Biochemical and Biophysical Research Communications **288,** 22–33 (2001)
doi:10.1006/bbrc.2001.5718, available online at http://www.idealibrary.com on IDEAL®

Molecular Cloning and Characterization of the Mouse E2F6 Gene

Zoulika Kherrouche, Agnès Begue, Dominique Stehelin, and Didier Monté[1]
Institut de Biologie de Lille, UMR 8526 CNRS/Institut Pasteur de Lille, 1 rue Calmette, BP 447, 59021 Lille Cedex, France

Received September 3, 2001

E2F6 is the most recently identified member of the E2F family. In this study, the murine E2F6 gene was cloned and found to consist of eight exons. Analysis of its 5′ flanking region revealed two transcription start sites. The proximal promoter region contained no TATA or CAAT box. We also identified a novel E2F6 mRNA containing the alternative exon 2. The E2F6 mRNAs are highly expressed during mouse embryogenesis and are present in all adult tissues examined. Moreover, E2F6 shows a unique expression pattern in synchronized mouse embryonic fibroblasts. E2F6 expression rapidly increases during the G0–G1 transition, reaching its higher level in mid-G1, and remains relatively constant thereafter. These findings suggest that E2F6 may contribute to the regulation of events throughout the cell cycle. Isolation of the murine E2F6 gene is a step toward generation of genetically modified mouse models that will help to understand the functions of E2F6.

Key Words: **E2F; mouse; gene; promoter; cell cycle; alternative splicing.**

Precise control of cellular proliferation is essential for normal development and for prevention of proliferative diseases such as cancer. Transition through the cell cycle requires an interplay of transcription factors that coordinately induce or repress gene expression in a temporally defined manner. Over the past decade, the E2F transcription factor has emerged as a central component of this regulation pathway. The E2F transcription factor was first identified as the factor that bound specifically to an element in the adenovirus E2 promoter (1). E2F is an heterodimeric factor, containing a subunit encoded by the E2F family of genes and a subunit encoded by the DP family of genes. To date, six E2F (E2F1–6) genes and two DP (DP1–2) genes have been identified in mammalian cells and the E2F activity found in cell extracts is the collective activity of many different forms of E2F complexes (2, 3). Previous studies have detailed the distinct expression pattern of the E2F1-5 family genes (1). In quiescent cells, E2F1-2 expression is tightly associated with cell proliferation as a function of negative autoregulatory control of these genes. That is, the presence of E2F binding elements within their promoters allows repression of transcription in quiescent cells. Particularly striking is the distinct pattern of E2F4–5 expression. During cell cycle progression, the expression of these genes is less affected but the regulation seems to be related to the cellular context. In REF-52 and NIH-3T3 cells, mRNA levels of these genes remain relatively constant throughout the cell cycle (1, 4). In contrast, E2F4–5 expression is rapidly enhanced in murine embryonic fibroblasts (MEF) or human keratinocytes during G0/G1 transition and decreases thereafter (5). The E2F3 expression is more complex and, in fact, appears to combine the patterns for the other E2F species (4, 6). The E2F3a mRNA accumulates with kinetics similar to those of E2F1–2, whereas E2F3b mRNA level remains constant throughout the cell cycle. To date, no data are available on E2F6 expression during cellular proliferation. In addition to the transcriptional regulation of some E2F species, the E2F transcription factor activity is also subjected to several levels of control including cellular localization, phosphorylation, targeted degradation and binding by pRb family members (pRb, p107, p130) (2, 3, 7). Interaction of pRb proteins with the E2F transactivation domain, not only blocks E2F transcriptional activity, but also forms a complex that actively represses transcription of cell cycle genes. However, it has not been determined whether the most important role of pRb is the blocking of transcriptional activation by E2F or the formation of the active pRb–E2F repressor complex. E2F can be divided into three subgroups, on the basis of both sequence homology and functional properties. The first subclass contains E2F1-3, which exhibit high transcriptional activity and can drive quiescent cells to S phase (2, 3). The second subclass contains E2F4–5. These factors are poor transcriptional activators and they are unable to induce quies-

[1] To whom correspondence and reprint requests should be addressed. Fax: 3330320871111. E-mail: didier.monte@ibl.fr.

0006-291X/01 $35.00

cent cells to enter S phase. E2F6 represents the third subclass (8–11). It lacks not only the sequences required for transcriptional activity, but also the pocket protein binding domain. In other regions, E2F6 shares significant homology with other E2F proteins and in association with DP is able to bind to consensus E2F sites (TTTCGCGC) but presents a preference for the TTTCCCGC E2F recognition site (11). Little information is available as to the physiological role of E2F6. However, it inhibits activation by the other E2F members through a mechanism that involves promoter competition, it increases the percentage of certain cells in S phase and, when over-expressed, inhibits S phase entry of quiescent cells. The possibility that E2F6 may also be an active repressor is controversial and would be cell- and/or promoter-specific. Recently, Trimarchi *et al.* reported that E2F6 was at least a component of the mammalian Bmi1-containing polycomb complex *in vivo* (12). These results suggest that E2F6′s ability to repress the transcription of E2F-responsive genes depends on its capacity to recruit this known transcriptional repressor complex. These observations also raise the possibility that E2F6 will play a key role, beyond E2F regulation, in mediating the changes in transcriptional regulation that are essential for normal developmental patterning. They are in agreement with the high expression level of E2F6 observed during embryogenesis (13). Nevertheless, the functional implication of E2F6 in repressive activity of polycomb complexes are still missing.

There is currently no report concerning E2F6 regulation, interaction with the cell cycle machinery and its precise physiological functions. To better understand these points, we initiated a study on the mouse E2F6 homologue (also termed EMA) as a prerequisite for homologous recombination approach. In this report, we describe the cloning and sequencing of the full-length E2F6 cDNA, its gene structure, the molecular analysis of its promoter region, and its expression during the cell cycle progression after serum starvation.

MATERIALS AND METHODS

Library screening and sequence analysis. A mouse ovary λgt11 cDNA expression library was screened with a [α-^{32}P]dCTP (Amersham) labeled probe corresponding to the 5′ 130-bp E2F6 cDNA published. This fragment was generated by PCR using 1F6 and 2F6 oligonucleotides using the AmpliTaq XL kit (Perkin–Elmer). After three purification steps, phages were isolated as described (14) and the inserts were subcloned into the pBluescript II SK+ vector (Stratagene). Using these clones, the full-length and the splice E2F6 coding sequences were amplified using 3F6 and 4F6 primers and subcloned into PCRII-Topo vector (Invitrogen) to generated respectively TA–E2F6-b and TA–E2F6-a constructs.

A mouse genomic λDash DNA library, whose DNA was obtained from Male 129 cells, was screened with the *Eco*RI fragment of TA–E2F6-a. Four positive clones (A/B/C/D) were purified and subcloned into pZero vector (Invitrogen). The complete sequences of exons and partial sequences of introns were determined on both strands by the automatic sequencing PCR procedure with fluorescent primer coupled (Prism TM ready reaction, Dye Deoxy terminator kit from Applied Biosystems). For PCRs the following primers were used: 1F6, 5′-GAACCTGAGGCCTCCGCGGTCA-3′; 2F6, 5′-CTCCACCCACGGAGCACT-3′; 3F6, 5′-GATCGGATCCGTATGAGTCAGCAGCGGACGGCGCGG-3′; 4F6, 5′-ACTCTGAGTTCTTCACTCACTGAGG-3′; 5F6, 5′-TATGTAACCTATCAGGATATTCACGG-3′; 6F6, 5′-TATTCCATCATTGGTTTTACCATTT-3′; 7F6, 5′-TTTCTCATGGACACATACTG-3′; 8F6, 5′-AGTAGGTTTTCCACGTTGATGG-3′; 9F6, 5′-CCGCAAGCTTGCTGTCCCGCGCCCCTC-3′; 10F6, 5′-GGCAGATCTGACGAGATCCAT-3′; 11F6, 5′-CGGGGTACCCCTTAAGATGGTATC-3′; and 12F6, 5′-TTCCCGGCCCCTTCAGCGCCGCCCGCTCCACCCA-3′.

Constructs. A 2200-bp fragment of mouse E2F6 DNA just upstream the ATG (position −1944 to +256) was amplified by PCR using 11F6 and 9F6 oligonucleotides that contained *Kpn*I and *Hin*dIII restriction sites respectively for convenient subcloning. This fragment was first cloned into PCRII-Topo to generated TA–E2F6promo and then subcloned into the luciferase reporter plasmid pGL3-basic (Promega) by *Bst*EII digestion blunting with Klenow to generated (−1944/+55)-pGL3. The fidelity of the inserts was verified by sequence analysis. Additional deletion mutants were created by digesting the initial construct with *Msc*I to generated (−1360/+55)-pGL3, with HindIII to generated (−333/+55)-pGL3 and with *Sma*I to generated (−75/+55)-pGL3.

Cell lines and transfections. RK13 (rabbit kidney) and HepG2 (human hepatocarcinoma) cells were grown in 5% CO_2 at 37°C in DMEM (Dulbecco's modified Eagle medium) plus 10% heat-inactivated fetal calf serum. MEF (mouse embryonic fibroblast) were grown in the same medium supplemented with 1% nonessential amino acids (Gibco).

Transfections were performed in 12 well plates using 2 μl of Ex gen 500 (Euromedex) with 200 ng reporter vector (pGL3-basic, 5′-E2F6-pGL3). Cells were harvested 24 h after transfection and lysed in the Promega reporter lysis buffer. Transfection efficiencies were normalized to an internal β-galactosidase control. Luciferase and β-galactosidase activities are measured on a Berthold luminometer. Experiments were repeated at least three times for reproducibility.

5′-RACE and RNase protection. The mouse heart Cap-site cDNA commercial kit was used to conduct 5′-RACE experiments. Amplifications were performed according to the manufacturer's instructions (Eurogentec). The primary PCR were realized with 1RC primer and the E2F6 specific antisense primers (9F6 or 7F6) followed by a secondary PCR using 2RC and the antisense specific E2F6 primers (2F6 or 8F6). PCR products were analyzed, purified by agarose gel electrophoresis, cloned into PCRII-Topo vector and sequenced.

The *Bgl*II, *Bam*HI, and *Afl*III fragments of TA-E2F6promo clone were transcribed *in vitro* using T7 RNA polymerase in the presence of [α-^{32}P]CTP to generated respectively 388 bp 5′-*Bgl*II, 446 bp 5′-*Bam*HI and 729 bp 5′-*Afl*III probes. RNase protection assays on total embryo RNA or MEF poly(A)$^+$ RNA were performed using the RPA III kit (Ambion) according to the manufacturer's conditions. The sequence ladder was generated with the 12F6 primer using B clone as a template.

RNA isolation, Northern blots, and RT-PCR experiments. Total RNAs from mouse tissues and cultured cells were extracted using RNA PLUS reagent (Quantum). RNAs were then fractionated by electrophoresis on denaturing 2.2 M formaldehyde–1% agarose gels and transferred to Hybond-N membranes (Amersham) by blotting in 10× SSC. Filters were hybridized to ^{32}P-labeled probes at 42°C and washed in stringent conditions. The E2F6 cDNA probe was the *Eco*RI fragment of TA–E2F6-a construct containing the full ORF. The other probes were generated by PCR, subcloned into PCRII-Topo vector and the identity of the PCR fragments was verified by sequencing.

The primers pairs employed in PCR are as follows: Cyclin A (CyA-5′-GGCGGATCCGCGATGCCGGGCACCTCG and CyA-3′-GGCTCTAGACACACTTAGTGTCTCTGGTGG); Cyclin E

(CyE-5′-GGCGGATCCGCCATGCCAAGGGAGAGAGAC and CyE-3′-CGCCGCTAGCGCAGCTTCTGGAGCACTC); E2F1 (E2F1-5′-CTATGGAAGAGGACCAACTGTCAC and E2F1-3′-ACACACACTGTACAACATCCTTCC); and E2F4 (E2F4–5′-TCCTTTGAGCCCATCAAAGCAGAC and E2F4-3′-CAGTCCCTAGTTGGCTCCAGGCC). The cDNA probe of acidic ribosomal phosphoprotein P0 (ARPP-P0) was used as control and generated using ARPP-P0-5′-GGCATCACCACGAAAATCTCCA and ARPP-P0-3′-GGTTGCTTTGGCCGGGATTAGTC primers.

For RT experiments, 2 μg of total RNA was reverse-transcribed 50 min at 42°C with Superscript II (Gibco-BRL) by using an oligo(dT) primer. One microliter of the reaction mix RT was used as template for PCR amplification in a 100-μl total volume with *Taq* Gold Star (Eurogentec). PCR conditions were 94°C for 2 min; 30 cycles of 94°C for 40 s, 60°C for 40 s, 72°C for 1 min; and 72°C for 10 min. The combination of primers that were used to generate the PCR products are the couples 3F6 and 10F6.

MEF poly(A)$^+$ RNA were purified using oligo(dT)-cellulose columns (Pharmacia).

RESULTS

Probe Synthesis and Identification of an Alternative Splice of Mouse E2F6

RT-PCR using a set of primers encompassing the open reading frame of murine E2F6 generated two products (846 and 912 bp) from adult mouse heart total RNA (data not shown). These products were cloned and fully sequenced. Nucleotide sequencing revealed that the 846 bp product corresponded to the published E2F6 cDNA (AF032131) (8), while the second product displayed a 66 bp insertion, which could correspond to an alternative exon of E2F6. This alternative form of E2F6 is called E2F6-b.

Isolation and Characterization of the Mouse E2F6 Gene

To confirm that the E2F6-b specific sequence was derived from an exon contained within the E2F6 gene, a λ-Dash mouse genomic library was screened with the full-length E2F6 c-DNA probe (TA-E2F6-a, *Eco*RI digested). Four partially overlapping clones were isolated and subcloned into the pZero vector (Invitrogen). The lacking intron 5 was obtained by PCR using 5F6 and 6F6 primers on mouse genomic DNA (Fig. 1A). The E2F6 genomic structure was further characterized by restriction mapping and nucleotide sequence analysis (Figs. 1 and 2). This analysis revealed that the E2F6-b specific 66-bp fragment was present as an exon within the genomic sequence of E2F6 and corresponded to exon 2 (Fig. 1B).

The E2F6 gene assembled into a 18-kb organization that contained 7 introns and 8 exons. Intron lengths were determined by PCR with the intron-spanning specific primers using mouse genomic DNA and phage clones (Fig. 1C). The exon/intron boundaries sequences are shown in Table 1. Of the 7 splice donor/acceptor sites, 6 contained consensus GT/AG dinucleotides; only the splice donor site in exon 5, which had the TA sequence and the splice acceptor site in exon 6 which had the CA sequence, varied from the consensus pattern (15). The open reading frame for the E2F6 protein begins in the first exon and encodes a protein of 272 amino acids with a molecular mass of approximately 30.8 kDa. Exon 8 contained the termination codon followed by 1387 bp of 3′-untranslated sequence containing a polyadenylation signal AAATAA situated 40 nucleotides upstream from the poly(A) tail (Fig. 2). The alternatively spliced isoform of E2F6 contained an additional exon between exon 1 and 3; this additional exon introduces an in-frame termination codon. In addition, the last three nucleotides of exon 2 introduced a new initiation codon on the same frame (Fig. 1B). A translational initiation with this putative second ATG, which does not perfectly match with a consensus Kozak sequence (16), could result in a N-terminal truncated E2F6 protein of 237 residues with all known functional subdomains of the molecule. The amino acid sequence of ΔNt-E2F6 predicts a protein of approximately 26.8 kDa.

Transcriptional Start Site Determination

To identify the transcription start site, we used several complementary strategies, including 5′-RACE, RNase protection assays and cDNA cloning.

A screening was carried out on a mouse ovary cDNA library using the 130-bp 5′-E2F6 probe (see Materials and Methods). Among the cDNAs isolated and size-characterized, the longest clone in 5′ was subcloned and sequenced. The start site of this clone was reported as a solid star on Fig. 4. To obtain further 5′ sequences, 5′-RACE experiments were performed on a mouse heart Cap site cDNA (Eurogentec). In our first experiments, we used as the primary E2F6 antisense primer the 7F6, which is located 230 bp downstream the initiation codon, followed by a secondary E2F6 nesting primer 8F6 located 110 bp downstream the ATG. In a second experiment the primary E2F6 antisense primer was 9F6 located on the ATG, followed by 2F6 located 110 bp upstream from the ATG. The RACE products obtained using this commercial kit were all shorter than our isolated cDNA and corresponded roughly to the previously described E2F6 cDNA (AF032131). We failed to obtain longer 5′-sequences. Moreover comparison of the 5′ flanking region of E2F6 with sequences deposited in EST database revealed identity with the mouse GenBank EST BE333479. This unpublished EST is 458 bp length and ends 7 bp upstream our isolated cDNA. The start site of this clone was reported as an open star on Fig. 4.

To determine whether this region represented the transcriptional initiation site or a strong premature retrotranscription arrest during the synthesis of the cDNAs, we performed RNase protection assays. To this aim a 2200-bp genomic 5′ fragment upstream from the

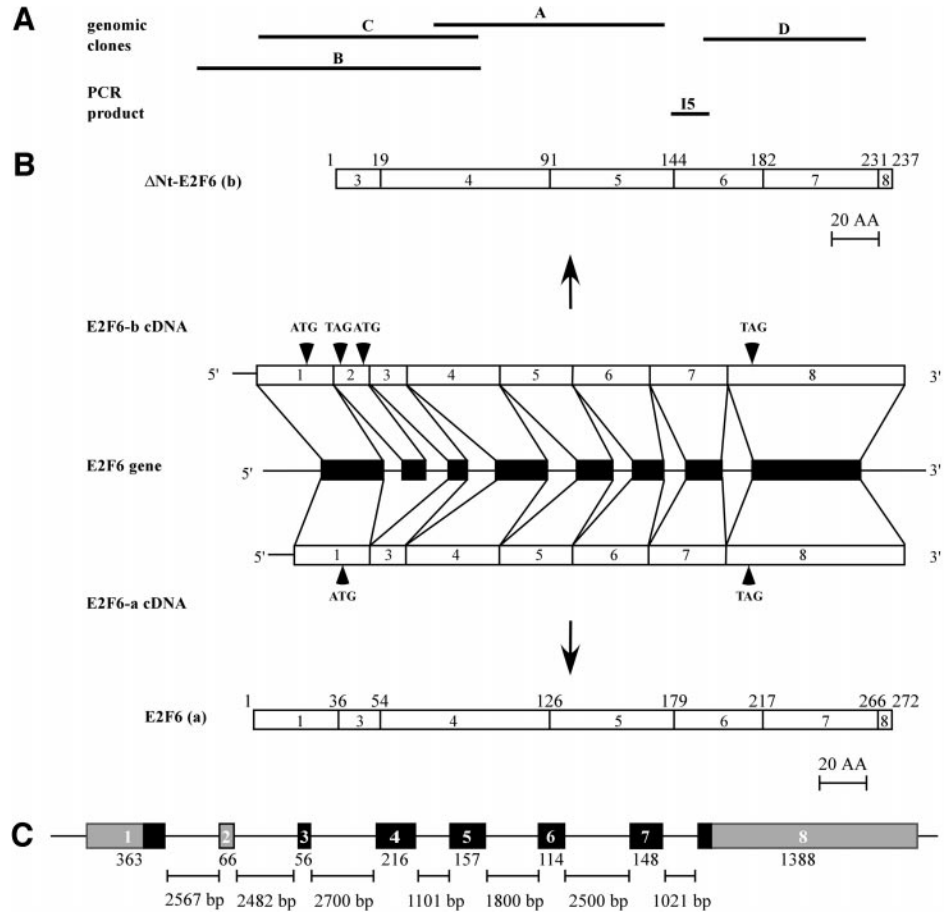

FIG. 1. Genomic structure of the mouse E2F6 gene. (A) Schematic diagram of the four isolated genomic phage clones encompassing the mouse E2F6 gene and the PCR product encompassing the fifth intron (I5). (B) The mouse E2F6 genomic structure. Exons 1 to 8 are represented by solid boxes, and the introns in between are represented by lines. The full size and the splice variant E2F6 cDNAs are represented at the top and at the bottom, respectively. Arrowheads indicate positions of the initiation and termination codons. The potential protein products represented with amino acids numbers are indicated above. The size of exon 1 is numbered from the transcription start site (see below). (C) Physical map of the gene. The open reading frame of E2F6 is indicated by black shading and the untranslated region by gray shading. The lengths of exons and introns are indicated.

ATG was subcloned into the PCRII-Topo vector. The *Bgl*II fragment of this clone containing 290 bp of E2F6 5′ sequence upstream from the ATG, was *in vitro* transcribed using T7 RNA polymerase in the presence of [α-^{32}P]CTP to generate the 5′-*Bgl*II probe. Hybridization of this riboprobe to total mouse embryonic RNA followed ribonucleases A and T1 digestion, allowed us to detect two protected fragments corresponding to 240 and 255 nucleotides (Fig. 3, lane 7). By contrast, these protected fragments were not observed when yeast RNAs were submitted to the same experimental conditions (lane 6). We noticed the presence of a larger fragment (~370 bp) in both the sample and the control lane (lane 6 and 7), probably due to nondigested probe. To confirm our result a second experiment was performed on poly(A)$^+$ RNA extracted from mouse embryonic fibroblasts (MEF). Two longest 5′-E2F6 riboprobes were generated using *Bam*HI and *Afl*III digestions of our plasmid clone. After hybridization of these riboprobes to cellular RNA, digestion with 1/80 and 1/50 dilution of the mixed ribonucleases A and T1 were used to avoid undigested fragments. The size of the protected fragments corresponded well with our first experiment. These sizes were reported on the sequence as solid arrows (Fig. 4).

The size of the longest protected fragment places the transcription start site 34 bp upstream of our longest cDNA clone and 27 bp upstream the GenBank EST BE333479. We have thus fixed the start sites of E2F6 gene 256 and 241 bp upstream from the ATG.

Promoter Analysis

As shown in Fig. 4, the sequence surrounding the putative transcription start sites does not contain a TATA nor a CAAT box. This region is relatively GC rich (70%) and contains several potential transcription factors binding sites (TESS program, TRANSFAC database). Among them, we have indicated sites implicated in cell cycle regulation, such as Myc and E2F

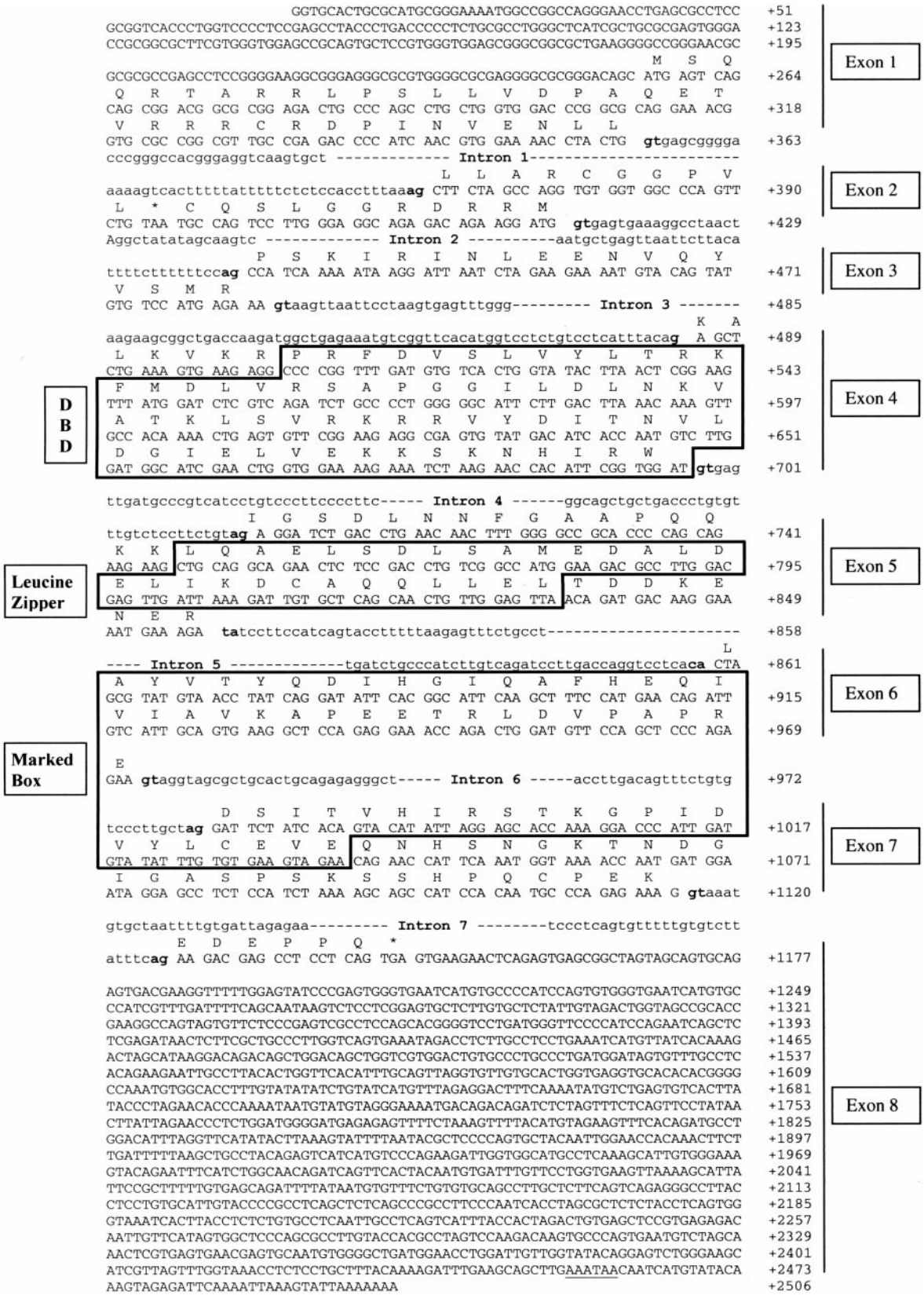

FIG. 2. Genomic sequence of E2F6 exons in uppercase and partial intron sequence in lowercase letters. Index on the right corresponds to the E2F6 cDNA sequence. Single-letter codes show the deduced E2F6 peptide sequence. The polyadenylation signal is underlined. Domains responsible for DNA binding and dimerization (leucine zipper, marked box) are boxed.

TABLE 1

Splice Donor and Acceptor Sequences of Mouse E2F6

Exon	Size (bp)	Splice acceptor	Splice donor	Intron size (bp)
1	363		AACCTACTG**gt**gagcggggacccgggccacgg	2567[a]
2	66	ttatttttctctccacctttaa**ag** CTTCTAGCC	AGAAGGATG**gt**gagtgaaaggcctaactaggc	2482[a]
3	56	ttcttacattcttctcttcccc**ag** CCATCAAAA	CATGAGAAA**gt**aagttaattcctaagtgagtt	2700[b]
4	216	atggtcctctgtcctcatttac**ag** AGCTCTGAA	TAGGTGGAT**gt**gagttgatgcccgtcatcctg	1101[a]
5	157	ccctgtgtttgtctccttctgt**ag** AGGATCTGA	AATGAAAGA**ta**tccttccatcagtacctttt	1800[b]
6	114	tcagatccttgaccaggtctca**ca** CTAGCGTAT	CCCAGAGAA**gt**aggtagcgctgcactgcagag	2500[b]
7	148	gacagtttctgtgtcccttgct**ag** GATTCTATC	CAGAGAAAG**gt**aaatgtgctaattttgtgatt	1021[a]
8	1386	cagtgtttttgtgtcttatttc**ag** AAGACGAGC		

Note. For each intron–exon boundary, exon (uppercase letters) and intron (lowercase letters) sequences are shown.

[a] The size of this intron was determined by sequencing.

[b] The size of this intron was estimated by restriction enzyme analysis and confirmed by PCR.

binding sites. At present, it is not clear whether any of these binding sites are important for the promoter activity.

To determine whether bona fide start sites of transcription have been detected, fragments surrounding this genomic region were fused to the luciferase gene into the promoter-less reporter plasmid pGL3-basic (Promega). A 2000-bp (−1944/+55) fragment of the 5′ flanking sequence of the E2F6 gene was tested for its ability to act as a functional promoter in transient transfection assays in HepG2 and RK13 cell lines (Fig. 5). This construct exhibited an approximately 15- and 8-fold increase in luciferase expression above the promoter-less vector in HepG2 and RK13 cells respectively. To narrow the minimal promoter region, deletion constructs were generated. These constructs displayed a higher activity in the HepG2 cell line, but followed a similar profile in RK13 cells. Deletion of sequences 5′ to −1360 did not affect the luciferase activities compared with the 2000-bp fragment. By contrast, the construct deleted of fragment (−1944/−333) initiated transcription of the reporter gene more efficiently than the extended fragment by 36- and 18-fold in HepG2 and RK13 cells respectively, indicating that potential repressor elements may reside further upstream of the basal promoter. However, more detailed studies will be required to identify the components involved in regulating transcription of the E2F6 gene. Deletion at nucleotides −75 resulted in activity similar to that of the −333-bp fragment, implying that the minimal sequence necessary for basal transcription of the E2F6 gene in HepG2 and RK13 cells maps to approximately −75 bp/+55. Accordingly the −333/−75 fragment was inactive (data not shown) and no activity was detected upon transfection of the fragments in an antisense orientation (Fig. 5).

Expression of Mouse E2F6 mRNAs

To provide clues about the gene regulation and potential function of E2F6 *in vivo,* we first investigated the distribution of E2F6 in adult tissues by Northern blot analyses. The full-length E2F6 coding sequence was used as probe. The results presented in Fig. 6A revealed the detection of one mRNA transcript of approximately 2.8 kb. All tissues analyzed expressed this transcript, but the highest level was detected in heart, muscle and brain of adult mice, according to the relatively low amounts of RNA loaded for these tissues. Lower signals can be detected in the other tissues tested (liver, lung, intestine, colon, kidney, testis, spleen, thymus, and ovary).

As a second step toward understanding developmental regulation of E2F6, we determined whether there was a specific developmental period in the mouse during which E2F6 was more importantly expressed. Analysis of total tissue RNA of staged mouse embryos demonstrated that E2F6 was expressed at high level from embryonic day 9 to embryonic day 18. At 11 days postcoitum, the strongest expression levels were detected. This was followed by a downregulation and then the level of E2F6 did not significantly change in the period from day 12 to day 18, when normalized with the ARPP-P0 control (Fig. 6B).

To investigate the cell cycle regulation of E2F6, mouse embryonic primary fibroblasts (MEF) at p3 were brought to quiescence by serum starvation for 3 days and stimulated to grow by the addition of fresh medium with 10% serum. Samples were then collected at successive 3-h intervals and RNA was prepared and analyzed by Northern blotting (Fig. 6C). The cell cycle progression upon serum addition, as indicated at the bottom of Fig. 6C, was determined by flow cytometry. As previously described, E2F1, Cyclin E and Cyclin A expression patterns were closely related to each others as a consequence of their common regulation via E2F binding sites present in their promoters. In contrast, E2F4 and E2F6 showed different expression profiles indicating that these genes are probably not regulated by E2F species during MEF cycling. The expression profile of E2F4 was in agreement with that observed in

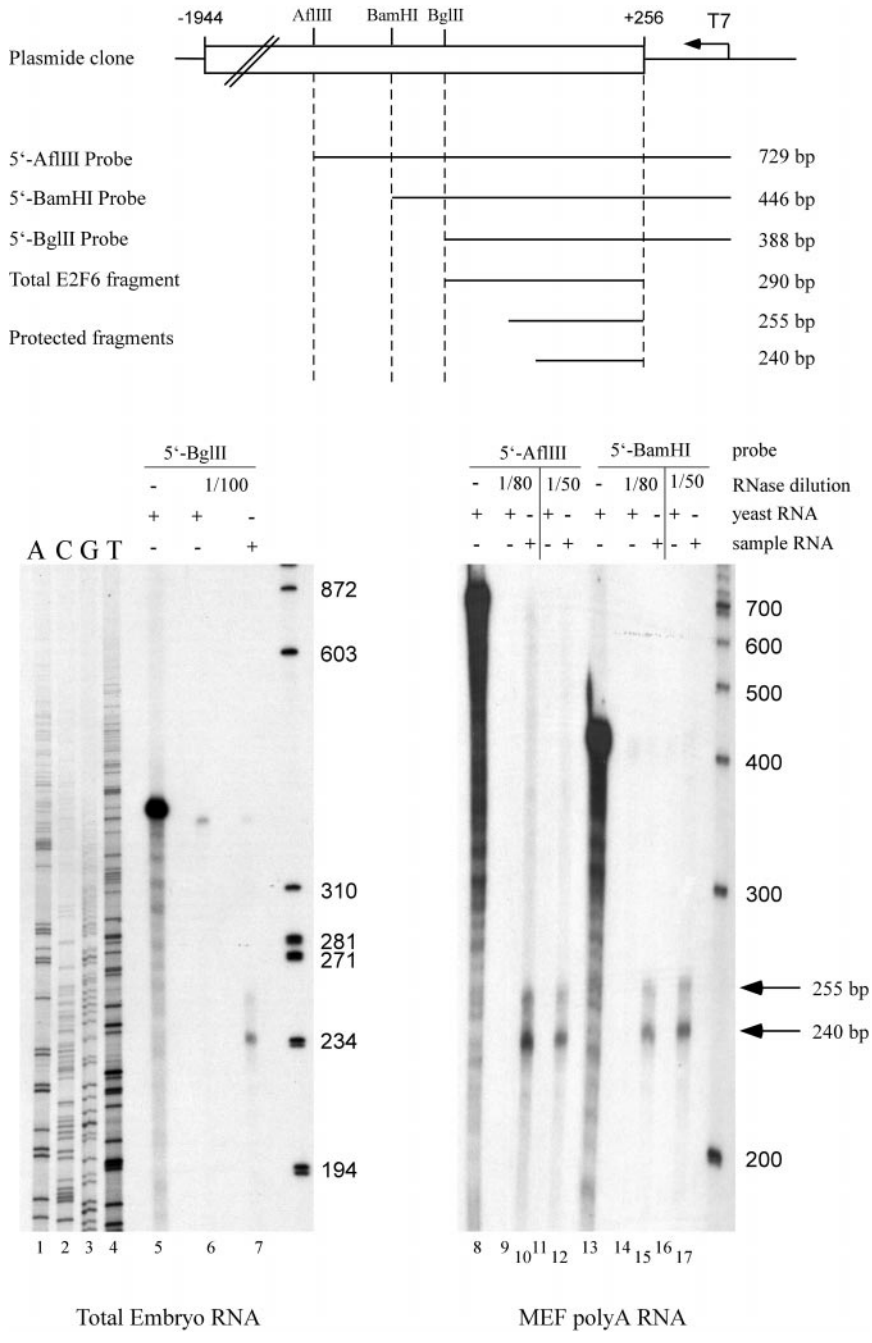

FIG. 3. Analysis of E2F6 transcription start site by RNase protection assay. (A) Schematic representation of the probes used for RNase protection analyses. Antisense RNA probes were transcribed from the TA-E2F6 promo construct. (B) RNase protection analysis. (Left) RNase protection of the *Bgl*II probe on total mouse embryonic RNA. (Right) Assays were performed using the *Bam*HI probe and *Afl*III-probe on MEF poly(A)$^+$ RNA. Integrity of probes is shown in lanes 5, 8, and 13. Negatives controls containing yeast tRNA are shown in lanes 6, 9, 11, 14, and 16. The size of the protected fragments (lanes 7, 10, 12, 15, and 17) are indicated by arrows. Lanes 1–4, sequencing reactions were run as size marker.

human keratinocytes but differed slightly from that observed in NIH-3T3 and REF-52. E2F6 show a unique expression pattern during MEF cycling. Quiescent fibroblasts contained low amounts of the E2F6 mRNA. Upon serum stimulation, E2F6 mRNA increased rapidly, reaching its higher level at mid-G1 at 9 h and remaining relatively constant thereafter.

We next examined whether E2F6 alternatively spliced transcripts are normally expressed in mouse tissues. To analyze the expression of the E2F6-a and E2F6-b isoforms, we designed PCR primers to set up a semiquantitative RT-PCR assay for separately detecting the isoforms as amplified bands of different sizes. Primers (3F6 and 10F6) were positioned so as to encompass introns 1

FIG. 4. Nucleotide sequence of the 5′-flanking region of the E2F6 gene. Index on the left referred to the +1 transcription start site based on the largest protected product from the RNase protection analysis. Arrowheads indicate transcription start sites determined by RNase protection assays. The open star indicates the 5′ end of the BE333479 EST, while the solid star indicates the 5′ end of the longest E2F6 cDNA obtained by library screening. The transcribed sequence is in bold. Several potential cis-acting elements determined by computer analysis are underlined. Restriction cloning sites are boxed.

and 2 in order to avoid amplification of potentially contaminating genomic DNA and the amplified fragments showed two sizes which could easily be detected on an agarose gel electrophoresis. As shown in Fig. 6D, the two isoforms of E2F6 were expressed in all mouse tissues tested, but we noticed a predominant expression of full-length E2F6-b mRNA compared to E2F6-a splice variant. This was particularly obvious in testis RNA.

During mouse development or cell cycle progression, a similar dynamic expression profile was observed for

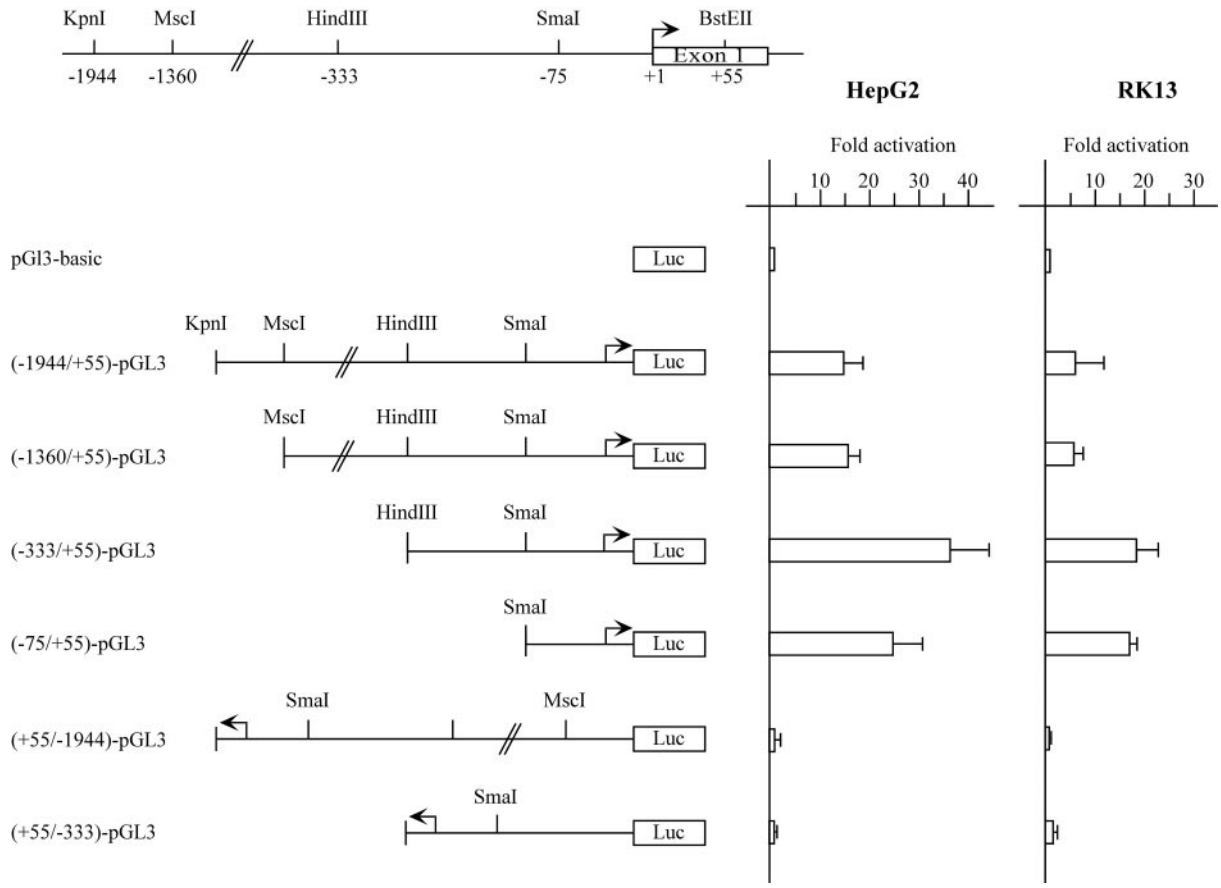

FIG. 5. Schematic diagram of E2F6 constructs in pGL3-basic vector tested for promoter activity. Promoter constructs were tested in HepG2 and RK13 cell lines. The transcriptional activity was expressed as a percentage of the luciferase activity observed for pGL3-basic. Data are means ± SE of at least three independent experiments.

both E2F6-a and E2F6-b transcripts, corresponding to those previously described, with always the splice variant significantly down-regulated (data not shown). Together these data indicate prominent expression of E2F6-b transcripts in both adult mouse tissues, during embryogenesis and cell cycle progression. Further investigations would be necessary to determine whether changes in E2F6 transcripts levels correlate with distinct functional properties.

DISCUSSION

We have reported here the cloning and sequencing of a large segment of the murine E2F6 gene as well as its 5′-flanking region. The gene that has been previously localized to chromosome 12 (17), spans over 18 kb of DNA and contains eight exons. The size of the introns was relatively small. As previously described for human E2F1 (18), structural domains of murine E2F6 are not perfectly delineated by separate exons. This organization is also found for human E2F2 and E2F3 (unpublished results). Conversely, E2F4 and E2F5 genomic organization that are very close to each other, are totally different (19). The murine E2F6 gene expresses two alternative splice variants. The first isoform, which correspond to the previously described E2F6 mRNA, was devoid of exon 2 (8). These two mRNA isoforms are co-expressed in a wide variety of tissues and the expression level of exon 2 containing mRNA is always higher. At the present time, the significance of the exon 2 containing second isoform is not clear and will be further investigated.

We have also cloned the promoter region of the murine E2F6 gene and performed an initial characterization of the features that support basal transcription. Two major transcription start sites were identified by RNase protection. These start sites are in agreement with the RACE products obtained as well as with the GenBank ESTs. An interesting finding is the observation that the murine E2F6 promoter belongs to the family of TATA-less promoters. There is also no CAAT box upstream of the transcriptional start site and this sequence is relatively GC-rich (70%). Basal promoter function for E2F6 was demonstrated by comparison of relative luciferase activities of specific constructs in various cells. As indicated by our results, the 75-bp

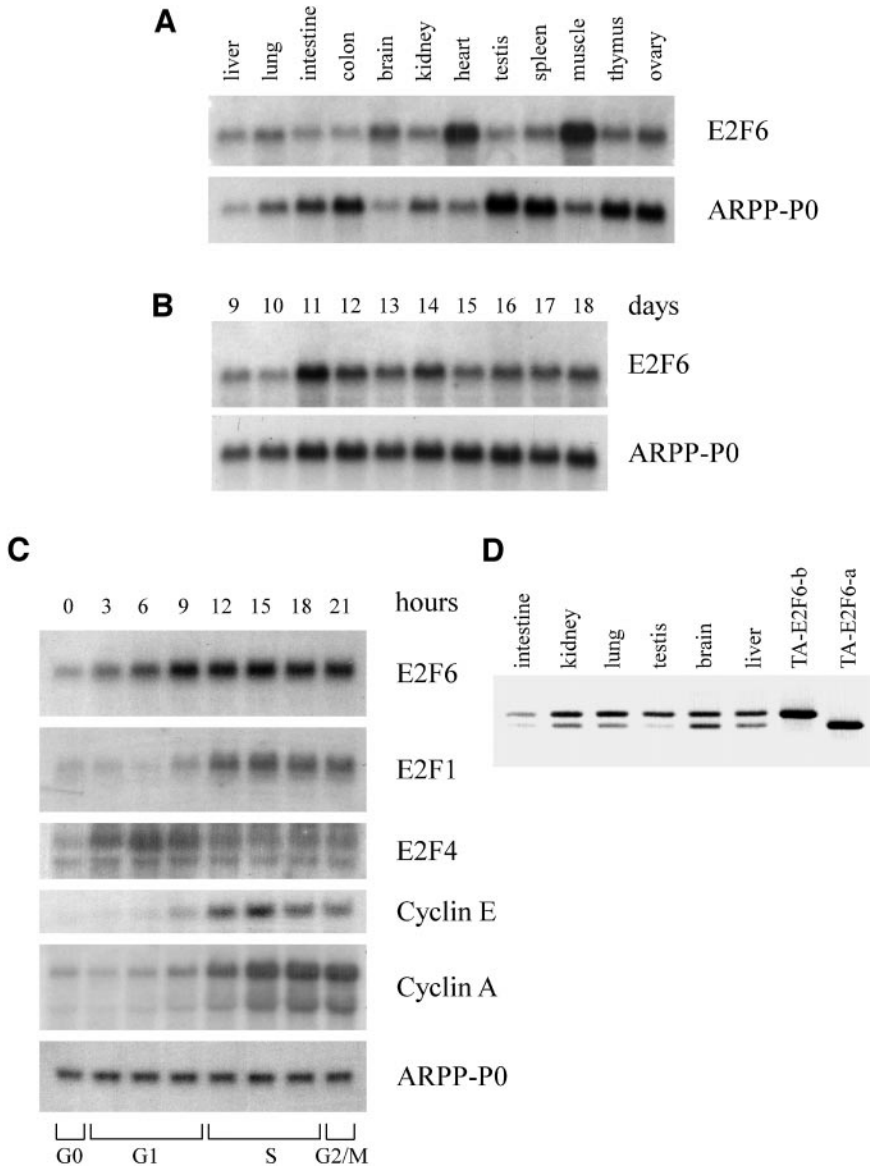

FIG. 6. Expression patterns of E2F6. (A) Northern blot analysis of total RNA isolated from indicated mouse tissues, screened with the full-length E2F6 coding sequence probe. As a control for the quality of the RNA, the blots were rehybridized with a probe that is specific for ARPP-P0 sequence. (B) Expression of E2F6 during mouse embryogenesis. Total RNA prepared from various stages of mouse embryonic (E9 to E18) development. As a control for the quality of the RNA, the blots were rehybridized with a probe that is specific for ARPP-P0 sequences. (C) Expression of E2F6 following stimulation of cell growth. MEF cells were brought to quiescence by serum starvation and then stimulated by the addition of serum. Cells were harvested at the indicated times after serum addition and total RNA was prepared from each time point. The transition from G1 to S phase, as indicated at the bottom of the figure, was determined by flow cytometry. For comparison the same blots were probed for expression of the E2F1, E2F4, cyclin A, cyclin E, as well as ARPP-P0 as a loading control. (D) Expression of E2F6 isoforms in adult mouse tissues. RT-PCR products were amplified from total RNAs of multiple mouse tissues using an oligo(dT) primer. PCR amplifications of E2F6 products were carried out with 3F6 and 10F6 primers. Amplifications on TA-E2F6-a and TA-E2F6-b constructs were realized as size control.

segment upstream from the E2F6 transcription start site drives the highest levels of basal expression. This segment contains potential consensus binding sites for several transcription factors (v-myb, Sp1, and USF). Interestingly, we found also one potential E2F consensus binding site in the 5′-untranslated region of the E2F6 gene. Despite the fact that E2F1-3 were able to enhance the luciferase activity of different E2F6 promoter constructs, we did not detect any variation of E2F6 mRNA level upon infection of various cell-lines with recombinant adenovirus expressing E2F1 or E2F2 (data not shown). These results are in agreement with those recently obtained in human by Helin's group, which used ER-E2F chimeric construct (20).

Currently, we have no explanation for this discrepancy between transfection and infection experiments. In this context, it has been previously reported that several genes described as being regulated by the E2Fs were only mildly affected or not affected at all by (direct) E2F activity (21, 22). There are numerous possible reasons for the lack of an E2F effect on these genes, but one consideration is of particular interest. The experiments demonstrating the functional role of E2F consensus binding site in these genes all involved transient expression experiments that take the promoter out of its chromosomal context. Since the accessibility of DNA binding sites for transcription factors is, to a large degree, regulated by the chromatin structure, it is likely that the untimely expression of the E2F proteins does not allow binding to all potential E2F binding elements. Moreover, despite the presence of an E2F DNA binding consensus site, this site may not be occupied due to interference by, for instance, other transcription factors or nucleosomes. However, we can not exclude that the potential E2F sites present in the E2F6 gene may only be functional in particular conditions or cell types. Moreover, the E2F6 mRNA expression pattern did not follow that of E2F1-2–3a nor that of E2F4–5 when MEF exit from quiescence upon serum addition. Thus, E2F6 shows a unique pattern of expression in synchronized MEF. Its expression is maximal as soon as mid-G1 phase and remains relatively high thereafter, in particular during S phase. This pattern of expression leads us to suppose that E2F6 may contribute to the regulation of events throughout the cell-cycle, and particularly during S phase. In this context, ectopic expression of E2F6 in U-2OS cells leads to accumulation of cells in S phase probably by delaying the exit from S phase (11).

The broad expression pattern of E2F6 suggests its important contribution to the regulation of E2F activity *in vivo*. Moreover, its high level of expression during embryogenesis and its recent implication in polycomb mediated repression permit to suppose that E2F6 is important for developmental patterning. The interaction of human E2F6 with polycomb complexes occurred via its marked-box domain (12). These results are in agreement with the mapping by Gaubatz *et al.* of the human E2F6 repression domain to its C-terminal portion that encompasses the marked-box domain (10). Mouse and human E2F6 marked-box domains are very close to each other. So, it is possible to consider that polycomb complexes are also involved in murine E2F6 repression activity via this domain. Unfortunately, murine E2F6 repressive activity has been mapped to its N-terminal part that does not encompass the marked-box domain (8). Altogether, these results indicate that repressive activity of E2F6 could be supported by different domains depending on the cell type, the promoter context, or the cellular process. For example, the E2F6 repressive activity via polycomb complexes would be important during development and the E2F6 N-terminal mediated repression would be important for cell-cycle regulation in adult tissues. Moreover, depending on the cell-type and/or promoter, E2F6 does not only act as an active repressor but also as an E2F competitor. This point permit us to suppose that in some cellular contexts E2F6 will also be able to enhance promoter activity by displacing E2F-pocket proteins repressor complexes. Further works will be necessary to address these different points. In particular functional analysis of the implication of polycomb complexes in E2F6 repressive activity and phenotypic analysis of E2F6-deficient mice will be very important to better understand E2F6 functions.

ACKNOWLEDGMENTS

The genomic E2F6 sequences have been deposited in the GenBank database under Accession Nos. AF393244, AF393245, AF393246, AF393247, AF393248, and AF393249. We thank J. Coll, L. Coutte, J. L. Baert, and Y. de Launoit for critical reading of the manuscript and B. Quatannens for FACS analysis. This study was supported by grants from the Centre National de la Recherche Scientifique (CNRS), the Institut Pasteur de Lille (IPL), and the Association pour la Recherche sur le Cancer (ARC).

REFERENCES

1. Slansky, J. E., and Farnham, P. J. (1996) Introduction to the E2F family: Protein structure and gene regulation. *Curr. Top. Microbiol. Immunol.* **208,** 1–30.
2. Black, A. R., and Azizkhan-Clifford, J. (1999) Regulation of E2F: A family of transcription factors involved in proliferation control. *Gene* **237,** 281–302.
3. Lavia, P., and Jansen-Durr, P. (1999) E2F target genes and cell-cycle checkpoint control. *Bioessays* **21,** 221–230.
4. Leone, G., Nuckolls, F., Ishida, S., Adams, M., Sears, R., Jakoi, L., Miron, A., and Nevins, J. R. (2000) Identification of a novel E2F3 product suggests a mechanism for determining specificity of repression by Rb proteins. *Mol. Cell. Biol.* **20,** 3626–3632.
5. Sardet, C., Vidal, M., Cobrinik, D., Geng, Y., Onufryk, C., Chen, A., and Weinberg, R. A. (1995) E2F-4 and E2F-5, two members of the E2F family, are expressed in the early phases of the cell cycle. *Proc. Natl. Acad. Sci. USA* **92,** 2403–2407.
6. He, Y., Armanious, M. K., Thomas, M. J., and Cress, W. D. (2000) Identification of E2F-3B, an alternative form of E2F-3 lacking a conserved N-terminal region. *Oncogene* **19,** 3422–3433.
7. Dyson, N. (1998) The regulation of E2F by pRB-family proteins. *Genes Dev.* **12,** 2245–2262.
8. Morkel, M., Wenkel, J., Bannister, A. J., Kouzarides, T., and Hagemeier, C. (1997) An E2F-like repressor of transcription. *Nature* **390,** 567–568.
9. Trimarchi, J. M., Fairchild, B., Verona, R., Moberg, K., Andon, N., and Lees, J. A. (1998) E2F-6, a member of the E2F family that can behave as a transcriptional repressor. *Proc. Natl. Acad. Sci. USA* **95,** 2850–2855.
10. Gaubatz, S., Wood, J. G., and Livingston, D. M. (1998) Unusual proliferation arrest and transcriptional control properties of a newly discovered E2F family member, E2F-6. *Proc. Natl. Acad. Sci. USA* **95,** 9190–9195.

11. Cartwright, P., Muller, H., Wagener, C., Holm, K., and Helin, K. (1998) E2F-6: A novel member of the E2F family is an inhibitor of E2F- dependent transcription. *Oncogene* **17,** 611–623.
12. Trimarchi, J. M., Fairchild, B., Wen, J., and Lees, J. A. (2001) The E2F6 transcription factor is a component of the mammalian Bmi1-containing polycomb complex. *Proc. Natl. Acad. Sci. USA* **98,** 1519–1524.
13. Palena, A., Mangiacasale, R., Magnano, A. R., Barberi, L., Giordano, R., Spadafora, C., and Lavia, P. (2000) E2F transcription factors are differentially expressed in murine gametes and early embryos. *Mech. Dev.* **97,** 211–215.
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. Burset, M., Seledtsov, I. A., and Solovyev, V. V. (2000) Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res.* **28,** 4364–4375.
16. Kozak, M. (1995) Adherence to the first-AUG rule when a second AUG codon follows closely upon the first. *Proc. Natl. Acad. Sci. USA* **92,** 2662–2666.
17. Peterfy, M., Phan, J., Oswell, G. M., Xu, P., and Reue, K. (1999) Genetic, physical, and transcript map of the fld region on mouse chromosome 12. *Genomics* **62,** 436–444.
18. Neuman, E., Sellers, W. R., McNeil, J. A., Lawrence, J. B., and Kaelin, W. G., Jr. (1996) Structure and partial genomic sequence of the human E2F1 gene. *Gene* **173,** 163–169.
19. Schwemmle, S., and Pfeifer, G. P. (2000) Genomic structure and mutation screening of the E2F4 gene in human tumors. *Int. J. Cancer* **86,** 672–677.
20. Muller, H., Bracken, A. P., Vernell, R., Moroni, M. C., Christians, F., Grassilli, E., Prosperini, E., Vigo, E., Oliner, J. D., and Helin, K. (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev.* **15,** 267–285.
21. Vigo, E., Muller, H., Prosperini, E., Hateboer, G., Cartwright, P., Moroni, M. C., and Helin, K. (1999) CDC25A phosphatase is a target of E2F and is required for efficient E2F-induced S phase. *Mol. Cell. Biol.* **19,** 6379–6395.
22. Helin, K. (1998) Regulation of cell proliferation by the E2F transcription factors. *Curr. Opin. Genet. Dev.* **8,** 28–35.