Structure and Expression of the ATFa Gene*

Jean Goetz, Bruno Chatton, Marie-Geneviève Mattei, and Claude Kedinger§

From the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/Université Louis Pasteur, F-67404 Illkirch Cedex Communauté Urbaine de Strasbourg, France and the 3Faculté de Médecine de la Timone, INSERM U-406, Boulevard Jean Moulin, 13385 Marseille Cedex, France

The human ATFa proteins belong to the ATF/CREB family of transcription factors. We have previously shown that they mediate the transcriptional activation by the largest E1a protein and can heterodimerize with members of the Jun/Fos family. ATFa proteins have also been found tightly associated with JNK2, a stress-activated kinase. We now report on the structure of the ATFa gene, which mapped to chromosome 12 (band 12q13). Sequence analysis revealed that ATFa isoforms are generated by alternative splice donor site usage. A minimal promoter region of ~200 base pairs was identified that retained nearly full transcriptional activity. Binding sites for potential transcription factors were delineated within a GC-rich segment by DNase I footprinting. Expression studies revealed that ATFa accumulates in the nuclei of transfected cells, and the nuclear localization signal was defined next to the leucine zipper domain. As revealed by hybridization with mouse ATFa sequences, low levels of ATFa mRNAs were ubiquitously distributed in fetal or adult mice, with enhanced expression in particular tissues, like squamous epithelia and specific brain cell layers. The possible significance of coexpression of ATFa, ATF-2, and Jun at similar sites in the brain is discussed.

Gene expression in eukaryotic cells is regulated by sequence-specific binding of transcription factors to cis-acting DNA elements. One example of such regulatory elements is the ATFa/CRE motif (TGACGCTCA), which was originally identified in several adenovirus promoters as the ATF-binding site and in some cellular promoters as the CRE. Many cDNAs encoding ATF/CRE-binding proteins have now been cloned (for reviews, see Refs. 1–3); the binding and trans-activation properties of all these proteins have been extensively studied. They belong to the family of "b-Zip" (basic-leucine zipper) proteins, which have a basic region in their C termini, rich in positively charged amino acids, flanked by a leucine zipper motif, and responsible both for sequence-specific DNA binding and protein dimerization (4, 5). Many of these ATFa/CREB factors have been found heterodimerized with members of the AP1 family (6–10), giving new combinations of factors with different binding capacities, which may increase their response pattern. The ATFa/CRE-binding proteins include factors implicated in the cAMP response pathway (CREB, CREM, and ATF-1) (3) as well as closely related factors that mediate promoter induction by the largest product of the adenovirus E1a oncogene, CRE-BP1 (also called ATFa) (11–13), and the ATFa proteins (14, 15).

Four variants of ATFa (ATFa0, ATFa1, ATFa2, and ATFa3) have been characterized that have indistinguishable properties and differ only by short peptide motifs (15, 16). Although these ATFa proteins had no detectable intrinsic trans-activation function, a transcriptional activation domain was identified after removal of C-terminal sequences. This cryptic activation domain comprises an essential N-terminal zinc-binding element, which, together with sequences located within the 170 C-terminal residues of the ATFa proteins, also contributes to the interaction of ATFa with the E1a oncprotein. As expected, both N- and C-terminal ATFa elements are instrumental in mediating E1a responsiveness of target promoters (15). In vitro and in vivo experiments have demonstrated that the ATFa proteins can associate with c-Jun or c-Fos proteins and that they bind TPA response element sequences only when heterodimerized with members of the Jun family (10, 17).

We also found that ATFa proteins were tightly associated with protein kinase activities and that this binding occurred both in vitro and in a cellular environment. We have shown that these kinase activities have properties related to specific members of the mitogen-activated protein kinase family (18). One of these kinases is the c-Jun N-terminus-associated kinase 2 (JNK2), a kinase that binds to N-terminal regions of ATF-2 and ATFa (18–21). It has been suggested that ATFa plays important roles in the early events of signaling pathways, as phosphorylation of this factor by the c-Jun N-terminus-associated kinases increases its trans-activating properties and permits the early activation of the c-Jun promoter (22). Despite their structural similarity to ATF-2, it is presently unclear whether the ATFa proteins serve similar functions in the cell, particularly because they display no transcriptional activity on their own.

As a step toward understanding the cellular functions of the ATFa proteins and their connections with the ATF-2 and c-Jun factors, we undertook the cloning of the genomic ATFa sequences and the study of their expression. The structure and chromosomal location of the ATFa gene and promoter were determined. Our results indicate that the different ATFa isoforms arise by alternative splicing through differential splice donor usage and that these proteins are targeted to the cell nucleus by a bipartite nuclear localization sequence that is part of the basic region of the protein. While low levels of ATFa

* This work was supported by grants and/or fellowships (to J.G.) from CNRS, INSERM, the Centre Hospitalier Universitaire Régional de Strasbourg, the Human Frontier Science Program, the Association pour la Recherche sur le Cancer, and the Ligue Nationale contre le Cancer.

§To whom correspondence should be addressed: IGBMC, B.P. 163, 67404 Illkirch, France. Tel.: 33-88-65-34-46; Fax: 33-88-65-32-01; E-mail: kedinger@igbmc.u-strasbg.fr.

1 The abbreviations used are: ATFa, activating transcription factor; CRE, cAMP-responsive element; CREB, cAMP-responsive element-binding protein; CRE-BP1, cAMP-responsive element-binding protein 1; AP1, activating protein 1; bp, base pair(s); kb, kilobase pair(s); PCR, polymerase chain reaction; CAT, chloramphenicol acetyltransferase; GST, glutathione S-transferase; NLS, nuclear localization signal; PBS, phosphate-buffered saline; b-Zip, basic leucine zipper.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) Y07712.

This paper is available on line at http://www-jbc.stanford.edu/jbc/29589
The ATFa Gene

mRNAs appeared to be expressed rather ubiquitously in the adult mouse, enhanced expression was observed in particular tissues during embryonic development.

**Experimental Procedures**

**Human Gene Mapping**—In situ hybridization was carried out on chromosome preparations obtained from phytohemagglutinin-stimulated human lymphocytes. A recombinant bacteriophage clones containing an ATFa cDNA fragment (1449 bp) was tritium-labeled by nick translation to a specific activity of ~1.6 × 10⁶ dpm/µg and used as probe. The hybridization procedure was carried out essentially as described previously (23).

**Genomic Cloning**—A human placental genomic DNA library (constructed by J. Capel under the auspices of DSM 121) was screened by hybridization with a random-primed 32P-labeled ATFa cDNA probe using conventional cloning methods (24). DNA inserts were recovered from positive recombinant phages by restriction digestion and subcloned into the pBluescript SK vector. Nucleotide sequencing was performed by the chain termination method (68) using synthetic oligonucleotide primers derived from appropriate regions of the known ATFa cDNA sequence.

**Mouse ATFa cDNA Fragment Cloning**—A 300-bp cDNA fragment was amplified by reverse transcription-PCR from poly(A)⁺ RNA from mouse P19 cells (25). The oligonucleotide primers used for the reaction bracketed a region unique to ATFa (when compared with other members of the family), spanning positions +803 to +1102 of the human ATFa cDNA and located just 5' to the basic-leucine zipper motif (b-Zip). The resulting 300-bp fragment was inserted into the EcoRV site of pBluescript and sequenced, revealing a 92% identity to its human homolog at the nucleotide level.

**Recombinant Plasmid Construction**—Subcloning of the ATFa gene promoter region was performed as follows. An 11-kb SacI insert was recovered from a positive recombinant phage and cloned into pBluescript SK. A 6.5-kb ApaI fragment containing the first exon, upstream sequences, and part of the first intron of the ATFa gene was directly subcloned into the Apal site of pBluescript SK. This recombinant plasmid was further processed by excising an XbaI fragment comprising the most 5'-portion of the ATFa gene upstream sequences and was mutated by oligonucleotide-directed mutagenesis to remove the intronic sequences and to create a SacI restriction site at the end of the first exon. A 2-kb SacI fragment containing the first exon and upstream sequences of the ATFa gene (between positions −197 and +83, with respect to the transcription start site) was then subcloned, in both orientations, in front of the chloramphenicol acetyltransferase (CAT) reporter gene of a promoterless vector (pBLCAT6) (26), generating the −197/+83 CAT and +83/−197 CAT recombinants. From the −197/+83 CAT recombinant, a series of additional deletions were constructed by restriction digestion (giving rise to −82/+83 CAT and −212/+83 CAT), by PCR amplification of specific promoter fragments followed by subcloning into the same host vector (giving rise to −23/+83 CAT and +8/+83 CAT), or by progressive exonuclease III-directed unilateral deletions of upstream sequences between positions −197 and −797 (generating −178/+83 CAT, −191/+83 CAT, −1467/+83 CAT, −1211/+83 CAT, −1051/+83 CAT, and −797/+83 CAT). The above-mentioned 2-kb SacI fragment was also subcloned in both orientations into a blunt-ended XbaI site at position −9 of a rabbit β-globin reporter gene of the PAL4 vector (27), generating the −197/+83 Glob and +83/−197 Glob recombinants.

**Immunoblot (Western Blot) Analysis**—Whole cell extracts from adult mouse tissues and nuclear extracts from 11.5–17.5-day-old mouse embryos were prepared (33, 34). After SDS-polyacrylamide gel electrophoresis, the ATFa proteins were revealed by immunoblotting essentially as described (10) with a monoclonal antibody directed against bacterially expressed ATFa3 (monoclonal antibody 1A7; used at 1:5000). Specific protein-antibody complexes were revealed using a peroxidase-conjugated rabbit anti-mouse IgG secondary antibody (Jackson ImmunoResearch Laboratories, Inc.) followed by treatment with the ECL detection system (Amersham Corp.) according to the manufacturer's instructions.

**In Situ Hybridization**—The organs collected from adult mice were frozen in 2-methylbutane maintained on dry ice. A 300-bp mouse ATFa cDNA fragment cloned into pBluescript (see above) was transcribed in vitro using T7 RNA polymerase in the presence of [35S]-CTP to generate an antisense riboprobe. Hybridization was carried out on cryosections as described (35), providing a higher signal-to-noise ratio, although with a somewhat poorer histology. Emulsion autoradiography was for 4 months.

**Immunocytochemistry**—Monolayer COS-7 cells (grown on glass coverslips) were transfected by calcium phosphate coprecipitation with 1 µg of recombinant vector or pBc parental vector, together with 3.5 µg of double-stranded pBluescript carrier DNA tube. The medium was changed after 20 h, and the day after, the cells were washed in PBS and treated for 4 min at room temperature with 2% formaldehyde in PBS. After fixation, the cells were permeabilized by two treatments with PT buffer (PBS, 0.1% Triton X-100) for 10 min. After a 30-minute treatment with blocking solution (2% nonfat dry milk in PBS), the cells were incubated for 1 h at room temperature with an anti-GST primary monoclonal antibody (EuroMedex, Paris, France) diluted in PT buffer, washed three times in PT buffer, incubated with a Cy3-conjugated goat anti-mouse secondary antibody for 1 h at room temperature, and washed in PT buffer. The preparations were mounted in glycerol/PBS (4:1) containing 5% propyl gallate, and observations were made with a fluorescence microscope (Nikon Microphot FXA).
The ATFa Gene—A human genomic library was screened at elevated stringency with the largest cDNA of the ATFa family. A number of positive clones were picked, characterized by Southern blot analysis, and subcloned. Partial sequence determination of the genomic clones and comparison with the previously established cDNA sequence (14) allowed us to position the introns and to establish the gene structure. As depicted in Fig. 2, the ATFa gene (~60 kb in length) comprises a leader (A) and 11 exons (B–L), all of which are <200 bp, except exon L, which exceeds 8000 bp and contains a large untranslated region of nearly 8 kb. The translation initiation and termination codons of the ATFa open reading frame are in exons B and L, respectively. The lengths of the 11 introns have not been precisely determined, but were estimated to have a mean size of ~4 kb each. The sequences at the exon/intron junctions were in very good agreement with the consensus sequences around both splice donor and acceptor sites (Fig. 3A).

Four ATFa cDNA variants were originally isolated from human cDNA libraries (14–16), differing from each other by specific sequence elements within the coding region (Fig. 3B): ATFa3 corresponds to the largest cDNA resulting from the juxtaposition of all elements boxed in Fig. 2; ATFa2 lacks a 33-bp element (element d); ATFa1 lacks both element d and an additional 63-bp element (element e); and ATFa0 lacks element d and exons P–I. The existence of cytoplasmic transcripts corresponding to these ATFa variants was confirmed by PCR-mediated analysis of HeLa cell poly(A)+ RNA (16), demonstrating the physiological relevance of these variants. These observations indicate that the various isoforms most likely result from differential splicing of a unique ATFa primary transcript. Thus, exons P–I correspond to alternative exons. Furthermore, examination of the nucleotide sequences at the 5’-border of elements d and e revealed strong homologies to the consensus splice donor sequence (Fig. 3C), suggesting that these sequences constitute alternative splice donor sites (DS1) that can be used in place of the donor sites located at the 3’-border of these elements (DS2). Nothing is known at present about the control of this splice donor site usage and the underlying mechanism.

The Promoter of the ATFa Gene—As a first step toward the delineation of the promoter region of the ATFa gene, we determined the transcription start sites. An RNase protection experiment performed with HeLa cell poly(A)+ RNA (Fig. 4B) clearly identified one minor and two major start sites, arbitrarily denoted −43, −22, and +1, respectively, within the 5’-region of the ATFa leader (exon A). Transcription initiation around position −22 seems to be less precise, as indicated by the slight scattering of the start sites in this area. The position of this initiation region is compatible with the coordinates of the most 5’-extending ATFa cDNA originally isolated (14), starting at position +8.

Examination of the nucleotide sequences located upstream of the transcription start sites did not reveal a typical TATA box. Instead, a significantly higher GC content (70%) was apparent between positions +1 and −140, which, together with the relative spreading of initiation sites, is a characteristic of promoters of housekeeping genes (37). To test for the presence of promoter sequences within the sequences located upstream of the ATFa leader, a vector harboring ATFa sequences between positions −1917 and +83 was transfected into HeLa cells. To ensure stable RNA synthesis, the ATFa sequences were linked to a promotereless rabbit β-globin gene. Specific transcription was investigated by S1 nuclease mapping and reverse transcriptase primer extension of total RNA prepared from the transfected HeLa cells. As shown in Fig. 4C, transcripts that

---

**FIG. 1. Metaphase chromosome analysis by in situ hybridization.** Two partial human metaphase spreads with the specific site of hybridization of the ATFa cDNA probe to chromosome 12 are shown on the left. Arrowheads indicate silver grains on Giemsa-stained chromosomes after autoradiography (top); chromosomes with silver grains were subsequently identified by R-banding (bottom). An ideogram of the human G-banded chromosome 12 is shown on the right. It illustrates the chromosomal distribution of labeled sites as deduced from quantitative analysis of autoradiograms.

**FIG. 2. Structure of the human gene encoding the ATFa proteins.** The gene structure, as deduced from partial nucleotide sequence analysis of inserts from positive genomic clones (aligned below, with respective lengths in kb), is schematically depicted. Exons (A–L) and two alternative exonic sequences (d and e) are represented as open and closed boxes, respectively, with corresponding lengths (in bp) above (exon L has not been sequenced entirely, but its size has been estimated electrophoretically). Transcription initiation sites, as determined in Fig. 4, are indicated. The coordinates (with respect to the + 1 transcription start site) of the translation initiation and termination codons as well as those of the extremities of each exon within the largest cDNA (ATFa3), are also given.

**RESULTS**

**Chromosomal Localization of the ATFa Gene**—To determine the genomic localization of the ATFa sequences, in situ hybridizations on human metaphase chromosomes were performed (Fig. 1). In the 100 metaphase cells examined after hybridization, 199 silver grains were associated with chromosomes, and 67 of these (33.6%) were located on chromosome 12. The grain distribution was not random since 89.5% (60/67) of the grains counted on chromosome 12 mapped to position 12q12–14, with a maximum in band 12q13. It is noteworthy that the gene encoding the related ATF-1 factor (36) also maps within this locus.

**The ATFa Gene**—A human genomic library was screened at elevated stringency with the largest cDNA of the ATFa family. A number of positive clones were picked, characterized by Southern blot analysis, and subcloned. Partial sequence determination of the genomic clones and comparison with the previously established cDNA sequence (14) allowed us to position the introns and to establish the gene structure. As depicted in Fig. 2, the ATFa gene (~60 kb in length) comprises a leader (A) and 11 exons (B–L), all of which are <200 bp, except exon L, which exceeds 8000 bp and contains a large untranslated region of nearly 8 kb. The translation initiation and termination codons of the ATFa open reading frame are in exons B and L, respectively. The lengths of the 11 introns have not been precisely determined, but were estimated to have a mean size of ~4 kb each. The sequences at the exon/intron junctions were in very good agreement with the consensus sequences around both splice donor and acceptor sites (Fig. 3A).

Four ATFa cDNA variants were originally isolated from human cDNA libraries (14–16), differing from each other by specific sequence elements within the coding region (Fig. 3B): ATFa3 corresponds to the largest cDNA resulting from the juxtaposition of all elements boxed in Fig. 2; ATFa2 lacks a 33-bp element (element d); ATFa1 lacks both element d and an additional 63-bp element (element e); and ATFa0 lacks element d and exons P–I. The existence of cytoplasmic transcripts corresponding to these ATFa variants was confirmed by PCR-mediated analysis of HeLa cell poly(A)+ RNA (16), demonstrating the physiological relevance of these variants. These observations indicate that the various isoforms most likely result from differential splicing of a unique ATFa primary transcript. Thus, exons P–I correspond to alternative exons. Furthermore, examination of the nucleotide sequences at the 5’-border of elements d and e revealed strong homologies to the consensus splice donor sequence (Fig. 3C), suggesting that these sequences constitute alternative splice donor sites (DS1) that can be used in place of the donor sites located at the 3’-border of these elements (DS2). Nothing is known at present about the control of this splice donor site usage and the underlying mechanism.

**The Promoter of the ATFa Gene**—As a first step toward the delineation of the promoter region of the ATFa gene, we determined the transcription start sites. An RNase protection experiment performed with HeLa cell poly(A) RNA (Fig. 4B) clearly identified one minor and two major start sites, arbitrarily denoted −43, −22, and +1, respectively, within the 5’-region of the ATFa leader (exon A). Transcription initiation around position −22 seems to be less precise, as indicated by the slight scattering of the start sites in this area. The position of this initiation region is compatible with the coordinates of the most 5’-extending ATFa cDNA originally isolated (14), starting at position +8.

Examination of the nucleotide sequences located upstream of the transcription start sites did not reveal a typical TATA box. Instead, a significantly higher GC content (70%) was apparent between positions +1 and −140, which, together with the relative spreading of initiation sites, is a characteristic of promoters of housekeeping genes (37). To test for the presence of promoter sequences within the sequences located upstream of the ATFa leader, a vector harboring ATFa sequences between positions −1917 and +83 was transfected into HeLa cells. To ensure stable RNA synthesis, the ATFa sequences were linked to a promotereless rabbit β-globin gene. Specific transcription was investigated by S1 nuclease mapping and reverse transcriptase primer extension of total RNA prepared from the transfected HeLa cells. As shown in Fig. 4C, transcripts that

---

**J. Goetz, unpublished observation.**
initiated at positions +1, −22, and −43 were detected in cells transfected with the ATFa/globin vector in which the ATFa sequences were in the sense orientation (lanes 5 and 7). No transcription could be detected from a vector harboring the ATFa sequences in the antisense orientation with respect to the globin sequences (lanes 6 and 8). These results not only confirm the positions of the start sites of the ATFa gene, but indicate that the isolated ATFa 2-kb upstream sequences carry sufficient information to direct efficient transcription from the transfected construct.

To more precisely delineate the essential promoter region of the ATFa gene, larger deletions were created at the 5′-side of the −1917/+83 ATFa gene fragment. The truncated segments were inserted in front of the bacterial CAT coding sequences, and their capacity to direct CAT transcription was assessed by measuring ATFa sequence-dependent CAT activity of cells transfected with the recombinant ATFa/CAT vectors (Fig. 5). In agreement with the S1 nuclease and reverse transcriptase mapping experiments (Fig. 4C), no CAT activity could be detected after transfection of a recombinant in which the −1917/+83 ATFa fragment was linked in the antisense orientation to the CAT sequences. By contrast, substantial enzymatic activity was measured when the ATFa promoter region was inserted in the sense orientation. Deletion of ATFa sequences between positions −1917 and −212 resulted in only modest effects on promoter activity, as reflected by the small fluctuations of relative CAT activities from one recombinant to the other: these activities varied by no more than 1.5-fold above or below the activity of the recombinant carrying the largest ATFa fragment. Further deletion to position −62 reduced ATFa promoter activity 5−6-fold compared with that of the −1917/+83 or −212/+83 fragment. Finally, removal of the next 40 bp (to position −23) had the most dramatic effect on promoter activity, reducing it to undetectable levels. Together, these results clearly indicate that important ATFa promoter elements are located essentially between positions −212 and −62 and between positions −62 and −23.

Putative protein-binding sites within the minimal ATFa promoter fragment were mapped by the DNase I footprinting assay with HeLa cell extracts (Fig. 6). Areas protected against the nuclease as well as hypersensitive sites were observed on both DNA strands (Fig. 6). Interestingly, one cluster of protected and hypersensitive sites was located between positions −131 and −83 (i.e. within the −212−62 element defined by deletion analysis), and a second, weaker area of interactions was identified between positions −56 and −33 (i.e. within the −62−23 element defined above). The nature of the proteins interacting with these sequence elements is presently unknown, but their crucial contribution to promoter activity is clearly illustrated by the deleterious effect of the successive deletion of the corresponding recognition sites (Fig. 5). Interestingly, a computer search for known protein-binding sequences revealed potential target sites for several factors within the −131−83 upstream protected region (see Fig. 6). On the other hand, no obvious candidate was found for the −56−33 area, a region that is flanked by two putative AP2-
binding sites (Fig. 6).

Nuclear Localization of the ATFa Protein—Immunofluorescent staining experiments with ATFa-specific antibodies show that ATFa proteins localize to the nucleus when overexpressed in transfected HeLa cells. In an attempt to map the peptidic elements (NLS) mediating ATFa nuclear localization, a selected series of ATFa fragments fused to GST were expressed in COS-7 cells after transfection of the corresponding vectors (Fig. 7A). The GST fusion proteins were revealed by indirect immunofluorescence with a monoclonal antibody directed against the GST moiety (Fig. 7A, panel a), whereas it was clearly directed to the nucleus when fused to the NLS of the SV40 large T antigen (panel b). Similarly, the full-length ATFa3 protein relocalized GST to the nucleus (Fig. 7B, panel 1), while deletion of the ATFa sequence C-terminal to residue 293 abolished nuclear accumulation of the chimeric protein (panel 2). A truncated ATFa version lacking the 326 N-terminal residues retained full nuclearization activity (Fig. 7B, panel 3), but derivatives with deletions extending to positions 348 and 368 partially or completely lost this activity (panels 4 and 5, respectively). In agreement with the ability of an ATFa segment spanning residues 327–377 to direct GST to the nucleus (Fig. 7B, panel 6), these results indicate that elements located between residues 327 and 368 are critically involved in the nuclear transportation of the ATFa protein. A comparison with reported NLS sequence compilations points to several elements (underlined in Fig. 7C) located within or flanking the basic region of the ATFa b-Zip motif and roughly spanning residues 327–368.

Tissue Distribution of ATFa Gene Expression—We first examined the expression of mRNAs from various members of the ATFa family. As shown in Fig. 8, Northern blot analysis of HeLa cell transcripts (using specific probes of equivalent radioactivities) revealed that the ATFa-specific RNAs (~9.5 kb in size) were the least abundant compared with the levels of the ATF-1, CRE-BP1/ATF-2, and CREB transcripts (~2.9, 5, and 5.5 kb in size, respectively, in agreement with reported data) (39–41). The relative intensities of the corresponding signals were ~1: 20:5:20, respectively. If translation efficiencies and protein stabilities are comparable in each case, this suggests that the ATFa and CRE-BP1/ATF-2 proteins are minor species within HeLa cells. A survey of a series of transformed human cell lines (originating from colon, cervix, breast, promyelocytes, chorion, endometrium, or fetal kidney) confirmed that the steady-state levels of ATFa transcripts were reproducibly low in each cell line tested (Fig. 9A, lanes 1–8).

To determine ATFa expression in normal cells, we chose to examine the level of ATFa-specific RNAs in adult mouse organs. As in human cells, a 9.5-kb transcript was detected with similar intensities in various organs (Fig. 9A, lanes 9–12; and data not shown), suggesting that there is no striking tissue specificity of ATFa expression in adults.

To detect ATFa transcripts at the cellular level, in situ hybridization was performed on mouse cryosections (Fig. 10) using a 35S-labeled antisense riboprobe. A control probe derived from the same DNA template, but transcribed in the sense orientation, was used on adjacent histological sections and showed no preferential labeling (data not shown). Weak and diffuse signals were detected in several adult organs (ovary, testis, liver, gut, kidney, adrenal gland, spleen, lymph node, thymus, skeletal muscle, and blood vessels) (data not shown). Interestingly, however, enhanced labeling was detected in several squamous epithelia such as those of the vagina (Fig. 10E), the left part of the stomach (Fig. 10D), or the tongue (not shown). Note that the labeling was restricted to the basal and suprabasal cell layers, i.e. the two layers adjacent to the basement membrane. Increased labeling was also seen in pseudostratiﬁed epithelia such as those of the epididymis (Fig. 10F), ductus deferens (Fig. 10G), seminal vesicle (Fig. 10H), and large bronchi (not shown). Strong expression was detected in the mycardium of the ventricles and in the atria (Fig. 10B) as well as in laryngeal muscle (Fig. 10C) and tracheal cartilage (not shown). On brain sections, specific signals were seen in the Purkinje cells of the cerebellum as well as in the limbic lobe (hippocampus and dentate gyrus) and piriform cortex (Fig. 10A).

In situ hybridization on embryo sections at various developmental stages showed weak and ubiquitous ATFa signals throughout the embryo. However, expression above the basal level was clearly detected in the seminiferous tubules of the fetal testis (at 14.5 days post-coitum) and cranial nerve and dorsal root ganglia (at 14.5 and 16.5 days post-coitum, respectively) and also in the squamous to glandular transition area of the stomach epithelium (at 16.5 days post-coitum) (data not shown).

Finally, the relative levels of ATFa protein in mouse organs

\[ \text{this activity (panels 4 and 5, respectively). In agreement with} \]

Finally, the relative levels of ATFa protein in mouse organs...
and in staged whole mouse embryos were determined by Western blot analysis (Fig. 9B). The monoclonal antibody (1A7) used in these experiments was directed against a peptide spanning residues 296–307, a region unique to the ATFa proteins and sharing no significant homology with ATF-2, the most closely related protein. This antibody revealed weak but specific signals in the various adult mouse organs tested, reflecting low levels of ATFa protein (Fig. 9B) and therefore confirming the rather ubiquitous expression suggested by the RNA distribution studies (Figs. 9A and 10). Strikingly, although a weak signal could be detected on longer exposures, the ATFa protein content of the heart (Fig. 9B, lane 11) did not seem to parallel the specific RNA signal observed by in situ hybridization (Fig. 10B). Whether this reflects a reduced translation of the ATFa mRNAs in this tissue remains to be established. More important, overall ATFa expression was maximal in mouse embryos younger than 13.5 days and progressively decreased from day 14.5 onward (Fig. 9B, lanes 1–6). This observation suggests that the ATFa proteins may play essential roles in early embryonic development.

**DISCUSSION**

**The ATFa Sequences Localize within a “Hot” Genomic Area—**

The human genomic sequences encoding the ATFa proteins have been localized on chromosome 12. Interestingly, these sequences map within a locus (12q12–14) that harbors a number of genes playing essential functions in the control of cell homeostasis and proliferation: the hoxC cluster (42), an adeny cyclase gene (43), the cdk2 and cdk4 genes (44), a member of the ras proto-oncogene family (rap1B) (45), the MDM2 gene (46), the ATFa gene (36), the C/EBP-related CHOP or GADD153 gene (47, 48), as well as the genes of the retinoic acid receptor-γ (49), vitamin D receptor (50), and Sp1 (51) transcription factors. Whether the clustering of genes encoding such regulatory proteins is fortuitous or unique in the human genome remains to be established. Strikingly, a number of chromosomal abnormalities within this portion of chromosome 12, mainly translocations, have been found in a variety of human neoplasia, including melanoma, liposarcoma, lipoma, bladder carcinoma, hemangioepithelioma, lung lymphoma, uterine leiomyoma, myelodysplasia, acute non-lymphoblastic leukemia, hairy cell leukemia, fibrosarcoma, and breast tumors. One translocation (t(12; 22)(q13; q12)) (52) that is recurrently found in malignant melanoma of soft tissues (or clear cell sarcoma) has been shown to result in the fusion of the b-Zip domain of ATFa (12q13) to the N-terminal part of the RNA-binding protein EWS. Another translocation (t(12; 16)(q13; p11)) (53) has been characterized in myxoid liposarcoma, where the full-length coding sequence of the CHOP gene (12q13) is linked to the sequences encoding the N-terminal portion of the FUS gene (also designated as TLS), which also encodes an RNA-binding

---

**Fig. 5. ATFa promoter analysis.** The ATFa promoter sequences that have been fused to the CAT reporter sequences in the pBLCAT6 vector and transfected in HeLa cells are schematically depicted on the left, with the numbers referring to ATFa coordinates and arrows oriented in the direction of transcription. Relative CAT activities of cell extracts from three to six independent transfection experiments were determined and are plotted on the right, with average values and standard deviations.
Finally, intragenic rearrangements of the \(\text{HMGI-C}\) gene (12q14–15), which encodes a member of the high mobility group family of DNA-binding proteins, have been associated with the genesis of pulmonary chondroid hamartomas (55). It will be of interest to determine the molecular basis of the other observed malignant transformation phenotypes and to examine whether they are linked to alterations of the regulatory proteins encoded within this portion of chromosome 12. In any case, the multiplicity of alterations in the q12–14 region of chromosome 12 suggests that this region is particularly recombinogenic or mutation-prone.

**The ATFa Isoforms: Structure and Function—Analysis of the ATFa genomic sequence revealed that two of the ATFa isoforms (ATFa1 and ATFa2) differ from the largest variant (ATFa3) by selection of alternative spliced donor sites, one being located between elements D and d and the other between elements E and f (Fig. 3). None of these three variants exhibits transcriptional activity when assayed as intact proteins on adenoviral or artificial promoters (10, 15), although some activity has been observed on the E-selectin promoter (16). In fact, deletion experiments revealed that a potent activation domain of the ATFa protein, which is comprised within the 80 most N-terminal residues of the protein (10), is masked by the 100 most C-terminal residues, located beyond the b-Zip domain (10). The possibility therefore exists that the ATFa factors may...**

**FIG. 6.** DNase I footprint on the upstream region of the ATFa gene. Left, the \(-153/ +83\) ATFa double-stranded DNA probe, 5'-end-labeled either on the transcribed (lanes 1–8) or nontranscribed (lanes 9–14) strand, was preincubated alone (–) or in the presence of HeLa whole cell extract (+), before DNase I treatment, as described under “Experimental Procedures.” Sequence tracks (A and C) of the transcribed strand of the probe fragment are given in lanes 1 and 2. Protected sequences are indicated by closed and shaded rectangles for strong and weak protections, respectively, and hypersensitive sites are indicated by closed and shaded triangles for strong and weak hypersensitivities, respectively. Transcription start sites are also shown with arrows. Right, shown is the nucleotide sequence of the human ATFa promoter region between positions −218 and +82. The major start sites are indicated by arrows. DNase I-protected sequences and hypersensitive sites are marked on both strands (closed and open rectangles and triangles, as in the left panel). Putative binding sites (Genetics Computer Group computer program) for known transcription factors (named above) are enclosed in shaded boxes. End points of the relevant promoter deletion mutants used in the CAT assays are indicated.

**FIG. 7.** Sequences involved in the nuclear addressing of ATFa. A, a schematic representation is shown of the p8B inserts expressing the GST protein (line a) or GST fusions with the SV40 T antigen NLS (line b) or with the ATFa sequences (lines 1–6). B, ~36 h following transfection of these vectors into COS-7 cells, localization of the chimeric proteins was assessed by immunofluorescence microscopy. Bars (20 \(\mu\)m) refer to the four upper and lower micrographs, respectively. C, the sequence element (residues 327–377 of ATFa3) exhibiting nuclear localization properties is shown, with conserved NLS motifs underlined (38, 60). The basic region and first leucine of the b-Zip element are shaded.
be activated in the cell upon unmasking of this activation domain by specific protein interactions.

No functional difference between ATFa1, ATFa2, and ATFa3 has so far been detected despite the rather peculiar amino acid composition of the peptides encoded by the d (rather basic, ARSRTVAKKLV) and e (Ser/Pro-rich, VDSSPPDSPASSPC-SPPLKEK) alternative elements as deduced from their nucleotide sequence (Fig. 3C). All three ATFa variants have previously been shown (i) to mediate the E1a responsiveness of specific genes by recruiting the E1a product to the corresponding promoters (15), (ii) to heterodimerize with members of the Jun/Fos family and thereby to modulate their DNA binding and trans-activation properties (10), and (iii) to strongly interact with a protein kinase closely related to JNK2 (18).

ATFa0, another variant of the ATFa family, is similar to ATFa2, but lacks residues 144–320. This variant has been reported to act as a dominant inhibitor on the E-selectin promoter (16). Although this result is difficult to explain on the basis of the ATFa functional organization (see above), it may be related to potential interactions between ATFa0 and NF-xB on this particular promoter (56).

The ATFa NLS Domain—Deletion and fusion protein analyses allowed the identification of the NLS element involved in the nuclear transportation of the ATFa proteins. This element, located between residues 327 and 368 (see Fig. 7C), appears to be composed of two subdomains, which cooperate in ATFa nuclearization. One is located within the basic region of the b-Zip domain (RNR- - - RCR-KRK) and is found at similar positions in several members of the AP1 family, including c-Jun, c-Fos, and the Epstein-Barr viral factor Zta (57, 58). Interestingly, the invariant cysteine of this element has been suggested to participate in a redox-mediated control of AP1 activity (57, 59). The other element, which is less conserved, is situated just N-terminal to the basic region, and its structure follows the rule initially proposed (60). This latter element, which is by itself unable to direct a heterologous protein to the nucleus, increases the efficiency of the second element. Similar sequences are also found in members of the Fos family (c-Fos, Fra1, and Fra2) and in the Zta protein, but not in c-Jun (58, 61).

Tissue-specific Expression of ATFa—ATFa transcripts were detected at low levels in all human cell lines and mouse organs tested (Fig. 9A). Furthermore, ATFa expression was confirmed at the protein level (Fig. 9B) in mouse embryo (from stage 11.5 days post-coitum onward) and in most adult organs studied. The results of in situ hybridization studies (Fig. 10) were in close accordance with the low level ubiquitous pattern of expression of ATFa. However, expression above the basal level was observed in some tissues of both embryos and adult animals, most particularly in squamous (stomach and vagina) or pseudostratified (epididymis, ductus deferens, and seminal vesicle) epithelia and in specific areas of the brain (Purkinje cells, hippocampus, dentate gyrus, and piriform cortex). Together, these results suggest that the ATFa proteins could play a critical role during embryogenesis and in adult organ function, although their precise contribution remains unclear. The preferential expression of ATFa in specific epithelial structures may explain the tropism of adenovirus infection that is selectively directed toward epithelial cells of the respiratory tract.

Ubiquitous expression has also been noticed for the related CRE-BP1/ATF-2 protein, with increased levels in the central nervous system (41). Interestingly, the strongest expression was observed in the hippocampus and in the dentate gyrus, as in the case of the ATFa proteins, suggesting that both types of factors are involved in signal transduction in brain. Previous studies (62) showed that members of the Jun family are expressed in the piriform cortex and the hippocampus (c-Jun, JunB, and JunD) and in the Purkinje cell layer of the cerebellum (JunB). Such a colocalization of Jun and ATFa proteins is interesting, particularly in view of their capacity to heterodimerize (10, 17).

Homozygous mutations have been introduced in the genes encoding the related CREB (63) and ATF-2 (64) proteins, with drastically different phenotypes: deficiency in long-term
memory for the CREB mutants and skeletal and central nervous system developmental abnormalities in the case of the ATF-2-disrupted mice. These results, indicating that particular members of the ATF family have specific effects, suggest that defects in these proteins are not or are only partially compensated by other members of the family. Using this approach, discrimination between each ATFa variant might perhaps be possible.

The ATFa Promoter Region—Examination of the 1.9-kb 5′-flanking sequences of the ATFa gene showed putative target sites for several transcription regulators, which included Sp1, AP1, AP2, PEA3, E2F, Myc, and NF-κB. Transfection experiments in HeLa cells revealed a 3-fold stimulation of ATFa promoter activity by c-Jun and c-Fos expression vectors (data not shown). In agreement with these results, 12-O-tetradecanoylphorbol-13-acetate or UV light treatments of HeLa cells resulted in similar levels of ATFa promoter activation, suggesting that AP1 may contribute, at least to some extent, to the control of ATFa expression. Our observation that expression of c-Ets or v-Ets vectors had no significant effects on ATFa promoter activity (data not shown) ruled out any major involvement of the PEA3 recognition element. In contrast to the promoter region of the ATF-3 gene (65), which harbors a consensus TATA box and an ATF/CRE site, the ATFa promoter exhibits neither of these elements. It is therefore unlikely that ATFa expression would be synergistically stimulated by cotransfection of c-Jun and ATF-2 vectors, as is the case for ATF-3 (65).

By contrast, in keeping with their similar expression pattern, the ATFa promoter region more closely resembles that of the CRE-BP1/ATF-2 gene, which contains no canonical TATA box, but several Sp1-binding sites (66). Further experiments, including site-directed mutagenesis of specific elements, will help to clarify the physiological relevance of each of the potential ATFa promoter-binding sites.

Acknowledgments—We thank C. Haus for excellent technical assistance; J.-M. Garnier for construction of the genomic library; and R. V. Gopalakrishnan for critical reading of the manuscript and, together with A. Bahr, M. Vigneron, J. Acker, P. Lutz, and H. Boeuf, for helpful discussions and the gift of materials. We also thank M. Gaire for the initial contribution to genomic cloning; P. Dolle´, P. Gorry, and D. De´cimo for expert advise on mouse in situ hybridization; Y. Lutz for antibody production and help with the immunofluorescence experiments; and J. Ménissier de Murcia for the gift of the GST-SVTNLS plasmid. We are grateful to the staff of the cell culture, chemistry, and artwork facilities for providing help and material.

REFERENCES
1. Brindle, P. K., and Montminy, M. R. (1992) Curr. Opin. Genet. & Dev. 2, 199–204
2. Hagmeyer, B. M., Angel, P., and Vandam, H. (1995) Bioessays 17, 621–629
3. Sassone-Corsi, P. (1996) Annu. Rev. Cell Dev. Biol. 11, 355–377
4. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) Science 240, 1759–1764
5. Kezariades, T., and Ziff, E. (1989) Cancer Cells (Cold Spring Harbor) 1, 71–76
