Mammary Gland Expression of Mouse Mammary Tumor Virus Is Regulated by a Novel Element in the Long Terminal Repeat

WEI QIN, TATYANA V. GOLOVKINA, TAO PENG, IRENE NEPOMNASCHY, VALERIA BUGGIANO, ISABEL PIAZZON, and SUSAN R. ROSS

Department of Microbiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6142, and Division of Experimental Medicine, Institute of Hematological Research, Academy of Medicine, Buenos Aires, Argentina

Received 27 March 1998/Accepted 9 October 1998

Mouse mammary tumor virus (MMTV) infects both lymphoid tissue and lactating mammary gland during its infectious cycle, but some endogenous MMTVs are transcribed only in lymphoid cells. We found a lymphoid cell-specific endogenous MMTV that was converted to a milk-borne, infectious virus through recombination with an exogenously transmitted MMTV. The changed expression pattern correlated with the alteration of a single base pair in the long terminal repeat of the lymphoid cell-specific virus. Transgenic mice with the element from either the milk-borne or lymphoid cell-specific virus upstream of the chloramphenicol acetyltransferase reporter gene showed the same pattern of expression as the virus from which the regulatory sequences were derived. Electrophoretic mobility shift assays with mammary cell extracts showed that the site from the milk-borne virus was preferentially bound by a prolactin-inducible factor that poorly bound the altered site from the lymphoid cell-specific virus. The complex that formed on the milk-borne virus-specific oligonucleotide supershifted with anti-Stat5b antibody. Mice lacking either Stat5a or Stat5b had dramatically reduced levels of MMTV transcripts in mammary gland but not in lymphoid tissue. Thus, a member of the STAT family of transcription factors is involved in the tissue-specific expression of mouse mammary tumor virus in vivo. This is the first example of the involvement of a member of the STAT family of transcription factors in the control of tissue-specific expression.

Mouse mammary tumor virus (MMTV) is a retrovirus that is either inherited as an endogenous virus or acquired as an exogenous virus through milk-borne infection. MMTV has been used as a model for the study of the regulation of gene transcription since the discovery that its expression was induced by glucocorticoid hormones in vivo and in tissue culture cells (reviewed in reference 61). Indeed, the first evidence that mammalian transcription factors interacted with specific DNA sequences (termed glucocorticoid response elements [GREs]) came from studies of how glucocorticoid receptors (GR) induced MMTV expression (61). The ability of glucocorticoids and progesterone to stimulate viral transcription is critical for MMTV transmission to subsequent generations, since as a result of this stimulation, virus production dramatically increases during pregnancy and lactation (5).

A number of additional transcription factors, including NF-1, Oct-1, and TFIID, are involved in the regulation of MMTV gene expression (10, 40, 54). Moreover, as expected for a virus transmitted through milk, there are sequences within the long terminal repeat (LTR) of the virus that confer mammary gland-specific expression, termed the mammary gland enhancer (9, 27, 40, 41, 47) (Fig. 1A). Transgenic mouse studies in which this enhancer was linked to the heterologous simian virus 40 promoter indicated that it directed expression to lactating and virgin mammary gland that was no longer lactation responsive (41). Inclusion of the GREs in the transgene restored lactation-induced expression. Several transcription factors, including AP-2 (56) and NF-1 or related factors (27, 40), have been shown to bind to this region.

In addition to mammary gland cells, lymphoid cells transcribe MMTV (9, 17, 22, 26) and shed virus particles (12). MMTV expression in these cells is critical to the virus life cycle, since infected B cells in newborn pups present an MMTV-encoded superantigen (Sag) to cognate T cells (3); this activation of T lymphocytes by Sag is requisite step in the virus’s ultimate transmission to the mammary gland (14). Sag proteins interact with all T cells that express particular β chains of the T-cell receptor (38). This interaction is determined by the sequence of the C-terminal amino acids of this type II membrane glycoprotein and different MMTVs associate with different Vβ-bearing T cells, depending on the sequence of this hypervariable region (1, 4, 8, 44). Little is known about what positively determines expression of the virus in lymphoid cells and where the regulatory regions controlling this expression lie, although it has recently been shown that a nuclear matrix bound transcription factor called SATB1 negatively regulates MMTV transcription in lymphocytes (31).

Recently, several milk protein genes, such as those encoding the whey acidic protein (WAP), β-casein, and α-lactoglobulin proteins, were shown to have a common transcription-regulatory element that bound a factor in lactating mammary gland (7, 29, 50). This factor was purified and shown to be a member of the STAT family of transcription factors (58), and it was subsequently called STAT5. STAT5 is activated in response to prolactin through phosphorylation by the Janus (JAK) family of tyrosine kinases (24, 49, 59). There are two STAT5 genes in mice, called Stat5a and Stat5b, that presumably arose through gene duplication. It is unclear whether their function is duplicated as well (32). Although STAT5 was thought to be important for the tissue-specific expression of milk protein genes, mutation of the STAT consensus binding sites in the WAP
or β-lactoglobulin gene promoter had only subtle effects on linked marker gene expression in lactating transgenic mice (7, 30). However, in mice lacking Stat5a, WAP expression was down-regulated, but little or no effect was seen on β-casein, α-lactoglobulin, or WDNM1 (a milk protein) RNA levels (33). Similarly, it has been reported that in Stat5b knockout mice, there is transcription of milk protein genes (56). It is now clear that Stat5a and Stat5b are expressed in wide variety of tissues and are activated by a number of different cytokine or growth factor receptors (24, 49, 59).

We recently found that while most endogenous MMTVs are expressed in both mammary gland and lymphoid tissue (ML expression), others are expressed only in lymphoid cells (L expression) (17, 46). Because this was true of proviruses at various chromosomal locations, it was unlikely that this tissue-specific expression was the result of position effects and was more likely due to sequences within the viruses. We also discovered that a lymphoid cell-specific endogenous MMTV, Mtv-7, could be converted to a mammary gland-expressing exogenous virus through retroviral recombination with a mammary gland-expressing exogenous virus (16).

We show here that there is a single base pair difference in a transcription factor binding site in the LTR between the lymphoid-specific MMTVs and those expressed both in mammary gland and lymphocytes. All MMTVs could be grouped accord-

Fig. 1. (A) Map of the MMTV LTR. Depicted are the binding sites for Oct-1 and NF-1 transcription factors, as well as the GRES. MGE refers to the region mapped as a mammary gland enhancer at the 5′ end of the LTR. The region with homology to STAT binding sites (bp 519 and 528) is also shown. HR denotes the coding region for the hypervariable domain of the Sag. (B) LTRs of the constructs used to create transgenic mice. The filled box represents the MMTV (C3H) LTR; the open box represents the Mtv-7 LTR. The Mtv7/C3H has the regions from bp 1 to 631 from the MMTV (C3H) LTR, and the C3H/Mtv7 LTR has the regions from bp 1 to 631 from the MMTV (C3H) LTR and from 632 to 1280 from the Mtv7 LTR. The wide stripes represent the STAT region from MMTV (C3H); the narrow stripes represent the STAT region from Mtv7. (C) Pictorial representation of the Mtv7/BALB14 naturally occurring recombinant LTRs. Two classes of recombinants are shown, those that had a recombination event between the HR and STAT binding sites of BALB14 and those with a recombination event between the HR and STAT binding sites of Stat5a and Stat5b knockouts mice and controls were kind gifts from L. Henninghausen and J. Ihle, respectively.

Transgene construction and CAT assays. The Mtv7-LTR was subcloned from plasmid pMO-BC (4), a kind gift from Brigitte Huber, and inserted into the pCATnuc plasmid (Promega Biotech, Inc., Madison, Wis.) to make the Mtv7 transgene. The MMTV (C3H) LTR was derived from the pLTR plasmid (36). To create the Mtv7/C3H and C3H/Mtv7 transgenes, the LTR-containing plasmids were restricted with Eco0109I, which releases a fragment containing the region from bp 632 to 1280 in both the Mtv7 and C3H LTRs (numbering is according to reference 6); the equivalent Eco0109I fragment from the different LTR was inserted in its place (Fig. 1).

CAT assays were carried out as previously described (47). CAT specific activities are given as counts per minute per minute of reaction time per milligram of protein.

RNA analysis. RNA was isolated by using QuickPrep kits (Pharmacia Biotech, Inc., Uppsala, Sweden) and then subjected to Northern blot analysis and hybridization with an MMTV envelope gene probe. The blot was subsequently stripped and reprobed to a mouse β-actin probe, to control for RNA integrity. Quantitation was performed with a PhosphorImager (Molecular Dynamics, Inc., Sunnyvale, Calif.). The LTR-CAT mice, containing the MMTV (C3H) LTR, were previously described (47). Tissues from the lactating mammary glands and spleens of Stat5a and Stat5b knockout mice and controls were kind gifts from L. Henninghausen and J. Ihle, respectively.

Nuclear extract preparation. The cells were harvested and pelleted by centrifugation for 5 min in a tabletop centrifuge at 1,000 rpm at room temperature. Nuclear extracts were prepared as previously described (11, 58). All steps were carried out at 4°C. Phosphatase inhibitors (1 mM sodium fluoride and 100 μM sodium orthovanadate) and protease inhibitors (leupeptin, pepstatin A, and aprotonin [5 μg of each per ml] and 0.5 mM phenylmethylsulfonyl fluoride) were included in the buffers at all steps. Protein concentration was measured by the colorimetric method (Bio-Rad Inc., San Rafael, Calif.).

Preparation of oligonucleotide probes. [γ-32P]ATP end-labeled oligonucleotide probes were used for gel mobility shift assays, supershift assays, and competition experiments. The double-stranded oligonucleotides were end labeled with 150 μCi of [γ-32P]ATP (6,000 Ci/mmol; NEN) in the presence of 10 U of T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.) at 37°C for 1 h. The labeled oligonucleotides were purified by passage through ProbeQuant G-50 microcolumns (Pharmacia Biotech). The sequences of the oligonucleotides used for [γ-32P]ATP labeling (upper strands only) are as follows: β-casein probe, 5′-GGAGTCATTGGAATTTAAAGGAG-3′ (59); Fcγ-R1 probe, 5′-GTAATTTCCCCAGAGAAGAAGAC-3′ (23); C3H probe, 5′-CTCAACCTCAGTTGAAGAACAAGTT-3′; Mtv7 probe, 5′-CTCAACCTCAGTTGAAGAACAAGTT-3′; NF-Eb probe, 5′-GGATCCTTGGTGACCAGTTA-3′ (60).
TABLE 1. Sequences of different MMTV LTRs between bp 519 and 528

<table>
<thead>
<tr>
<th>Virus</th>
<th>Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMTV(CH)</td>
<td>C T C A A T T G A A</td>
<td>ML</td>
</tr>
<tr>
<td>Mtv-1, -3, -6</td>
<td>C T C A A T T G A A</td>
<td>ML</td>
</tr>
<tr>
<td>Mtv-43</td>
<td>C T C A A T T G A A</td>
<td>ML</td>
</tr>
<tr>
<td>Mtv-7, -9</td>
<td>C T C A A T T G A A</td>
<td>L</td>
</tr>
<tr>
<td>BALB14</td>
<td>C T C A A T T G A A</td>
<td>L</td>
</tr>
<tr>
<td>REC 1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>C T C A A T T G A A</td>
<td>ML</td>
</tr>
<tr>
<td>REC 2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>C T C A A T T G A A</td>
<td>ML</td>
</tr>
<tr>
<td>STAT consensus&lt;sup&gt;d&lt;/sup&gt;</td>
<td>T T C N (N) G A A</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

<sup>a</sup> Sequences are numbered according to the system of Brandt-Carlson et al. (6). Double underlining denotes change from the MMTV(CH) sequence.
<sup>b</sup> Recombinant between Mtv-7 and BALB14 with breakpoint within the Stat site (Fig. 1C).
<sup>c</sup> Recombinant between Mtv-7 and BALB14 with breakpoint 3′ of the Stat site (Fig. 1C).
<sup>d</sup> Taken from reference 49

Gel mobility shift, supershift, and competition experiments. For binding studies, the reaction mixtures contained 10 μg of protein extract, 10 μg of poly(dI-dC) (Sigma), and 0.5 ng of 32P-end-labeled oligonucleotide probe (2 × 10<sup>6</sup> cpm) in D<sub>2</sub>O, built (20 mM HEPES [pH 7.9], 0.25 M NaCl, 0.2 M Na<sub>2</sub> EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride) in a total reaction volume of 20 μl. Reaction mixtures were incubated on ice for 30 min and loaded onto 5% polyacrylamide gels containing 22.5 mM Tris-HCl, 22.5 mM borate, and 0.5 mM EDTA. Electrophoresis was for 2 h at 20 mA at room temperature. After electrophoresis, the gel was transferred to 3MM paper, dried, and exposed to Kodak XAR-5 film overnight at 70°C. For the supershift assays, 10 μg of anti-STAT antibodies (Santa Cruz Biotechnology, Inc., Santa Cruz, Calif.) were used. The following polyclonal rabbit anti-mouse antibodies were tested: polyclonal anti-Stat5a (anti-carcinob tyrosine [SC1081]), anti-Stat5b (anti-carcinob tyrosine of the p80 protein [SC835]), anti-Stat4 (anti-aminob tyrosine [SC848]), and anti-Stat6 (anti-carcinob tyrosine [SC981]). Antibodies were incubated with the nuclear extract for 30 min at room temperature prior to addition of the oligonucleotide probe. Reaction mixtures were loaded on a 5% polyacrylamide gel and run at room temperature as described above. For the competition assays, nuclear extracts were incubated with the specified amounts of unlabeled oligonucleotide competitors and the labeled oligonucleotide probe for 30 min on ice. To determine relative affinities, 100- and 800-fold molar excesses of the unla beled oligonucleotide competitors were used to inhibit the binding of a constant amount of nuclear extract to the probes.

RESULTS

Comparison of sequences of different MMTVs. Endogenous MMTVs can be divided into two groups, ML and L, based on their expression in mammary gland (17, 46). For example, Mtv-7 and -9 are expressed only in lymphoid tissue, while the exogenous viruses MMTV(CH) and MMTV(SW) are expressed by several endogenous MMTVs (Mtv-1, -3, -6, and -43) are expressed in both lymphoid and mammary gland cells (Table 1). The LTRs of the different endogenous MMTVs are remarkably homologous to each other outside the superantigen hypervariable coding region (6). There is only one additional region of dissimilarity among the different proviruses, between bp 519 and 528. As shown in Table 1, there are several nucleotide changes in this region. For example, all of the viruses shown except MMTV(CH) have a G·C base pair at position 523, and some viruses [MMTV(SW), Mtv-7, and Mtv-43] have CA·GT rather than a TG·AC at positions 525 to 526. Significantly, all MMTVs that are expressed only in lymphoid tissue (i.e., Mtv-7 and -9) have a T·A→G·C change at position 520. Although MMTV(SW), Mtv-7, and Mtv-43 all interact with V<sub>β</sub>6-bearing T cells and have almost identical Sag hypervariable regions (6), only MMTV(SW) and Mtv-43 are expressed in both mammary gland and lymphocytes (46). This finding indicated that the presence of a T·A base pair at position 520 was important for mammary gland expression. The functional importance of this T·A residue for viral expression in mammary gland was also demonstrated by the analysis of novel MMTVs generated by recombination between Mtv-7 and an exogenous virus, BALB14 (16). While the Mtv-7 Sag interacts with V<sub>β</sub>6·T cells, the BALB14 Sag causes the stimulation or deletion of V<sub>β</sub>4·T cells. However, because Mtv-7 is not expressed in mammary gland, it is not transmitted as an exogenous virus in milk. The novel recombinant MMTVs all retained the Mtv-7 Sag hypervariable region and therefore interacted with V<sub>β</sub>6-bearing T cells. Moreover, these recombinant viruses were all expressed in mammary gland and transmitted through milk to offspring. Upon examination of the LTR sequences, we found that the new viruses had acquired (through retroviral recombination) sequences 5′ of the Mtv-7 Sag hypervariable region from the BALB14 virus (Fig. 1C). At least five different types of recombinants were identified, four of which had the entire region between 519 and 528 of BALB14 as well as sequences upstream (Table 1 and Fig. 1C, REC 2). Significantly, the fifth type of recombinant had a breakpoint within this region, such that the T·A base pair at position 520 (and sequences 5′ of this nucleotide) was acquired from BALB14, but the region 3′ of this nucleotide came from Mtv-7 (Table 1 and Fig. 1C, REC 1). This selection of the T·A base pair by viruses that gained the capacity to express in mammary gland provided strong biological evidence that this nucleotide change was important to this tissue-specific expression.

Tissue-specific expression is determined by the LTR. We and others have previously reported that the LTR from the MMTV(CH) milk-borne virus directs expression of linked transgenes to both mammary gland and lymphoid tissues (9, 22, 52). To further show that the sequences within the region from nucleotides 518 to 528 were required for mammary gland-specific expression, we constructed three types of transgenic mice. One set of mice had the Mtv-7 LTR (Mtv-7; four independent lines), the second had a hybrid LTR (Mtv7/CH) with sequences upstream of bp 632 from Mtv-7 and sequences downstream of this site from the MMTV(CH) LTR (Fig. 1B, Mtv7/CH) (one line), and the third had a hybrid LTR (CH/ Mtv7) with sequences upstream of bp 632 from MMTV(CH) and downstream from Mtv-7 (Fig. 1B, CH/Mtv7) (three lines). All three LTRs were inserted in the same position in a CAT expression vector. The mice were sacrificed, and CAT assays were performed with extracts prepared from various tissues. In addition, the DNA of these mice was examined by semiquantitative Southern blot analysis, to obtain an approximate estimate of the transgene copy number present in the offspring derived from different founder animals (not shown).

All low-copy-number (≤10 copies) Mtv-7 and Mtv7/CH transgenic strains had high levels of CAT activity only in lymphoid tissues, such as spleen, thymus, lymph nodes and Peyer’s patches; shown in Fig. 2A are the results for three different Mtv7/CH lines (10, 12, and 25) and one Mtv7/CH line. Low levels of CAT activity were detected in the mammary glands of all strains, and virtually no transgene expression was found in the liver. In contrast, mice bearing the CH/Mtv7 LTR expressed the transgene at highest levels in mammary gland, with lower levels in lymphoid tissues (Fig. 2B). The other two strains of CH/Mtv7 mice had similar patterns of expression (data not shown). This high level of expression in mammary gland and lower level in lymphoid tissues was also similar to what we previously reported for mice with the MMTV(CH) LTR driving the CAT transgene (Fig. 2B) (41, 47) and to mice containing other transgenes under the control of this LTR (9, 22, 52). Thus, sequences upstream of position 528 that were derived from Mtv-7 appeared to result in preferential high-level expression in lymphoid tissues or, conversely, prevented high-level expression...
expression in mammary gland. Importantly, there are no other sequence differences between the MMTV(C3H) and Mtv-7 LTRs within the 1 to 632 region that could account for this change in expression. Indeed, the Mtv-7 LTR is 100% homologous within this region to several MMTVs that are transcribed in mammary tissue (6).

In contrast to the mice described in the preceding paragraph, high-copy-number (50 to 500 copies) Mtv-7/transgenic mice showed very high levels transgene expression in many but not all tissues; CAT activity levels in liver were always low (for example, Fig. 2C). The regulation of tissue-specific expression apparently could be overcome when large numbers of the transgene were integrated into the mouse genome.

**STAT consensus sequence binding.** The results presented above provided strong biological evidence that the T3A3G3C change in the lymphotropic LTRs was responsible for the lack of expression of these viruses in mammary gland. Comparison of these sequences between gp 519 and 528 to those recognized by known transcription factors revealed that they were similar, although not identical, to those bound by STAT family members (Table 1). If the STAT-like sequence in the LTR was a recognition site for a transcription factor present in mammary gland, it was possible that the nucleotide change in the Mtv-7 LTR prevented binding and thus expression in this tissue.

To determine whether there was any factor in mammary gland cells that bound to this consensus sequence, nuclear extracts were prepared from NMuMG cells and used in electrophoretic mobility shift assays (EMSA) with oligonucleotides containing the STAT-like region from MMTV(C3H) (ML), Mtv-7 (L), and Mtv-43 (ML) LTRs (Table 1). The STAT consensus sequences from the rat β-casein gene, which binds STAT5 (59), and FcγRI gene, which binds several STAT factors(23), were also used with these extracts.

As can be seen in Fig. 3, all of the oligonucleotide probes were bound by factors in this extract, and the complexes formed migrated as doublets in all cases. However, when the C3H and Mtv-43 probes were used, the upper band of the doublet was the predominant complex formed, whereas with the Mtv-7 and FcγRI probes, the two complexes were of almost equal intensity. The bands shifted with the casein probe were of almost equal intensity and appeared to migrate differently than those complexes formed on the ML or L oligonucleotides. These results indicated that there was differential binding of a nuclear factor to the ML consensus sequence found in the

![FIG. 3. Analysis of STAT-like DNA-protein complexes in NMuMG cells by EMSA. NMuMG nuclear extracts (10 μg) were incubated with the indicated 32P-labeled probes. The free probe was run off the gel.](http://jvi.asm.org/.../index.html)
LTRs of those MMTVs that were expressed in mammary gland.

To further confirm that a unique complex was formed on the MMTV(C3H) STAT site, competition experiments were performed. When the MMTV(C3H) ML oligonucleotide was used as a probe, the β-casein oligonucleotide competed only the faster-migrating complex; the Mtv-7 L oligonucleotide also competed this same complex (Fig. 4). In contrast, the unlabelled ML oligonucleotide competed both slow- and fast-migrating complexes. An unrelated oligonucleotide, the recognition sequence for the NF-κB transcription factor, did not compete this binding. When the Mtv-43 ML oligonucleotide was used as a competitor, the more slowly migrating complex was also unaffected, similar to what was seen with MMTV (C3H) ML oligonucleotide (not shown).

These results indicated that there was a factor that uniquely bound to the ML consensus sequence and that the altered base pair in the L consensus sequence prevents its binding. The more slowly migrating complex that formed on the ML oligonucleotide was also apparently different from the one which recognized the rat β-casein STAT site. In contrast, because the faster-migrating complex that formed on the L and ML sequences showed reciprocal competition (Fig. 4 and not shown) and was competed by the β-casein oligonucleotide the factors that make up this complex may recognize all three sequences.

Stat5b is found in a prolactin-inducible complex on the ML sequence. Prolactin induces transcription of several milk protein genes in lactating mammary gland, presumably through the phosphorylation and activation of the DNA binding activity of STAT5 (58). Transcription of MMTV also increases during lactation, when prolactin levels are elevated; however, much, if not all, of this induction is attributable to glucocorticoid hormone induction (41). Although the STAT consensus sequences in the MMTV LTR differed somewhat from those found in milk protein and other genes that have previously been shown to be transcriptionally regulated by STAT5 and other members of this family (Fig. 1), it was possible that prolactin would induce the activity of the factors that bound to the ML- or L-STAT sequence. To test this, nuclear extracts were prepared from NMuMG cells that were treated with prolactin and used in EMSAs within the MMTV(C3H) ML-STAT or Mtv-7 L-STAT probe. Prolactin treatment resulted in induction of the more slowly migrating complex that formed on the ML-STAT (Fig. 5, lanes 2 and 4) but not the L-STAT (lane 8) probe. This was further evidence that the factor(s) recognizing the ML-STAT sequence differed from that bound to the L-STAT sequence.

To determine whether Stat5a, Stat5b, or any other family members were present in the prolactin-inducible complex, we used antibodies directed against four different proteins, Stat5a, Stat5b, STAT4, and STAT6, in EMSA supershift assays. Antisera against the C terminus of Stat5b, but not Stat5a, supershifted the complex formed on the ML-STAT but not the L-STAT probe (Fig. 5) only in prolactin-treated extracts (not shown), indicating that the upper band contains Stat5b. None of the other antisera supershifted or disrupted either of these complexes (not shown).

These results indicated that Stat5b or a related factor was involved in the mammary gland-specific transcription of MMTVs containing the ML-STAT sequence and that the altered nucleotide present in the L-STAT sequence prevented this binding. Although a small amount of the more slowly migrating complex did form on the L-STAT probe, there was probably an insufficient amount of Stat5b present to be detected in this assay. Moreover, because the supershift is incomplete even on the ML-STAT, it is possible that the antisera cross-reacted with a Stat5b-related protein that was present in the extract or that it did not efficiently bind Stat5b in the complex. It is also possible that Stat5b is only a small component of the complex formed on the ML-STAT oligonucleotide and thus the antisera does not efficiently supershift it.

MMTV transcription is prolactin inducible. Since the complex that bound specifically to the ML-STAT region of the MMTV(C3H) LTR was prolactin inducible, we examined whether viral transcription was affected by this hormone. A normal murine mammary gland cell line transfected with an infectious molecular clone of MMTV, called NMgCl1 (51), was used for this analysis. The cells were grown to confluence, cultured in 1% serum for 48 h, and then induced with prolactin or dexamethasone for 14 to 16 h. Northern blot analysis was performed with a probe specific for the MMTV envelope gene (37). As seen in Fig. 6, there was an approximately two- or fivefold induction of MMTV RNA in cells grown in the presence of prolactin or dexamethasone, respectively. Moreover, prolactin acted synergistically with dexamethasone to further increase transcript levels, since there was a ninefold induction...
in cells grown with both hormones. Thus, prolactin increased the production of steady-state levels of MMTV RNA, probably through activation of the JAK-STAT pathway.

**MMTV transcription in Stat5b and Stat5a knockout mice.** If Stat5b is required for high-level mammary gland expression of MMTV, then mice lacking this transcription factor should have little or no viral RNA in this tissue. RNA was prepared from the mammary glands and lymphoid tissue of mice with targeted mutation of the Stat5b gene, and RNase protection analysis was carried out to examine the level of endogenous MMTV expression. The level of Mtv-17 RNA was dramatically reduced in the Stat5b knockout mice in comparison to a wild-type mouse derived from the same cross (Fig. 7A). Mtv-17 is normally highly expressed in the mammary gland but not lymphoid tissue of inbred mice containing this endogenous provirus. A similar decrease in another endogenous MMTV expressed in mammary gland, Mtv-3, was also seen (not shown). In contrast, the Stat5b knockout mice had normal levels of lymphoid tissue-expressed proviruses Mtv-3 (Fig. 7B) and Mtv-9 (not shown).

We also examined the expression of endogenous MMTVs in mice lacking the Stat5a transcription factor. Interestingly, the level of Mtv-17 and -3 in the mammary gland was also dramatically reduced in these animals (Fig. 7A). Expression of lymphoid tissue-expressed Mtv-3 and -9 was not affected in the Stat5a-negative mice (Fig. 7B and data not shown).

**DISCUSSION**

MMTV is unique among the murine and perhaps other retroviruses in encoding a Sag whose activity it uses as part of its infection pathway. Because this protein causes profound deletion of cognate T cells, any mouse that is infected with an exogenous virus encoding a Sag with the same specificity as its endogenous loci cannot be infected with this virus. As a result, mice with such endogenous viruses do not acquire exogenous MMTV in the mammary gland, nor do they develop mammary tumors (21). Thus, the reduced expression of MMTV may be due to a lack of Stat5b in these mice.

Although MMTV transcripts were diminished in the mammary glands of the Stat5a and 5b knockout mice, they were not affected in lymphoid tissue. This finding argues that although these transcription factors are found and function in lymphoid tissue, they are not required for the transcription of MMTV in mammary gland.
cytes, other factors play a dominant role in these cells. Indeed, the T→A→G→C change at bp 520 does not affect lymphocyte transfection of MMTV, also indicating that control of expression does not rely on STAT5 factors in these cells. Interestingly, in mice with large copy numbers of the L-STAT-containing transgene, we did detect high levels of both lymphoid and mammary gland tissue expression. This could be due to integration site effects or perhaps because the factors that control transcription through the ML-STAT sequences also weakly interact with the L-STAT sequence; the large number of L-STAT copies present in the same region of the chromosome may alter the kinetics of binding of such factors, thereby allowing transcription. Indeed, the upper complex of the doublet formed on the ML-STAT probe was also weakly found on the L-STAT probe (Fig. 3 and 5).

Stat5b may play an important role in the control of gene expression in the mammary gland that is not solely dependent on prolactin induction. In support of this possibility, tissue culture transfection experiments have shown that Stat5b but not Stat5a could activate gene expression in the absence of this hormone (32). We also found that the complex formed on the ML-STAT oligonucleotide was present in the absence of prolactin, although prolactin both induced the factor(s) bound to this oligonucleotide and caused increased MMTV transcription in a mammary gland tissue culture cell line. Moreover, although MMTV RNA levels increase dramatically in response to lactation in vivo, due at least in part to increased glucocorticoids and progesterone and perhaps to prolactin, MMTV transcription also occurs in nonlactating mammary gland (41). This finding indicates that MMTV expression in the mammary gland is dependent on a transcription factor that is active in the absence of prolactin. Since a number of other growth factors activate Stat5b (as well as Stat5a), it is possible that activation of this factor in the mammary gland occurs predominantly through cytokines other than prolactin.

This is the first example of a STAT consensus sequence affecting a gene’s tissue-specific expression, rather than just its ability to be induced by a growth factor or mitogen. Indeed, mutation of the WAP or β-lactoglobulin STAT consensus sequence decreased but did not eliminate mammary gland expression in transgenic mice (7, 30). The STAT consensus sequence in all milk protein genes differ from that found in MMTV (see Table 1). For example, there is a C at position 519 in the MMTV sequence instead of the T that is usually seen in the binding site for all STAT factors. Whether this change affects the interaction of Stat5b or some other factor with the MMTV sequence is not yet known. Although none of the WAP or β-lactoglobulin constructs used to make transgenic mice had a mutation in an equivalent position in the STAT sequence, substitution of a C for a T in this position in the STAT sequence from the Ly6E/A gene completely abolished binding of STAT1α and STAT3 (28).

How is tissue specificity affected by Stat5b, since this factor is expressed in many tissues and induced by different cytokines (24, 49)? It has been shown recently by a number of groups that STAT factors form complexes with a number of transcription factors, including SPI (34), γRF (20), c-Jun (48), and the GR (53). One possibility is that mammary gland-specific expression is achieved by a complex formed between Stat5b and an as yet undiscovered factor and that this heterologous complex is required to activate MMTV transcription in this tissue. The recognition site for this unknown factor could be near or overlapping that of Stat5b or at some other location within the LTR. For example, we and others have shown previously that sequences between bp 28 and 207 in the MMTV LTR are important for mammary gland-specific expression (9, 27, 40, 41, 47). Alternatively, immediately adjacent to the STAT binding site in MMTV is a region with high homology to the rat β-casein gene (Fig. 8). Indeed, 18 of 21 bp are identical between the two genes in this region. It is possible that this region is a recognition site for an as yet unidentified factor that interacts with Stat5b and that this complex controls transcription from MMTV and perhaps the β-casein gene.

Using transient transfection assays of plasmids bearing the cloned GR and prolactin receptor (PrlR), STAT5, and a β-casein reporter construct, Stocklin and colleagues showed that prolactin interfered with glucocorticoid induction (53). Because they also found that the STAT5 and GR proteins complexed with each other, they proposed that the GR sequestered STAT5 and acted as a negative regulator of transcription. We showed here that glucocorticoids and dexamethasone showed synergistic activation of MMTV transcription in a stably transfected cell line. The differences in our results could be due to the level of expression of the GR, PrlR, or STAT5 protein; most likely the endogenous GR, PrlR, and STAT5 levels in NMuMG cells are much lower than those achieved in transient transfection. Similarly, it is possible that these transcription factors interact differently with each other in the context of a stably integrated transcription unit than on an unintegrated plasmid. In support of this possibility, it has recently been suggested that this mechanism of negative regulation, termed squelching, may occur in transient transfection assays but not on genomic DNA (43). Finally, it is also possible that these two transcription factors form a complex that is distinct from that formed on the β-casein gene or do not directly interact together at all on the MMTV regulatory sequences.

In the β-casein gene, the region thought to control STAT5-regulated expression also contains a duplicated STAT element and a YY1 binding site (Fig. 7). It has been shown that this YY1 binding site prevents STAT5 binding and represses β-casein gene transcription (39, 45). There is no YY1 consensus sequence adjacent to the STAT binding site in the MMTV LTR, and so there may indeed be different regulation of MMTV and the β-casein gene. However, SATB1, which binds to the MMTV LTR at two positions (bp 688 to 717 and 908 to 927), does repress MMTV transcription (31). Since both SATB1 and YY1 are nuclear matrix binding proteins (19, 42), these factors could function similarly in conjunction with STAT factors in the regulation of MMTV and β-casein transcription, respectively.

The majority of MMTVs are expressed in lymphoid tissue in addition to mammary gland. It has been reported that cytokines such as interleukin-2 and interleukin-5, as well as lipopolysaccharide, a B-cell mitogen, induce MMTV transcription in B cells (26, 35). These cytokines and lipopolysaccharide induce transcription of a number of genes through activation of the JAK-STAT pathway (13, 23, 25, 55). Interestingly, the casein (18) and WAP (30) genes have been reported to be
transcribed in T lymphocytes at low levels. Determination of whether the STAT sequences in the MMTV LTR function to direct expression in lymphoid cells or to mediate response to cytokines awaits analysis similar to that described here for the mammary gland.

ACKNOWLEDGMENTS

We thank the members of our labs for helpful discussions, Mitch Lazar for critical reading of the manuscript, and Bernadette van den Hoogen and Marta de Olano Vela for excellent technical assistance. Plasmid pMT-BC was a gift from Brigitte Huber. The tissues from the Stat5a and Stat5b knockout mice were a kind gift from Lothar Henningshausen and James Ihle, respectively.

This work was supported by NCI grants CA45954 (S.R.R.) and grants from CONICET and Fundación Antorchas (I.P.).

REFERENCES


6. Dignam, J. D., R. M. Lebovitz, and R. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mamma-


19. Meier, V., and B. Groner. 1994. The nuclear factor YY1 participates in repression of the β-casein gene promoter in mammary epithelial cells and is counteracted by mammary gland factor during lactogenic hormone induc-


