Molecular cloning and expression of the 32-kDa subunit of human TFIID reveals interactions with VP16 and TFIIB that mediate transcriptional activation

RICHARD D. KLEMM*, JAMES A. GOODRICH, SHARLEEN ZHOU, AND ROBERT TJIAN†

Howard Hughes Medical Institute, Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, CA 94720-3202

Contributed by Robert Tjian, March 8, 1995

ABSTRACT Transcription factor TFIID consists of TATA binding protein (TBP) and at least eight TBP-associated factors (TAFs). As TAFs are required for activated but not basal transcription, we have proposed that TAFs act as coactivators to mediate signals between activators and the basal transcription machinery. Here we report the cloning, expression, and biochemical characterization of the 32-kDa subunit of human (h) TFIID, termed hTAF132. We find that hTAF132 is the human homologue of Drosophila TAF140. In vitro protein–protein interaction assays reveal that as observed with Drosophila TAF140, hTAF132 interacts with the C-terminal 39-amino acid activation domain of the acidic transactivator viral protein 16 (VP16) as well as with the general transcription factor TFIIB. Moreover, a partial recombinant TFIID complex containing hTAF132 was capable of mediating in vitro transcriptional activation by the VP16 activation domain. These findings indicate that specific activator–coactivator interactions have been conserved between human and Drosophila and provide additional support for the function of these interactions in mediating transcriptional activation.

Regulated transcription of protein encoding genes in eukaryotes by RNA polymerase II requires a multitude of factors including transcriptional activators that bind DNA to enhance transcription in a gene-specific manner and basal factors (transcription factors TFIIB, TFIID, TFIIE, TFIIH, and TFIIF) that are required for promoter-specific initiation (for reviews, see refs. 1 and 2). Of these factors, TFIID specifically recognizes and binds to “TATA boxes” found within many promoters. We now know that TFIID is a multisubunit complex containing the TATA binding protein (TBP) and at least eight TBP-associated factors (TAFs; refs. 3–5; for review, see ref. 6). It was found that recombinant TBP can substitute for TFIID in directing basal levels of transcription in vitro (7, 8). However, transcription driven by TBP alone is not responsive to transcriptional activators. Thus TAFs appear to represent a unique class of factors in that they are not necessary to direct basal transcription from TATA-containing promoters but are required for transcriptional activation.

This observation (8) led to the proposal that some of the TAFs in the TFIID complex would directly contact activators and basal factors thereby mediating signals from activators to the basal transcription machinery and, furthermore, that different classes of activation functions (such as acidic, glutamine-rich, proline-rich, or isoleucine-rich) would contact distinct TAFs within the TFIID complex. This laboratory has now obtained (9–13) evidence for a number of interactions between TAFs and transcriptional activators and between TAFs and other basal transcription factors. Direct evidence for the function of TAFs as coactivators was obtained when recombinant TAFs and TBP were assembled into partial TFIID complexes that could support transcriptional activation in vitro by a number of activators (11, 12). For example, a partial complex consisting of the 250-kDa subunit of human TFIID (hTAF112), Drosophila (d) TBP, dTAF110, and dTAF110 can replace a crude TFIID fraction in mediating transcriptional activation by Sp1 (11). However, all of our studies examining the activity of partial TFIID complexes have used Drosophila coactivators. It was therefore important to test the generality of the specific TAF–activator interactions in mediating transcriptional activation by using the human subunits.

We have previously reported the cloning and characterization of dTAF110 (10). This 40-kDa subunit of dTFIID was found to specifically interact with the C-terminal 39-amino acid activation domain of viral protein 16 (VP16) (termed VP16C) and with the basal transcription factor TFIIB. Here we extend our analysis of coactivators for VP16 by cloning and characterizing the human homologue of dTAF110, hTAF132.‡ We have expressed recombinant hTAF132 in several forms and have tested their ability to interact with the activation domains of VP16 and with TFIIB. In addition, we have assembled partial TFIID complexes consisting of recombinant hTBP and hTAFs (including hTAF132) and tested the ability of these complexes to mediate transcriptional activation in vitro by a GAL4–VP16C fusion protein. Our studies demonstrate a functional conservation between dTAF110 and hTAF132 as coactivators for the C-terminal VP16 activation domain and provide further support for the role of TAFs in potentiation of transcriptional activation.

MATERIALS AND METHODS

Purification of hTFIID and Cloning of hTAF132. The TFIID fraction was purified from HeLa nuclear extracts by phosphocellulose chromatography followed by immunopurification on an anti-TBP column. Antibody affinity columns were made with ~200 µg of affinity-purified polyclonal TBP antibodies crosslinked to 200 µl of protein A-Sepharose with dimethyl pimelimidate (ref. 14, pp. 522–523). The phosphocellulose TFIID fraction (1 M KCl eluate) was passed over a mock column of protein A-Sepharose treated with crosslinker to reduce nonspecific background. The flowthrough of the mock column was then loaded on the anti-TBP column multiple times. After stringent washing, TFIID was eluted with 50 mM glycine (pH 2.5) containing 0.15 M NaCl and then subjected to reversed-phase HPLC on a C8 column as described (9). Peak fractions were collected and digested with the

Abbreviations: GST, glutathione S-transferase; TBP, TATA binding protein; TAF, TBP-associated factor; TAF132, etc., 32-kDa subunit of TFIID, etc.; VP16, viral protein 16; h, human; d, Drosophila; HA, hemagglutinin.

*Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.
†To whom reprint requests should be addressed.
‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. U21858).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.
protease Lys-C. The resulting peptides were separated by reversed-phase HPLC and microsequenced.

The amino acid sequence of one of the resulting peptides had a high degree of similarity to dTAFI40 (KDMGITYEPRVINQMLEAFARYVTILDDLK) and was used to design degenerate 17-mer oligonucleotides for PCR. A first-strand cDNA library was made from poly(A)+ mRNA isolated from HeLa cells and used as a template for PCR. A PCR product of the correct size was subcloned into a T vector (15) and sequenced. The DNA sequence of the PCR product was found to encode the correct amino acid sequence of the peptide. A 48-mer oligonucleotide was synthesized and used to screen a λgt10 library derived from HeLa-cell mRNA. The EcoRI fragments from five positive clones were characterized by restriction analysis and DNA sequencing. One of the five contained a short fragment that by sequencing was determined to be an N-terminal truncation of hTAFI32. The other four were identical by restriction mapping and contained the full-length cDNA of hTAFI32. Two of the clones were completely sequenced to arrive at the final sequence.

Expression of Recombinant hTAFI32 and Generation of Antibodies. An Nde I site was introduced at the 5' end of the coding region by site-directed mutagenesis for simplified cloning into various expression vectors. For expression of recombinant hTAFI32 in *Escherichia coli*, the cDNA was subcloned into pET19b (His-tagged; Novagen) and pSETA (untagged; Invitrogen). In addition, a C-terminal deletion of hTAFI32 was created by PCR such that the ATG start codon was contained in an Nde I site and a stop codon (TAA) was engineered after the codon for amino acid 181 of the protein. This truncated cDNA was subcloned into pET19b for subsequent expression of His-hTAFI32(N181). Proteins were expressed in *E. coli* (BL21DE3) by using the T7 expression system. For production of glutathione S-transferase (GST) fusion proteins, the hTAFI32 cDNA was subcloned into pGEX-2TK (Pharmacia) by using an EcoRV site 6 nt upstream of the start codon. Antisera against hTAFI32 was raised by injection of purified His-hTAFI32 (100 μg) into rabbits. Suitable antibodies were obtained after four injections.

GST Protein–Protein Interaction Assays. GST fusion proteins were expressed and purified from *E. coli* TG2 as described (10). In general, each assay contained 1 μg of GST fusion protein immobilized on 10 μl of glutathione-Sepharose beads (Pharmacia). Target proteins (100–200 ng) were incubated with GST and GST fusion protein beads generally for 1–2 h in TEGM buffer containing 0.1% Nonidet P-40 and NaCl at concentrations noted in the figure legends. The incubations were performed in 100 μl at 4°C with constant rotation (orbital mixing). After the incubation, the beads were washed at least four times with 10–15 ml volumes of buffer (same as load buffer). Protein was eluted from the beads in SDS sample buffer, resolved by SDS/PAGE, and blotted to a nitrocellulose membrane. Membranes were probed as described in the figures.

Assembly of Partial TFII D Complexes. Partial TFII D complexes were assembled as described by Chen et al. (11). A hemagglutinin (HA)-epitope-tagged version of hTAFI250 was expression in SF9 cells from a recombinant baculovirus as described (11). hTAFI70 was expressed in *E. coli* by using the pET expression system. hTBP was expressed in *E. coli* and purified to homogeneity as described (16). His-hTAFI32 (N181) was expressed in *E. coli* and purified for nondenaturing gel filtration. The complexes were assembled on protein A-Sepharose beads to which anti-HA antibodies had been crosslinked. Each assembly reaction mixture contained 20 μl of anti-HA beads and was incubated first with SF9-cell extract containing recombinant HA–hTAFI250 (40 μg) in TGEMC (TGEM containing 1% CHAPS) containing 0.4 M NaCl. The mixture was mutuated for 4 h at 4°C. The beads were then washed twice with 200 μl of TGEMC containing 1.0 M NaCl and four times with TGEMC containing 0.2 M NaCl. The HA–hTAFI250 beads were then used for the assembly of the remaining factors in the following order: (i) hTAFI70 (*E. coli* sonicate containing ~20 μg of hTAFI70), (ii) His-hTAFI32(N181) (20 μg), and (iii) hTBP (20 μg). Each incubation was performed for 4 h at 4°C with constant rotation, and was followed by four 200-μl washes with TGEMC containing 0.2 M NaCl. The complexes were eluted in 100 μl of TGEMC containing 0.2 M NaCl and HA-epitope peptide at 1 mg/ml. Elutions were performed for 1.5 h at 4°C followed by a 20-min incubation at room temperature. The eluted material was stored in small aliquots at −80°C.

In Vitro Transcription. Transcription reactions were performed in a partially fractionated HeLa system. Briefly, a HeLa nuclear extract was chromatographed over phosphocel-lulose as described (17). The 0.1 M KCl flowthrough fraction containing TFIIA was applied to a DE-52 column, and the TFIIA activity was eluted in a 0.35 M KCl step, followed by dialysis to 0.1 M KCl. The phosphocellulose 0.5 M fraction (P.5) containing the remaining basal factors, except TFIIID, was dialyzed to 0.1 M KCl. The phosphocellulose 1.0 M KCl fraction was dialyzed to 0.1 M KCl and applied to a DE-52 column. The flowthrough of this column, containing upstream factor stimulatory activity (USA) (18), was passed back over DE-52 to further deplete the fraction of TFIIID. TFIIID was eluted from the DE-52 column with a 0.3 M KCl step, followed by dialysis to 0.1 M KCl. Transcription reactions were performed in a volume of 20 μl containing 2.5 μg of P.5 fraction, 0.5 μg of USA fraction, 0.8 μg of TFIIA fraction, and 250 ng of pUC18 as a nonspecific competitor DNA template. In addition, each reaction mixture contained 100 ng of supercoiled plasmid DNA template consisting of five GAL4 binding sites upstream of the adenovirus Elb TATA box fused to a 371-bp G-less cassette (28) in a pSP72 backbone. GAL4–VP16c and GAL–(1–147) were expressed in *E. coli*, purified as described (10), and included in transcription reaction mixtures (25 ng) where indicated. The final reaction conditions and incubation times were as described (19).

RESULTS AND DISCUSSION

Molecular Cloning of hTAFI32. hTFIID was purified from HeLa-cell nuclear extracts by conventional chromatography and anti-hTBP affinity chromatography (4, 20). Two hundred liters of HeLa cells yielded ~10 pmol of TFII D for microsequencing. Several peptides were found to have sequence similarity to dTAFI40 (10). A PCR-based strategy was used to isolate cDNAs encoding these peptides. Several cDNAs were isolated and their DNA sequences were determined. The cDNA sequences revealed a transcript of 1143 bp and an open reading frame of 264 aa. Five peptides determined by microsequencing were found in the predicted protein sequence. The open reading frame encodes a protein with a predicted molecular mass of 28 kDa, and MS analysis of the His-tagged recombinant protein produced in *E. coli* confirmed the molecular mass (calculated Mr, 31,607; measured Mr, 31,608 ± 4). However, the endogenous TAF migrates in SDS/PAGE with an apparent molecular mass of 32 kDa (see below). Consequently, this subunit of hTFIID has been named hTAFI32. Comparison of hTAFI32 to dTAFI40 shows strong conservation in an N-terminal region and at the carboxy terminus (Fig. 1). However, the C-terminal halves of the *Drosophila* and human proteins are highly divergent. Throughout the entire length of these two proteins, there is 40% identity and 48% sequence similarity. A higher degree of similarity (57% identity and 69% similarity) is seen in the N-terminal 145 aa. Data base searches were performed by using BLAST (21). As with the dTAFI40 (10), hTAFI32 has sequence similarity to histone H3. This similarity was conserved in the well-conserved
The hTAF1132. hTAFs gel silver-stained hTAF,132 of found by Western blot analysis indicates that they are likely to represent the human homologues of dTAF130 and dTAF30a, and an as yet uncharacterized TAF, respectively.

**hTAFr32 Interacts with the C-Terminal 39-aa VP16 Activation Domain.** The herpes virus transactivator VP16 has been reported to contain two separable activation domains (10, 23–25). We have reported (10) that dTAFr40 interacts specifically with the C-terminal VP16C activation domain. We were therefore interested in determining whether VP16C would also interact specifically with hTAFr32. A GST fusion to hTAFr32 was expressed in E. coli and immobilized on glutathione-Sepharose beads. This affinity resin was then tested for its ability to bind hybrid GAL4–VP16 activators that contained the DNA binding domain of GAL4–(1–147) fused to different portions of VP16 (Fig. 3A): aa 412–490 (VP16NC), 412–456 (VP16n), and 452–490 (VP16c). These three hybrid proteins have been shown (10, 23, 24, 26) to activate transcription in vitro and in vivo. As shown in Fig. 3B, GST–hTARF32 bound both GAL4–VP16NC and GAL4–VP16c but did not bind GAL4–(1–147) and GAL4–VP16N. All of the hybrid activators failed to bind control GST beads. These results correspond exactly to those found for dTARF40 (10), thus showing a conservation of interactions between TAFs and activators across species. Moreover, these results further demonstrate the specificity of the interactions between the VP16C activation domain and these two TAFs since the similarly sized and charged VP16NC activation domain failed to interact with either the d- or hTAFs.

Because the N-terminal region of dTAFr40 and hTAFr32 were highly conserved, it seemed likely that this region was responsible for making contact with VP16C. As an initial test of this hypothesis, we have created C-terminal deletions removing 56 aa from dTAFr40 and 83 aa from hTAFr32. We have demonstrated (10) that dTAFr40(N222) was capable of interacting with immobilized GST–VP16NC. Here we show that both hTAFr32(N181) and dTAFr40(N222) interact efficiently with immobilized GST–VP16NC and GST–VP16c but not with control GST beads (Fig. 3C). Apparently, much of the nonconserved C termini of hTAFr32 and dTAFr40 are dispensable for binding VP16, whereas the conserved N-terminal region contains the interface important for recognition of VP16C.

**hTAFr32 Interacts with hTAFr70 and TFII.B.** Interactions between dTAFr40 and two additional proteins, dTAFr60 (27) and the basal transcription factor TFII.B (10), have been detected. The first of these interactions appears to be responsible for recruitment of dTAFr40 to the TFII.D complex, and the latter of the interactions may play a functional role in mediating transcriptional activation by VP16. As shown in Fig. 4A, immobilized His-TAFr32(N181) can specifically retain hTAFr70 from an E. coli extract. We believe that this interaction between hTAFr32 and hTAFr70 may actually function to recruit hTAFr32 to the TFII.D complex (see below).

To test for interactions between hTAFr32 and TFII.B, we performed an affinity chromatography experiment similar to those described above. Immobilized GST–hTAFr32 resin was incubated with TFII.B. Specifically bound proteins were resolved by SDS/PAGE, transferred to nitrocellulose, and probed with anti-TFII.B serum. As shown in Fig. 4B, hTAFr32 binds efficiently to GST–hTAFr32 but no detectable TFII.B bound to control GST beads. Thus the TFII.B interaction domain also appears to be conserved between hTAFr32 and dTAFr40. Moreover, these protein–protein interaction studies establish that several sets of interactions including TAF–TAF,
recombinant dTBP 

lized HA-hTAFII250 

with antiserum raised between factor, and TAF-basal activation transcriptional 

truncated TAFs 

VP16, retained on GAL4-VP16NC 

GAL4-VP16C 

binds GAL4-VP16NC and GAL4-VP16c to GST-hTAFII32 is evident in lanes 8 and 12, respectively. GAL4-(1–147) and GAL4–VP16N are not retained on the GST–hTAFII32 resin (lanes 6 and 10, respectively). (C) GST–VP16 binds C-terminal truncations of hTAFII32 and dTAFII40. GST–VP16C, GST–VP16NC, and GST beads were incubated with His-hTAFII32(N181) and dTAFII40(N222) in TGEM containing 0.2 M NaCl. The bound proteins were analyzed by Western blot analysis with antisera raised against either hTAFII32 (lanes 1–4) or dTAFII40 (lanes 5–8). Lanes 1 and 5 show the inputs of the assays. Both truncated TAFs bind GST–VP16C (lanes 2 and 6) and GST–VP16NC (lanes 3 and 7) but not control GST beads (lanes 4 and 8).

TAF–basal factor, and TAF–activator contacts are conserved between Drosophila and humans.

A Partial Recombinant TFII D Complex Containing hT AF II 32 Mediates Activation by GAL4–VP16C. The reconstitution of partial TFII D complexes that are functional in mediating transcriptional activation by specific activators has been reported (11, 12). These complexes were assembled by using immobilized HA–hTAFII250 as a building block upon which purified recombinant dTBP and dTAFs were added. However, the assembly of transcriptionally active TBP–TAF complexes has not been achieved with hTFIID subunits. Also, we were particularly interested in testing whether a partial complex containing hTAFII32 could mediate transcriptional activation by GAL4–VP16C. Pilot experiments with hTAFII32 and dTAFII40 indicated that these TAFs do not directly bind to HA–hTAFII250 nor to a complex containing HA–hTAFII250 and TBP (data not shown). Thus, it seemed likely that an additional subunit of TFII D would be required to bridge hTAFII32 to HA–hTAFII250. Protein–protein interaction experiments revealed that hTAFII70 had the potential to serve as a bridging factor, in that it could interact with hTAFII250 (27) and hTAFII32 (Fig. 4.4). We thus set out to assemble complexes from these four subunits of hTFIID.

The four recombinant TFII D subunits hTBP, HA–hTAFII250, hTAFII70, and His-hTAFII32(N181) were expressed and assembled into three partial TFII D complexes. All three complexes contained HA–hTAFII250 and hTBP. The triplex contained in addition hTAFII70, while the quadruplex included hTAFII70 and His-hTAFII32(N181). A silver-stained gel of the eluted complexes revealed that HA–hTAFII250 and hTBP were present at similar levels in all three complexes (Fig. 5A Upper), while hTAFII70 and His-hTAFII32(N181) appeared to be in substoichiometric amounts relative to HA–hTAFII250 and hTBP. We confirmed the presence of hTAFII70 and His-hTAFII32(N181) by Western blot analysis (Fig. 5A Middle and Lower).

As shown in Fig. 5B, the quadruplex was capable of mediating transcriptional activation by GAL4–VP16C (lanes 7 and 8). Thus, although the assembled quadruplex contained substoichiometric amounts of hTAFII70 and His-hTAFII32(N181), we were able to observe a significant level of transcriptional activation. Interestingly, the triplex seemed to direct a low but reproducible level of stimulation by GAL4–VP16C, and an interaction between hTAFII70 and GST–VP16C has been detected (C. Thut and R.T., unpublished data). But under the conditions of these transcrip-
Fig. 5. hTAFIS2 functions as a coactivator for GAL4-VP16c. (A) Recombinant partial TFIID complexes. Three complexes were assembled that all contained HA-hTAFIS250 and hTBP. In addition, the triplex contained hTAFIS70 and the quadruplex contained hTAFIS70 and His-hTAFIS32(N181). (Upper) Silver stain of the complexes resolved by SDS/PAGE in 8.5% gels. HA-hTAFIS250 and hTBP can be seen in all three lanes; however, the other two subunits are not visible in this silver stain. Western blot analysis was performed to establish the presence of hTAFIS70 (Middle) and His-hTAFIS32(N181) (Lower). (B) A partial TFIID complex containing hTAFIS32 mediates transcriptional activation by GAL4-VP16c. In vitro transcriptional activation by GAL4-VP16c was tested in a TFIID-dependent system using a TFIID fraction (lanes 1 and 2), the duplex (lanes 3 and 4), the triplex (lanes 5 and 6), and the quadruplex (lanes 7 and 8). In addition, the quadruplex was tested for GAL4-(1-147) transcriptional activation (lanes 9 and 10). The data shown are from a single representative experiment, but lanes 1 and 2 were exposed to film three times longer than lanes 3–10.