The Intronless and TATA-less Human TAF_{II}55 Gene Contains a Functional Initiator and Downstream Promoter Element*

Tianyuan Zhou‡ and Cheng-Ming Chiang§

From the Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106-4935

Running Title: Regulatory Elements and the Core Promoter of the hTAF_{II}55 Gene

* This work is supported in part by grant # 9950106N from the American Heart Association and in part by grants GM59643 and CA81017 from the National Institutes of Health.

‡ Graduate student on leave from the Department of Biochemistry, University of Illinois at Urbana-Champaign.

§ A Pew Scholar in the Biomedical Sciences and a Mt. Sinai Health Care Foundation Scholar.

To whom correspondence should be addressed: Department of Biochemistry, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, OH 44106-4935. Tel.: 216-368-8550; Fax: 216-368-3419; E-mail: c-chiang@biochemistry.cwru.edu.

The abbreviations used are: TFIID, transcription factor IID; TBP, TATA-binding protein; TAF_{II}55, a 55-kDa TBP-associated factor found in TFIID; hTAF_{II}55, human TAF_{II}55; Sp1, specificity protein 1; AP2, activator protein 2; Inr, initiator element; DPE, downstream promoter element.
ABSTRACT

Human TAFII55 (hTAFII55) is a component of the multisubunit general transcription factor TFIID and has been shown to mediate the functions of many transcriptional activators via direct protein-protein interactions. To uncover the regulatory properties of the general transcription machinery, we have isolated the hTAFII55 gene and dissected the regulatory elements and the core promoter responsible for hTAFII55 gene expression. Surprisingly, the hTAFII55 gene has a single uninterrupted open reading frame and is the only intronless general transcription factor so far identified. Its expression is driven by a TATA-less promoter which contains a functional initiator and a downstream promoter element, as illustrated by both transfection assays and mutational analyses. Moreover, this core promoter can mediate the activity of a transcriptional activator that is artificially recruited to the promoter in a heterologous context. Interestingly, in the promoter-proximal region there are multiple Sp1-binding sites juxtaposed to a single AP2-binding site, indicating that Sp1 and AP2 may regulate the core promoter activity of the hTAFII55 gene. These findings implicate that a combinatorial regulation of a general transcription factor-encoding gene can be conferred by both ubiquitous and cell type-specific transcriptional regulators.
INTRODUCTION

Studies on eukaryotic promoters have identified several core promoter elements, which are characteristic DNA sequences required for promoter function. The TATA box is an A/T-rich sequence located approximately 25 to 30 nucleotides upstream of the transcription start site. It contains a consensus sequence, TATA(A/T)A(A/T), whose recognition by the TATA-binding protein (TBP) subunit of TFIID nucleates the formation of a preinitiation complex (1-3). A second core promoter element, the initiator (Inr), contains a pyrimidine-rich core sequence, PyPyA_{+1}N(T/A)PyPy, surrounding the transcription start site (4). The Inr is capable of directing accurate transcription initiation either alone or in conjunction with a TATA box or other core promoter elements (5-10). Several protein factors, including the TAFII150/CIF150 component of TFIID (11-17), RNA polymerase II (6), TFII-I/SPIN/BAP-135 (18-21), USF (22) and YY1 (23), have been implicated in Inr function. However, the nucleation pathways of these Inr-targeting proteins have not yet been defined.

The downstream promoter element (DPE), which is located 28-34 nucleotides downstream of the transcription start site in many Drosophila TATA-less promoters (9, 10, 24), has a consensus sequence, (A/G)G(A/T)CGTG, and can be recognized by the dTAFII60 and dTAFII40 components of Drosophila TFIID (9, 24). This finding suggests that TFIID is likely to be the DPE-binding factor. Interestingly, negative cofactor 2 (NC2 or Dr1-Drap1), initially characterized as a TBP-inhibitory activity with a TATA-containing promoter (25-28), has recently been shown to facilitate transcription from DPE-driven promoters (29). It seems that
TFIID and NC2, two of the DPE-acting factors, may work synergistically through the DPE, although their functional relationship remains to be elucidated. Another upstream core promoter element, (G/C)(G/C)(G/A)CGCC, was identified through binding site selection as a GC-rich sequence recognized by TFIIB (30). This TFIIB recognition element (BRE) is located immediately upstream of the TATA box and can be used to modulate preinitiation complex assembly in eukaryotic cells (30) as well as in Archaea (31). Analysis of the promoter database reveals that 57% of the Drosophila core promoters do not contain a TATA box and the DPE occurs in approximately 40% of the Drosophila promoters (10). Although such statistical data is not yet available for the human genome, it appears that the promoters of human housekeeping genes, oncogenes, growth factors, and transcription factors often lack a TATA box (32). In addition, many natural promoters contain distinct combinations of core promoter elements whose differential utilization plays an important role in regulating gene expression in a spatial, temporal or lineage-specific manner (13, 33, 34).

Human TAFII55 (hTAFII55) was first identified as an RNA polymerase II-specific TBP-associated factor (TAFII) in TFIID (35, 36) and, like many other TAFIIs, was also detected in the TBP-free-TAFII-containing complex (TFTC) (37). However, TAFII55 is not present in some other TAFII-containing complexes, such as human PCAF (38) and yeast SAGA complexes (39), suggesting that TAFII55 has unique properties distinct from its role as a structural component of TFIID and of TFTC. This idea is further substantiated by the finding that TAFII55 can interact with many transcription factors, including Sp1, YY1, USF, CTF, adenovirus E1A and HIV-1 Tat (35), and can also mediate vitamin D3 and thyroid hormone
receptor activation in a ligand-independent manner (40), consistent with a coactivator role of TAFII55 in transcriptional regulation. Moreover, TAFII55 may be implicated in mRNA 3'-end processing, as it shows strong affinity towards the human cleavage-polyadenylation specificity factor (CPSF) (41).

TAFII55 homologues have also been identified in several organisms. The mouse homologue, mTAFII55, is 95% identical to its human counterpart (42) and the *Saccharomyces cerevisiae* homologue, yTaf67, is essential for cellular viability2 (43). Recently, the *Schizosaccharomyces pombe* homologue of yTaf67, Ptr6p [poly(A)+ RNA transport], was shown to be involved in nucleocytoplasmic transport of mRNAs during a genetic screen for mutants that accumulate mRNAs in the nucleus (44). Moreover, proteins that share high sequence homology with hTAFII55 have also been identified in *Caenorhabditis elegans* (GenBank accession number CEF54F7) and *Drosophila melanogaster* (GenBank accession number AF017096). The chromosomal location of the hTAFII55 gene has been mapped to 5q31, where chromosomal mutations have been associated with stomach adenocarcinoma (45), suggesting that hTAFII55 or other genes localized in this region may act as an oncogene.

Interestingly, Northern blot analysis showed that hTAFII55 is differentially expressed in various human tissues3. In addition, we observed that in a HeLa-derived cell line that conditionally expresses FLAG-tagged hTAFII55, the overall level of the induced tagged protein and the

---

2 S.-Y. Wu and C.-M. Chiang, unpublished data.
3 C.-M. Chiang, unpublished data.
endogenous untagged hTAFII55 protein remains constant (46). This indicates a tight regulation over hTAFII55 expression in vivo. In order to understand the regulation of hTAFII55 gene expression and to gain further insight into the regulatory pathways of general transcription factor-encoding genes, we dissected the cis-acting elements and trans-acting factors that regulate the expression of the hTAFII55 gene. Our studies indicate that hTAFII55 gene expression is combinatorially regulated by both ubiquitous and cell type-specific transcription factors. Moreover, we have characterized the core promoter elements of the hTAFII55 gene, which surprisingly contains a single uninterrupted open reading frame whose expression is driven by a TATA-deficient promoter with a functional initiator and a DPE. Collectively, these findings uncover unusual features of hTAFII55 gene structure and regulatory properties which are significantly different from other general transcription factor-encoding genes.

**EXPERIMENTAL PROCEDURES**

*Isolation of Human TAFII55 Genomic Clones*—A human genomic library, derived from the HT1080 human fibrosarcoma cell line and cloned in the λ-DASH II vector (Stratagene), was screened with a 32P-labeled DNA fragment spanning the first 474 nucleotides (cut between the HpaI and EcoRI sites) of the hTAFII55 cDNA (35). From ~1 x 10^6 plaque forming units, 12 positive clones were isolated. The inserts were individually cloned into the NotI site of pBS-SK (+) (Stratagene). A clone, pBS/3′-8, which contains an insert of approximately 17 kb including
regions 5’ and 3’ of hTAFII55, was manually sequenced (GenBank accession number AF349038).

**Plasmid Constructions**—A 1459-bp genomic DNA fragment that extends 1436 bp upstream and 23 bp downstream of the 5’ end of the hTAFII55 cDNA (35) was amplified by polymerase chain reaction (PCR) from pBS/3’-8 using an upstream *Kpn*I site-containing primer (5’ CATTCTGGTACCAGGCACTGGGACAC 3’) and a downstream *Bgl*II site-containing primer (5’ AGCGCGAGATCTTGCGAGAGG 3’). The amplified DNA fragment was then cloned into pGL2-Basic (Promega) between the *Kpn*I and *Bgl*II sites. The resulting construct was denoted pGL2-TAF55(-1372/+87).

A series of hTAFII55 promoter deletion constructs, including pGL2-TAF55(-128/+87), pGL2-TAF55(-99/+87), pGL2-TAF55(-71/+87), pGL2-TAF55(-55/+87), pGL2-TAF55(-26/+87), pGL2-TAF55(-128/+36), pGL2-TAF55(-71/+36), pGL2-TAF55(-55/+36), and pGL2-TAF55(-26/+36), were similarly made in pGL2-Basic by using primer pairs with introduced *Kpn*I and *Bgl*II sites at their 5’ and 3’ ends, respectively. The numbers in the deletion constructs indicate the boundaries of the inserts relative to the transcription start site.

The plasmids pGL2-TAF55(-748/+87) and pGL2-TAF55(-281/+87) were created by first cleaving pGL2-TAF55(-1372/+87) with *Sca*I or XbaI, filling in the XbaI-digested end with Klenow enzyme, and releasing the inserts with *Bgl*II. The promoter-containing fragments were then cloned into pGL2-Basic between the *Bgl*II site and the Klenow filled-in *Xho*I site to generated pGL2-TAF55(-748/+87) and pGL2-TAF55(-281/+87), respectively. The plasmid pGL2-TAF55(-161/+87) was generated by cloning a PCR fragment, amplified with an upstream
primer spanning -161 to -144 and the same downstream BgIII site-containing primer ending at +87, between the BgIII site and the Klenow filled-in XhoI site of pGL2-Basic. Similarly, the plasmid pGL2-TAF55(-1372/-140) was made by inserting a PCR fragment, amplified with the same upstream KpnI site-containing primer ending at -1372 and a downstream primer spanning -157 to -140, between the KpnI site and the Klenow filled-in XhoI site of pGL2-Basic.

Promoter constructs containing nucleotide substitutions in the sequence motifs of Sp1, AP2, Inr, and DPE (denoted by asterisks) were individually generated by PCR amplification with primer pairs spanning the mutated nucleotides according to the QuickChange Site-Directed Mutagenesis Protocol (Stratagene). The plasmids pGL2-TAF55(-71/+36)Sp1*-60, pGL2-TAF55(-71/+36)AP2*, pGL2-TAF55(-71/+36)Sp1*-60/AP2*, pGL2-TAF55(-71/+36)Sp1*-20, pGL2-TAF55(-26/+36)Inr*, pGL2-TAF55(-26/+36)DPE*, and pGL2-TAF55(-26/+36)Inr*DPE* were constructed in the backbone of pGL2-TAF55(-71/+36) or pGL2-TAF55(-26/+36) using primer pairs containing the introduced mutations as shown in Figures 4B and 5A. For five Gal4-binding site-containing constructs, the SacI-PstI fragment of pG5HMC2AT (47) with 5 Gal4-binding sites was first cloned into pBS-SK (+) between SacI and PstI sites to generate pBS-5Gal, from which the SmaI-KpnI fragment was isolated and cloned into pGL2-TAF55(-26/+36), pGL2-TAF55(-26/+36)Inr*, pGL2-TAF55(-26/+36)DPE*, pGL2-TAF55(-26/+36)Inr*DPE* and pGL2-Basic at the same enzyme-cutting sites to create pGL2-5Gal(-26/+36)WT, pGL2-5Gal(-26/+36)Inr*, pGL2-5Gal(-26/+36)DPE*, pGL2-5Gal(-26/+36)Inr*DPE* and pGL2-5Gal, respectively. All constructs were confirmed by restriction enzyme digestion and DNA sequencing.

The HIV-1 promoter construct pGL2-HIV(-167/+80) was created by transferring the
XhoI-HindIII fragment, which contains the HIV-1 promoter region spanning -167 to +80, from p-167 (48) into the same enzyme-cutting sites in pGL2-Basic. The pBS-TAF55(-128/+87) plasmid used to generate riboprobe for RNase protection analysis was created by subcloning the Smal-HindIII fragment from pGL2-TAF55(-128/+87) into the same enzyme-cutting sites of pBS-SK (+). The other plasmids, pHIV+58 (49), pGL7072-161 (50), pSGVP (51), pSG424 (52) and pGL2-Control (Promega) have already been described.

Transient Transfection and Reporter Gene Analysis—C-33A cells, which are derived from a human cervical carcinoma, were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum in a humidified 5% CO2 incubator at 37°C. Transient transfection was carried out in C-33A cells with 4 µg of each reporter plasmid, either alone or in conjunction with varying amounts of the Gal4-VP16-expressing plasmid (pSGVP) supplemented with the cloning vector (pSG424) to a total of 1 µg, using the calcium phosphate precipitation method as described (53). The transfected cells, after rinsing twice with 1x PBS, were collected 24 hours post transfection by rubber policeman and resuspended in 100 µl of T250E5 buffer (250 mM Tris-HCl, pH 7.6 and 5 mM EDTA). Cell lysates were then prepared by three cycles of freezing and thawing in liquid nitrogen and 37°C water bath. Following centrifugation at 4°C for 10 min, 2 µl of the supernatant was mixed with 350 µl of luciferase buffer (25 mM HEPES, pH 7.8, 5 mM ATP, 15 mM MgSO4) with luciferase assays conducted by automatically injecting 100 µl of 0.2 mM luciferin (Analytical Luminescence Laboratory) into the samples and measuring the luminescence for 12 sec after an initial 2 sec delay, using a Monolight 2010 luminometer (Analytical Luminescence Laboratory).
Transfection and reporter gene assays were performed independently at least four times, each in duplicate.

_In Vitro Transcription and Primer Extension_—*In vitro* transcription was performed with HeLa nuclear extracts and analyzed by primer extension as described (50). The Luc-5 primer (5’ CTCTTCATAGCCTTATGCAG 3’) and the Luc-1 primer (5’ TCTTTATGTGGTGGCGGTCT 3’) that anneal respectively to nucleotides 151-170 and 81-100 of pGL2-Basic were used for examining products derived from hTAFII55 promoter-containing constructs, whereas a CAT primer (5’ CAACGGTGTTATCCAGTG 3’) that anneals to nucleotides 4936-4953 of pSV2CAT (54) was used for determining the product derived from pHIV+58. All the primer extension products were analyzed on an 8 M urea, 5% Long Ranger (FMC) polyacrylamide gel together with the dideoxynucleotide sequencing products generated with the phosphorylated forms of the corresponding primers.

_RNase Protection Assay_—Total cellular RNA was prepared from eight 100-mm plates of 80% confluent C-33A cells by guanidinium thiocyanate-phenol extraction method using 8 ml of Trizol reagent (Life Technologies) according to the manufacturer’s instructions. Poly (A)+RNA was isolated by first passing heat-treated total cellular RNA, after mixing with an equal volume of 2x loading buffer, through a 1-ml oligo (dT)-cellulose (Amersham Pharmacia Biotech) column, which was pre-equilibrated with 1x loading buffer (20 mM Tris-HCl, pH 7.6, 0.5 M LiCl, 1 mM EDTA and 0.1% SDS). The flowthrough fraction was collected, denatured at 65°C for 5 min, chilled on ice, and loaded again onto the column. This process was repeated for two additional times. The column was then washed with 6-8 column volumes (CV) of 1x loading buffer and stained with ethidium bromide for visualization of the RNA bands.
buffer. Poly (A)+ RNA was eluted with 1 CV of elution buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA and 0.05% SDS) for a total of three times, precipitated with ethanol, and finally resuspended in DEPC-treated water.

An antisense riboprobe, corresponding to +87 to -128 of the hTAFII55 promoter region with flanking polylinker sequences, was synthesized by transcribing the BamHI-linearized pBS-TAF55(-128/+87) template with 2 units of T7 RNA polymerase in the presence of 2 µCi/µl [α-32P]CTP, 10 µM CTP, 0.1 mM ATP, UTP and GTP, 40 mM Tris-HCl (pH 8.0), 8 mM MgCl2, 50 mM NaCl, 30 mM DTT, 1 unit/µl RNasin (Promega) and 2 mM spermidine in a 25-µl mixture. The reaction was conducted at 37°C for 60 min. The riboprobe was then separated on a 4% polyacrylamide-8M urea gel, eluted from the gel slice in elution buffer (0.5 M ammonium acetate, 10 mM magnesium acetate, 0.1% SDS and 1 mM EDTA), extracted with phenol/chloroform, precipitated with ethanol, and finally dissolved in 50 µl of 1x hybridization buffer (40 mM PIPES, pH 6.7, 0.4 M NaCl and 1 mM EDTA). RNase protection assay was carried out as described previously (55) with minor modifications. Briefly, ~5 x 10^5 cpm of the

in vitro-synthesized riboprobe was mixed with 3 µg of poly (A)+ RNA in a 30-µl reaction mixture containing 80% formamide in final 1x hybridization buffer, overlaid with mineral oil, heated at 90°C for 10 min, and hybridized at 58°C overnight. The hybridization reaction was then quenched on dry ice and incubated with 350 µl of RNase solution containing 14 µg of RNase A (Sigma), 50 units of RNase T1 (Pharmacia), 0.3 M NaCl, 10 mM Tris-HCl (pH 7.5) and 5 mM EDTA at 30°C for 60 min. The ribonucleases were degraded by adding 50 µg of Proteinase K (USB) and 5 µl of 10% SDS and incubated for another 15 min at 37°C. The
protected fragments were purified by phenol/chloroform extraction, precipitated twice with ethanol, and finally analyzed on a 5% polyacrylamide-urea gel with a DNA sequencing ladder loaded in parallel as size markers. The migration differences between the protected RNA fragments and the DNA size markers were adjusted using undigested riboprobe as standard.

RESULTS

A 1.4-kb Genomic Fragment Preceding the 5' End of the Human TAFII55 cDNA Sequence Has Intrinsic Promoter Activity—To understand the regulatory properties of general transcription factor-encoding genes, we isolated a dozen genomic clones from an HT1080 human fibrosarcoma genomic library using probes derived from hTAFII55 cDNA (35). One of the isolated clones, 3'-8, containing the entire ORF and flanking regions was completely sequenced (17,042 bp, GenBank accession number AF349038). The hTAFII55 gene, which encodes a component of the eukaryotic core promoter-binding factor TFIID, encompasses the complete cDNA sequence of hTAFII55, suggesting that it is an intronless gene (Fig. 1A). This finding is surprising, given the fact that the human general transcription factor-encoding genes so far identified, including TFIIA (α/β and γ), TFIIB, TFIIEα, TFIIEβ, the RAP30 and RAP74 subunits of TFIIF, components (p89, p80, p62, p52, p44, p34, CDK7, cyclin H and MAT1) of TFIIH, TBP and other TAFII's in TFIID, all have introns4. The possibility that our hTAFII55 genomic DNA was derived from retrotransposition of the hTAFII55 cDNA was excluded for the

4 for locus information, see website: http://www.ncbi.nlm.nih.gov/locuslink/
following reasons. First, a poly (A) tail sequence found at the 3’ end of the hTAFII55 cDNA (35) is absent in all of our genomic clones. The hTAFII55 sequences identified in the genomic clones and the cDNA diverge at the 3’ cleavage site where poly (A) addition occurs (data not shown), indicating that reverse transcription and retroviral insertion are unlikely to be involved in generating the genomic copy. Second, all of our independent clones which extended beyond the 3’ end of the hTAFII55 cDNA had identical sequences and all of them lack a poly (A) tail.

Third, a portion of the BAC clone 249h5 (GenBank accession number AC005618), derived from human chromosome 5, has a nearly identical nucleotide sequence with that of our 3’-8 genomic clone. Fourth, the human protocadherin gamma A1 gene sequence found at the 5’ end of the 3’-8 clone is also present in human chromosome 5, indicating the authenticity of our isolated hTAFII55 genomic sequence. Lastly, the entire sequence of our 3’-8 clone is also found in the just-deposited human genome databases (56, 57). Taken together, the absence of a poly (A) tail in our genomic clones and the co-localization of the entire genomic sequence in a single chromosomal locus exclude the possibility that our genomic clones are artifacts and further confirm that the hTAFII55 gene is indeed devoid of introns, an unusual feature distinct from all the other general transcription factor-encoding genes so far identified.

To identify a functional promoter in the isolated hTAFII55 gene, we cloned a 1.4-kb genomic DNA fragment that extends 1436 bp upstream and 23 bp downstream of the 5’ end of the hTAFII55 cDNA (35) into pGL2-Basic (Fig. 1A). The promoter activity of the resulting construct, pGL2-TAF55(-1372/+87), was examined by luciferase assays in a human cervical carcinoma-derived C-33A cell line following transient transfection. As shown in Fig. 1B, the 1.4-kb
genomic fragment of hTAFII55 has promoter activity that is stronger than those exhibited by HIV-1, human papillomavirus type 11 (HPV-11) and SV40.

**Mapping the Transcription Start Site of the hTAFII55 Promoter**—In order to locate the transcription start site of the hTAFII55 gene, we first performed *in vitro* transcription with HeLa nuclear extracts, using pGL2-TAF55(-1372/+87). The *in vitro*-synthesized transcripts were then detected by primer extension analysis (Fig. 2A). To minimize artifacts caused by spurious primer annealing, we used two primers, Luc-1 and Luc-5, which anneal to different positions of the transcript and are expected to generate approximately 150-nt and 210-nt products, respectively. When either primer was used, the transcription start site was mapped to the same position in the genomic sequence (nucleotide 12,849), which was designated +1 (Fig. 2A, lanes 1 and 3, indicated by an arrow). Several signals of less intensity corresponding to +3 to +6 positions were also detected. The presence of multiple minor transcription start sites in a TATA-less promoter is not uncommon (9, 24, 58-60; and see below). A control template with the TATA-containing HIV-1 promoter was mapped to the same start site as previously determined (Fig. 2A, lane 5; Ref. 49). The transcripts derived from the hTAFII55 and HIV-1 promoters are RNA polymerase II-specific, since the addition of a low concentration (2 µg/ml) of α-amanitin, which inhibits the activity of RNA polymerase II, completely abolished the specific signals (Fig. 2A, compare lanes 1 and 2, 3 and 4, and 5 and 6).

A transcription signal detected in the Luc-5 experiment (Fig. 2A, lane 3, indicated by an asterisk) may result from a premature termination of reverse transcriptase during primer extension, since it lies within the isolated cDNA region (36). Another transcription signal
detected in the Luc-1 experiment (Fig. 2A, lane 1, shown with an arrowhead) is located at -57, surrounded by GC-rich sequences. It might represent an alternative transcription start site or a spurious transcript caused by non-specific initiation of RNA polymerase II in vitro. To distinguish between these two possibilities, we isolated endogenous poly (A)+ RNA from C-33A cells and determined the transcription start site using RNase protection assay with an antisense RNA probe that spans nucleotides from +87 to -128 relative to the transcription start site (Fig. 2B). A correct initiation at +1 would give rise to a protected fragment of 87 nucleotides. In contrast, were transcription initiated from -57, a 144-nt protected fragment would be detected. As shown in Fig. 2B, only an 87-nt protected fragment, corresponding to the start site mapped in vitro, was observed when the $^{32}$P-labeled riboprobe was hybridized with poly (A)+ RNA, but not with tRNA (lanes 1 and 2). The absence of a 144-nt protected fragment suggests that the start site detected at -57 is an artifact caused by non-specific initiation of RNA polymerase II in vitro. Therefore, we concluded that nucleotide 12,849 in the 3'-8 genomic clone is the major transcription start site of the hTAFII55 promoter both in vivo and in vitro.

**Transcription Factors Potentially Regulate hTAFII55 Promoter Activity**—A search for transcription factors potentially regulating hTAFII55 gene expression was performed using the MatInspector program. We found putative binding sites for STAT-1, MEF2, E2F, Sp1, AP2, AREB6 and E47 in the promoter-proximal region (Fig. 3A). Obviously, no TATA box is located between -25 and -30; but, instead, there are consensus Inr and DPE sequences surrounding the transcription start site and spanning +30 to +35, respectively. This inspection reveals that an intrinsic TATA-less promoter is used for hTAFII55 gene expression, which is likely regulated by both ubiquitous and cell type-specific transcription factors.

---

5 http://transfac.gbf.de/TRANSFAC/index.html
Sequences for Sp1, AP2, AREB6 and E47 Binding Are Important for hTAFII55 Gene Expression

To define the transcription factor-binding sites that were important for hTAFII55 gene expression, we made a series of 5’ deletion constructs and tested their promoter activity following transfection into C-33A cells. As shown in Fig. 3A, deletions progressing to -99 that removed the STAT-1, MEF2, and E2F sites showed no significant reduction in promoter activity in C-33A cells. Further deletion of the region from -99 to -71, which contains no known transcription factor-binding sites, resulted in ~ 2-fold decrease in promoter activity. Interestingly, deletion up to -55, that removes a putative Sp1-binding site centering on -60, caused another 2- to 3-fold reduction. An additional deletion to -26, that eliminates an overlapping Sp1- and AP2-binding site at -50, resulted in an extra 8-fold reduction. In contrast, an upstream fragment spanning -1372 and -140 showed no promoter activity, further confirming our results of start site mapping.

A similar conclusion was also obtained by in vitro transcription and primer extension assays performed with HeLa nuclear extracts using similar hTAFII55 deletion constructs. As shown in Fig. 3B, while a series of 5’ deletions up to -128 did not markedly affect promoter functioning, the -71/+87 construct reduced approximately 50% of the promoter activity. Further deletions to -55 and -26 significantly decreased the signal intensity. However, the transcription start site initiating at +1 was still detectable after longer exposure. This analysis suggests that both Sp1 and AP2 sites are important for hTAFII55 gene expression. Moreover, both in vitro and in vivo assays indicated that the -26/+87 construct, although it shows much weaker promoter activity compared to that from the 1.4-kb genomic fragment, could still direct reporter
gene expression, suggesting that the critical core promoter elements essential for hTAFII55 promoter activity are retained in this short region (see below).

To examine whether the downstream region containing putative AREB6- and E47-binding sites are also critical for hTAFII55 gene expression, we created several 3' deletion constructs removing the sequence between +36 and +87, and tested their promoter activity in C-33A cells by transfection assays. As shown in Fig. 4A, promoter constructs deleted to +36 reduced reporter gene activity approximately 3 fold. This result reveals that the sequences downstream of the transcription start site, including the AREB6- and E47-binding sites, contribute to hTAFII55 promoter activity. Nevertheless, for core promoter activity, the region between +36 and +87 seems dispensable. The -55/+36 construct still maintained promoter activity sufficient to drive reporter gene expression.

From the deletion analysis, it appears that Sp1 and AP2 likely play an important role in optimizing hTAFII55 promoter activity, which could be conferred by a small DNA fragment spanning -26 to +36. To test this hypothesis, we created the -26/+36 promoter construct and compared its promoter activity with several 5' deletion constructs all ending at +36 as well as with new constructs containing nucleotide substitutions in the DNA-binding sites for Sp1 and AP2. As shown in Fig. 4B, nucleotide substitutions at the -60 Sp1-binding site reduced promoter activity 2-3 fold (compare 2nd and 3rd constructs), consistent with the result from the 5' deletion constructs (see Fig. 3A, compare -71/+87 and -55/+87 constructs). However, mutations introduced at the AP2-binding site showed only 10-20% reduction of reporter activity (Fig. 4B, compare 2nd and 4th constructs, and 3rd and 5th constructs). Interestingly, mutations
at the -20 Sp1-binding site showed the same activity as that of the wild-type construct (Fig. 4B, compare 2\textsuperscript{nd} and 6\textsuperscript{th} constructs). This finding indicates that different promoter-proximal Sp1-binding sites contribute unequally to hTAF\textsubscript{II}55 promoter activity. As expected, the -26/+36 construct still retains promoter activity (Fig. 4B). A similar result was also obtained with \textit{in vitro} transcription assays performed with HeLa nuclear extracts (data not shown).

\textit{The hTAF\textsubscript{II}55 Core Promoter is TATA-less with Functional Inr and DPE Sequences}—The finding that a DNA fragment spanning -26 to +36 still retained hTAF\textsubscript{II}55 promoter activity and that mutations introduced at the -20 Sp1-binding site had no effect on reporter activity (Fig. 4B) suggested that the Inr and DPE motifs present in this region were likely to be the functional modules driving hTAF\textsubscript{II}55 gene expression. To test this, we introduced mutations in the Inr and DPE, either individually or in combination, and tested the activity of the core promoter constructs using transfection assays in C-33A cells. As shown in Fig. 5A, mutations in the Inr and the DPE reduced promoter activity approximately 33 fold and 5 fold, respectively, whereas double mutations essentially abolished the promoter function.

To verify that the Inr and the DPE identified in the hTAF\textsubscript{II}55 promoter can function as independent promoter modules, we introduced 5 Gal4-binding sites into the wild-type and mutated hTAF\textsubscript{II}55 core promoter constructs, and tested promoter activity by cotransfection with a Gal4-VP16-expression plasmid. As shown in Fig. 5B, expression of Gal4-VP16 significantly enhanced wild-type (WT) hTAF\textsubscript{II}55 promoter activity in a dose-dependent manner. In contrast, Gal4-VP16 had little, if any, effect on constructs containing Inr mutation (Inr\textsuperscript{*}) or Inr and DPE double mutations (Inr*DPE\textsuperscript{*}). The heterologous promoter with the DPE mutation (DPE\textsuperscript{*}) showed a slight response to Gal4-VP16. This result demonstrates that the Inr and the DPE derived from the hTAF\textsubscript{II}55 promoter are indeed core promoter elements which can mediate the activity of a transcriptional activator artificially recruited to the promoter in a heterologous context.
DISCUSSION

In this report, we describe the detailed characterization of regulatory elements and core promoter critical for the expression of the human TAF\(\Pi\)55 gene, which encodes a component of the general transcription factor TFIID. Sequencing of our isolated hTAF\(\Pi\)55 genomic clones and mapping of the transcription start site reveal that the hTAF\(\Pi\)55 gene is intronless, a feature distinct from the other general transcription factor-encoding genes so far identified\(^4\) (56, 57, 61-63). Furthermore, expression of hTAF\(\Pi\)55 is driven by a TATA-less promoter with a functional Inr and the DPE active in both homologous and heterologous promoter contexts.

**Intronless Genes**—In higher eukaryotes, most genes contain introns. Compared to 96% intronless genes in *S. cerevisiae*, there are 17% intronless genes in *D. melanogaster* and merely 6% of genes in mammals without introns (64). One family of intronless genes encodes histones, which are comparatively small, abundantly expressed and highly conserved in sequence (65). Another family encodes G-protein-coupled receptors (66). Since intronless genes such as those encoding hsp70 (67), c-jun (68) and interferon-\(\alpha\) (69) do not require post-transcriptional splicing, they may be expressed more efficiently and are believed to be involved in immediate response to extracellular signals. On the other hand, many viruses that undergo reverse transcription during the replication cycle have evolved special mechanisms to specifically facilitate export of intronless gene products to the cytoplasm and inhibit the splicing process (70, 71). Considering the hTAF\(\Pi\)55 gene has multiple STAT- and E2F-binding sites in its promoter-proximal region, it seems probable that hTAF\(\Pi\)55 is involved in integrating
extracellular signals to the general transcription machinery. This point of view is further supported by the finding that hTAFII55 can interact with many transcription factors (35) and can also mediate the functions of several nuclear hormone receptors (40).

Core Promoter Elements—One interesting property of the hTAFII55 promoter sequence is that it has no cognate TATA box or even AT-rich sequences within the promoter-proximal 80 nucleotides upstream of the transcription start site. Instead, it contains a consensus Inr that overlaps the transcription start site and the DPE core sequence (GACGGA) from +30 to +35. Both the Inr and the DPE are critical for hTAFII55 core promoter function, as illustrated by both transfection assays (Fig. 5) and in vitro transcription analysis (data not shown). The Inr is clearly protected by proteins present in nuclear extracts (data not shown), consistent with the functional importance of the Inr (Fig. 5). Although mutations at the DPE did not completely abolish promoter function, these constructs displayed dramatic decreases in promoter activity. We speculate that the incomplete destruction of the Sp1-binding site at the -20 region in the -26/+36 promoter-based constructs might partially compensate for the loss of the DPE function. It is also likely that the Inr and the DPE of the hTAFII55 gene as well as the -20 Sp1-binding site are differentially utilized in different cell types. Therefore, the DPE may be more important in some cells than in others. Nevertheless, this study is the first demonstration of a functional DPE in a human promoter following the initial report on hIRF (24).

MED-1 (Multiple start site Element Downstream) in many TATA-less promoters and Downstream Core promoter Element (DCE) in the TATA-containing human β-globin promoter are additional examples of downstream elements which function in concert with the Inr and
appear to affect TFIID binding (58, 72). Cellular proteins, such as TFIID and NC2, have been reported to act through these downstream promoter elements. However, it is not clear whether the downstream sequences are essential for promoter activity merely to affect preinitiation complex formation or whether they are also involved in promoter clearance and the formation of a highly processive RNA polymerase II elongation complex (73).

Our present study provides convincing evidence that the -26/+36 sequence can serve as an independent core promoter module, which can be further activated by a transcriptional activator in the context of a heterologous promoter (Fig. 5B). The hTAFII55 core promoter could thus be a model to further analyze molecular mechanisms of transcription initiation on TATA-less promoters.

Transcription Factors Binding to the Promoter-proximal Region—Our study also details transcriptional regulation of hTAFII55 promoter activity by Sp1 and AP2 proteins, a phenomenon commonly observed in mammalian TATA-less promoters (74, 75). Sp1 is a well-characterized ubiquitous transcription factor whose binding sites are found in numerous promoters that regulate both ubiquitous and tissue-specific genes (60, 76-78). Mice with homozygous deletions of the Sp1 gene show severe developmental defects and die early during embryogenesis, suggesting that Sp1 is essential for embryonic development (79). Our study also indicates the importance of Sp1 in regulating TAFII55 gene expression. First, deletion of the GC-rich Sp1-binding sequences resulted in significant reduction in hTAFII55 promoter activity (Fig. 3). Second, point mutations introduced at the Sp1-binding sites resulted in similar decreases in activity (Fig. 4). Third, DNase I footprinting shows direct binding of purified Sp1
to its cognate DNA sequences (data not shown).

In contrast, AP2 is a cell type-specific transcription factor important in retinoid-controlled morphogenesis and differentiation, especially in neural crest-derived cell lineages and epithelial cells (80). AP2 responds to at least two different signal transduction pathways, the phobol ester/PKC signaling and the PKA pathway (81). AP2 has a spatially and temporally restricted expression pattern in murine embryos and shows significant expression levels in adult skin and urogenital tissues (80). We found that Sp1 and AP2 proteins, whose binding sites are closely positioned on the hTAFII55 promoter, could bind simultaneously to the promoter (data not shown). This finding suggests that Sp1 and AP2 can regulate the hTAFII55 promoter in a combinatorial manner, although they do not appear to function synergistically in C-33A cells (Fig. 4B). We estimate by quantitative Western blotting analysis that C-33A cells have ~100 fg of Sp1 and less than 5 fg of endogenous AP2 per cell (data not shown). It is likely that the relatively low level of AP2 proteins in C-33A cells cannot confer significant activator function on hTAFII55 expression. AP2 may function as a more potent transcription activator in keratinocytes or in the neural crest lineage in which it is expressed in high levels. We will further clarify the role of AP2 in regulating hTAFII55 gene expression by overexpressing AP2 in C-33A cells or by performing transfection assays in different cell types.

The discovery of many potential transcription factor-binding sites in the hTAFII55 promoter-proximal region raises several interesting issues. Many TATA-less genes involved in DNA replication and cell cycle control have been reported to contain E2F- and Sp1-binding sites (77). Colocalization studies of cells at different stages of the cell cycle indicate that Sp1
may physically and functionally associate with E2F (77, 82). It remains to be investigated whether the E2F proteins also functionally interact with Sp1 on the hTAFII55 promoter. Intriguingly, the presence of a MEF2-binding site may account for the preferential expression of the TAFII55 mRNA in skeletal muscle, as revealed in Northern blotting analysis\(^3\). Although transfection assays and \textit{in vitro} transcription carried out in human cervical cancer cell lines (HeLa or C-33A) did not reveal the functional importance of E2F-, STAT-, and MEF2-binding sites in the hTAFII55 promoter, we cannot exclude the possibility that these factors are either limiting in our cells or they require additional cellular factors to support activator function. These interesting possibilities remain to be addressed in the future.

\textit{Acknowledgments}\textemdash We are grateful to R. G. Roeder for pHIV+58 and p-167; M. C. Thomas for pGL7072-161; J. Lasky for cloning of the hTAFII55 gene; S. Y. Hou, K. Kitiphongspattana, D. Vichugsananon and S.-Y. Wu for assistance in sequencing the hTAFII55 genomic clones; P. de Haseth, S. Y. Hou and S.-Y. Wu for helpful discussions; and C. Croniger, S. Y. Hou, P. M. MacDonald, D. McPheeters, D. Samols, M. Snider and S.-Y. Wu for comments on the manuscript.

\textbf{REFERENCES}

1. Breathnach, R., and Chambon, P. (1981) \textit{Annu. Rev. Biochem.} \textbf{50}, 349-383

2. Hernandez, N. (1993) \textit{Genes Dev.} \textbf{7}, 1291-1308
3. Burley, S. K., and Roeder, R. G. (1996) *Annu. Rev. Biochem.* **65**, 769-799

4. Smale, S. T. (1997) *Biochim. Biophys. Acta* **1351**, 73-88

5. Smale, S. T., and Baltimore, D. (1989) *Cell* **57**, 103-113

6. Carcamo, J., Buckbinder, L., and Reinberg, D. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 8052-8056

7. Weis, L., and Reinberg, D. (1992) *FASEB J.* **6**, 3300-3309

8. Javahery, R., Khachi, A., Lo, K., Zemie-Gregory, B., and Smale, S. T. (1994) *Mol. Cell. Biol.* **14**, 116-127

9. Burke, T. W., and Kadonaga, J. T. (1996) *Genes Dev.* **10**, 711-724

10. Kutach, A. K., and Kadonaga, J. T. (2000) *Mol. Cell. Biol.* **20**, 4754-4764

11. Verrijzer, C. P., Yokomori, K., Chen, J.-L., and Tjian, R. (1994) *Science* **264**, 933-941

12. Verrijzer, C. P., Chen, J.-L., Yokomori, K., and Tjian, R. (1995) *Cell* **81**, 1115-1125

13. Hansen, S. K., and Tjian, R. (1995) *Cell* **82**, 565-575

14. Kaufmann, J., Verrijzer, C. P., Shao, J., and Smale, S. T. (1996) *Genes Dev.* **10**, 873-886

15. Kaufmann, J., Ahrens, K., Koop, R., Smale, S. T., and Müller, R. (1998) *Mol. Cell. Biol.* **18**, 233-239

16. Martinez, E., Ge, H., Tao, Y., Yuan, C.-X., Palhan, V., and Roeder, R. G. (1998) *Mol. Cell. Biol.* **18**, 6571-6583

17. Chalkley, G. E., and Verrijzer, C. P. (1999) *EMBO J.* **18**, 4835-4845

18. Roy, A. L., Meisterernst, M., Pognonec, P., and Roeder, R. G. (1991) *Nature* **354**, 245-248

19. Roy, A. L., Du, H., Gregor, P. D., Novina, C. D., Martinez, E., and Roeder, R. G. (1997) *EMBO J.* **16**, 7091-7104
20. Grueneberg, D. A., Henry, R. W., Brauer, A., Novina, C. D., Cheriyath, V., Roy, A. L., and Gilman, M. (1997) Genes Dev. 11, 2482-2493

21. Yang, W., and Desiderio, S. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 604-609

22. Du, H., Roy, A. L., and Roeder, R. G. (1993) EMBO J. 12, 501-511

23. Seto, E., Shi, Y., and Shenk, T. (1991) Nature 354, 241-244

24. Burke, T. W., and Kadonaga, J. T. (1997) Genes Dev. 11, 3020-3031

25. Meisterernst, M., and Roeder, R. G. (1991) Cell 67, 557-567

26. Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S., and Reinberg, D. (1992) Cell 70, 477-489

27. Goppelt, A., Stelzer, G., Lottspeich, F., and Meisterernst, M. (1996) EMBO J. 15, 3105-3116

28. Mermelstein, F., Yeung, K., Cao, J., Inostroza, J. A., Erdjument-Bromage, H., Eagelson, K., Landsman, D., Levitt, P., Tempst, P., and Reinberg, D. (1996) Genes Dev. 10, 1033-1048

29. Willy, P. J., Kobayashi, R., and Kadonaga, J. T. (2000) Science 290, 982-985

30. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., and Ebright, R. H. (1998) Genes Dev. 12, 34-44

31. Qureshi, S. A., and Jackson, S. P. (1998) Mol. Cell 1, 389-400

32. Zhang, M. Q. (1998) Genome Res. 8, 319-326

33. Novina, C. D., and Roy, A. L. (1996) Trends Genet. 12, 351-355

34. Ren, B., and Maniatis, T. (1998) EMBO J. 17, 1076-1086

35. Chiang, C.-M., and Roeder, R. G. (1995) Science 267, 531-536

36. Lavigne, A.-C., Mengus, G., May, M., Dubrovskaaya, V., Tora, L., Chambon, P., and
Davidson, I. (1996) *J. Biol. Chem.* **271**, 19774-19780

37. Wieczorek, E., Brand, M., Jacq, X., and Tora, L. (1998) *Nature* **393**, 187-191

38. Ogryzko, V. V., Kotani, T., Zhang, X., Schlitz, R. L., Howard, T., Yang, X.-J., Howard, B. H., Qin, J., and Nakatani, Y. (1998) *Cell* **94**, 35-44

39. Grant, P. A., Schieltz, D., Pray-Grant, M. G., Steger, D. J., Reese, J. C., Yates, J. R., and Workman, J. L. (1998) *Cell* **94**, 45-53

40. Lavigne, A.-C., Mengus, G., Gangloff, Y.-G., Wurtz, J.-M., and Davidson, I. (1999) *Mol. Cell. Biol.* **19**, 5486-5494

41. Dantonel, J.-C., Murthy, K. G., Manley, J. L., and Tora, L. (1997) *Nature* **389**, 399-402

42. Wu, S.-Y., Thomas, M. C., Hou, S. Y., Likhite, V., and Chiang, C.-M. (1999) *J. Biol. Chem.* **274**, 23480-23490

43. Moqtaderi, Z., Yale, J. D., Struhl, K., and Buratowski, S. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 14654-14658

44. Shibuya, T., Tsuneyoshi, S., Azad, A. K., Urushiyama, S., Ohshima, Y., and Tani, T. (1999) *Genetics* **152**, 869-880

45. Purrello, M., Di Pietro, C., Viola, A., Rapisarda, A., Stevens, S., Guermah, M., Tao, Y., Bonaiuto, C., Arcidiacono, A., Messina, A., Sichel, G., Grzeschik, K.-H., and Roeder, R. (1998) *Oncogene* **16**, 1633-1638

46. Wu, S.-Y., and Chiang, C.-M. (1996) *BioTechniques* **21**, 718-725

47. Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A., and Roeder, R. G. (1993) *EMBO J.* **12**, 2749-2762

48. Rosen, C. A., Sodroski, J. G., and Haseltine, W. A. (1985) *Cell* **41**, 813-823
49. Kato, H., Horikoshi, M., and Roeder, R. G. (1991) *Science* **251**, 1476-1479

50. Hou, S. Y., Wu, S.-Y., Zhou, T., Thomas, M. C., and Chiang, C.-M. (2000) *Mol. Cell. Biol.* **20**, 113-125

51. Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988) *Nature* **335**, 563-564

52. Sadowski, I., and Ptashne, M. (1989) *Nucleic Acids Res.* **17**, 7539

53. Chiang, C.-M., Broker, T. R., and Chow, L. T. (1991) *J. Virol.* **65**, 3317-3329

54. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) *Mol. Cell. Biol.* **2**, 1044-1051

55. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., and Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035-7056

56. International Human Genome Sequencing Consortium. (2001) *Nature* **409**, 860-921

57. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J. etc. (2001) *Science* **291**, 1304-1351

58. Ince, T. A., and Scotto, K. W. (1995) *J. Biol. Chem.* **270**, 30249-30252

59. Boam, D. S., Davidson, W. I., and Chambon, P. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **270**, 19487-19494

60. Benson, L. Q., Coon, M. R., Krueger, L. M., Han, G. C., Sarnaik, A. A., and Wechsler, D. S. (1999) *J. Biol. Chem.* **274**, 28794-28802

61. Chalut, C., Gallois, Y., Poterszman, A., Moncollin, V., and Egly, J.-M. (1995) *Gene* **161**, 277-282

62. Scheer, E., Mattei, M.-G., Jacq, X., Chambon, P., and Tora, L. (1995) *Genomics* **29**, 269-272

63. van der Knaap, J. A., Borst, J. W., van der Vliet, P. C., Gentz, R., and Timmers, H. T. M.
(1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 11827-11832

64. Lewin, B. (2000) *Genes VII*, pp. 37-65, Oxford University Press, New York, NY

65. Kedes, L. H. (1979) *Annu. Rev. Biochem.* **48**, 837-870

66. Gentles, A. J., and Karlin, S. (1999) *Trends Genet.* **15**, 47-49

67. Wu, B., Hunt, C., and Morimoto, R. (1985) *Mol. Cell. Biol.* **5**, 330-341

68. Hattori, K., Angel, P., Le Beau, M. M., and Karin, M. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 9148-9152

69. Nagata, S., Mantei, N., and Weissmann, C. (1980) *Nature* **287**, 401-408

70. Huang, Z.-M., and Yen, T. S. (1995) *Mol. Cell. Biol.* **15**, 3864-3869

71. Trubetskoy, A. M., Okenquist, S. A., and Lenz, J. (1999) *J. Virol.* **73**, 3477-3483

72. Lewis, B. A., Kim, T. K., and Orkin, S. H. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7172-7177

73. Rougvie, A. E., and Lis, J. T. (1988) *Cell* **54**, 795-804

74. Mavrothalassitis, G. J., Watson, D. K., and Papas, T. S. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 1047-1051

75. Weis, L., and Reinberg, D. (1997) *Mol. Cell. Biol.* **17**, 2973-2984

76. Chang, D. J., Paik, Y. K., Leren, T. P., Walker, D. W., Howlett, G. J., and Taylor, J. M. (1990) *J. Biol. Chem.* **265**, 9496-9504

77. Karlseder, J., Rotheneder, H., and Wintersberger, E. (1996) *Mol. Cell. Biol.* **16**, 1659-1667

78. Pipaón, C., Tsai, S. Y., and Tsai, M.-J. (1999) *Mol. Cell. Biol.* **19**, 2734-2745

79. Marin, M., Karis, A., Visser, P., Grosveld, F., and Philipsen, S. (1997) *Cell* **89**, 619-628

80. Mitchell, P. J., Timmons, P. M., Hebert, J. M., Rigby, P. W., and Tjian, R. (1991) *Genes
Dev. 5, 105-119

81. Imagawa, M., Chiu, R., and Karin, M. (1987) Cell 51, 251-260

82. Lin, S.-Y., Black, A. R., Kostic, D., Pajovic, S., Hoover, C. N., and Azizkhan, J. C. (1996)

Mol. Cell. Biol. 16, 1668-1675
FIGURE LEGENDS

Fig. 1. Identification of the human TAFII55 promoter. A, schematic diagram of the 3’-8 genomic clone containing the hTAFII55 gene. The complete nucleotide sequence of the 3’-8 genomic clone, which contains a total of 17,042 base pairs (bp) represented by the thick bar with the hTAFII55 open reading frame (ORF) indicated by an open box, is deposited to GeneBank with accession number AF349038. The positions of the isolated hTAFII55 cDNA (35, 36) and its corresponding mRNA, relative to the genomic clone, are indicated by a thin line and a thick line, respectively. The 1459-bp genomic fragment encompassing the 5’ end of the hTAFII55 cDNA upstream of a luciferase reporter gene as in pGL2-TAF55(-1372/+87) is also depicted.

B, the 1.4-kb genomic fragment upstream of the 5’ end of the human TAFII55 cDNA has intrinsic promoter activity. Human cervical carcinoma C-33A cells were transiently transfected with reporter plasmids pGL2-TAF55(-1371/+87), pGL-HIV+80, pGL7072-161 or pGL2-Control, which contains the hTAFII55 promoter from -1372 to +87 (-1372/+87), the HIV-1 promoter from -167 to +80 (-167/+80), the HPV-11 promoter spanning 7072-7933/1-161 (7072/161) or the SV40 promoter/enhancer (Pro/Enh), respectively. The pGL2-Basic plasmid (vector), used for constructions of the above mentioned reporter plasmids, was also included as control. The luciferase activity was determined as described under "Experimental Procedures", and normalized to that of the hTAFII55 promoter construct.
Fig. 2. **Mapping of the transcription start site of the hTAF\(\text{I}\)55 promoter.** *A*, primer extension analysis of the *in vitro*-synthesized hTAF\(\text{I}\)55 transcript. *In vitro* transcription was conducted with HeLa nuclear extracts using the hTAF\(\text{I}\)55 promoter-containing construct pGL2-TAF55(-1372/+87) or the HIV-1 promoter-containing construct pHIV+58, in the absence (-) or presence (+) of 2 \(\mu\)g/ml of \(\alpha\)-amanitin. Two primers, Luc-1 and Luc-5, whose relative positions and expected sizes of primer extension products are indicated at the bottom, were used to determine the start site of the hTAF\(\text{I}\)55 gene, whereas a CAT primer that anneals to the chloramphenicol acetyltransferase reporter gene was used for mapping the start site of the HIV-1 promoter. DNA sequencing ladders, prepared from the phosphorylated forms of the corresponding primers and DNA templates as employed for *in vitro* transcription, were included for the assignment of the transcription start sites (indicated by arrows). The DNA sequences surrounding the transcription start sites are shown on the left of each panel with a bent arrow pointing to the major start site at +1 and solid square dots indicating relative intensities of the transcription signals. Two reproducible transcription signals, mapped to an upstream (indicated by an arrowhead) or downstream (indicated by an asterisk) location of the hTAF\(\text{I}\)55 cDNA, are marked on the right of the panels. *B*, RNase protection analysis of *in vivo* hTAF\(\text{I}\)55 transcripts. RNase protection assays were performed by first hybridizing *in vitro*-synthesized antisense riboprobe of 285 nucleotides (nt) spanning -128 to +87 with endogenous poly (A)\(^+\) RNA isolated from C-33A cells or with tRNA. RNase A and RNase T1 were then added to digest the single-stranded region. The protected fragments, along with a DNA size marker (A, C, G and T) and the original riboprobe (-) used to adjust the migration difference between DNA and RNA,
were then analyzed on a 5% polyacrylamide-8M urea gel and visualized after exposure to an X-ray film. The positions of the major protected fragment (87 nt) and the riboprobe are indicated respectively by arrows.

**Fig. 3.** **Promoter sequences containing Sp1- and AP2-binding sites are important for hTAFII55 promoter activity.** A, reporter gene assays performed in C-33A cells with hTAFII55 promoter constructs that sequentially remove potential transcription factor-binding sites. Transient transfection and reporter gene assays were performed as described under "Experimental Procedures" using plasmids containing the hTAFII55 promoter sequences with the indicated boundaries. The pGL2-Basic plasmid containing no insert was also used for transfection as control. Luciferase activity was normalized to that of the full-length promoter construct, pGL2-TAFII55(-1372/+87), and presented in the bar graph with error bars showing standard deviation. B, *in vitro* transcription and primer extension analysis of hTAFII55 promoter deletion constructs with HeLa nuclear extracts. *In vitro* transcription was performed in HeLa nuclear extracts with plasmids containing the hTAFII55 promoter sequences as indicated. The *in vitro*-synthesized transcripts were then mapped by primer extension using Luc-1 primer and analyzed on a 5% polyacrylamide-urea gel. The products with correct transcription start sites at +1 are indicated by an arrow. DNA sequencing ladders, prepared from pGL2-TAF55(-1372/+87) using the 5’-phosphorylated Luc-1 primer and [α-35S]dATP, were included to determine the position of the correctly initiated transcripts. Schematic diagrams of the promoter deletion constructs used for both experiments are drawn on the left, with potential transcription
factor-binding sites marked in boxes.

**Fig. 4.** The hTAF\(_{II55}\) promoter is regulated by proteins targeting the AREB6-, E47, Sp1- and AP2-binding sites in C-33A cells. A, reporter gene assays performed in C-33A cells with hTAF\(_{II55}\) promoter constructs with or without the AREB6- and E47-binding sites. Transient transfection and reporter gene assays were performed as described under "Experimental Procedures" using plasmids containing the hTAF\(_{II55}\) promoter sequences with the indicated boundaries. The pGL2-Basic plasmid containing no insert was also used for transfection as control. Luciferase activity was normalized to that of the full-length promoter construct, pGL2-TAF55(-1372/+87), and presented in the bar graph with error bars showing standard deviation. B, nucleotide substitutions in the Sp1- and AP2-binding sites reduce hTAF\(_{II55}\) promoter activity in C-33A cells. Transient transfection and reporter gene assays were performed with plasmids containing the hTAF\(_{II55}\) promoter sequences with the indicated boundaries. Asterisks indicate mutations introduced at specific protein-binding motifs in the plasmids. The nucleotides changed in each motifs are denoted at the bottom.

**Fig. 5.** Inr and DPE are both important core promoter elements for hTAF\(_{II55}\) gene expression. A, mutations at the Inr and the DPE reduce hTAF\(_{II55}\) promoter activity. Transient transfection and reporter gene assays were performed as described under "Experimental Procedures" using plasmids containing either wild-type or mutated nucleotides at the Inr and/or
the DPE of the hTAFII55 promoter fragment spanning -26 to +36. The pGL2-Basic plasmid (vector) containing no insert was also used for transfection as control. Luciferase activity was normalized to that of the wild-type promoter construct, pGL2-TAF55(-26/+36), and presented in the bar graph with error bars showing standard deviation. Asterisks and "X" indicate mutations introduced at specific protein-binding motifs in the plasmids. The nucleotides changed in each motifs are denoted at the bottom. B, the Inr and DPE modules of the hTAFII55 core promoter can mediate transcriptional activation in a heterologous promoter context. Transient transfection was performed in C-33A cells by cotransfecting different amounts of the Gal4-VP16-expressing plasmid (pSGVP), together with either wild-type (WT) or mutated reporter constructs driven by 5 Gal4-binding sites as indicated.
Figure 1A, Zhou and Chiang
Relative Luciferase Activity

Figure 1B, Zhou and Chiang
Figure 2A, Zhou and Chiang
Figure 3A, Zhou and Chiang

Relative Luciferase Activity

-1372/+87: 100.0%
-281/+87: 112.9%
-128/+87: 88.6%
-99/+87: 87.4%
-71/+87: 51.2%
-55/+87: 24.3%
-26/+87: 3.4%
-1372/-140: 0.1%
vector: 0.0%
Figure 3B, Zhou and Chiang
Figure 4A, Zhou and Chiang
Figure 4B, Zhou and Chiang
Figure 5A, Zhou and Chiang

Relative Luciferase Activity

-26/36

Inr*(-26/+36)
DPE*(-26/+36)
Inr*DPE*(-26/+36)
vector

-26

CCTCCCAGCCCGTGGTCGACGTACCTCTCGTTTTGTGCTGGTGGGACCGAGCCCGACGGAG

+36

Inr*: gtgg
DPE*: tcat
Figure 5B, Zhou and Chiang
The intronless and TATA-less human TAFII55 gene contains a functional initiator
and downstream promoter element
Tianyuan Zhou and Cheng-Ming Chiang

J. Biol. Chem. published online May 4, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M102875200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts