Cloning and characterization of the rat *apobec-1* gene: a comparative analysis of gene structure and promoter usage in rat and mouse[1]

Ken-Ichi Hirano, Jing Min, Toru Funahashi, and Nicholas O. Davidson[2]

Department of Medicine, University of Chicago, 5841 S. Maryland Avenue, Chicago, IL 60637

Abstract ApolipoproteinB (apoB) mRNA editing involves a C to U deamination of the nuclear apoB mRNA and occurs in mammalian small intestine and in the liver of certain species. This reaction is mediated by a multicomponent enzyme complex that includes a catalytic subunit, apobec-1. Apobec-1 mRNA is widely expressed in the rat and mouse and is subject to tissue-specific regulation. In order to understand the basis for the species- and tissue-specific pattern of *apobec-1* gene expression we have cloned and characterized the rat chromosomal *apobec-1* gene. We demonstrate its structural organization and regulation in comparison to that of the mouse *apobec-1* gene. The rat *apobec-1* gene spans 16 kb and includes one untranslated (exon A) and five translated exons (exons 1–5). The mouse *apobec-1* gene contains eight exons, of which the first three (exons A, B, C) are untranslated. Independent approaches demonstrated three distinct clusters of transcription initiation sites in both species, including exon A, the distal region of exon 1, and a separate group in the proximal region of exon 1. These transcription start sites generate three distinct mRNA species whose proportions differ in a tissue-specific fashion. Promoter-luciferase reporter constructions using regions flanking exon A and exon 1 of the rat *apobec-1* gene identified two functional regions upstream of exon 1 that independently promote luciferase expression in transfected hepatoma and colon cancer cells. These data serve as a basis for an understanding of the regulation of apobec-1 gene expression, in particular the mechanisms that serve to restrict its expression to the gastrointestinal tract in higher mammals.—**Hirano, K-I., J. Min, T. Funahashi, and N. O. Davidson.** Cloning and characterization of the rat *apobec-1* gene: a comparative analysis of gene structure and promoter usage in rat and mouse. *J. Lipid Res.* 1997. **38:** 1103–1119.

Supplementary key words RNA editing • apolipoprotein B • hepatic protein section

ApolipoproteinB (apoB) is an abundantly expressed gene product, transcribed in mammalian liver and small intestine, where it serves an indispensable role in the assembly and secretion of triglyceride-rich lipoproteins (reviewed in ref. 1). Among the mechanisms demonstrated to be of importance in the regulation of apoB gene expression is the tissue-specific control of post-transcriptional RNA editing, in which a single C to U change alters a glutamine (CAA) codon and generates a translational UAA stop (reviewed in refs. 2, 3). This reaction is responsible for the production of two distinct isomorphs of apoB from a single nuclear gene, thereby expanding the genetic repertoire of the organism.

ApoB mRNA editing occurs in the mammalian small intestine and in the liver of certain species, including the rat and mouse, but not others, notably humans (2–4). The production of a truncated species of apoB, (apoB-48), has important implications for the metabolism of apoB-containing lipoproteins. ApoB-100, the full length form of the protein, contains 4536 residues and functions as a ligand for the low density lipoprotein (LDL) receptor by virtue of a region of basic amino acids in the carboxyl terminus (1). ApoB-48, by contrast, contains the amino-terminal 2152 residues and lacks the LDL receptor-binding region (1–3). The catabolism of apoB-48-containing lipoproteins (chylomicrons and chylomicron remnants) thus differs from that of LDL and involves a receptor-dependent process, most plausibly the low density lipoprotein-related protein (LRP) whose activity in the uptake of apoB-48-containing particles reflects the content of apolipoprotein E (5). A further biological distinction between the two

Abbreviations: apobec-1, apolipoprotein B mRNA editing enzyme, catalytic polypeptide #1; RT-PCR, reverse transcription-polymerase chain reaction; apoB, apolipoprotein B; RACE, rapid amplification of cDNA ends.

[1]Genbank accession numbers for the sequence information contained in this report are as follows: Rat exon A and its 5′ flanking region, U72885; Rat intron A and exon 1, U72886; Mouse exon A and its 5′ flanking region, U72887; Mouse exon 1 and its 5′ flanking region, U72888.

[2]To whom correspondence should be addressed.

different isomorphs of apoB concerns their relative propensity to promote the formation of a highly atherogenic lipoprotein particle referred to as Lp[a]. ApoB-100 is absolutely required for the assembly of Lp[a], a process involving a single unpaired cysteine residue in the extreme carboxyl terminus of apoB (6, 7). By corollary, apoB-48, which lacks this particular region of the protein, is incapable of assembling into a covalent linkage with apo[a], a functional limitation that precludes Lp[a] formation (1–3).

ApoB mRNA editing is mediated by a multicomponent enzyme complex which has been referred to as an editosome (8). This complex contains a catalytic subunit apobec-1, a 27-kDa protein that demonstrates an absolute requirement for additional, complementation factor(s) in order to mediate apoB RNA editing (9). Apobec-1 is an RNA-specific cytidine deaminase with homology to other members of the cytidine/deoxycytidine deaminase family, particularly within the domain HVE–PCXXC proposed to coordinate zinc binding and catalysis (10, 11). Several studies have established the importance of these and other residues in terms of the RNA editing activity of apobec-1 (11–15). In addition, apobec-1 was determined to have apoB RNA binding activity as well as the capacity to function as a cytidine deaminase on a monomeric nucleoside substrate (11, 14, 15). Apobec-1 is expressed in virtually all tissues of the rat and mouse, including such sites as the kidney and spleen which contain trivial amounts of apoB mRNA (16, 17). Apobec-1 distribution in humans and rabbits, by contrast, appears to be confined to the gastrointestinal tract, principally the small intestinal enterocyte (12, 18, 19). At present, virtually nothing is known about the mechanisms that regulate *apobec-1* gene expression in these species.

Several studies have demonstrated that apoB mRNA editing is regulated in the rat and mouse liver in response to a variety of developmental, hormonal, nutritional, and other cues (20–27). Some of these regulatory events are linked to alterations in the abundance of apobec-1 mRNA (16). For instance, the changes in rat hepatic apoB mRNA editing observed in response to fasting and refeeding a high-carbohydrate diet were shown to be accompanied by a 3- to 5-fold induction of apobec-1 mRNA abundance (16). By contrast, the increases noted in apoB mRNA editing, after administration of thyroid hormone and in response to ethanol feeding, were not accompanied by changes in apobec-1 mRNA abundance (16, 26). Studies of the developmental regulation of apoB mRNA editing in the rat liver and small intestine suggested an intermediate response, demonstrating a general trend towards increasing apobec-1 mRNA abundance with the increase in apoB mRNA editing, although no strict temporal concordance was demonstrated (16). Taken together, the collective information emerging from studies in rats suggests that tissue-specific control of apoB mRNA editing can be exerted through alterations in the composition of the editing enzyme complex, an effect that can be mediated either through changes in apobec-1 abundance or the activity of complementation factors (27). While little information is available concerning *apobec-1* gene regulation in mice, gene targeting has demonstrated that the *apobec-1* gene product is absolutely required for apoB mRNA editing, there being no functional duplication in any tissue examined (28–30). The regulation of *apobec-1* gene expression in the rat and mouse thus serves as a focal point for an understanding of some important elements of the tissue-specific control of apoB mRNA editing.

In order to pursue the regulation of *apobec-1* gene expression, we have characterized genomic clones spanning both the rat and mouse *apobec-1* genes. The mouse *apobec-1* gene contains 8 exons, of which the first 3 (exons A, B, C) are untranslated. The rat *apobec-1* gene contains 6 exons, of which only the first (exon A) is untranslated. We have determined the presence of multiple transcription start sites in both species, including one within exon A and two distinct groups of start sites within exon 1. We also demonstrate a complex, tissue-specific pattern of alternative promoter usage in both the rat and mouse *apobec-1* genes and, finally, demonstrate in vitro promoter activity using 5′ flanking sequences from the rat gene to localize distinct regions that promote expression in intestine-derived cells.

MATERIALS AND METHODS

Isolation of genomic clones

A rat genomic library, prepared from adult female Sprague Dawley rat liver, was purchased from Clontech (cat #RL 1005A). Approximately three million recombinants were screened at high stringency using a combination of rat apobec-1 cDNA (a *Sal I/Xba*I fragment, spanning nucleotides $+1 \rightarrow +833$) and a mouse *apobec-1* genomic fragment obtained by PCR using primers P5 and P6, listed below, which span a region of the untranslated exon and its 5′ flanking region, from nucleotides $-72 \rightarrow +161$. Three overlapping clones were obtained that were characterized by Southern blotting and restriction mapping. In addition, primer pairs MEPR 1F and MEPR 2R were used to amplify a genomic fragment of 220 bp from mouse DNA and these primers

were sent to Genome Systems, St. Louis, MO, for the purposes of obtaining a mouse ES C129 P1 clone, which was characterized by restriction digestion and Southern blotting. The application of this clone in the successful gene targeting of mouse *apobec-1* has been recently reported (28). Sequence comparison was undertaken using the Genetics Computer Group (GCG) software.

Amplification of the 5′ end of the rat and mouse apobec-1 cDNA (5′ RACE)

First strand cDNA synthesis was undertaken using 1 μg total RNA from mouse and rat small intestine and liver with oligo dT priming, using the following incubation conditions, which represent minor modifications of the recommendations provided with the Marathon cDNA amplification kit (cat # K1802-1, Clontech): reverse transcription was performed at 42°C for 60 min in 20 μL (final volume) buffer containing 6 mM $MgCl_2$, 1 mM dNTP, 1 mM oligo dT, and 50 units Superscript reverse transcriptase (Life Technologies). The reaction mixture was then transferred to a tube containing 24 units *E. coli* DNA polymerase I, 5 units *E. coli* DNA ligase, and 1 unit *E. coli* RNase H, in a final volume of 80 μL second strand buffer (Clontech). PCR amplification was undertaken using the adapter primer (AP1) and gene-specific primer P2 (below) under the following conditions: initial denaturation for 3 min at 95°C followed by 30 cycles of 30 sec at 95°C, 30 sec at 55°C and ninety seconds at 72°C. The amplicon was analyzed by agarose electrophoresis and Southern transfer, followed by hybridization with an internal primer (P1, below). The fragment was purified from the agarose gel and cloned into a TA vector (pGEM T, Promega) for sequencing.

Primer extension analysis

The relevant primer, as indicated in the appropriate figure legend, was gel-purified and end-labeled by T4 polynucleotide kinase. Ten fmol of labeled primer was annealed to 50 μg of total RNA (or 10 μg of tRNA) by heating the mixture to 65°C for 90 min in 15 μL of hybridization buffer (0.15 M KCl, 10 mM Tris-HCl, pH 8.3, and 1 mM EDTA), after which the tubes were allowed to cool slowly to room temperature. After annealing, 0.9 μL of 1 M Tris-HCl, pH 8.3, 0.9 μL of 0.5 M $MgCl_2$, 0.25 μL of 1 M DTT, 6.75 μL of 1 mg/ml actinomycin D, 1.33 μL of 5 mM dNTPs, 20 μL H_2O, and 5 units of Superscript reverse transcriptase (Life Technologies, Inc.) were added to each tube and the mixture was incubated for 1 h at 42°C. After RNase A treatment (500 μg/ml RNase A, 37°C, 30 min), phenol/chloroform extraction and ethanol precipitation, the pellet was resuspended in 3 μL of loading buffer, denatured at 95°C for 3 min and subjected to denaturing electrophoresis, using either 6 or 8% polyacrylamide (Sequegel, National Diagnostics). Gels were dried and exposed to X-ray film (XAR, Kodak) at −80°C with one intensifying screen.

RNase protection assay

The relevant fragments were generated by restriction digestion of subgenomic clones or by PCR as indicated in the figure legends. All the templates for riboprobe synthesis were fully sequenced and subcloned into pGem 3Zf(+) (Promega), linearized and radiolabeled antisense cRNA synthesized using α-^{32}P-UTP (3000 Ci/mmol, NEN) and either T7 or SP6 RNA polymerase (Promega). These transcripts demonstrated specific activities in the range of 1.7–2.6 × 10^9 cpm/μg. Total RNA (50 μg) was coprecipitated with 4 × 10^4 cpm of the appropriate riboprobe and redissolved in 30 μl of 80% formamide, 40 mM PIPES (pH 6.4), 0.4 M NaCl, and 5 mM EDTA. After denaturation at 90°C for 5 min, the mixture was incubated at 42°C for 16 h. The annealed mixture was incubated at 37°C for 30 min with RNase (A/T1 mixture, Ambion) and the protected fragments were precipitated by Solution Dx (Ambion) and separated on 6 or 8% denaturing polyacrylamide gels.

Reverse transcription-coupled polymerase chain reaction (RT-PCR) amplification analysis of apobec-1

Primers were selected to span intron-exon junctions and used to delineate the tissue-specific patterns of alternate splicing in both rat and mouse. Conditions for RT-PCR used ~1 μg total RNA and used downstream priming (31) with primer P2 and 50 units superscript reverse transcriptase at 42°C for 60 min. After heat denaturation at 95°C for 5 min, PCR was conducted using 35 cycles of 95°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for ninety sec. A final extension of 10 min at 72°C was used for the last cycle. To facilitate amplicon identification in the tissue surveys, α-^{32}P d-CTP (3000 Ci/mmol, NEN, 0.3 μl per reaction) was included in the PCR mixture as previously described (30) and the product mixture was analyzed by agarose electrophoresis and fluorography. The major products obtained with mouse and rat liver RNA were cloned into a TA vector (Promega) and multiple independent clones were sequenced.

Construction of promoter-reporter DNA constructs for in vitro transfection

The following constructions were tested: **REXA,** a ~3.0 kb *Hin*d III-*Spe* I subgenomic fragment of rat

apobec-1 containing ~2.7 kb of 5′ flanking DNA upstream of exon A, 111 bp of exon A, and 185 bp of intron A; **RIN1+REX1,** a 4206 bp *Spe* I-*Ava* I subgenomic fragment spanning 3895 bp of intron A and 311 bp of exon 1 of rat *apobec-1;* **REX1** a 1745 bp *Eco*RI-*Ava*I subgenomic fragment including the distal 1434 bp of intron A and 311 bp of exon 1. These constructions were blunt-ended by Klenow enzyme and subcloned into the blunt-ended *Bgl* II site of pGL2-Basic (Promega) to be tested for promoter activity. Serial deletions (REX1-A to REX1-H) were generated from the REX1 construct by exonuclease III digestion (Double Strand Nested deletion kit, Pharmacia) and mapped by restriction digestion and sequencing. As a positive control, cells were transfected with the pGL2-Control Vector (Promega), which contains an SV 40 promoter upstream from the luciferase gene of pGL2-Basic. All transfections were conducted along with pCMVβGal, in which beta-galactosidase is driven by the cytomegalovirus promoter, as an internal control.

Transient transfection and luciferase assay

All the cell lines described were obtained from the American Type Culture Collection. Five × 10^5 CaCo-2 cells were added per well to 6-well plates in minimum essential medium as previously described (30) with 1 mM sodium pyruvate, 0.1 mM nonessential amino acids, and 20% fetal bovine serum. CaCo-2 cells were transiently transfected at times determined visually to coincide with their reaching 90–95% confluence. HeLa cells were cultured in 6-well plates in minimum essential medium supplemented with Earle's salt, 0.1 mM nonessential amino acid, and 10% fetal bovine serum. Mouse hepatoma (Hepa-6), rat hepatoma (McA RH7777), and human hepatoma cell lines (HepG2) were cultured in 6-well plates in Dulbecco's modified Eagle's medium with 10% fetal bovine serum. All cultures were maintained at 37°C with 5% CO_2. Cells were co-transfected using 10 μL per well of Lipofectamine® (Life Technologies, Inc.) with 1.5 μg pGL2 or the appropriate reporter construct and 0.15 μg of the internal reference plasmid pCMV-βGal. Cells were exposed to lipid-DNA complexes for 5 h, as recommended by the manufacturer, without antibiotics and then fed their complete growth medium. Forty-eight hr after transfection, the cells were lysed in 200 or 300 μL reporter lysis buffer (Promega), subjected to one cycle of freeze-thaw, and microcentrifuged. Luciferase activity was measured in 20 μL of the cell extract in a luminometer (Monolight 2010, Analytical Luminescence Laboratory) and the results were standardized for beta-galactosidase activity which was determined by luminometry in a 5-μL aliquot of the same extract using the Galacto-Light chemiluminescent reporter assay system (Tropix). All assays were undertaken in triplicate and the results represent the mean ±SD from two or three independent transfections. The data are presented as relative luciferase activity (RLA) where the luciferase activity from cells transfected with pGL-2 Basic is assigned a value of 1.0.

Oligonucleotides

All oligonucleotides were purchased from GIBCO-BRL and purified, where indicated, by polyacrylamide gel electrophoresis. The following oligonucleotides were used:

AP1; 5′-CCATCCTAATACGACTCACTATAGGGC-3′

AP2; 5′-ACTCACTATAGGGCTCGAGCGGC-3′

MEPRIF; 5′-CAGAAAGATACTTTCGTCCGAACAC-3′
(Translated mouse apobec-1 cDNA nt +215 to +239)

MEPR2R; 5′-GTCATGATCTGGATAGTCACACCGCT-3′
(Translated mouse apobec-1 cDNA nt +434 to +409)

P1; 5′-CTGTCTCGGAACTCATCTTGC-3′
(Rat/Mouse Exon 1 nt +16 to −5)

P2; 5′-GGTCAAAGAAGACTTCAAACTCG-3′
(Rat apobec-1, translated region nt +85 to +63)

P4; 5′-GGATCAACAGCTACAGGGCCTGTCTCGGA ACTCAT-3′
(Rat Exon A nt +77 to +111)

P5; 5′-GCTGAGAGTCCTGACATGCTGGTA-3′
(Rat 5′ flanking region of Exon A nt −72 to −49)

P6; 5′-TTCTTCCTGATCTGATTTCAAG-3′
(Mouse Exon A nt +161 to +140)

P7; 5′-GGCTTAGTATGTAACATCTCAGATT-3′
(Rat 5′ flanking region of Exon A nt −195 to −171)

P8; 5′-TTCTTCCTGATCTAATTTCAAG-3′
(Rat Exon A nt +111 to +90)

P9; 5′-CTGTCTCGGAACTCATCTTGCTCTCTGTG TCTCTG-3′
(Rat Exon 1 nt +16 to −19)

P10; 5′-CACACCTGAGGAAGCAAAGTCCGGCA-3′
(Rat Exon 1 nt −95 to −70)

P12; 5′-ATCTGGATTTCCGTCCTCCCTGGAGAGAG CC-3′
(Rat Exon 1 nt −218 to −248)

P14; 5′-GCCTACTCCCGCTACAGAACCAC-3′
(Rat Exon A nt +22 to +44)

P15; 5′-TCCGTGCTGGCTTCTCTCCCACTCT-3′
(Rat 5′ flanking region of Exon 1 nt −366 to −342)

P16; 5′-GGATCAACAGCTACAGGGCCTGTCTCGGA ACTCAT-3′
(Mouse translated apobec-1 cDNA nt +35 to +1)

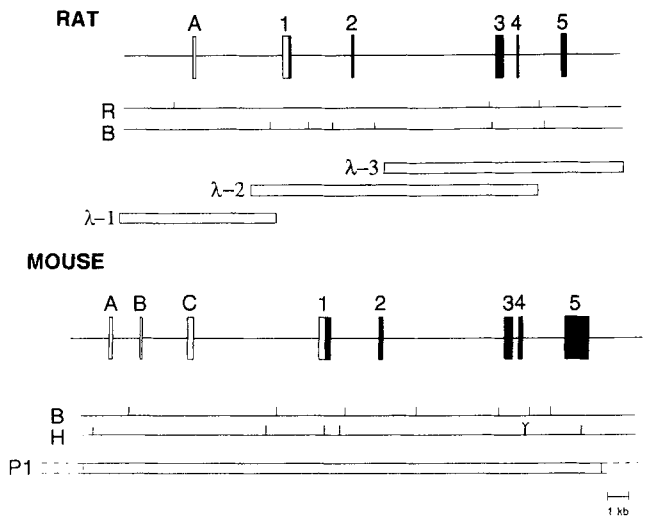

Fig. 1. Structural organization and partial restriction map of the *apobec-1* gene in rat and mouse. Filled and open boxes indicate the coding and non-coding exons, respectively. R, *EcoR* I; B, *Bam*H I; H, *Hind* III; P1, P1 clone obtained from a mouse ES/C129 library.

RESULTS

Structural organization of the rat and mouse apobec-1 genes

As illustrated in **Fig. 1,** three independent lambda clones were isolated from a rat genomic library and their restriction pattern was determined. In addition, all the exons, the exon-intron junctions, and the indicated 5′ flanking regions upstream of both the rat and mouse exon A and exon I were sequenced (see below). The rat *apobec-1* gene is shown in alignment with the mouse gene, the latter determined by analysis of a single P1 clone. The rat *apobec-1* gene spans 15 kb and contains one untranslated exon (designated exon A) and five translated exons, referred to as 1–5. The mouse *apobec-1* gene spans 24 kb and contains three untranslated exons (designated exons A, B, and C) and five translated exons, as previously demonstrated (17). Exon A demonstrates 90% sequence homology between the rat and mouse. However, sequence analysis of the entire (rat) intron A (~4 kb) demonstrated that regions homologous to mouse exons B and C were absent from the rat *apobec-1* gene (data not shown).

The exon sizes for rat *apobec-1* vary from 28 bp (exon 2) to 398 bp (exon 3), as illustrated in **Table 1.** The intron-exon boundaries occur at identical positions within the coding region of both the rat and mouse genes and all the splice acceptor sites were found to conform to the gt-ag rule. We confirmed the presence of alternative splice-acceptor sites in exon 1 of both the rat and mouse genes (17). The intron-exon organization of the mouse *apobec-1* gene is illustrated in **Table 2.** The findings differ in several respects from other recent reports of the genomic organization of mouse *apobec-1* (17, 32). These differences include the sizes of exon A (161 bp) and exon 1 (271 bp), which reflect differences in the assignment of transcription start sites, as will be detailed below. The discrepancy in the size of intron 3 [~0.5 kb versus 1.4 kb (17)] is unexplained.

Delineation of the 5′ untranslated regions of the rat and mouse apobec-1 mRNA species

Transcription start sites in rat and mouse exon A. Primer extension and RNase protection both demonstrated a single, major start site in exon A of the rat *apobec-1* gene using liver RNA (**Fig. 2A, B**). These data were independently confirmed by cloning and sequencing the products of 5′ RACE (data not shown).

Primer extension and RNase protection yielded complementary results also for mouse apobec-1 mRNA, both approaches demonstrating the presence of three major start sites in exon A using mouse liver RNA (Fig. 2A, C). Other minor bands, of undetermined significance, are detectable in both panels. The current data differ from previous findings which demonstrated two major transcription initiation sites for mouse liver RNA, neither of which correspond to those found in the current study (17).

Transcription start sites in rat exon 1. A combination of primer extension and RNase protection was used to demonstrate a single distal transcription start site in exon 1 of the rat gene, as indicated schematically in **Fig. 3A.** These approaches revealed a single, major start site at nucleotide −328 in liver, kidney, and small intestine (Fig. 3B, C).

The same combination of primer extension and RNase protection mapping, shown schematically in **Fig. 4A,** defined three major transcription start sites in the proximal region of rat exon 1 (Fig. 4B, C), corresponding to nucleotides −44, −55, and −84, respectively. Other minor bands of undetermined significance were also detectable. Eight independent clones were sequenced from the rat intestine 5′ RACE reactions, of which two revealed identical sequence and a suggested start site at nucleotide −84 in exon 1. Three additional clones with sequence identity suggested a start site at nucleotide −44 of exon 1. The remaining three clones demonstrated a variety of shorter products (data not shown).

Transcription start sites in mouse exon 1. Analysis of mouse small intestinal RNA using both primer extension and RNase protection (data not shown) demonstrated four proximal transcription start sites within exon 1, at positions −34, −48, −66, and −76, respec-

TABLE 1. Organization of rat *apobec 1* gene

Exon Number	Exon Size (bp)	Sequence at Exon-Intron Junction				Codon Interrupted	Intron Size (kb)
		5′ Splice Donor		3′ Splice Acceptor			
		↓		↓			
A	111	CAGGAAGAA	gtaaggaa	attacatag	CATTCATAT[a]		4.0
				ccctgcag	GTGGCCGT[a]		
1	344[b]	GAG ACA G	gtaaag	ggtttacag	GC CCT GTA	Gly[6]	3.2
	173[c]						
	122[c]						
2	28	ACT CTG AG	gtaagt	atgtttcag	G AGA AGA	Arg[15]	6.5
3	398	AGA CAA G	gtgaag	gtttctttag	AG TCT GGC	Glu[148]	0.7
4	119	TC ATT TTA	gtaag	ctttgtttcag	GGA CTT	Leu[187]/Gly[188]	2.0
5	230						

[a]Alternative 3′ splice acceptor sites, beginning of exon 1.
[b]Size of mRNA from distal transcription start site in exon 1 (see also Fig. 5A).
[c]Two different sized mRNA species resulting from alternative 3′ splice acceptor sites (see also Fig. 5A).

tively. Primer extension and RNase protection revealed at least one distal transcription start site in exon 1 of the mouse *apobec-1* gene, corresponding to nucleotide −255 (data not shown).

Distribution of apobec-1 mRNA species in the rat and mouse (Overview)

The major species of apobec-1 mRNA in the rat are summarized in **Fig. 5A.** Only the major splice variants are shown for simplicity, as these represent greater than 95% of the total population found in most tissues of the rat (see also below). Similarly, the major species of apobec-1 mRNA in the mouse are summarized in Fig. 5B. Again, for simplicity, the major splice variants have been illustrated, representing generally greater than 90% of the total population in this species (see also below).

Tissue-specific distribution of apobec-1 mRNA species in the rat: RNase protection reveals three distinct populations. RNase protection was used to determine the relative distribution of different apobec-1 mRNA species in multiple adult rat tissues. A riboprobe template was constructed to span the distal portion of rat exon A and a continuous stretch from exon 1c to exon 3 (**Fig. 6A**). RNase protection demonstrated five distinct fragments, corresponding to the patterns of transcription initiation proposed above. As illustrated in Fig. 6A, the presence of a 282 nt protected fragment (F1), indicates that the transcript contains the distal portion of exon A, the prototype for which would be hepatic apobec-1 mRNA (Fig. 6B). The absence of fragment F1 in apobec-1 mRNA from the small intestine, coupled with the presence of fragments F2–F5 (Fig. 6B) is consistent with the prediction that transcription in this tissue arises predominantly from exon 1, including the distal portion of untranslated RNA from this exon. Conversely, apobec-1 mRNA from the stomach demonstrates a pattern indicating transcription predominantly from exon A, while the mRNA from the colon demonstrates elements of both exon A- and exon 1-derived transcripts, as evi-

TABLE 2. Organization of mouse *apobec 1* gene

Exon Number	Exon Size (bp)	Sequence at Exon-Intron Junction				Codon Interrupted	Intron Size (kb)
		5′ Splice Donor		3′ Splice Acceptor			
		↓		↓			
A	161	CAGGAAGAA	gtaagg	cttctcag	GACTGCCC		2.2
B	68	ATGAGCCAG	gtatgg	ccccag	GACTGCTA		3.6
C	146	TCACACAG	gtaagc	acatag	TACAGATC[a]		6.6
				ctgcag	GTGGCTGT[a]		
1	271[b]	GAG ACA G	gtaaag	tttcag	GC CCT GTA	Gly[6]	3.0
	174[c]						
	118[c]						
2	28	CT CTG AG	gtaaat	ttccag	G AGA AGA	Arg[15]	5.8
3	398	GAG CAA G	gtaaa	ctttag	AG TATTG	Glu[148]	0.5
4	119	C ATT TTA	gtaagt	tttcag	GGA CTT C	Leu[187]/Gly[188]	1.9
5	1144						

[a]Alternative 3′ splice acceptor sites, beginning of exon 1.
[b]Size of mRNA from distal transcription start site in exon 1 (see also Fig. 5B).
[c]Two different sized mRNA species resulting from alternative 3′ splice acceptor sites (see also Fig. 5B).

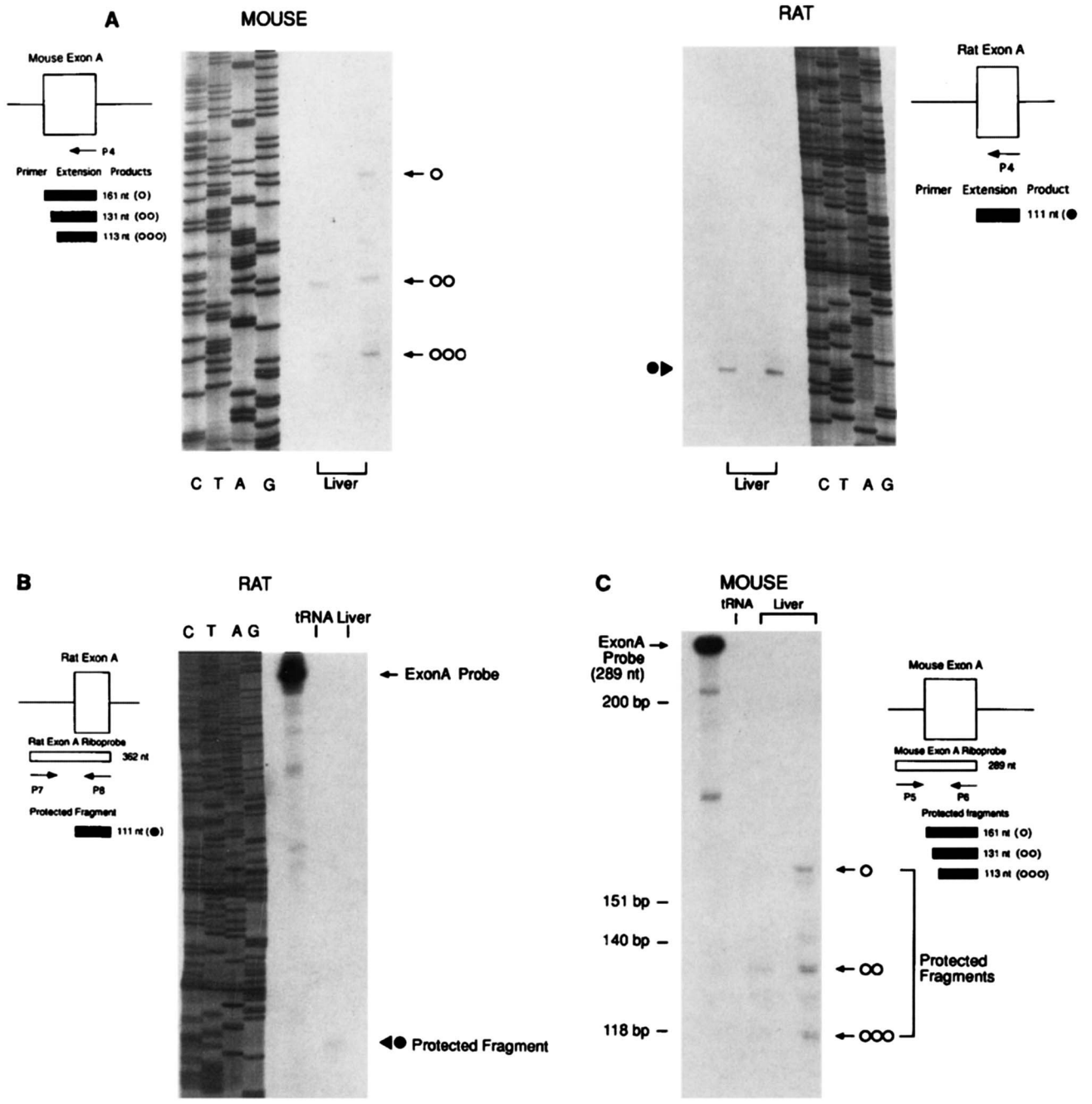

Fig. 2. Determination of the transcription start sites in exon A of rat and mouse *apobec-1* genes. A: Primer extension assay. A ^{32}P-end-labeled antisense oligonucleotide (P4) corresponding to nucleotide +77 to +111 of rat exon A was hybridized with 50 μg of mouse (left panel) or rat (right panel) liver total RNA and extended with reverse transcriptase. Products were separated by 6% PAGE. Lanes C, T, A, and G represent dideoxynucleotide chain-termination sequencing reactions performed with the same primer used in primer extension. Transcription start sites of rat and mouse *apobec-1* are indicated by closed and open circles, respectively. B: 5′ end mapping of rat exon A by RNase protection assay. The template for riboprobe synthesis was made by PCR with primers P7 and P8. 50 μg of rat total liver RNA, or yeast tRNA control, was hybridized with the ^{32}P-labeled riboprobe and the protected fragment analyzed by 6% PAGE. Lanes C, T, A, and G represent a dideoxynucleotide chain-termination sequencing reaction performed with the P8 primer. The protected fragment is indicated with the filled circle. C: 5′ end mapping of mouse exon A by RNase protection assay. The template for riboprobe synthesis was made by PCR with primers P5 and P6. Fifty μg of mouse total liver RNA, or yeast tRNA control, was hybridized with the ^{32}P-labeled riboprobe and the protected fragment analyzed by 6% PAGE. The protected fragments are indicated by open circles.

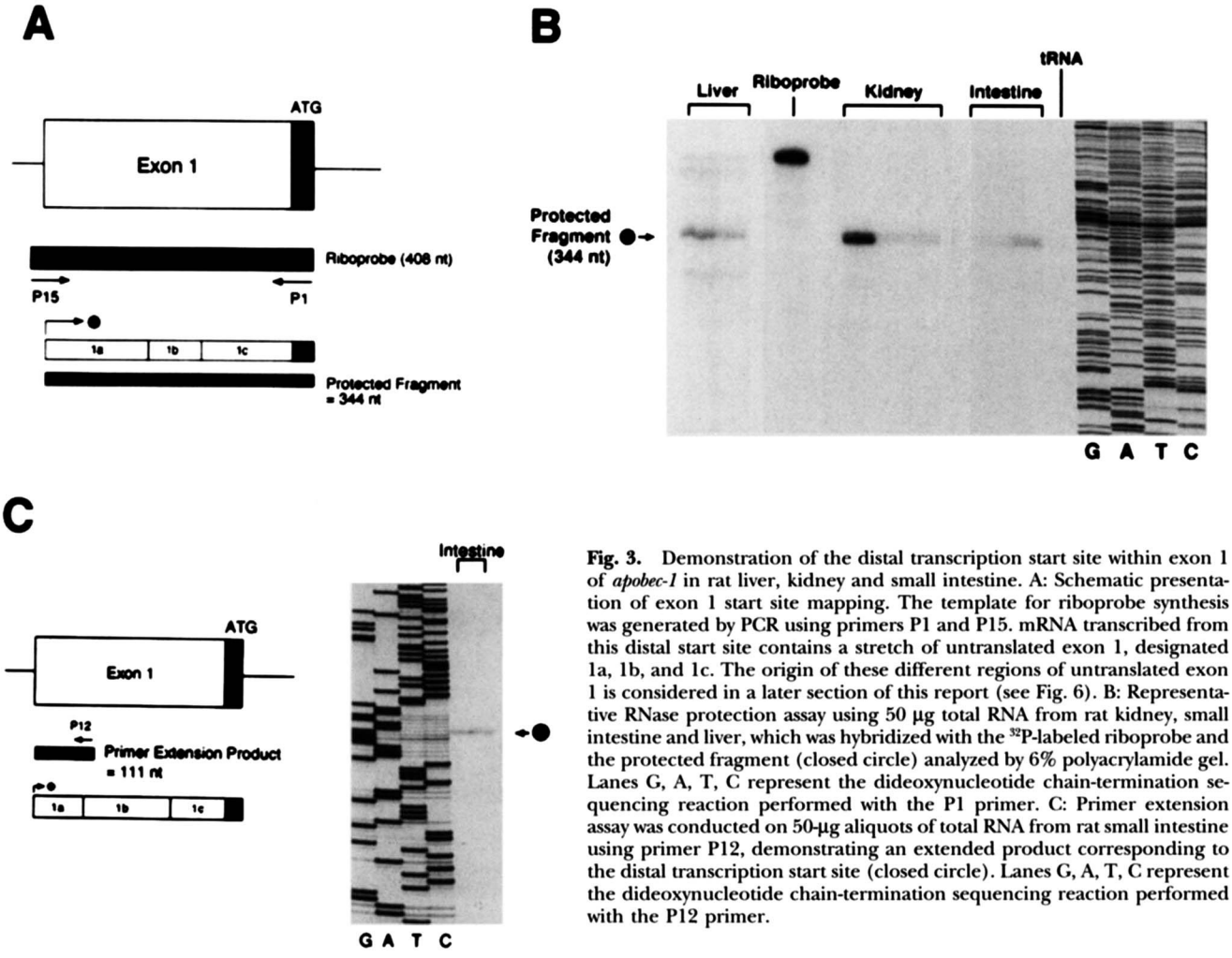

Fig. 3. Demonstration of the distal transcription start site within exon 1 of *apobec-1* in rat liver, kidney and small intestine. A: Schematic presentation of exon 1 start site mapping. The template for riboprobe synthesis was generated by PCR using primers P1 and P15. mRNA transcribed from this distal start site contains a stretch of untranslated exon 1, designated 1a, 1b, and 1c. The origin of these different regions of untranslated exon 1 is considered in a later section of this report (see Fig. 6). B: Representative RNase protection assay using 50 μg total RNA from rat kidney, small intestine and liver, which was hybridized with the ^{32}P-labeled riboprobe and the protected fragment (closed circle) analyzed by 6% polyacrylamide gel. Lanes G, A, T, C represent the dideoxynucleotide chain-termination sequencing reaction performed with the P1 primer. C: Primer extension assay was conducted on 50-μg aliquots of total RNA from rat small intestine using primer P12, demonstrating an extended product corresponding to the distal transcription start site (closed circle). Lanes G, A, T, C represent the dideoxynucleotide chain-termination sequencing reaction performed with the P12 primer.

denced by the presence of all five fragments (Fig. 6B). Laser scanning densitometry of the protected fragments was used to calculate the distribution of apobec-1 mRNA species originating in exon A, versus the distal or proximal regions of exon 1, using appropriate corrections for the length of protected fragments. The results are illustrated in Fig. 6C, which demonstrates that small intestinal apobec-1 mRNA arises almost exclusively from exon 1 while hepatic apobec-1 mRNA arises from exon A. The remaining tissues demonstrated differences in promoter usage. Interestingly, stomach demonstrated a predominance of transcripts arising in exon A, while colon demonstrated a pattern closer to that of small intestine (Fig. 6C). Other extrahepatic, non-intestinal tissues such as spleen demonstrated a predominance of transcripts originating in exon A (Fig. 6B, C).

Exon 2 skipping is a minor component in regulating rat and mouse apobec-1 mRNA species. Our recent findings have indicated that ~50% of human small intestinal apobec-1 mRNA undergoes alternative splicing of exon 2, which results in a frame shift and creates a novel truncated peptide (33). We examined the presence of a similar, alternative splicing event in RNA derived from multiple tissues of both the rat and mouse. Using primers complementary to regions in exon A and in exon 3 of the rat and mouse *apobec-1* genes, a screening strategy was conducted as indicated schematically in **Fig. 7A** and **Fig. 8A.** The results, illustrated in Fig. 7B and Fig. 8B, suggest that generally less than 10% of mRNA species in the rat and mouse (a possible exception being mouse heart) contain an exon 2-skipped form. Thus, the distinctive alternative splicing pattern found in human small intestine represents a quantitatively minor source of variability of apobec-1 mRNA species in the rat and mouse. Note that the relatively weak signal from small intestinal RNA samples in this RT-PCR analysis most likely reflects the observation, alluded to above, that the

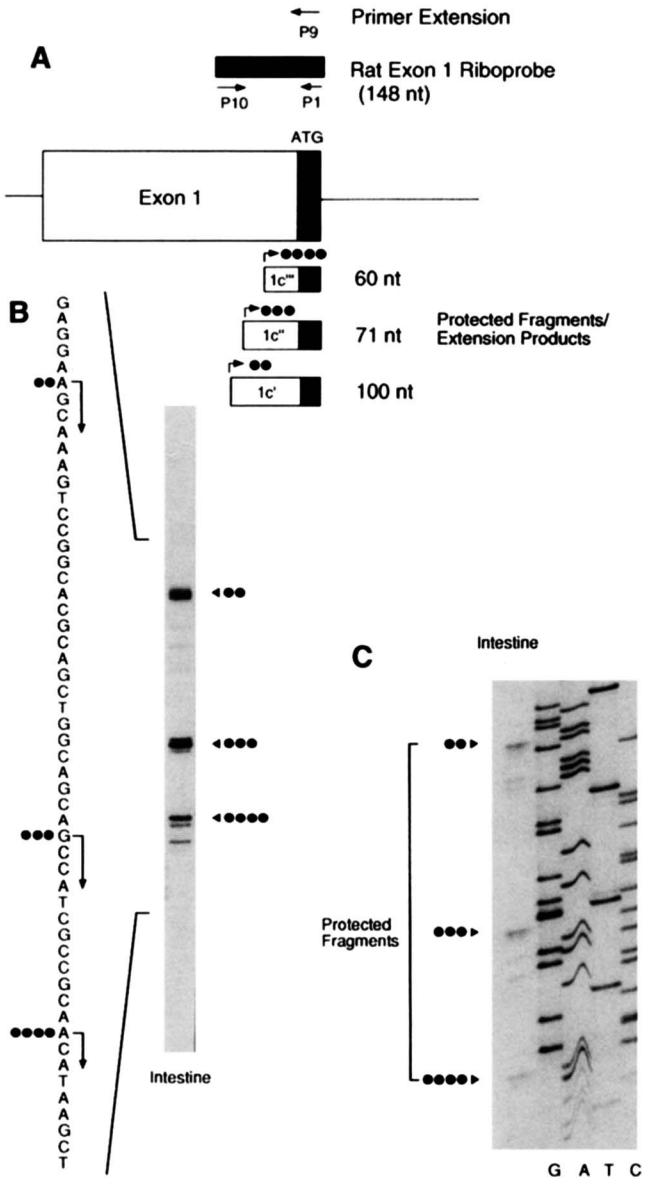

Fig. 4. Demonstration of proximal transcription start sites within exon 1 of *apobec-1* in rat small intestine. A: Schematic presentation of exon 1 start site mapping. Primer extension was performed with antisense primer P9, corresponding to nucleotide +1 to +35 of rat *apobec-1* cDNA. The exon 1 riboprobe template was generated by PCR using primers P1 and P10. Three major extension products/protected bands are demonstrated, corresponding to transcription initiation sites at exon 1c′, 1c″ and 1c‴. B: Primer extension assay results using 50 μg total RNA from rat small intestine, hybridized to the end-labeled primer P9 and extended with reverse transcriptase. Primer extension products were separated by 8% PAGE. The corresponding nucleotide sequence is shown and the closed circles indicate major transcription start sites. C: RNase protection mapping of major transcription start sites in exon 1 of rat small intestine. The template forriboprobe synthesis was made by PCR with primers P1 and P10. Fifty μg of rat total intestinal RNA was hybridized with the ^{32}P-labeled riboprobe and the protected fragment was analyzed by 8% polyacrylamide gel. The protected fragments are indicated by closed circles. Lanes G, A, T, and C represent a dideoxynucleotide chain-termination sequencing reaction performed with the P1 primer.

majority of apobec-1 mRNA species originates from exon 1 and not exon A.

Comparative analysis of 5′ flanking regions upstream of exon A and exon 1 in the mouse and rat *apobec-1* gene

Studies of exon A. The nucleotide sequence of ~1 kb of the 5′ flanking region upstream of exon A was obtained and compared for the rat and mouse. The most proximal 400 bp is shown in **Fig. 9A.** Interspecies nucleotide homology in this region was 78% as shown in Fig. 9A. A variant TATA box with the sequence TAAA, previously identified in the mouse apobec-1 gene (17) is present at nucleotides −50 to −47 (17). The corresponding region in the rat gene, from nucleotides −88 to −85 contains the sequence TGAA and no other region in the flanking region of either species was found to contain a canonical TATA box. A putative CAT box (17), with the sequence CACT, was identified in both the mouse and rat genes, at positions −85 to −82 and −107 to −104, respectively. There were two nucleotide differences in the mouse sequence from a previous study, at positions +85 and +86 in the untranslated exon A, where a C C dinucleotide was found instead of T T (17). A number of potential binding motifs to known transcription factors were identified (34–40), including sites for PEA3 (34, 35), AP2 (36), AP1 (37), TRE (37), Sp1 (38), glucagon (39), insulin (40), and an insulin response element (41) (Fig. 9A).

Studies of intron A and the region flanking exon 1. The entire 4 kb intron A in rat apobec-1 was sequenced and compared to the mouse sequence over the proximal ~1.2 kb upstream of exon 1. There was a single nucleotide sequence difference in the mouse sequence found in comparison to the previous report, namely a C at position −206 instead of a T (17). The overall nucleotide homology (rat/mouse) within this region was 61%; the first 560 bp upstream of the initiator methionine demonstrated 84% homology (Fig. 9B) while the distal ~700 bp demonstrated only 39% homology. A putative TATA box was identified in the mouse gene, at position −288 to −285, as previously noted (17). The corresponding region in the rat gene, from nucleotides −312 to −309 contains the sequence TTTA and no other region in the flanking region of the rat *apobec-1* gene was found to contain a canonical TATA box (Fig. 9B). Both the rat and mouse genes contained a putative CAT box with the sequence CCACT, in the mouse from nucleotides −330 to −326 and in the rat from nucleotides −349 to −344 (Fig. 9 B). A number of potential binding motifs to known transcription factors were identified (42–47) in addition to the numerous PEA3 sites also found in exon A, above, including sites for GATA (42),

c/EBP (43), NF-1 (44), HNF-5 (45), GCN4 (46), and OCT-1 (47) (Fig. 9B). In addition, there was a direct repeat of 35 nucleotides (boxed motif, Fig. 9B) from −214 to −180 and from −179 to −145, found only in the rat promoter. This region is AT-rich (67%) and contains a potential binding site for OCT-1 (47), but the functional significance of this adjacent repeat structure is unknown.

Functional analysis of the 5′ flanking regions of the rat *apobec-1* gene

In order to determine the functional importance of the 5′ regions flanking rat exon A and exon 1, three different promoter constructs were assembled (**Fig. 10A**), and their relative ability to direct the expression of a heterologous reporter gene was determined in several different cell types. Among the cell types selected were three hepatoma lines, Hepa-6 and McA-7777, derived from mouse and rat hepatoma, respectively (both of which express endogenous apobec-1), and HepG2 cells, a human hepatoma-derived line that does not express apobec-1 mRNA and does not edit endogenous apoB mRNA (48). The human colon cancer-derived cell line CaCo-2 was used to examine intestinal apobec-1 expression and, finally, HeLa cells were used as an epithelial cell control. Expression of luciferase was detectable at levels 3- to 9-fold above control (pGL2-Basic) in all cells using the REXA construct, indicating ubiquitous promoter activity (Fig. 10B). The [RIN1 + REX1] and REX1 constructs both produced high levels of luciferase activity in CaCo-2 cells, at levels 20- to 30-fold above control and 3- to 6-fold higher than with the REXA construct (Fig. 10 B). In addition, a high level of luciferase activity was noted with the [RIN1 + REX1] and REX1 constructs when transfected into McA7777 cells (Fig. 10B). The activity of these constructs, although lower than in McA7777 cells, was also demonstrable in both Hepa and HepG2 cells.

In order to further define the promoter region that may regulate reporter activity in McA7777 and Caco-2 cells, a series of nested deletions was prepared. The results indicate that luciferase activity was maintained, in both cell types, at levels ~30- to 40-fold above control with constructs REX1 and REX1-A (Fig. 10C). Deletion of the region from −1415 to −1053 resulted in a ~30–50% decrease in luciferase activity in both cell types (Fig. 10C). Deletion of the region from −674 to −523 produced a 67% increase in luciferase activity in Caco-2 cells without changing the activity noted in McA7777 cells (Fig. 10C). Conversely, deletion of the region from −262 to −407 produced an 81% increase in luciferase activity in McA7777 cells, without changing the activity noted in Caco-2 cells (Fig. 10C). These data imply some cell-type specificity mediated by the proximal ~700 bp region flanking exon 1. Deletion of the region from −262 to −137 led to a fall in luciferase activity to ~11-fold above control in both cell types (Fig. 10C). Further deletion of the region between −137 and −48 resulted in total loss of activity (Fig. 10C), suggesting that the −137 construct (REX1-G), which contains binding sites

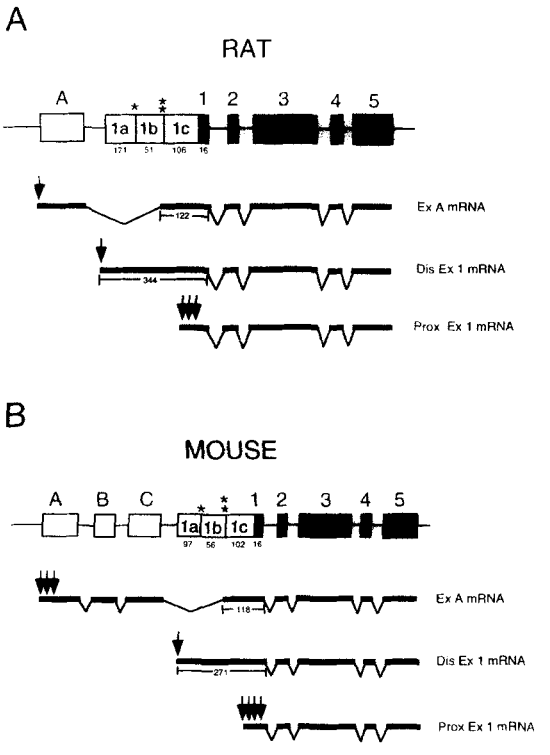

Fig. 5. Schematic representation of the origin of the major apobec-1 mRNA species in rat and mouse tissues. A: Major species of apobec-1 mRNA in the rat. Open and closed boxes denote untranslated and translated exons, respectively. Arrows and asterisks denote transcription start sites and alternative splice acceptor sites, respectively. Rat exon 1 contains untranslated regions, designated exon 1a (171 bp), 1b (51 bp) and 1c (106 bp), and also contains 16 bp of translated sequence. Note that the majority of apobec-1 transcripts initiating in exon A do not contain exon 1a or 1b, the result of alternative splicing. mRNA transcribed from the distal transcription start site in exon 1 (Dis Ex 1 mRNA) contains the entire sequence of exon 1 (344 bp). Within 1c, there are three major proximal transcription start sites (three arrows), which produce mRNA species with relatively short 5′ untranslated regions, referred to as exon 1c′ (84 nt), exon 1c″ (55 nt), and exon 1c‴ (44 nt) (see also Fig. 4A). B: Major species of apobec-1 mRNA in the mouse. Open and closed boxes denote untranslated and translated exons, respectively. Arrows and asterisks denote transcription start sites and alternative splice acceptor sites, respectively. Mouse exon 1 contains untranslated regions, designated exon 1a (97 bp), 1b (56 bp), 1c (102 bp), and also contains 16 bp of translated sequence. Note that the majority of apobec-1 transcripts initiating in exon A (three arrows) do not contain exon 1a or 1b, the result of alternative splicing. mRNA transcribed from the distal transcription start site in exon 1 (Dis Ex 1 mRNA) contains the entire sequence of exon 1 (271 bp). Within exon 1c there are four proximal transcription start sites (four arrows), which produce mRNA species with relatively short 5′ untranslated regions, referred to as exon 1c′ (76 nt), exon 1c″ (66 nt), exon 1c‴ (48 nt), and exon 1c⁗ (34 nt) (data not shown).

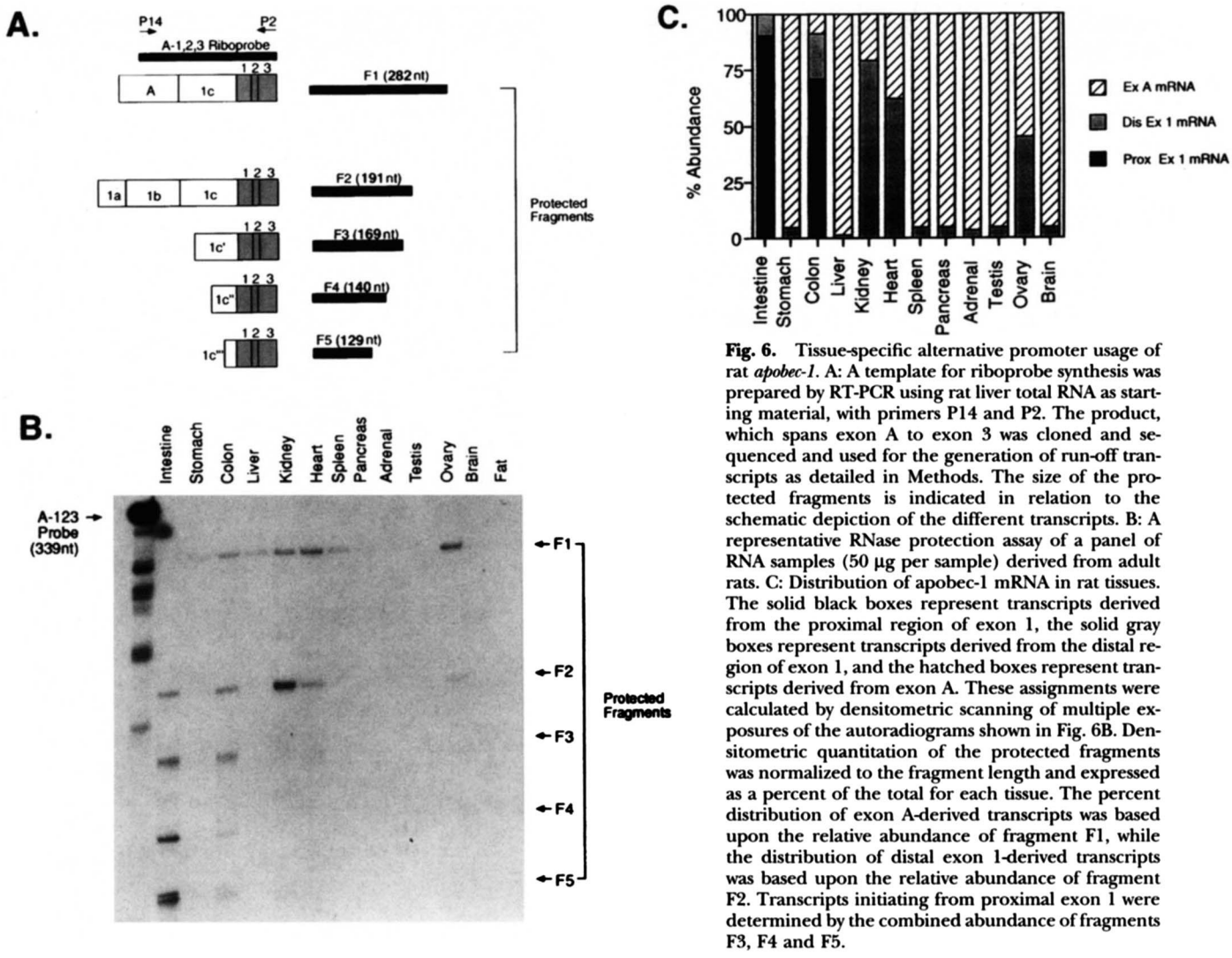

Fig. 6. Tissue-specific alternative promoter usage of rat *apobec-1*. A: A template for riboprobe synthesis was prepared by RT-PCR using rat liver total RNA as starting material, with primers P14 and P2. The product, which spans exon A to exon 3 was cloned and sequenced and used for the generation of run-off transcripts as detailed in Methods. The size of the protected fragments is indicated in relation to the schematic depiction of the different transcripts. B: A representative RNase protection assay of a panel of RNA samples (50 μg per sample) derived from adult rats. C: Distribution of apobec-1 mRNA in rat tissues. The solid black boxes represent transcripts derived from the proximal region of exon 1, the solid gray boxes represent transcripts derived from the distal region of exon 1, and the hatched boxes represent transcripts derived from exon A. These assignments were calculated by densitometric scanning of multiple exposures of the autoradiograms shown in Fig. 6B. Densitometric quantitation of the protected fragments was normalized to the fragment length and expressed as a percent of the total for each tissue. The percent distribution of exon A-derived transcripts was based upon the relative abundance of fragment F1, while the distribution of distal exon 1-derived transcripts was based upon the relative abundance of fragment F2. Transcripts initiating from proximal exon 1 were determined by the combined abundance of fragments F3, F4 and F5.

for several transcription factors, including AP1 (37), TRE (37), GCN-4 (46), and PEA-3 (34, 35) (see Fig. 9B), may function as a minimal essential (proximal) promoter region. In order to determine whether the distal region of exon 1, from which transcription initiation was demonstrated in rat small intestine, kidney, and liver (Fig. 3, above), could function in the absence of the proximal promoter region, a fusion construct was created in which the region from −1778 to −282 was fused directly to luciferase (REX1-ΔF, Fig. 10C), thereby eliminating the region of exon 1 from nucleotides −281 to −33. Transfection of the REX1-ΔF fusion construct yielded levels of luciferase activity in Caco-2 cells that were ~30% higher than observed above with the parental construct (REX1). By contrast, luciferase activity in McA7777 cells with the REX1-ΔF fusion construct was ~50% lower than found with REX1 (Fig. 10C). Taken together, these data suggest that the region from nucleotides −281 to −33 may confer limited cell-type specificity. More importantly, however, the findings suggest that the region from −1778 to −282 (containing the distal transcription start site) demonstrates functional activity in both cell types that is independent of the proximal promoter region.

DISCUSSION

These studies provide information concerning the comparative gene structure of *apobec-1* in the rat and mouse, two species for which considerable information exists concerning the tissue-specific distribution and physiological regulation of this gene product. As alluded to above, there are many distinctive features of apoB mRNA editing in the rat and mouse which together serve as an initiation point for the investigations

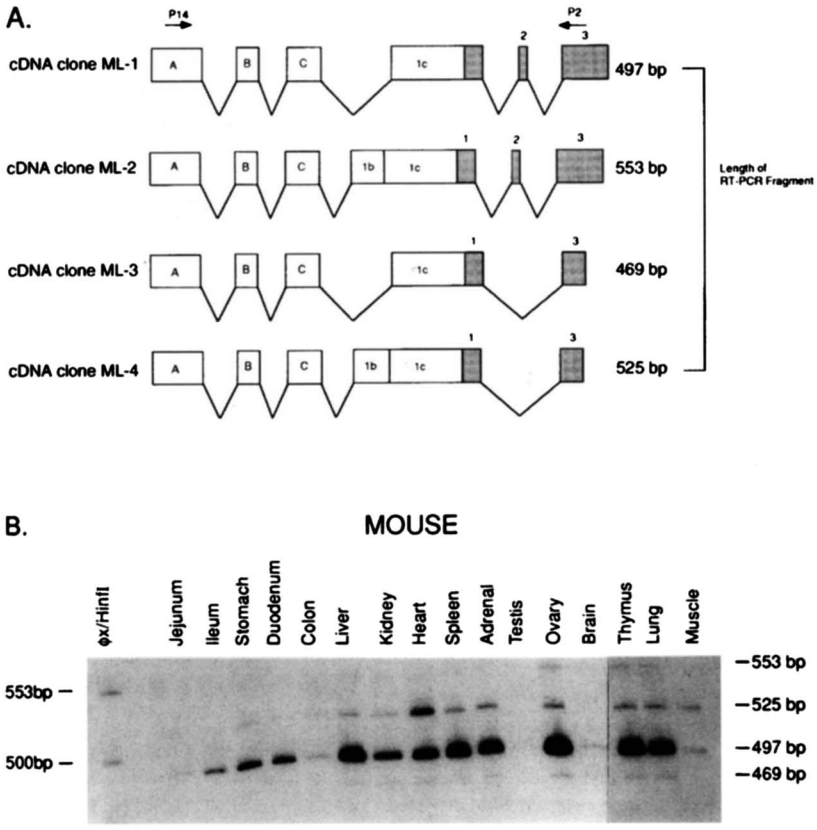

Fig. 7. Schematic representation and tissue distribution of alternative splicing in the mouse apobec-1 gene. A: Schematic representation of mouse liver transcripts. RT-PCR was performed using primer P14, which is specific for exon A and P2, which is specific for exon 3. The PCR fragments were subcloned and sequenced. Four different types of clones (ML-1, 2, 3, and 4) were identified. B: Tissue distribution of the splice variants of mouse apobec-1 mRNA is illustrated using RT-PCR with the indicated sources of total RNA. Radioactive dCTP was included for visualization of the products.

reported. Prominent among these are the observations that apoB mRNA editing is subject to developmental, metabolic, hormonal, and nutritional regulation in the liver of rats and mice (20–27). An additional feature of interest in regard to apoB mRNA editing in the rat and mouse is the fact that apobec-1 mRNA is widely distributed, including tissues with virtually no detectable apoB mRNA (16, 17). These findings imply some presumed advantage to the organism, an implication which must be reconciled with the impression that gene-targeted *apobec-1* −/− mice have no apparent phenotype with regard to intestinal or hepatic lipoprotein secretion (28, 29). Knowledge of the gene structure and regulatory elements that control the tissue-specific expression of apobec-1 may thus enhance our understanding of some of these issues.

Comparative analysis of the structural genes encoding rat and mouse apobec-1

A number of important organizational features of these homologs were found to be shared. Both genes contain untranslated mRNA at the 5′ end of the transcript. Mouse *apobec-1* was previously shown to contain three untranslated exons (17), a finding confirmed in the present study. The most 5′ of these untranslated exons in the mouse (referred to in this report as exon A) exhibits 90% sequence homology to the corresponding region in the rat *apobec-1* gene. However, no region homologous to exons B and C in the mouse was found in the ~4 kb rat intron A. The promoter region upstream of exon A in both the rat and mouse genes exhibit considerable homology (78%) over the ~1 kb for which sequence information was obtained. Another shared feature of the promoter region upstream of exon A is the absence of a consensus TATA box or CAAT box, which, considered together with the GC-rich region flanking the transcription start sites and the presence of multiple, potential Sp-1 binding sites (38), are features of mammalian housekeeping genes, many of which also have multiple start sites (49, 50). In considering possible alternative transcriptional regulatory regions in these promoters, the sequence flanking the transcription start site of rat exon A from −2 to +6 was noted to be pyrimidine-rich, GCTT(+1)TCTCT. However, this region lacks the canonical features of an initiator element (5′-YYCA(+1)YYYYY), a regulatory motif present in some TATA-less genes with a wide tissue distribution (49, 50). Other potential binding sites were found in the 5′ flanking region upstream of exon A, whose transcriptional activity has been well documented in different systems. Prominent among these are binding sites for PEA-3, a motif found in the promoter region of several developmentally regulated

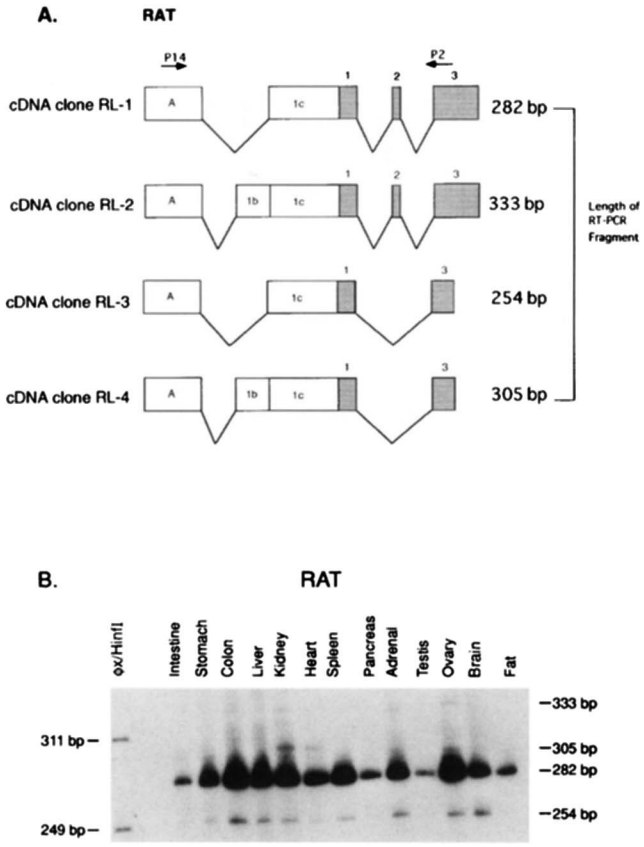

Fig. 8. Schematic representation and tissue distribution of alternative splicing in the rat apobec-1 gene. A: Schematic representation of rat liver transcripts. RT-PCR was performed using primers P14, which is specific for exon A and P2, which is specific for exon 3. The PCR fragments were subcloned and sequenced. Four different types of clones (RL-1, 2, 3, and 4) were identified. B: Tissue distribution of the splice variants of rat apobec-1 mRNA is illustrated using RT-PCR with the indicated sources of total RNA. Radioactive dCTP was included for visualization of the products.

genes (34, 35) and which was present in two sites in exon A of both the rat and mouse genes. Previous studies have established that the $p68^{c\text{-}ets\text{-}1}$ product of the ets oncogene can activate heterologous transcription through the PEA-3 motif, an effect that was dramatically magnified by AP-1, after coexpression of the c-Fos and c-Jun products (35). The presence of an AP-1 site in the rat exon A promoter, when considered along with the demonstration of multiple PEA-3 sites, lends weight to the possibility that this region may function as a regulatory domain. Convincing proof of this speculation, however, will await more formal examination.

Physiologic maneuvers previously demonstrated to modulate apoB mRNA editing and which also induce changes in apobec-1 mRNA abundance include fasting and refeeding a high carbohydrate diet to rats (16). Recent work has identified a carbohydrate-response element in the liver-type pyruvate kinase gene and also in the rat hepatic S 14 gene promoter, both of which consist of two CACGTG motifs separated by 5 base pairs (51, 52). A single copy of CACGTG motif was found in intron A of the rat apobec-1 gene, at position −1995 (K. Hirano and N. O. Davidson, unpublished observations). The required second copy of this motif, however, was absent from this region of the apobec-1 promoter (tggGaa) and consequently the importance of this domain in regard to the demonstrated responsiveness of apobec-1 gene expression to carbohydrate flux remains unknown. An additional observation was the presence of an insulin-response element (41) in both the rat and mouse exon A promoter region (Fig. 9A). In this regard, it is noteworthy that earlier studies demonstrated an increase in apoB mRNA editing in primary rat hepatocytes after insulin treatment (24). The question of whether insulin produces a transcriptional increase in apobec-1 gene expression in rat liver cells is clearly worthy of study. In addition, it will be important to define the mechanisms underlying the accumulation of apobec-1 mRNA in the liver of rats subjected to fasting and refeeding a high carbohydrate diet.

Tissue-specific expression of apobec-1 mRNA in rat and mouse: regulated alternative promoter usage and differential splicing

A central conclusion of these studies is that the expression pattern of apobec-1 mRNA in both the rat and mouse is a product of differential promoter usage and multiple transcription initiation sites, coupled with alternative splicing. The present findings demonstrate that there are at least three clusters of transcription initiation sites in both the rat and mouse *apobec-1* gene. This represents an important departure from the data of Nakamuta and colleagues (17). These authors found evidence of transcription initiation sites in exon A and in the distal region of exon 1, the latter suggested to be dominant in the mouse small intestine (17). In addition, however, these authors demonstrated the presence of a 102 nt sequence at the 5′ end of the ''intestinal-type'' mRNA which was not detectable in the ''liver-type'' species (see Figs. 5 and 7 of reference 17). Our results suggest an explanation for this observation, but do not support the contention that this sequence is intestine-specific. As summarized in Fig. 5 of the present study, our data suggest the presence of both distal and proximal transcription start sites within exon 1. Indeed, evidence for a cluster of proximal start sites was found for both rat and mouse *apobec-1* genes. Accordingly, we propose that the 102 nt sequence[3] alluded to in the

[3]Owing to differences in allocation of the distal transcription initiation site in exon 1, between the current study and that of Nakamuta and colleagues (17), we estimate that mouse exon 1a is 97 nt in length (Fig. 5).

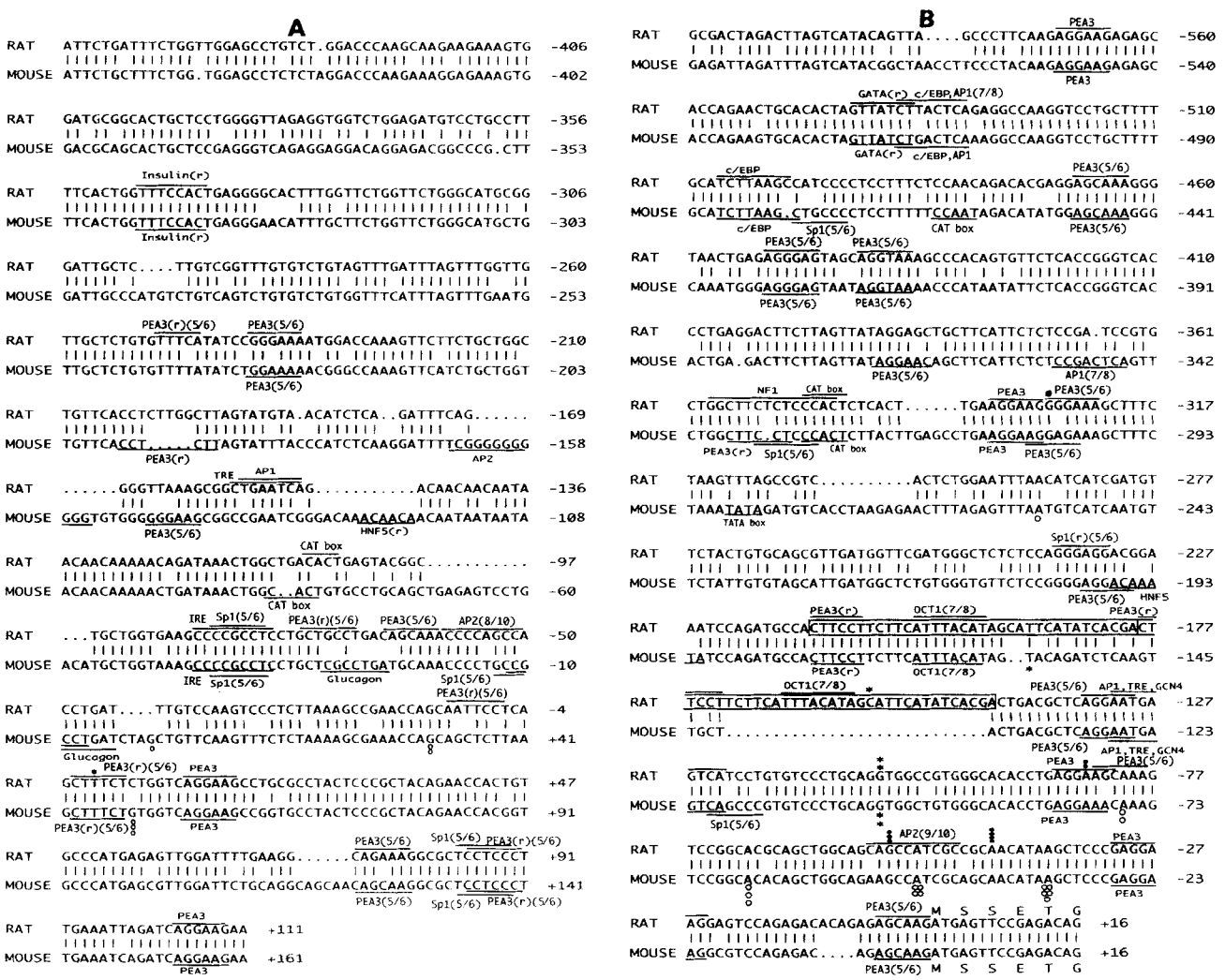

Fig. 9. Sequence alignment of mouse and rat apobec-1 gene flanking regions. A: 5′ flanking DNA sequence, upstream of and including the first untranslated exon (exon A). The locations of potential transcription factor binding sites are indicated by a solid line. The solid circle indicates the major transcription start site of rat exon A. The open circles indicate the different transcription start sites identified in exon A of the mouse *apobec-1* gene. B: 5′ flanking region upstream of exon 1 in the rat and mouse *apobec-1* gene. The locations of potential transcription factor binding sites are indicated by a solid line. A direct repeat of 35 nucleotides in the rat *apobec-1* gene is boxed. The solid circles indicate the different transcription start sites in rat exon 1. The open circles indicate the different transcription start sites identified in exon 1 of the mouse *apobec-1* gene. The asterisks indicate the sites of alternative splice acceptor sites.

study of Nakamuta and colleagues (17) corresponds to exon 1a, and results from utilization of the distal transcription start site in mouse exon 1. In particular, the data establish the presence of exon 1a in rat kidney, colon, heart, and ovary in addition to the small intestine (Fig. 6B). Our findings thus suggest that this untranslated region of exon 1 is transcribed in multiple tissues of both the rat (data above, Fig. 6) and mouse (K. Hirano and N.O. Davidson, unpublished observations) *apobec-1* genes. By contrast, the data suggest that apobec-1 mRNA originating from the proximal transcription start site in exon 1 is confined to the small intestine and colon in the rat.

We confirmed the presence of alternative splice acceptor sites within exon 1 of both the mouse and rat *apobec-1* genes, although, as suggested by Nakamuta and colleagues (17), the major splice variants account for more than 90% of the population of apobec-1 transcripts found in both species. In addition to these previously described splice variants, minor populations of mRNA (~5–10% total) were found in both the rat and mouse that arise from a novel, exon 2-skipped, splice variant which generates a frame-shift and produces a premature stop codon. As a result, it is anticipated that the variant mRNA species would not encode a functional protein, but firm conclusions on this point must await formal examination. Our recent results indicate that ~50% of the total population of apobec-1 mRNA

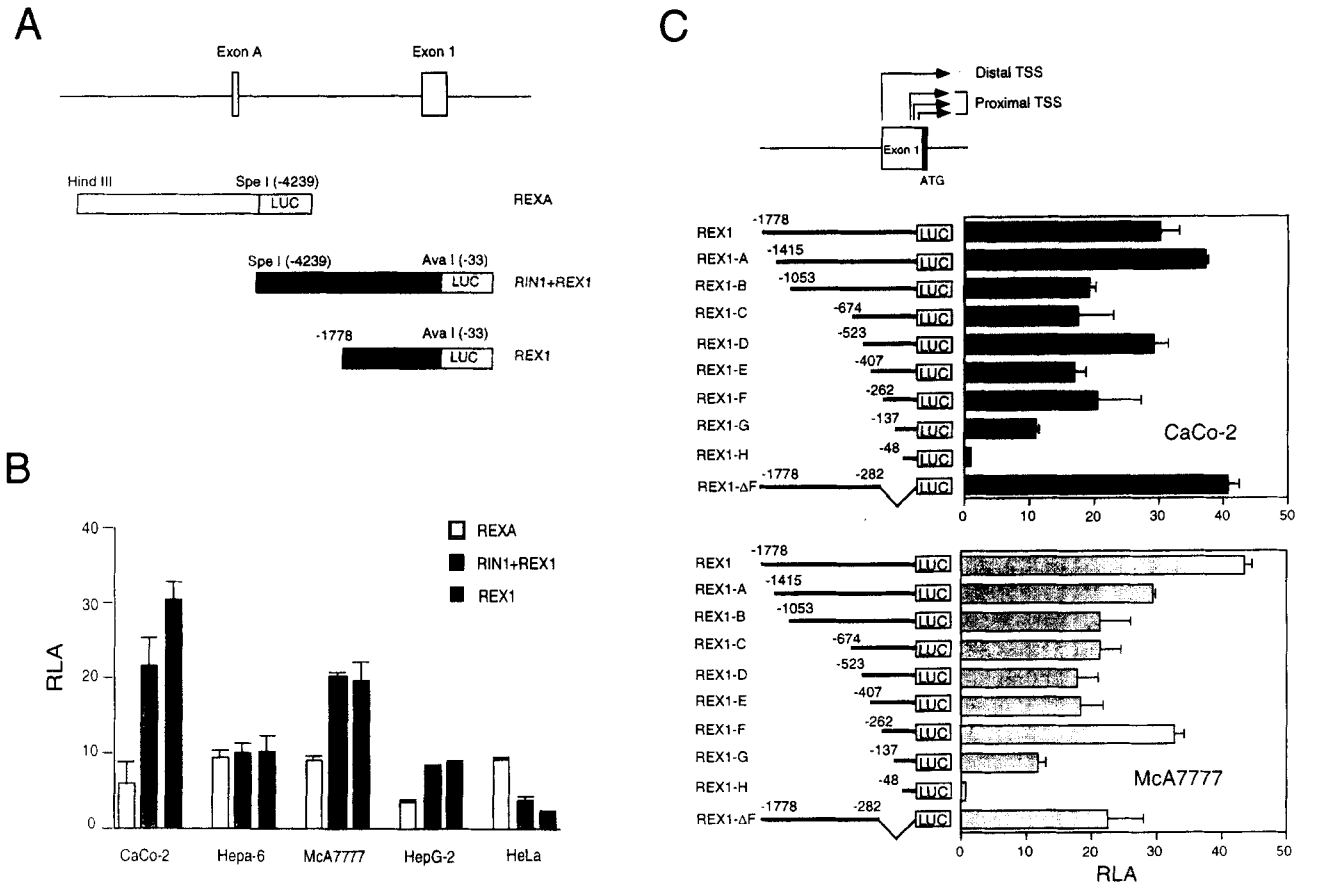

Fig. 10. Demonstration of promoter activity of rat *apobec-1*-luciferase fusion constructs. A: Schematic presentation of rat *apobec-1* gene illustrating 5′ flanking regions of exon A and exon 1 and the three constructs used to examine promoter activity. REXA contains ~3.0 kb of upstream sequence; RIN1 + REX1 contains a continuous region from −4239 to −33 (intron 1 and exon 1) while REX1 contains the proximal region of intron 1, from −1778 to −33. LUC, luciferase reporter gene. B: Reporter gene constructs were cotransfected into the indicated cell lines along with pCMV β-Gal as an internal control and luciferase activity determined 48 h later. The data (mean ± SD) are expressed as relative luciferase activity (RLA), normalized to the expression of β-Gal. For the purposes of comparison, luciferase activity in cells transfected with pGL-2 Basic is assigned a value of 1.0. C: Distinct promoter regions within intron A and exon 1 of the rat apobec-1 gene function in the regulation of intestinal expression. The parental construct REX1 was used as a starting point for serial deletions using exonuclease III digestion, resulting in the production of constructs REX1-A through REX1-H, each of which was sequenced. REX1-ΔF was constructed with an internal deletion from nucleotides −281 to −33 and religated into the original plasmid. These constructs were transfected into confluent Caco-2 and McA7777 cells and luciferase activity measured 48 h later. The data (mean ± SD of triplicate transfections) are expressed as relative luciferase activity (RLA), normalized to the expression of β-Gal, as detailed above. For the purposes of comparison, luciferase activity in cells transfected with pGL-2 Basic is assigned a value of 1.0. The location of transcription initiation sites within exon 1 is indicated schematically at the top of the figure as a reference point for the constructions examined.

in the human small intestine undergoes exon 2 skipping. The splice variant mRNA species is predicted to encode a novel truncated peptide of 36 residues which is detectable by immunocytochemical staining within small intestinal enterocytes (33). Conceivably, the modulated pattern of promoter usage and alternative splicing could function in the physiological regulation of apobec-1 gene expression and this possibility merits further attention. While yet untested, such a possibility could potentially provide a mechanism to limit the expression of apobec-1. Considered in the light of recent studies from Yamanaka and colleagues (53), which indicate that forced overexpression of apobec-1 in the liver of transgenic mice and rabbits is associated with the development of hepatomas, it is tempting to speculate that adaptive mechanisms may have evolved to restrict the expression of apobec-1.

The current studies also provide important preliminary information concerning the regulation of apobec-1 transcription in the rat. The data demonstrate that an unrestricted pattern of luciferase expression was promoted by a construction containing the region flanking exon A, a finding consistent with the widespread expression of apobec-1 mRNA in the rat (16). However, the ability of constructions containing ~1.8 kb flanking exon 1 to promote luciferase activity in McA7777 cells was somewhat

unexpected, as the evidence from RNase protection assay (Fig. 6) implies that the major population of apobec-1 mRNA arising in the rat liver is transcribed from exon A. Among the potential explanations for this observation is the absence of liver-specific repressor elements from the constructions used. While the use of constructions containing larger flanking regions upstream of exon A will help in the resolution of this issue, it bears emphasis that the interpretation of such transient transfection assays requires caution. In addition, while a defined region from the putative proximal promoter region of exon 1, containing 262 nt of flanking sequence upstream of the initiator codon, appears to contain a minimal functional promoter active in both Caco-2 and McA7777 cells, further work will be required to delineate the mechanisms responsible for intestine-specific expression. The evolutionary and functional significance of these regulatory motifs will be the focus of future reports.

This work was supported by grants from the National Institutes of Health (HL-38180, DK-42086, and HL-18577) to NOD.

Manuscript received 23 December 1996 and in revised form 21 February 1997.

REFERENCES

1. Young, S. G. 1990. Recent progress in understanding apolipoprotein B. *Circulation.* **82:** 1574–1594.
2. Scott, J., N. Navaratnam, S. Bhattacharya, and J. R. Morrison. 1994. The apolipoprotein B messenger RNA editing enzyme. *Curr. Opin. Lipidol.* **5:** 87–93.
3. Davidson, N. O., S. Anant, and A. J. MacGinnitie. 1995. Apolipoprotein B messenger RNA editing: insights into the molecular regulation of post-transcriptional cytidine deamination. *Curr. Opin. Lipidol.* **6:** 70–74.
4. Greeve, J., I. Altkemper, J. H. Dieterich, H. Greten, and E. Windler. 1993. Apolipoprotein B mRNA editing in 12 different mammalian species: hepatic expression is reflected in low concentrations of apoB-containing plasma lipoproteins. *J. Lipid Res.* **34:** 1367–1383.
5. Herz, J., and T. E. Willnow. 1995. Lipoprotein and receptor interactions in vivo. *Curr. Opin. Lipidol.* **6:** 97–103.
6. McCormick, S. P., J. K. Ng, S. Taylor, L. M. Flynn, R. E. Hammer, and S. G. Young, 1995. Mutagenesis of the human apolipoprotein B gene in a yeast artificial chromosome reveals the site of attachment for apolipoprotein[a]. *Proc. Natl. Acad. Sci. USA.* **92:** 10147–10151.
7. Callow, M. J., and E. M. Rubin. 1995. Site-specific mutagenesis demonstrates that cysteine 4326 of apolipoprotein B is required for covalent linkage with apolipoprotein [a] in vivo. *J. Biol. Chem.* **270:** 23914–23917.
8. Harris, S. G., I. Sabio, E. Mayer, M. F. Steinberg, J. W. Backus, J. D. Sparks, C. E. Sparks, and H. C. Smith. 1993. Extract-specific heterogeneity in high-order complexes containing apolipoprotein B mRNA editing activity and RNA-binding proteins. *J. Biol. Chem.* **268:** 7382–7392.
9. Teng, B., C. F. Burant, and N. O. Davidson. 1993. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science.* **260:** 1816–1819.
10. Navaratnam, N., J. R. Morrison, S. Bhattacharya, D. Patel, T. Funahashi, F. Giannoni, B. B. Teng, N. O. Davidson, and J. Scott. 1993. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J. Biol. Chem.* **268:** 20709–20712.
11. Navaratnam, N., S. Bhattacharya, T. Fujino, D. Patel, A. L. Jarmuz, and J. Scott. 1995. Evolutionary origins of apoB mRNA editing: catalysis by a cytidine deaminase that has acquired a novel RNA-binding motif at its active site. *Cell.* **81:** 187–195.
12. Yamanaka, S., K. S. Poksay, M. E. Balestra, G. Q. Zeng, and T. L. Innerarity. 1994. Cloning and mutagenesis of the rabbit apoB mRNA editing protein. A zinc motif is essential for catalytic activity, and noncatalytic auxiliary factor(s) of the editing complex are widely distributed. *J. Biol. Chem.* **269:** 21725–21734.
13. Driscoll, D. M., and Q. Zhang. 1994. Expression and characterization of p27, the catalytic subunit of the apolipoprotein B mRNA editing enzyme. *J. Biol. Chem.* **269:** 19843–19847.
14. MacGinnitie, A. J., S. Anant, and N. O. Davidson. 1995. Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity, *J. Biol. Chem.* **270:** 14768–14775.
15. Anant, S., A. J. MacGinnitie, and N. O. Davidson. 1995. apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, is a novel RNA-binding protein. *J. Biol. Chem.* **270:** 14762–14767.
16. Funahashi, T., F. Giannoni, A. M. DePaoli, S. F. Skarosi, and N. O. Davidson, 1995. Tissue-specific, developmental and nutritional regulation of the gene encoding the catalytic subunit of the rat apolipoprotein B mRNA editing enzyme: functional role in the modulation of apoB mRNA editing. *J. Lipid Res.* **36:** 414–428.
17. Nakamuta, M., K. Oka, J. Krushkal, K. Kobayashi, M. Yamamoto, W. H. Li, and L. Chan. 1995. Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (Apobec1) gene in mice. Structure and evolution of Apobec1 and related nucleoside/nucleotide deaminases. *J. Biol. Chem.* **270:** 13042–13056.
18. Hadjiagapiou, C., F. Giannoni, T. Funahashi, S. F. Skarosi, and N. O. Davidson. 1994. Molecular cloning of a human small intestinal apolipoprotein B mRNA editing protein. *Nucleic Acids Res.* **22:** 1874–1879.
19. Lau, P. P., H. J. Zhu, A. Baldini, C. Charnsangavej, and L. Chan. 1994. Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc. Natl. Acad. Sci. USA.* **91:** 8522–8526.
20. Wu, J. H., C. F. Semenkovich, S. H. Chen, W. H. Li, and L. Chan. 1990. Apolipoprotein B mRNA editing. Validation of a sensitive assay and developmental biology of RNA editing in the rat. *J. Biol. Chem.* **265:** 12312–12316.
21. Higuchi, K., K. Kitagawa, K. Kogishi, and T. Takeda. 1992. Developmental and age-related changes in apolipoprotein B mRNA editing in mice. *J. Lipid Res.* **33:** 1753–1764.
22. Davidson, N. O., L. M. Powell, S. C. Wallis, and J. Scott, 1988. Thyroid hormone modulates the introduction of a stop codon in rat liver apolipoprotein B messenger RNA. *J. Biol. Chem.* **263:** 13482–13485.
23. Davidson, N. O., R. C. Carlos, and A. M. Lukaszewicz. 1990. Apolipoprotein B mRNA editing is modulated by thyroid hormone analogs but not growth hormone administration in the rat. *Mol. Endocrinol.* **4:** 779–785.
24. Thorngate, F. E., R. Raghow, H. G. Wilcox, C. S. Werner,

M. Heimberg, and M. B. Elam. 1994. Insulin promotes the biosynthesis and secretion of apolipoprotein B-48 by altering apolipoprotein B mRNA editing. *Proc. Natl. Acad. Sci. USA.* **91:** 5392–5396.
25. Baum, C. L., B. B. Teng, and N. O. Davidson. 1990. Apolipoprotein B messenger RNA editing in the rat liver. Modulation by fasting and refeeding a high carbohydrate diet. *J. Biol. Chem.* **265:** 19263–19270.
26. Lau, P. P., D. J. Cahill, H. J. Zhu, and L. Chan. 1995. Ethanol modulates apolipoprotein B mRNA editing in the rat. *J. Lipid Res.* **36:** 2069–2078.
27. Inui, Y., F. Giannoni, T. Funahashi, and N. O. Davidson. 1994. REPR and complementation factor(s) interact to modulate rat apolipoprotein B mRNA editing in response to alterations in cellular cholesterol flux. *J. Lipid Res.* **35:** 1477–1489.
28. Hirano, K., S. G. Young, R. V. Farese, Jr., J. Ng, E. Sande, C. Warburton, L. M. Powell-Braxton, and N. O. Davidson. 1996. Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J. Biol. Chem.* **271:** 9887–9890.
29. Morrison, J. R., C. Paszty, M. E. Stevens, S. D. Hughes, T. Forte, J. Scott, and E. M. Rubin. 1996. Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. *Proc. Natl. Acad. Sci. USA.* **93:** 7154–7159.
30. Nakamuta, M., B. Hung-Junn Chang, E. Zsigmond, K. Kobayashi, H. Lei, B. Y. Ishida, K. Oka, E. Li, and L. Chan. 1996. Complete phenotypic characterization of apobec-1 knockout mice with a wild-type genetic background and a human apolipoprotein B transgenic background, and restoration of apolipoprotein B mRNA editing by somatic gene transfer of apobec-1. *J. Biol. Chem.* **271:** 25981–25988.
31. Giannoni, F., F. J. Field, and N. O. Davidson. 1994. An improved reverse transcription-polymerase chain reaction method to study apolipoprotein gene expression in Caco-2 cells. *J. Lipid Res.* **35:** 340–350.
32. Osuga, J., T. Inaba, K. Harada, H. Yagyu, M. Shimada, Y. Yazaki, N. Yamada, and S. Ishibashi. 1995. Cloning and structural analysis of the mouse apolipoprotein B mRNA editing protein gene. *Biochem. Biophys. Res. Commun.* **214:** 653–662.
33. Hirano, K., J. Min, T. Funahashi, D. A. Baunoch, and N. O. Davidson. 1997. Characterization of the human *apobec-1* gene: expression in gastrointestinal tissues determined by alternative splicing with production of a novel truncated peptide. *J. Lipid Res.* **38:** 847–859.
34. Martin, M. E., J. Piette, M. Yaniv, W. J. Tang, and W. R. Folk. 1988. Activation of the polyomavirus enhancer by a murine activator protein 1 (AP1) homolog and two contiguous proteins. *Proc. Natl. Acad. Sci. USA.* **85:** 5839–5843.
35. Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. Leprince, and D. Stehelin. 1990. The c-ets proto-oncogenes encode transcription factors that cooperate with c-Fos and c-Jun for transcriptional activation. *Nature.* **346:** 191–193.
36. Mitchell, P. J., C. Wang, and R. Tjian. 1987. Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. *Cell.* **50:** 847–861.
37. Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell.* **49:** 741–752.
38. Schmidt, M. C., Q. Zhou, and A. J. Berk. 1989. Sp1 activates transcription without enhancing DNA-binding activity of the TATA box factor. *Mol. Cell Biol.* **9:** 3299–3307.
39. Knepel, W., L. Jepeal, and J. F. Habener. 1990. A pancreatic islet cell-specific enhancer-like element in the glucagon gene contains two domains binding distinct cellular proteins. *J. Biol. Chem.* **265:** 8725–8735.
40. Ohlsson, H., and T. Edlund. 1986. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell.* **45:** 35–44.
41. Nasrin, N., L. Ercolani, M. Denaro, X. F. Kong, I. Kang, and M. Alexander. 1990. An insulin response element in the glyceraldehyde-3-phosphate dehydrogenase gene binds a nuclear protein induced by insulin in cultured cells and by nutritional manipulations in vivo. *Proc. Natl. Acad. Sci. USA.* **87:** 5273–5277.
42. Barnhart, K. M., C. G. Kim, and M. Sheffery. 1989. Purification and characterization of an erythroid cell-specific factor that binds the murine alpha- and beta-globin genes. *Mol. Cell Biol.* **9:** 2606–2614.
43. Costa, R. H., D. R. Grayson, K. G. Xanthopoulos, and J. E. Darnell, Jr. 1988. A liver-specific DNA-binding protein recognizes multiple nucleotide sites in regulatory regions of transthyretin, alpha 1-antitrypsin, albumin, and simian virus 40 genes. *Proc. Natl. Acad. Sci. USA.* **85:** 3840–3844.
44. Courtois, S. J., D. A. Lafontaine, F. P. Lemaigre, S. M. Durviaux, and G. G. Rousseau. 1990. Nuclear factor-I and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and *trans*-activate the human growth hormone gene. *Nucleic Acids Res.* **18:** 57–64.
45. Grange, T., J. Roux, G. Rigaud, and R. Pictet. 1991. Cell-type specific activity of two glucocorticoid responsive units of rat tyrosine aminotransferase gene is associated with multiple binding sites for C/EBP and a novel liver-specific nuclear factor. *Nucleic Acids Res.* **19:** 131–139.
46. Arndt, K., and G. R. Fink. 1986. GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′ TGACTC 3′ sequences. *Proc. Natl. Acad. Sci. USA.* **83:** 8516–8520.
47. Ho, I. C., N. K. Bhat, L. R. Gottschalk, T. Lindsten, C. B. Thompson, T. S. Papas, and J. M. Leiden. 1990. Sequence-specific binding of human Ets-1 to the T cell receptor alpha gene enhancer. *Science.* **250:** 814–818.
48. Giannoni, F., D. K. Bonen, T. Funahashi, C. Hadjiagapiou, C. F. Burant, and N. O. Davidson. 1994. Complementation of apolipoprotein B mRNA editing by human liver accompanied by secretion of apolipoprotein B-48. *J. Biol. Chem.* **269:** 5932–5936.
49. Penn, R. B., and J. L. Benovic. 1994. Structure of the human gene encoding the beta-adrenergic receptor kinase. *J. Biol. Chem.* **269:** 14924–14930.
50. Jbilo, O., J. P. Toutant, K. P. Vatsis, A. Chatonnet, and O. Lockridge. 1994. Promoter and transcription start site of human and rabbit butyrylcholinesterase genes. *J. Biol. Chem.* **269:** 20829–20837.
51. Shih, H. M., Z. Liu, and H. C. Towle. 1995. Two CACGTG motifs with proper spacing dictate the carbohydrate regulation of hepatic gene transcription. *J. Biol. Chem.* **270:** 21991–21997.
52. Towle, H. C. 1995. Metabolic regulation of gene transcription in mammals. *J. Biol. Chem.* **270:** 23235–23238.
53. Yamanaka, S., M. E. Balestra, L. D. Ferrell, J. Fan, K. S. Arnold, S. Taylor, J. M. Taylor, and T. L. Innerarity. 1995. Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc. Natl. Acad. Sci. USA.* **92:** 8483–8487.