MOLECULAR AND CELLULAR BIOLOGY, Feb. 1991, p. 886–893
0270-7306/91/020886-08$02.00/0

Vol. 11, No. 2

Structure and Expression of a Gene Encoding a Putative GTP-Binding Protein Identified by Provirus Integration in a Transgenic Mouse Strain

KATRIN MOOSLEHNER, URSULA MÜLLER, URSULA KARLS, LUTZ HAMANN, AND KLAUS HARBERS*

Heinrich-Pette-Institut für Experimentelle Virologie und Immunologie an der Universität Hamburg, Martinistrasse 52, 2000 Hamburg 20, Federal Republic of Germany

Received 24 July 1990/Accepted 8 November 1990

The *Mov-10* mouse strain was derived by infection of preimplantation embryos with the Moloney murine leukemia virus and carries one copy of the provirus in its germ line. Here we show that the provirus has integrated into an evolutionarily conserved gene that can code for a protein of 110 kDa containing the three consensus elements characteristic for GTP-binding proteins. The *Mov-10* locus was expressed in a variety of cell types, including embryonal carcinoma and embryonic stem cells. Transcription of the gene was down-regulated about 10-fold when F9 embryonal carcinoma cells are differentiated into parietal endodermlike cells and about 2-fold when they are differentiated into visceral endodermlike cells. High levels of *Mov-10* transcripts were also found at different stages of embryonal development and in the testes and thymus of adult animals. Expression was cell cycle controlled, with steady-state RNA levels significantly higher in growth-arrested than in growth-stimulated cells. The results suggest that the *Mov-10* locus has an important function in development and/or control of cell proliferation. The provirus was shown to have integrated into intron 1 of the gene without disrupting expression, indicating that integration into intronic sequences of a transcription unit does not necessarily affect transcription. This result together with previous results from the *Mov-13* mouse strain suggest that proviruses exert their mutagenic effect only by integration in specific sites, such as *cis*-regulatory DNA elements.

The introduction of foreign DNA into the mouse germ line by retroviral infection or microinjection can result in insertional mutations of genes with important roles in development (for reviews, see references 10 and 17). By using the transgene as a tag for molecular cloning, it is possible to isolate flanking cellular DNA sequences and thus to identify the mutated gene. Several retrovirus-induced mutations in mice occurring either spontaneously or through experimental manipulations have been described previously (20, 23, 33, 35, 36). A summary of published and unpublished data shows that about 5% of all proviral insertions in transgenic mice result in recessive lethal mutations. This rather high frequency has led to the suggestion that retroviruses integrate preferentially into transcribed DNA regions (10). In support of this hypothesis, we have recently shown that proviruses in at least three of five transgenic *Mov* mouse strains, which were derived by retroviral infection of preimplantation embryos with the Moloney murine leukemia virus (M-MuLV), had integrated in DNA regions transcribed in embryonal stem cells (27). In none of these mouse strains has provirus integration resulted in an altered phenotype, suggesting either that transcription of the respective gene is not blocked or that a block in expression does not interfere with normal development.

In the present report, we describe the structure and expression of one of these genes identified by provirus integration in the *Mov-10* mouse strain. We show that the gene encodes a protein with potential GTP-binding activity and that its expression is regulated during development and cell proliferation. Transcription is not affected by the insertion of the provirus in intron 1.

* Corresponding author.

MATERIALS AND METHODS

Mice. All mice were bred in our mouse colony at the Heinrich-Pette-Institut. The origin of the *Mov-10* mice has been described earlier (19). Embryos were obtained from timed matings; the day of the mating plug was counted as day 1 of gestation.

Cell culture. F9 cells were grown on gelatinized petri dishes in Dulbecco modified Eagle medium containing 10% fetal calf serum (FCS). Cells were induced to differentiate into parietal endodermlike cells by growth as monolayers in the presence of 5×10^{-8} M retinoic acid (RA), 10^{-4} M N^6,O^2-dibutyryl cyclic AMP (cAMP), and 10^{-4} M isobutyl methyl xanthine for 4 to 5 days unless stated otherwise (14). Differentiation to visceral endodermlike cells was achieved by growing small aggregates of 20 to 30 cells for 7 days in suspension in bacteriological petri dishes in the above medium containing 0.1 μM RA (14, 15). CCE cells (7), obtained from M. Evans, Cambridge, England, were maintained in Buffalo rat liver cell-conditioned medium as described previously (34). Mouse NIH 3T3 cells were routinely grown in Dulbecco modified Eagle medium supplemented with 10% FCS. To obtain quiescent cells, we incubated exponentially growing cells for 72 h in the presence of 1% FCS. For serum stimulation, quiescent cultures were incubated with 10% FCS for the indicated times. Before RNA isolation, the growth state of the cells was assessed in parallel cultures by a 1-h pulse with [^{3}H]thymidine (15 μC/ml; 20 Ci/mmol) followed by cell lysis and trichloroacetic acid precipitation of incorporated [^{3}H]thymidine.

Screening of cDNA library. A cDNA library prepared from the undifferentiated F9 embryonal carcinoma cell line constructed in λgt10 (16) was a gift from H. Shin, Whitehead Institute, Cambridge, Mass. About 5×10^5 plaques were

screened with genomic probes from the *Mov-10* locus described previously (27). Eight positive plaques were identified and plaque purified. Of these plaques, the largest one, designated p10-CD1, was further characterized.

DNA sequencing. The nucleotide sequence of cDNA clone p10-CD1 was determined by the dideoxy-chain termination method (31). Restriction fragments were subcloned into M13mp18 and M13mp19 bacteriophages. M13 17-bp universal primer along with synthetic oligonucleotides were used as primers. Both strands were sequenced completely.

RNA preparation and analysis. Total cellular RNA was isolated by the method of Auffray and Rougeon (2) and fractionated (15 μg per lane) by electrophoresis in 1% agarose slab gels containing formaldehyde and ethidium bromide (30). rRNA bands were used to confirm the integrity and amount of RNA in each lane. RNA was transferred onto Gene Screen Plus membranes (Dupont) and hybridized in the presence of 50% formamide at 42°C for 18 to 20 h as described by the manufacturers. ^{32}P-radiolabeled probes of the *Mov-10* cDNA clone p10-CD1 were prepared by the random priming procedure of Feinberg and Vogelstein (8). Final filter washes were done at 65°C with a solution of 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) and 1% sodium dodecyl sulfate.

Nucleotide sequence accession number. The nucleotide sequence shown in Fig. 1 has been assigned the EMBL accession number X52574.

RESULTS

***Mov-10* cDNA nucleotide sequence and predicted amino acid sequence.** The *Mov-10* locus was originally defined by the presence of an M-MuLV provirus in the *Mov-10* mouse strain (12, 19). Previous experiments indicated that the provirus was located within a transcription unit as genomic fragments on both sites of the integration site hybridized with RNA (27). Subfragments of a cosmid clone hybridized to a 3.6-kb RNA species in both the CCE embryonal stem cell line and the F9 embryonal carcinoma cell line (27). The same fragments were used to screen a λgt10 cDNA library prepared from RNA of undifferentiated F9 cells. Several positive plaques were isolated and characterized by restriction digestion. The largest cDNA insert was about 3.6 kb long and was recloned in the ribovector pSP65. This clone, designated p10-CD1, was shown by sequence and primer extension analysis to represent the total length of the *Mov-10* transcript and was used as a probe in all subsequent expression studies. The complete nucleotide sequence of p10-CD1 was determined and is shown in Fig. 1. It is about 3,528 bp long, with a major open reading frame of 3,012 bp beginning with a potential ATG initiator codon at nucleotide 171 and ending with a TGA terminator codon at position 3183, coding for a protein with a calculated molecular mass of about 110 kDa. Two other possible initiation codons are present at positions 339 and 351; however, the ATG at position 171 conforms best to the consensus for translation start sequences (21). The 3′ untranslated region does not contain the AATAAA polyadenylation signal usually located 10 to 30 nucleotides upstream of the poly(A) addition site (29). Instead, the sequences AAGAAA and AACAAA are 21 and 14 bp upstream of the terminal poly(A) tract, respectively, either of which might serve as a polyadenylation signal. Sequence analysis of a genomic clone from the 3′ end (data not shown) confirmed the absence of an AATAAA motif and showed that about 14 bp downstream of the poly(A) addition site there is a GT-rich sequence which is present in many genes and has been postulated to be involved in the cleavage and/or processing of primary transcripts (25). A search for similarities with protein sequences collected in the Leeds data base revealed that the gene product of the *Mov-10* locus belongs to the GTP-binding protein family. It contains the three peptide sequences GPPGTGKT, DEAG, and NGYD (separated by 113 and 69 residues, respectively), corresponding to the three consensus motifs found in GTP-binding proteins (4). The peptide structure suggests that the *Mov-10* gene product is not a membrane-associated protein.

Expression in F9 cells and differentiated derivatives of F9. We have previously shown that the *Mov-10* locus is expressed in the embryonal carcinoma cell line F9 (27). Transcription decreased by at least 10-fold after differentiation of F9 cells into parietal endodermlike cells and by about 2-fold after differentiation into visceral endodermlike cells. (Fig. 2A). Down-regulation of *Mov-10* mRNA after differentiation of F9 cells into parietal endoderm cells with RA and cAMP was studied in more detail, and the results are shown in Fig. 2B. A time course experiment revealed that after 96 h the *Mov-10*-specific RNA was no longer detectable. Differentiation in the presence of cAMP alone produced hardly any change in *Mov-10* expression, whereas RA alone resulted in a three-fold decrease. Exposure to both RA and cAMP was required to give maximal down-regulation of the *Mov-10* mRNA level. A characteristic feature of the differentiated F9 phenotype is that once established, it remains stable after removal of the inducers (38). *Mov-10* mRNA levels were similarly reduced upon incubation with RA and cAMP for 1 day followed by 3 days without inducers compared with continuous exposure to inducers (Fig. 2B). This result indicates that the down-regulation of *Mov-10* mRNA is associated with the differentiated phenotype and does not require the continuous presence of the inducers.

Expression in embryos and different tissues of adult mice. To determine whether the *Mov-10* locus is expressed in other cell types, we extracted RNA from embryos of different developmental stages and from different tissues of adult mice and assayed it by Northern (RNA) blot analysis. The integrity and amount of RNA in each lane was confirmed by staining RNA bands with ethidium bromide (data not shown). Furthermore, blots were rehybridized with an actin probe, which, however, hybridized to different extents with total RNA of various organs. The *Mov-10* locus was transcribed at all stages of fetal development (Fig. 3). The transcript has the same size as the one in F9 cells. In late embryos (days 17 to 19 of gestation) and in newborn mice, expression seemed to occur predominantly in fetal liver. *Mov-10* mRNA was also present in placenta of different developmental stages (data not shown). When RNA from different tissues of adult mice was analyzed, significant expression was found only in testes and thymus (Fig. 4). Expression in adult liver (Fig. 4) was significantly reduced when compared with expression in fetal or newborn liver (Fig. 3). The decline from fetal to adult levels occurred during the first week of postnatal development (data not shown). Low levels of *Mov-10* mRNA were detected in all tissues tested with the exception of muscle (Fig. 4) and skin (data not shown). In the latter tissues, transcripts were absent even after long exposures of Northern blots with poly(A)-enriched RNA or when an RNase protection assay was used. In summary, the results in Fig. 3 and 4 show that expression of the *Mov-10* locus is not restricted to embryonal stem cells but occurs also in other cell types. Furthermore, the expression is tissue and stage specific.

Expression at different stages of the cell cycle. Mouse NIH

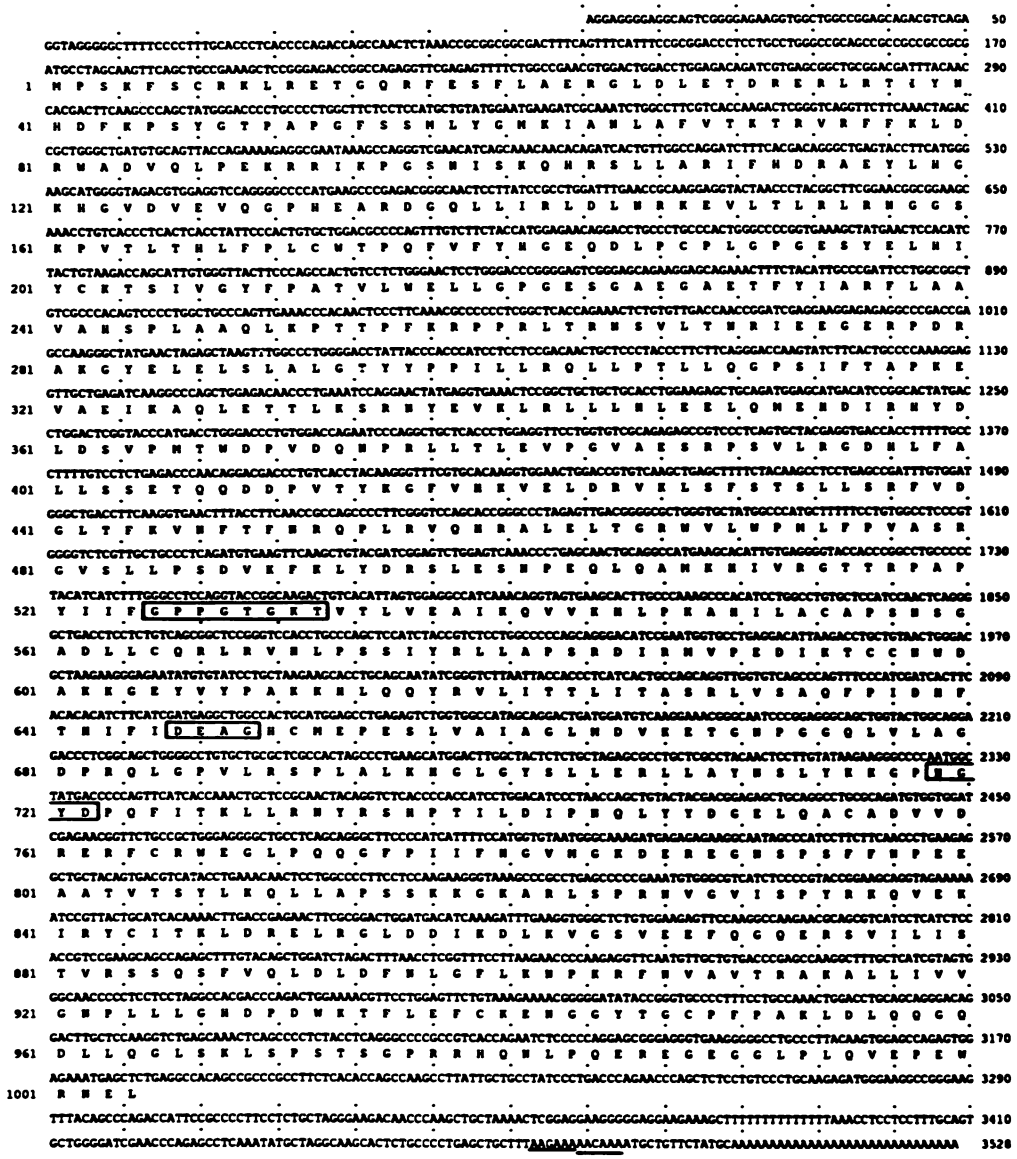

FIG. 1. Nucleotide sequence of the *Mov-10* cDNA clone p10-CD1 and the deduced amino acid sequence of the protein. The three peptide sequences corresponding to consensus motifs found in GTP-binding domains (4) are boxed. The sequences AAGAAA and AACAAA potentially used as polyadenylation signals are underlined. The numbers on the left refer to amino acid position, and those on the right refer to nucleotide position.

3T3 fibroblasts constitute a convenient system for studying cell cycle-dependent gene expression since they can be brought to a quiescent state (G0) by serum deprivation and, upon subsequent exposure to serum, they reenter the G1 phase and resume proliferation. The *Mov-10* RNA level decreased after serum stimulation and reached a minimum after 8 to 12 h, when it again increased and reached a steady-state level that was about the same as the level in arrested cells after 22 h (Fig. 5). Subsequent rehybridization of the filter with a probe corresponding to glyceraldehyde-3-phosphate dehydrogenase, whose gene expression does not vary significantly during the cell cycle, confirmed that similar amounts of RNA were present in different lanes. Therefore, our results suggest that the expression of the *Mov-10* locus is cell cycle controlled, with steady-state RNA levels higher in growth-arrested than in growth-stimulated cells.

Provirus integration and expression in *Mov-10* mice. The exon-intron structure of the *Mov-10* locus was determined by restriction, nuclease S1, and partial sequence analyses of genomic and cDNA clones. The provirus integration site with respect to the transcribed gene was mapped by restriction analysis of the previously described genomic clone pMov10, which contains the proviral insert (12), and the cosmid clone representing the wild-type allele (27). The total length of the gene is about 17 kb. Only the exon-intron distribution at the 5′ end is shown in Fig. 6. The provirus integrated into intron 1 about 8.5 kb downstream of the first exon-intron boundary in the opposite orientation relative to the transcriptional orientation of the *Mov-10* locus. We have

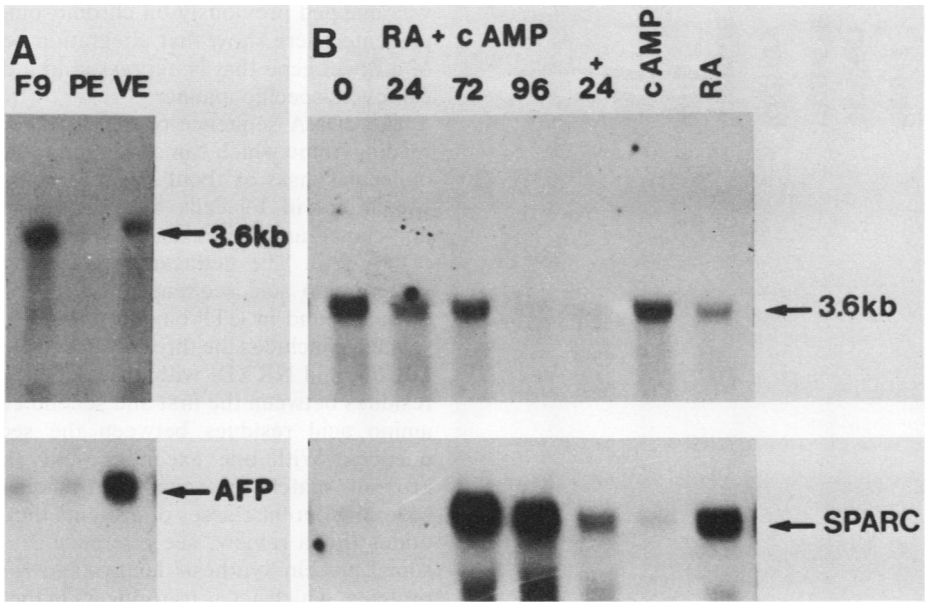

FIG. 2. Expression of *Mov-10* locus in F9 cells and its differentiated derivatives. (A) Total cellular RNA from undifferentiated F9 cells and F9 cells differentiated into parietal endodermlike cells (PE) and visceral endodermlike cells (VE) was analyzed by Northern blot hybridization with the *Mov-10* cDNA clone p10-CD1 as a probe. The filter was rehybridized with the α-fetoprotein gene (AFP) to confirm the differentiated state of the VE cells (6). (B) Effect of RA and cAMP on expression of the *Mov-10* locus from F9 cells cultured for indicated times (hours) in the presence of RA and cAMP, from F9 cells kept for 24 h in the presence of RA and cAMP followed by 72 h without inducers (lane labeled +24), and from F9 cells kept for 96 h in the presence of either cAMP or RA was analyzed by Northern blot hybridization. The filter was rehybridized with the SPARC gene to confirm the differentiated state of the F9 cells (24).

previously shown that provirus integration into intron 1 can result in a transcriptional block and lead to a recessive lethal mutation (11). It was therefore of interest to analyze transcription of the *Mov-10* locus in homozygous *Mov-10* mice. For this purpose, RNA was isolated from d15 embryos and from thymus and testes of adult homozygous *Mov-10* animals and of adult 129 mice, the parental strain from which the *Mov-10* strain was derived (19). The Northern blot shown in Fig. 7 shows that provirus integration did not affect the size and the amount of the *Mov-10*-specific transcript. The transcript from homozygous *Mov-10* mice did not hybridize with M-MuLV sequences, ruling out the possibility that it is a hybrid RNA consisting of cellular and viral sequences (data not shown). The observation that the pro-

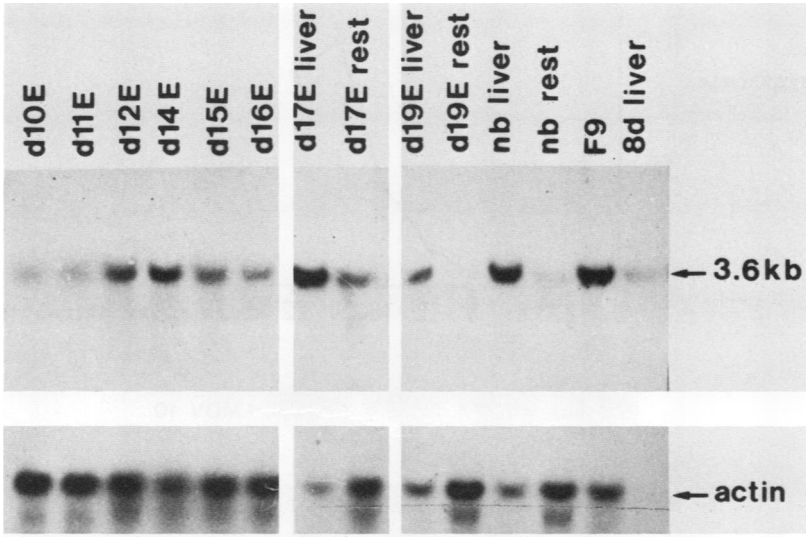

FIG. 3. Expression of *Mov-10* locus during embryonal development. Total cellular RNA isolated from whole embryos or from liver and remaining embryo at indicated days of development, from liver and remainder of newborn (nb) animals, from livers of 8-day-old animals, and, for comparison, from F9 cells was analyzed by Northern blot hybridization. The blot was subsequently rehybridized with the actin probe pAct-1 (5).

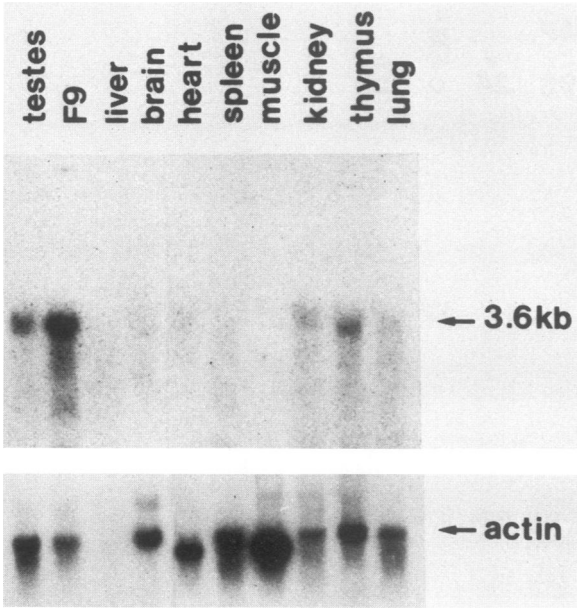

FIG. 4. Expression of *Mov-10* locus in different organs of adult mice. Total cellular RNA from indicated organs of 2-month-old animals and, for comparison, from F9 cells was analyzed by Northern blot hybridization. The blot was subsequently rehybridized with the actin probe.

virus did not interfere with transcription is in agreement with the fact that homozygous *Mov-10* mice do not show any phenotypic alteration during their entire life span (18).

DISCUSSION

***Mov-10* locus: cDNA and predicted protein.** The *Mov-10* mouse strain was derived by infection of preimplantation embryos with M-MuLV (19). The proviral integration site was mapped previously on chromosome 3 (28). The results presented here show that integration occurred into intron 1 of a novel gene that is expressed in a developmentally and cell cycle-specific manner.

The cDNA sequence of the gene contains a major open reading frame which can code for a protein with a calculated molecular mass of about 110 kDa. A protein of similar size was found in F9 cells by Western immunoblotting with polyclonal antibodies made from the in vitro-translated cDNA (9a). The deduced protein (Fig. 1) contains three short amino acid sequences that match the consensus sequence found in GTP-binding proteins (4). This consensus sequence includes the three sequence elements GXXXXGK, DXXG, and NKXD, with spacings of 40 to 170 amino acid residues between the first and second elements and 40 to 80 amino acid residues between the second and third sequences. With one exception, only known GTP-binding proteins match this consensus sequence (4). There are several different classes of proteins that bind guanine nucleotides (for a review, see reference 3). These include ribosomal protein synthesis factors and *ras-21* proteins and G proteins, which act as transducers in the signal pathway. The function of the protein described here is not known. Experiments are under way to characterize biochemically the GTP-binding and hydrolysis activity as well as its location in the cell.

A distinct feature of the *Mov-10* cDNA sequence is the absence of the sequence motif AATAAA, which usually serves as a polyadenylation signal and in most genes is located 10 to 30 nucleotides upstream of the poly(A) addition site (29). For reasons presented in the Results, we think that either AACAAA or AAGAAA is used as a polyadenylation signal. However, both sequences have previously been shown by mutational analysis to result in greatly reduced levels of polyadenylated mRNA (26, 39). Further studies are necessary to resolve this discrepancy and to study the possible role of these sequences in regulating the level of *Mov-10* transcripts.

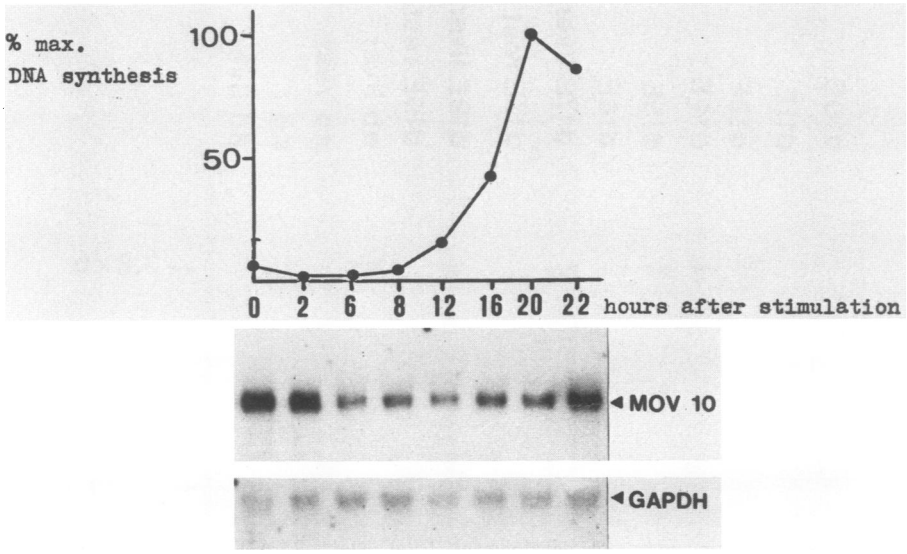

FIG. 5. Growth cycle regulation of *Mov-10* gene expression. Total cellular RNA was extracted from arrested cells and at indicated times after the addition of 10% FCS and was analyzed by Northern blotting. DNA synthesis was measured at the indicated times by [^{3}H]thymidine incorporation. The blot was subsequently rehybridized with the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) probe.

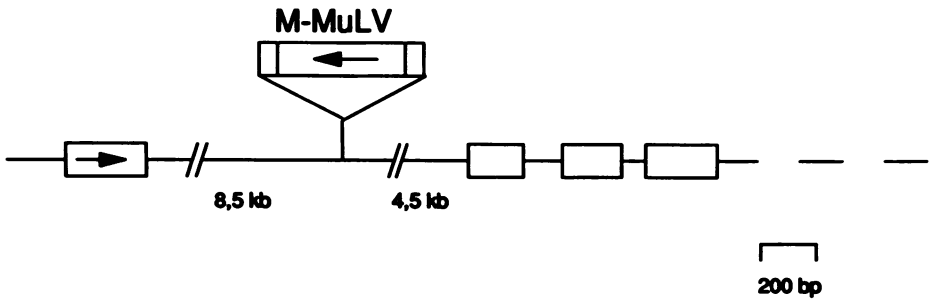

FIG. 6. Schematic representation of the exon-intron structure at the 5' end of the *Mov-10* gene. The first four exons (indicated by open boxes) and the integration site of the M-MuLV provirus are shown. Transcriptional orientations of the gene and the provirus are indicated by arrows.

Developmentally and cell cycle-regulated expression. The *Mov-10* locus can serve as a model system for studying gene regulation during the formation of early embryonic tissues. Its expression is decreased at least 10-fold upon differentiation to parietal endodermlike cells and about 2-fold upon differentiation to visceral endodermlike cells (Fig. 2). Down-regulation of *Mov-10* expression in parietal endodermlike cells appears to reflect the differentiated state of the cell, since it remains stable after removal of the inducers RA and cyclic AMP. As the level of transcription is reduced only after more than 72 h in the presence of the inducers, the *Mov-10* locus is not involved in early events of the differentiation process. As undifferentiated embryonal carcinoma cells are tumorigenic, whereas differentiated derivatives have lost this property (37), it is possible that expression of the *Mov-10* locus is associated with maintaining the rapidly growing transformed state.

Expression of the *Mov-10* locus is regulated during the cell cycle (Fig. 5) and reaches its highest level in growth-arrested cells. After induction of growth by serum stimulation, the expression is down-regulated (to a minimum after 8 to 12 h) and accumulates again during the S phase of synchronized cultures. Studies on the control of cell proliferation have mainly focused on positively regulated genes that are activated after stimulation of arrested cells (e.g., see reference 1). Very few studies so far have dealt with negatively regulated genes that are specifically expressed in the G0 phase and are down-regulated after growth stimulation. Some growth arrest-specific (*gas*) genes have recently been isolated from a subtraction cDNA library (32) and show cell cycle-controlled expression patterns similar to the one described here for the *Mov-10* locus. The *Mov-10* locus and its gene product may provide an additional tool to understand growth control.

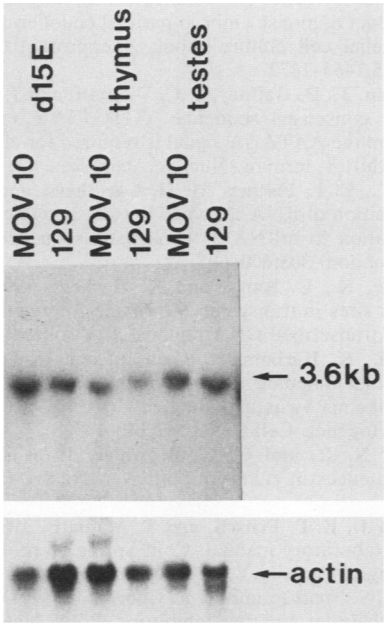

FIG. 7. Expression of *Mov-10* locus in homozygous *Mov-10* mice. Total RNA was isolated from whole 15-day-old embryos and from indicated organs of 2-month-old homozygous *Mov-10* and 129 mice and analyzed by Northern blotting. The blot was subsequently rehybridized with the actin probe.

Gene expression and provirus integration. Transcription of the *Mov-10* locus does not seem to be affected by the provirus insertion since no differences were observed when the amount and size of *Mov-10*-specific RNA from three different tissues of homozygous *Mov-10* and wild-type mice were compared (Fig. 7). These results suggest that processing of the primary transcript, including proper splicing of intron 1 containing the viral sequence, results in a mature mRNA indistinguishable from that derived from the wild-type allele. This conclusion is supported by the observation that homozygous *Mov-10* mice develop normally and show no phenotypic alteration (18). However, the experiments presented here do not rule out small alterations in mRNA structure and/or a transcriptional block in minor cell types.

When compared with the previously described *Mov-13* mouse strain, the present results show clear differences but also some similarities. In both *Mov* mouse strains the provirus has integrated into intron 1 of a gene, with the transcriptional orientation of the provirus opposite to that of the host gene (11). In the *Mov-13* mouse, provirus integration into intron 1 of the α1(I) collagen gene has led to a block in transcription resulting in a recessive lethal mutation (11, 13, 18, 33). However, recently we have shown that the *Mov-13* provirus affects transcription in a tissue-specific manner (22). In contrast to all other cell types, the mutated *Mov-13* allele is transcribed in odontoblasts and osteoblasts and the primary transcripts are spliced properly, resulting in a functional mature mRNA. These results suggest that odontoblasts and osteoblasts use *cis*-regulatory elements for expression of the α1(I) collagen gene which are different from those in other cells and which are not affected by the provirus (22). Similarly, we assume that the provirus in the *Mov-10* locus does not disrupt any regulatory DNA elements

and therefore does not interfere with transcription. The transcriptional analysis of the *Mov-10* and *Mov-13* loci shows that provirus integration into intronic sequences does not necessarily affect transcription and implies that proviruses exert their mutagenic effect only within a limited range of a chromosomal domain. From the results of the present study and previous studies of the *Mov-13* locus, we suggest that proviruses have to disrupt control elements to cause mutations. Support for this hypothesis comes from related work in other laboratories, where it was shown that retrovirus-induced mutations are clustered in a defined region in or near to the affected gene (e.g., see reference 9). For *Mov-10* and *Mov-13*, experiments are under way to define the *cis*-regulatory DNA elements and their positions relative to the proviral insert.

ACKNOWLEDGMENTS

We thank R. Jaenisch for making the *Mov-10* mouse strain available, H. Shin for kindly supplying the F9 cDNA library, and A. Schnieke for the initial screening of the library. We thank John Coadwell, Computer Department, AFRC Institute for Animal Physiology and Genetics Research, Babraham, United Kingdom, for the protein similarity search.

The work was supported by a grant from the Deutsche Forschungsgemeinschaft. The Heinrich-Pette-Institut is supported by Freie und Hansestadt Hamburg and Bundesministerium für Jugend, Familie, Frauen und Gesundheit.

REFERENCES

1. **Almendal, J. M., D. Sommer, J. Perera, J. Burckhardt, H. Macdonald-Bravo, and R. Bravo.** 1988. Complexity of the early genetic response to growth factors in mouse fibroblasts. Mol. Cell. Biol. **8**:2140–2148.
2. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy-chain messenger RNA's from total myeloma tumor RNA. Eur. J. Biochem. **107**:303–314.
3. **Bosch, L., B. Kraal, and A. Parmeggiani (ed).** 1989. The guanine nucleotide-binding proteins: common structural and functional properties. Plenum Publishing Corp., New York.
4. **Dever, T. E., M. J. Glynias, and W. C. Merrick.** 1987. GTP-binding domain: three consensus sequence elements with distinct spacing. Proc. Natl. Acad. Sci. USA **84**:1814–1818.
5. **Dodemont, H. J., P. Soriano, W. J. Quax, F. C. S. Ramaekers, J. A. Lenstra, M. A. M. Groenen, G. Bernardi, and H. Bloemendal.** 1982. The genes coding for the cytoskeletal proteins actin and vimentin in warm-blooded vertebrates. EMBO J. **1**:167–171.
6. **Dziadek, M., and G. K. Andrews.** 1983. Tissue specificity of alpha-fetoprotein messenger mRNA expression during mouse embryogenesis. EMBO J. **2**:549–554.
7. **Evans, M. J., and M. H. Kaufman.** 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature (London) **292**:154–156.
8. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **137**:266–270.
9. **Frankel, W. N., T. A. Potter, and T. V. Rajan.** 1989. Effect of proviral insertion on transcription of $B2m^b$ gene. J. Virol. **63**:2623–2628.

9a. **Gessner, A., H. Lothar, and K. Harbers.** Unpublished data.

10. **Gridley, T., P. Soriano, and R. Jaenisch.** 1987. Insertional mutagenesis in mice. Trends Genet. **3**:162–166.
11. **Harbers, K., M. Kuehn, H. Delius, and R. Jaenisch.** 1984. Insertion of retrovirus into the first intron of α1(I) collagen gene leads to embryonic lethal mutation in mice. Proc. Natl. Acad. Sci. USA **81**:1504–1508.
12. **Harbers, K., A. Schnieke, H. Stuhlmann, and R. Jaenisch.** 1982. Infectivity and structure of clones obtained from genetically transmitted Moloney leukemia proviral genomes. Nucleic Acids Res. **10**:2521–2537.
13. **Hartung, S., R. Jaenisch, and M. Breindl.** 1986. Retrovirus insertion inactivates mouse α1(I) collagen gene by blocking initiation of transcription. Nature (London) **320**:365–367.
14. **Hogan, B., F. Constantini, and E. Lacy.** 1986. Manipulating the mouse embryo: a laboratory manual, p. 262–267. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Hogan, B. L. M., A. Taylor, and E. D. Adamson.** 1981. Cell interactions modulate embryonal carcinoma cell differentiation into parietal or visceral endoderm. Nature (London) **291**:235–237.
16. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Constructing and screening cDNA libraries in λgt10 and λgt11, p. 49–78. *In* D. Glover (ed.), DNA cloning, vol. 1. A practical approach. IRL Press, Oxford.
17. **Jaenisch, R.** 1988. Transgenic animals. Science **240**:1468–1474.
18. **Jaenisch, R., K. Harbers, A. Schnieke, J. Löhler, J. Chumakov, D. Jähner, D. Grotkopp, and E. Hoffmann.** 1983. Germ line integration of Moloney murine leukemia virus at the Mov13 locus leads to recessive lethal mutation and early embryonic death. Cell **32**:209–216.
19. **Jaenisch, R., D. Jähner, P. Nobis, J. Simon, J. Löhler, K. Harbers, and D. Grotkopp.** 1981. Chromosomal position and activation of retroviral genomes inserted into the germ line of mice. Cell **24**:519–529.
20. **Jenkins, N. A., N. G. Copeland, B. A. Taylor, and B. K. Lee.** 1981. Dilute (d) coat colour mutation of DBA/2J mice is associated with the site of integration of an ecotropic MuLV genome. Nature (London) **293**:370–374.
21. **Kozak, M.** 1984. Compilation and analysis of sequences from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **11**:6199–6209.
22. **Kratochwil, K., K. von der Mark, E. J. Koller, R. Jaenisch, K. Mooslehner, M. Schwarz, K. Haase, J. Gmachl, and K. Harbers.** 1989. Retrovirus-induced insertional mutation in Mov13 mice affects collagen I expression in a tissue-specific manner. Cell **57**:807–816.
23. **Kuehn, M. R., A. Bradley, E. J. Robertson, and M. J. Evans.** 1987. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. Nature (London) **326**:295–298.
24. **Mason, I. J., A. Taylor, J. G. Williams, H. Sage, and B. L. M. Hogan.** 1986. Evidence from molecular cloning that SPARC, a major product of mouse embryo parietal endoderm, is related to an endothelial cell culture shock glycoprotein of Mr 43000. EMBO J. **5**:1465–1472.
25. **McLauchlan, J., D. Gaffney, J. C. Whitton, and J. B. Clements.** 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**:1347–1368.
26. **Montell, C., E. F. Fischer, M. H. Caruthers, and A. J. Berk.** 1983. Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3′ consensus sequence AAUAAA. Nature (London) **305**:600–605.
27. **Mooslehner, K., U. Karls, and K. Harbers.** 1990. Retroviral integration sites in transgenic Mov mice frequently map in the vicinity of transcribed DNA regions. J. Virol. **64**:3056–3058.
28. **Münke, M., K. Harbers, R. Jaenisch, and U. Francke.** 1986. Chromosomal mapping of four integration sites of Moloney murine leukemia virus including the locus of α1(I) collagen in mouse. Cytogenet. Cell Genet. **43**:140–149.
29. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′-Non-coding region sequences in eukaryotic mRNA. Nature (London) **263**:211–214.
30. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
31. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
32. **Schneider, C., R. M. King, and L. Philipson.** 1988. Genes specifically expressed at growth arrest of mammalian cells. Cell **54**:787–793.
33. **Schnieke, A., K. Harbers, and R. Jaenisch.** 1983. Embryonic

lethal mutation in mice induced by retrovirus insertion into the α1(I) collagen gene. Nature (London) **304**:315–320.

34. **Smith, A. G., and M. L. Hooper.** 1987. Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. Dev. Biol. **121**:1–9.
35. **Soriano, P., T. Gridley, and R. Jaenisch.** 1987. Retroviruses and insertional mutagenesis in mice: proviral integration at the Mov34 locus leads to early embryonic death. Genes Dev. **1**:366–375.
36. **Stoye, J. P., S. Fenner, G. E. Greenoak, C. Moran, and J. M. Coffin.** 1988. Role of endogenous retroviruses as mutagens: the hairless mutation of mice. Cell **54**:383–391.
37. **Strickland, S.** 1981. Mouse teratocarcinoma cells: prospects for the study of embryogenesis and neoplasia. Cell **24**:277–278.
38. **Strickland, S., K. K. Smith, and K. R. Marotti.** 1980. Hormonal induction of differentiation in teratocarcinoma stem cells. Generation of parietal endoderm by retinoic acid and dibutyryl cAMP. Cell **21**:347–355.
39. **Wickens, M., and P. Stephenson.** 1984. Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent mRNA 3′ end formation. Science **226**:1045–1051.