BA and Yamamoto KR (**1983**) DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. **Cell** 33, 489-499.

Chattopadhyay S, Whitehurst CE and Chen J (**1998**) A nuclear matrix attachment region upstream of the T cell receptor beta gene enhancer binds Cux/CDP and SATB1 and modulates enhancer-dependent reporter gene expression but not endogenous gene expression. **J Biol Chem** 273, 29838-29846.

Choi Y, Kappler JW and Marrack P (**1991**) A superantigen encoded in the open reading frame of the 3' long terminal repeat of mouse mammary tumour virus. **Nature** 350, 203-207.

Clausse N, Baines D, Moore R, Brookes S, Dickson C and Peters G (**1993**) Activation of both *Wnt-1* and *Fgf-3* by insertion of mouse mammary tumor virus downstream in the reverse orientation: a reappraisal of the enhancer insertion model. **Virology** 194, 157-165.

de Belle I, Cai S and Kohwi-Shigematsu T (**1998**) The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. **J Cell Biol** 141, 335-348.

Dickinson LA, Dickinson CD and Kohwi-Shigematsu T (**1997**) An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. **J Biol Chem** 272, 11463-11470.

Dickinson LA, Joh T, Kohwi Y and Kohwi-Shigematsu T (**1992**) A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. **Cell** 70, 631-645.

Dickson C and Peters G (**1983**) Proteins encoded by mouse mammary tumour virus. **Curr Top Micro Immunol** 106, 1-34.

Dudley J and Risser R (**1984**) Amplification and novel locations of endogenous mouse mammary tumor virus genomes in mouse T-cell lymphomas. **J Virol** 49, 92-101.

Dudley JP (**1999**) Mouse mammary tumor virus. In **Encyclopedia of Virology**. R.G. Webster and A. Granoff, eds. (London, UK: Academic Press Ltd.), 965-972.

Dufort D and Nepveu A (**1994**) The human Cut homeodomain protein represses transcription from the c-*myc* promoter. **Mol Cell Biol** 14, 4251-4257.

Elliott JF, Pohajdak B, Talbot DJ, Shaw J and Paetkau V (**1988**) Phorbol diester-inducible, cyclosporine-suppressible transcription from a novel promoter within the mouse mammary tumor virus *env* gene. **J Virol** 62, 1373-1380.

Ellis T, Gambardella L, Horcher M, Tschanz S, Capol J, Bertram P, Jochum W, Barrandon Y and Busslinger M (**2001**) The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. **Genes Dev** 15, 2307-2319.

Fletcher TM, Ryu BW, Baumann CT, Warren BS, Fragoso G, John S and Hager GL (**2000**) Structure and dynamic properties of a glucocorticoid receptor-induced chromatin transition. **Mol Cell Biol** 20, 6466-6475.

Golovkina TV, Chervonsky A, Dudley JP and Ross SR (**1992**) Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. **Cell** 69, 637-645.

Golovkina TV, Dudley JP and Ross SR (**1995**) Mouse mammary tumor viruses with functional superantigen genes are selected during in vivo infection. **Proc Natl Acad Sci U S A** 92, 4828-4832.

Golovkina TV, Dudley JP, Jaffe AB and Ross SR (**1998**) B and T cells are required for mouse mammary tumor virus spread within the mammary gland. **J Immunol** 161, 2375-2382.

Golovkina TV, Dzuris J, van den Hoogen B, Jaffe AB, Wright PC, Cofer SM and Ross SR (**1998**) A novel membrane protein is a mouse mammary tumor virus receptor. **J Virol** 72, 3066-3071.

Golovkina TV, Shlomchik M, Hannum L and Chervonsky A (**1999**) Organogenic role of B lymphocytes in mucosal immunity. **Science** 286, 1965-1968.

Gouilleux F, Wakao H, Mundt M and Groner B (**1994**) Prolactin induces phosphorylation of Tyr694 of Stat5 (MGF), a prerequisite for DNA binding and induction of transcription. **EMBO J** 13, 4361-4369.

Groner B, Herrlich P, Kennedy N, Ponta H, Rahmsdorf U and Hynes NE (**1982**) Delimitation of a DNA sequence which confers inducibility by glucocorticoid hormones. **J Cell Biochem** 20, 349-357.

Hager GL, Archer TK, Fragoso G, Bresnick EH, Tsukagoshi Y, John S and Smith CL (**1993**) Influence of chromatin structure on the binding of transcription factors to DNA. **Cold Spring Harbor Symp Quant Biol** 58, 63-71.

Haraguchi S, Good RA, Engelman RW, Greene S and Day NK (**1997**) Prolactin, epidermal growth factor or transforming growth factor-_ activate a mammary cell-specific enhancer in mouse mammary tumor virus-long terminal repeat. **Mol Cell Endocrinol** 129, 145-155.

Hawkins S, Kohwi-Shigematsu T and Skalnik DG (**2001**) The matrix attachment region-binding protein SATB1 interacts with multiple elements within the gp91^{phox} promoter and is down regulated during myeloid differentiation. **J Biol Chem** (276, 44472-44480.

Held W, Waanders GA, Shakhov AN, Scarpellino L, Acha-Orbea H and MacDonald HR (**1993**) Superantigen-induced immune stimulation amplifies mouse mammary tumor virus infection and allows virus transmission. **Cell** 74, 529-540.

Henrard D and Ross SR (**1988**) Endogenous mouse mammary tumor virus is expressed in several organs in addition to the lactating mammary gland. **J Virol** 62, 3046-3049.

Hilkens J, van der Zeijst B, Buijs F, Kroezen V, Bleumink N and Hilgers J (**1983**) Identification of a cellular receptor for mouse mammary tumor virus and mapping of its gene to chromosome 16. **J Virol** 45, 140-147.

Hsu CL, Fabritius C and Dudley J (**1988**) Mouse mammary tumor virus proviruses in T-cell lymphomas lack a negative regulatory element in the long terminal repeat. **J Virol** 62, 4644-4652.

Huang M, Lee JW and Peterson DO (**1993**) Functional redundancy of octamer elements in the mouse mammary tumor virus promoter. **Nucleic Acids Res** 21, 5235-5241.

Hynes N, van Ooyen AJ, Kennedy N, Herrlich P, Ponta H and Groner B (**1983**) Subfragments of the large terminal repeat cause glucocorticoid-responsive expression of mouse mammary tumor virus and of an adjacent gene. **Proc Natl Acad Sci U S A** 80, 3637-3641.

Janeway CA, Jr (**1991**) Selective elements for the V β region of the T cell receptor: *Mls* and the bacterial toxic mitogens. **Adv Immunol** 50, 1-53.

Jarvis CD, Germain RN, Hager GL, Damschroder M and Matis LA (**1994**) Tissue-specific expression of messenger RNAs encoding endogenous viral superantigens. **J Immunol** 152, 1032-1038.

Jonkers J and Berns A (**1996**) Retroviral insertional mutagenesis as a strategy to identify cancer genes. **Biochim Biophys Acta** 1287, 29-57.

Kang CJ and Peterson DO (**1999**) Identification of a protein that recognizes a distal negative regulatory element within the mouse mammary tumor virus long terminal repeat. **Virology** 264, 211-219.

Karapetian O, Shakhov AN, Kraehenbuhl JP and Acha-Orbea H (**1994**) Retroviral infection of neonatal Peyer's patch lymphocytes: the mouse mammary tumor virus model. **J Exp Med** 180, 1511-1516.

Kim MH and Peterson DO (**1995**) Oct-1 protein promotes functional transcription complex assembly on the mouse mammary tumor virus promoter. **J Biol Chem** 270, 27823-27828.

Kohwi-Shigematsu T, Maass K and Bode J (**1997**) A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. **Biochemistry** 36, 12005-12010.

Korman AJ, Bourgarel P, Meo T and Rieckhof GE (**1992**) The mouse mammary tumour virus long terminal repeat encodes a type II transmembrane glycoprotein. **EMBO J** 11, 1901-1905.

Kuhnel B, Buetti E and Diggelmann H (**1986**) Functional analysis of the glucocorticoid regulatory elements present in the mouse mammary tumor virus long terminal repeat. A synthetic distal binding site can replace the proximal binding domain. **J Mol Biol** 190, 367-378.

Kusk P, John S, Fragoso G, Michelotti J and Hager GL (**1996**) Characterization of an NF-1/CTF family member as a functional activator of the mouse mammary tumor virus long terminal repeat 5' enhancer. **J Biol Chem** 271, 31269-31276.

Kwan H, Pecenka V, Tsukamoto A, Parslow TG, Guzman R, Lin TP, Muller WJ, Lee FS, Leder P and Varmus HE (**1992**) Transgenes expressing the *Wnt-1* and *int-2* proto-oncogenes cooperate during mammary carcinogenesis in doubly transgenic mice. **Mol Cell Biol** 12, 147-154.

Langer SJ and Ostrowski MC (**1988**) Negative regulation of transcription in vitro by a glucocorticoid response element is mediated by a trans-acting factor. **Mol Cell Biol** 8, 3872-3881.

Lee FS, Lane TF, Kuo A, Shackleford GM and Leder P (**1995**) Insertional mutagenesis identifies a member of the *Wnt* gene family as a candidate oncogene in the mammary epithelium of *int-2/Fgf-3* transgenic mice. **Proc Natl Acad Sci U S A** 92, 2268-2272.

Lee HL and Archer TK (**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, 32-41.

Lee JW, Moffitt PG, Morley KL and Peterson DO (**1991**) Multipartite structure of a negative regulatory element associated with a steroid hormone-inducible promoter. **J Biol Chem** 266, 24101-24108.

Lee WT, Prakash O, Klein D and Sarkar NH (**1987**) Structural alterations in the long terminal repeat of an acquired mouse mammary tumor virus provirus in a T-cell leukemia of DBA/2 mice. **Virology** 159, 39-48.

Lefebvre P, Berard DS, Cordingley MG and Hager GL (**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 Y, Hively WP and Varmus HE (**2000**) Use of MMTV-*Wnt-1* transgenic mice for studying the genetic basis of breast cancer. **Oncogene** 19, 1002-1009.

Liu J, Bramblett D, Zhu Q, Lozano M, Kobayashi R, Ross SR and Dudley JP (**1997**) The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. **Mol Cell Biol** 17, 5275-5287.

Ludlow C, Choy R and Blochlinger K (**1996**) Functional analysis of *Drosophila* and mammalian cut proteins in flies. **Dev Biol** 178, 149-159.

MacArthur CA, Shankar DB and Shackleford GM (**1995**) *Fgf-8*, activated by proviral insertion, cooperates with the *Wnt-1* transgene in murine mammary tumorigenesis. **J Virol** 69, 2501-2507.

Majors J and Varmus HE (**1983**) A small region of the mouse mammary tumor virus long terminal repeat confers glucocorticoid hormone regulation on a linked heterologous gene. **Proc Natl Acad Sci U S A** 80, 5866-5870.

Marchetti A, Buttitta F, Miyazaki S, Gallahan D, Smith GH and Callahan R (**1995**) *Int-6*, a highly conserved, widely expressed gene, is mutated by mouse mammary tumor virus in mammary preneoplasia. **J Virol** 69, 1932-1938.

Mellentin-Michelotti J, John S, Pennie WD, Williams T and Hager GL (**1994**) The 5' enhancer of the mouse mammary tumor virus long terminal repeat contains a functional AP-2 element. **J Biol Chem** 269, 31983-31990.

Mertz JA, Mustafa F, Meyers S and Dudley JP (**2001**) Type B leukemogenic virus has a T-cell-specific enhancer that binds AML-1. **J Virol** ___ 2174-2184.

Michalides R and Wagenaar E (**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, Wagenaar E and Weijers P (**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, 823-830.

Michalides R, Wagenaar E, Hilkens J, Hilgers J, Groner B and Hynes NE (**1982**) Acquisition of proviral DNA of mouse mammary tumor virus in thymic leukemia cells from GR mice. **J Virol** 43, 819-829.

Miller CL, Garner R and Paetkau V (**1992**) An activation-dependent, T-lymphocyte-specific transcriptional activator in the mouse mammary tumor virus *env* gene. **Mol Cell Biol** 12, 3262-3272.

Mink S, Hartig E, Jennewein P, Doppler W and Cato AC (**1992**) A mammary cell-specific enhancer in mouse mammary tumor virus DNA is composed of multiple regulatory elements including binding sites for CTF/NFI and a novel transcription factor, mammary cell-activating factor. **Mol Cell Biol** 12, 4906-4918.

Mink S, Ponta H and Cato AC (**1990**) The long terminal repeat region of the mouse mammary tumour virus contains multiple regulatory elements. **Nucleic Acids Res** 18, 2017-2024.

Mok E, Golovkina TV and Ross SR (**1992**) A mouse mammary tumor virus mammary gland enhancer confers tissue-specific but not lactation-dependent expression in transgenic mice. **J Virol** 66, 7529-7532.

Moon NS, Berube G and Nepveu A (**2000**) CCAAT displacement activity involves CUT repeats 1 and 2, not the CUT homeodomain. **J Biol Chem** 275, 31325-31334.

Morley KL, Toohey MG and Peterson DO (**1987**) Transcriptional repression of a hormone-responsive promoter. **Nucleic Acids Res** 15, 6973-6989.

Morris VL, Rao TR, Kozak CA, Gray DA, Chan ECML, Cornell TJ, Taylor CB Jones RF and McGrath CM (**1991**) Characterization of *Int-5*, a locus associated with early events in mammary carcinogenesis. **Oncogene Res** 6, 53-63.

Mustafa F, Lozano M and Dudley JP (**2000**) C3H mouse mammary tumor virus superantigen function requires a splice donor site in the envelope gene. **J Virol** 74, 9431-9440.

Nakagomi K, Kohwi Y, Dickinson LA and Kohwi-Shigematsu T (**1994**) A novel DNA-binding motif in the nuclear matrix

attachment DNA-binding protein SATB1. **Mol Cell Biol** 14, 1852-1860.

Nandi S and McGrath CM (**1973**) Mammary neoplasia in mice. **Adv Cancer Res** 17, 353-414.

Nepveu A (**2001**) Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. **Gene** 270, 1-15.

Neufeld EJ, Skalnik DG, Lievens PM and Orkin SH (**1992**) Human CCAAT displacement protein is homologous to the *Drosophila* homeoprotein, cut. **Nat Genet** 1, 50-55.

Parks WP, Scolnick EM and Kozikowski EH (**1974**) Dexamethasone stimulation of murine mammary tumor virus expression: a tissue culture source of virus. **Science** 184, 158-160.

Pattison S, Skalnik DG and Roman A (**1997**) CCAAT displacement protein, a regulator of differentiation-specific gene expression, binds a negative regulatory element within the 5' end of the human papillomavirus type 6 long control region. **J Virol** 71, 2013-2022.

Payvar F, Wrange O, Carlstedt-Duke J, Okret S, Gustafsson JA and Yamamoto KR (**1981**) Purified glucocorticoid receptors bind selectively in vitro to a cloned DNA fragment whose transcription is regulated by glucocorticoids in vivo. **Proc Natl Acad Sci U S A** 78, 6628-6632.

Pryciak PM, Sil A and Varmus HE (**1992**) Retroviral integration into minichromosomes in vitro. **EMBO J** 11, 291-303.

Qin W, Golovkina TV, Peng T, Nepomnaschy I, Buggiano V, Piazzon I and Ross SR (**1999**) Mammary gland expression of mouse mammary tumor virus is regulated by a novel element in the long terminal repeat. **J Virol** 73, 368-376.

Quaggin SE, Heuvel GV, Golden K, Bodmer R and Igarashi P (**1996**) Primary structure, neural-specific expression, and chromosomal localization of *Cux-2*, a second murine homeobox gene related to *Drosophila cut*. **J Biol Chem** 271, 22624-22634.

Rabson AB and Graves BJ (**1997**) Synthesis and processing of viral RNA. In **Retroviruses**. J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), 205-261.

Reuss FU and Coffin JM (**1995**) Stimulation of mouse mammary tumor virus superantigen expression by an intragenic enhancer. **Proc Natl Acad Sci U S A** 92, 9293-9297.

Rijsewijk F, van Deemter L, Wagenaar E, Sonnenberg A and Nusse R (**1987**) Transfection of the *int*-1 mammary oncogene in cuboidal RAC mammary cell line results in morphological transformation and tumorigenicity. **EMBO J** 6, 127-131.

Ringold GM, Yamamoto KR, Bishop JM and Varmus HE (**1977**) Glucocorticoid-stimulated accumulation of mouse mammary tumor virus RNA: increased rate of synthesis of viral RNA. **Proc Natl Acad Sci U S A** 74, 2879-2883.

Ross SR, Hsu CL, Choi Y, Mok E and Dudley JP (**1990**) Negative regulation in correct tissue-specific expression of mouse mammary tumor virus in transgenic mice. **Mol Cell Biol** 10, 5822-5829.

Shackleford GM and Varmus HE (**1988**) Construction of a clonable, infectious, and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. **Proc Natl Acad Sci U S A** 85, 9655-9659.

Shackleford GM, MacArthur CA, Kwan HC and Varmus HE (**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*. **Proc Natl Acad Sci U S A** 90, 740-744.

Skalnik DG, Strauss EC and Orkin SH (**1991**) CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-*phox* gene promoter. **J Biol Chem** 266, 16736-16744.

Stewart TA, Pattengale PK and Leder P (**1984**) Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/*myc* fusion genes. **Cell** 38, 627-637.

Swanstrom R and Wills JW (1997) Synthesis, assembly, and processing of viral proteins. In **Retroviruses**. J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press), 263-334.

Telesnitsky A and Goff SP (1997) Reverse transcriptase and the generation of retroviral DNA. In **Retroviruses**. J.M. Coffin, S.H. Hughes, and H.E. Varmus, eds. (Cold Spring Harbor,N.Y.: Cold Spring Harbor Laboratory Press), 121-160.

Tufarelli C, Fujiwara Y, Zappulla DC and Neufeld EJ (**1998**) Hair defects and pup loss in mice with targeted deletion of the first cut repeat domain of the Cux/CDP homeoprotein gene. **Dev Biol (Orlando)** 200, 69-81.

Valarche I, Tissier-Seta JP, Hirsch MR, Martinez S, Goridis C and Brunet JF (**1993**) The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. **Development** 119, 881-896.

van Leeuwen F and Nusse R (**1995**) Oncogene activation and oncogene cooperation in MMTV-induced mouse mammary cancer. **Semin Cancer Biol** 6, 127-133.

Wakao H, Gouilleux F and Groner B (**1994**) Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. [erratum appears in EMBO J 1995 Feb 15;14(4):854-5]. **EMBO J** 13, 2182-2191.

Wang Z, Goldstein A, Zong RT, Lin D, Neufeld EJ, Scheuermann RH and Tucker PW (**1999**) Cux/CDP homeoprotein is a component of NF-muNR and represses the immunoglobulin heavy chain intronic enhancer by antagonizing the bright transcription activator. **Mol Cell Biol** 19, 284-295.

Weidhaas JB, Angelichio EL, Fenner S and Coffin JM (**2000**) Relationship between retroviral DNA integration and gene expression. **J Virol** 74, 8382-8389.

Wheeler DA, Butel JS, Medina D, Cardiff RD and Hager GL (**1983**) Transcription of mouse mammary tumor virus: identification of a candidate mRNA for the long terminal repeat gene product. **J Virol** 46, 42-49.

Withers-Ward ES, Kitamura Y, Barnes JP and Coffin JM (**1994**) Distribution of targets for avian retrovirus DNA integration in vivo. **Genes Dev** 8, 1473-1487.

Wrona TJ, Lozano M, Binhazim AA and Dudley JP (**1998**) Mutational and functional analysis of the C-terminal region of the C3H mouse mammary tumor virus superantigen. **J Virol** 72, 4746-4755.

Yanagawa S-I, Kakimi K, Tanaka H, Murakami A, Nakagawa Y, Kubo Y, Yamada Y, Hiai H, Kuribayashi K, Masuda T and Ishimoto A (**1993**) Mouse mammary tumor virus with

rearranged long terminal repeats causes murine lymphomas. **J Virol** 67, 112-118.

Zaret KS and Yamamoto KR (**1984**) Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. **Cell** 38, 29-38.

Zhu Q, Gregg K, Lozano M, Liu J and Dudley JP (**2000**) CDP is a repressor of mouse mammary tumor virus expression in the mammary gland. **J Virol** 74, 6348-6357.

Jaquelin P. Dudley Ph.D.
Section of Molecular Genetics and Microbiology, The University of Texas at Austin, 100 W. 24th Street, Austin, TX 78705; Telephone: (512) 471-8415; Fax: (512) 471-7088; E-mail: jdudley@uts.cc.utexas.edu