

# Cloning and Characterization of Human TAF20/15

MULTIPLE INTERACTIONS SUGGEST A CENTRAL ROLE IN TFIID COMPLEX FORMATION\*

(Received for publication, March 5, 1996, and in revised form, May 3, 1996)

Alexander Hoffmann<sup>‡</sup> and Robert G. Roeder<sup>§</sup>

From the Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York 10021

**TFIID is a multiprotein complex that plays a central role in the initiation and regulation of class II transcription. Transcription factor IID (TFIID) nucleates transcription initiation complex formation by direct core promoter binding and mediates the action of transcriptional activators, in part via direct interactions with them. Molecular studies of the TFIID complex have identified multiple subunits whose potential interactions can be recapitulated *in vitro* with recombinant polypeptides. Here we report the cloning of human TATA box binding protein (TBP)-associated factor 20 (TAF20) and the consequent identification of an additional, related TFIID subunit, human TAF15 (hTAF15). Multiple TAF20/15 interactions have been detected within native TFIID preparations and further analyzed with recombinant subunits. Along with the demonstration of a high affinity association between TAF20/15 and TBP, the present results suggest that hTAF20/15 may complement hTAF250 in directing the association of TAFs with TBP to form a TFIID complex. Finally, we present detailed mutagenesis studies that reveal multiple, distinct interaction surfaces on the presumed globular domain of hTAF20/15 and may be used, in conjunction with structural data, to model the architecture of the TFIID multiprotein complex.**

Cell-free transcription systems capable of accurate transcription initiation on a promoter-containing plasmid template (1, 2) have served as a basis for the identification and characterization of a number of human class II General Transcription Factors (GTFs)<sup>1</sup> (3, 4) (reviewed in Ref. 5). One of these, TFIID, was shown to act through the TATA box, a ubiquitous class II core promoter element required for stable preinitiation complex formation (6–8). Early characterization of the human

TFIID-core promoter interactions revealed a remarkably large footprint on the adenovirus major late core promoter (Ad2 MLP) extending from –47 to +35 and covering both TATA box and the transcription initiation site (7). On other, weaker promoters, however, the human TFIID-DNA interactions were found to be restricted to the TATA box, thus suggesting multiple modes of binding (8).

Studies on the adenovirus E4 promoter revealed an interesting connection between the alternative footprinting patterns: activators bound upstream of the TATA element induce the downstream extension of the footprint of prebound TFIID, which in turn correlates with increased recruitment of GTFs and activated transcription (9–11). These observations, as well as subsequent studies with affinity-purified TFIID (12), suggest that the TFIID-promoter complex is dynamic and capable of reversibly assuming at least two distinct conformations. Furthermore, direct interactions between TFIID and activators, first shown with USF on the Ad2 MLP (7), indicated a possible functional significance: namely, that activators binding to their cognate sites can modulate TFIID's interaction with DNA, thus increasing transcription initiation rates.

The cloning of the TATA box-binding subunit, TBP, from several organisms allowed immunoaffinity purification and molecular characterization of TFIID (reviewed in Ref. 13). This has included the identification of evolutionary conserved TBP-associated factors (TAFs), including  $\geq 13$  in human TFIID,  $\geq 9$  in *Drosophila*, and  $\geq 10$  in yeast (reviewed in Refs. 14–16), which are required, in addition to TBP, for co-activator function of TFIID (17–20) and for activator-independent (basal) transcription from TATA-less promoters (21). Molecular characterization of purified TFIID and recombinant subunits have largely substantiated the notion that interactions between activators and specific TFIID subunits, including both TBP and individual TAFs, indeed have functional relevance (reviewed in Refs. 14 and 15). Reconstituted partial TFIID complexes have provided direct evidence that distinct activators have differential TAF requirements (see Ref. 22 and references therein).

While activator-TFIID interactions may have a stabilizing effect on the preinitiation complex, little insight has been gained into the molecular mechanism of activation. Whether activators function by effecting enhanced TFIID recruitment to the promoter (12, 23, 24) or qualitative changes in TFIID-DNA interactions that lead to enhanced downstream factor recruitment (10, 11), possibly through direct TAF-GTF interactions (15, 25, 26), currently remains unclear. To gain insight into the mechanism of transcriptional regulation, attempts have been made to describe the structure of the TFIID complex in terms of protein interactions between its isolated components (see Refs. 27–29 and references therein). Thus the largest subunit (dTAF230/hTAF250) has been proposed to direct the assembly of other TAFs into a TFIID complex (30, 31), but in the absence of appropriate analyses of native TFIID an understanding of the structural organization of this multiprotein complex both

\* This work was supported in part by a Boehringer-Ingelheim Graduate Fellowship (to A. H.), by grants from the National Institutes of Health and the Human Frontiers Science Program (to R. G. R.), and by the Pew Trust to the Rockefeller University. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) U57693.

‡ Current address: Dept. of Biology, MIT, 77 Massachusetts Ave., Cambridge, MA 02139. Tel.: 617-252-1574; Fax: 617-253-2153; E-mail: alexhoff@MIT.edu.

§ To whom correspondence should be addressed: Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, 1230 York Ave., New York, NY 10021. Tel.: 212-327-7600; Fax: 212-327-7949.

<sup>1</sup> The abbreviations used are: GTF, general transcription factor; TF, transcription factor; TBP, TATA box binding protein; hTBP, human TBP; dTBP, *Drosophila* TBP; TAF, TBP-associated factor; hTAF, human TAF; dTAF, *Drosophila* TAF; Ad2 MLP, adenovirus major late core promoter; GST, glutathione *S*-transferase; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; EST, expressed sequence tag.

on and off the DNA remains elusive.

In this study, we describe the cloning of hTAF20, the identification of hTAF15, and a detailed molecular characterization of their interactions within the TFIID complex. These results support and extend a novel model concerning the architecture of the TFIID complex.

#### EXPERIMENTAL PROCEDURES

**Cloning of hTAF20/15**—TFIID was purified from f:TBP-expressing cell line (3–10) as described previously (32) and electrotransferred to a polyvinylidene difluoride membrane. Thermolysine-digested peptides were separated by reverse-phase high performance liquid chromatography and submitted to the Protein Sequencing Facility of the Rockefeller University for peptide sequencing. Peptides derived from hTAF20 yielded one usable sequence VQLHLERQXNMXIPGFG. A degenerate oligonucleotide 5'-GTICAACTICAYCTIGAACGCARIIIACATGIIIATICIGGITTGG-3' was used as an end-labeled probe to screen 10<sup>6</sup> phage of a human placental oligo(dT)-primed cDNA library (15) at moderate stringency (15% formamide, 42 °C) and resulted in six independent, but multiply isolated clones. Complete bidirectional sequencing with the aid of internal primers showed that they possessed ORFs containing the peptide sequence.

**Sequence Analysis**—Data base searches were done with the BLAST program and subsequent sequence alignments with MacVector (Kodak) which was also used for computer-aided secondary structure predictions (Chou-Fasman and Robson-Garnier algorithms).

**TAF20/15 Plasmid Constructions**—Expression plasmids for TAF20, TAF15 and derivatives thereof were constructed by polymerase chain reaction creating an *NdeI* site at the N-terminal end (utilizing an existing methionine residue or thereby creating a new one) and a *BglII* site at the C-terminal end following the natural or a newly created stop codon. Polymerase chain reaction-generated fragments were then inserted into bacterial expression plasmids at *NdeI* and *BamHI* sites and sequenced. Primers for polymerase chain reaction were designed to contain the bases to be mutated in their center with 10–12 bases extending on either side that are complementary to the TAF20 sequence at that position. (Due to the large number of oligonucleotides used, we have refrained from stating their exact sequences, but this information can be furnished upon request.)

Specifically, plasmids for Fig. 2A were TAF20-pET11a, TAF15-pET11a, and C109-pET11a allowing expression in the T7-mediated expression system (33). Antigens for rabbit immunizations were produced in 6HisT-pET11d (TAF20, TAF15, C65) and -T6His-pET11a (TAF20ΔC65) (34), which was generated with a C-terminal primer, creating an *XhoI* site for in-frame fusion to the C-terminal His tag.

GST fusion constructs (Fig. 4) were either made in pGEX-2T(+), a pGEX-2T (Pharmacia Biotech Inc.) derivative in which the polylinker was expanded to contain the sites *NcoI-NdeI-XhoI-HindIII-AflII-BglII-BamHI-SmaI-EcoRI*, or in newly constructed pGET3 or pGET11, which are GST-containing pET3 and pET11 vectors containing the cloning sites *NcoI-NdeI-XhoI-HindIII-AflII-BglII-BamHI*, and which are particularly suitable for the expression of genes that are toxic to bacteria as a result of the low uninduced level of expression in the T7-mediated system (33). Both vectors were constructed for general application, but are not commercially available.

**Generation of Antibodies against TAF20/15 and hTBP**—His-tagged TAF20/15 and hTBP constructs (see above) were used to express and affinity-purify the following polypeptides for immunization of two rabbits each: 6HisT-TAF20, 6HisT-TAF15, 6HisT-C65, TAF20Δ65-T6His, and 6HisT-hTBP (34). Rabbits were injected with 200 μg of protein (as an emulsion with Freund's adjuvant) once a month and bled twice a month for 18 months. Third month bleeds of α-hTBP, α-C65, and α-TAF20ΔC65 were affinity-purified with the respective antigen to yield more than 10 mg of pure antibody; 1 mg was bound to 1 ml of protein A-Sepharose (Pharmacia) and cross-linked with DMP for use in immunoprecipitation experiments.

**Immunologic Methods**—TFIID was immunoprecipitated (Figs. 2B and 3) from 1 ml of P11 0.85 M KCl fraction (32) by binding at 4 °C for 4 h to 20 μl of immunoaffinity resin in the presence of 0.2% Nonidet P-40 and 0.85 M KCl. Immunoprecipitates were washed five times at similar stringency and eluted with SDS sample buffer. Partial disruption of TFIID complexes bound to anti-TBP immunoaffinity columns was achieved by adding the detergent Sarkosyl to indicated concentration to the wash buffer.

Western blotting involved electrotransfer of SDS-PAGE-separated proteins in a Hoefer semidry apparatus to Schleicher & Schuell's BA85S (0.4 μm) or BA83S (0.22 μm) for proteins ≤20 kDa. Membranes

were blocked in 5% skim milk, Tris-buffered saline + 0.2% Tween 20, reacted to antibodies at various dilutions, and washed with Tris-buffered saline + 0.2% Tween 20. Signals were obtained with horseradish peroxidase-linked secondary antibodies developed by the ECL system (Amersham Corp.).

**Interaction Assays**—GST fusion proteins were expressed in bacteria XA90 or BL21 bacteria and sonicated in 0.1 volume of lysis buffer (34), separated from insoluble debris by ultracentrifugation, and stored at –80 °C until used. The volume of lysate employed in one interaction assay was titrated with lysate from bacteria containing the insertless vector to yield about 2 μg of fusion protein in a 200-μl reaction volume. Each purified tester protein was added (≈100 ng) to yield a reasonable signal with 1/20 of the input amount. Proteins were incubated in the presence of 10 μl of bed volume glutathione-Sepharose (Pharmacia) in CRC tubes (U. S. Biochemical Corp.) at 30 °C for 30 min in lysis buffer (0.5 M NaCl) including 0.2% Nonidet P-40 and 0.02% Sarkosyl, except for the TAF135 and TAF55 interactions which were tested in buffer containing 0.2 M NaCl, 0.2% Nonidet P-40 (no Sarkosyl) at 4 °C. Each lysate was washed five times in incubation buffer and eluted in 50 μl of SDS sample buffer. Ten-μl samples were subjected to electrophoreses on two identical gels each. One was stained with Coomassie Blue (Fig. 4A) to confirm that fusion proteins were present in equal amounts, while the other was subjected to Western analysis (Fig. 4, B–D and F) or autoradiography (Fig. 4E).

#### RESULTS

**Cloning of Human cDNAs Encoding TAF20 and TAF15**—Polypeptide components of affinity-purified TFIID (32) were separated by SDS-PAGE and subjected to microsequence analysis. One peptide sequence derived from a TAF with SDS-PAGE mobility of 20 kDa was used to design oligonucleotides to screen a human cDNA library. Overlapping clones were sequenced and revealed an open reading frame which encodes a 161-residue polypeptide (molecular mass = 17,913 daltons) that contains the natural TAF20-derived peptide sequence near its C terminus (Fig. 1A). Curiously, only the second ATG within the open reading frame is set within a good Kozak consensus (35), possibly indicating multiple translation initiation sites. Translation in reticulocyte lysates of mRNAs derived from a cDNA clone spanning the complete open reading frame resulted in the appearance of two polypeptides, one of 20 kDa and the other of 15 kDa, while mRNAs derived from truncated cDNAs containing the second but not the first in-frame ATG gave rise only to the 15-kDa polypeptide (data not shown).

To verify that the polypeptides encoded by the cloned cDNAs are in fact subunits of TFIID, antibodies were generated in rabbits against the full-length protein, or fragments thereof, that had been expressed in and purified from bacteria. Western blot analysis with high titer sera or antigen-purified antibodies confirmed the presence of 20- and 15-kDa immunoreactive bands in both crude (phosphocellulose, 0.85 M KCl) and affinity-purified (f:TBP-IID) TFIID preparations (Fig. 2A), but little in other phosphocellulose fractions (data not shown). As predicted, the immunoreactive polypeptides showed SDS-gel mobilities that were indistinguishable from bacterially expressed proteins initiated at the first and the second ATG in the cloned cDNA (Fig. 2A). These results confirm that two low molecular weight TAFs are generated from a single gene, whose derived cDNA sequence is presented in Fig. 1A. While Northern blot analysis showed a single hybridizing RNA species within the limits of resolution (data not shown), it is presently unclear whether utilization of the second downstream ATG is the result of alternative splicing. Comparison of the relative abundance of the two forms of this TAF in various human cell lines and mouse tissues by Western blot analysis did not reveal dramatic differences; the longer form, TAF20, is between 2- and 5-fold more abundant than the shorter TAF15 version in all cell types examined (data not shown).

Antigen-purified antibodies against the hydrophilic N-terminal part of TAF20/15 were immobilized on Sepharose beads



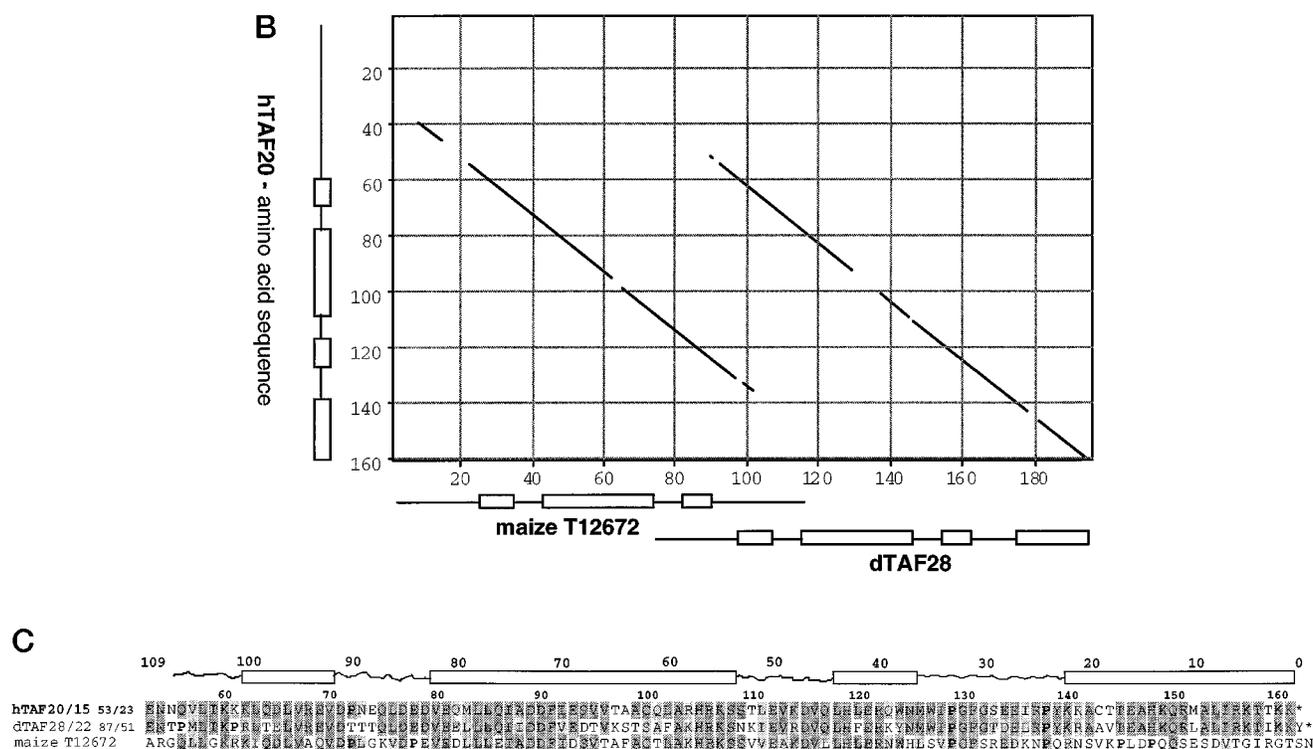


FIG. 1.—continued

a TFIID subunit or identify the related hTAF15.

**Sequence Analysis and Homologs**—Comparison of the cDNA sequence of TAF20/15 with sequences of other recently cloned TBP-associated proteins revealed high similarities to *Drosophila melanogaster* TAF28/22 (36), alternatively termed dTAF30 $\alpha$  (28). Fig. 1B presents a matrix comparison which demonstrates that the high degree of co-linear conservation between the two sequences is confined to the C-terminal 109 residues (65% identity, 79% similarity). The N termini, though similarly hydrophilic, are of different length and exhibit no sequence relationship. Interestingly, however, both human and *Drosophila* cells produce this TAF in a short and a long form (hTAF20/15 and dTAF28/22), suggesting a possible functional significance for alternate N-terminal extensions.

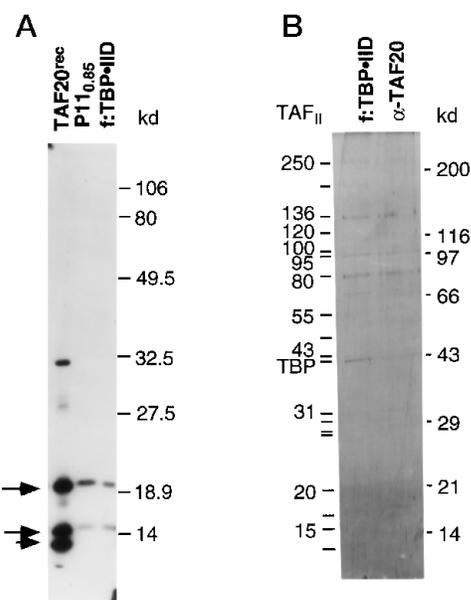
In addition, data base searches identified a short expressed sequence tag (EST) derived from maize (*Zea mays*) with sequence similarity to TAF20 (zEST00038). Sequencing the parental cDNA clone (GenBank<sup>TM</sup> T12672) revealed that the encoded protein is highly homologous along a 90 amino acid stretch (56% identity, 76% similarity) and therefore is likely to be the maize homolog of hTAF20/15.

Commonly used algorithms for secondary structure prediction of hTAF20/15 indicate relatively unstructured N-terminal tails attached to a globular domain, consisting of four  $\alpha$ -helices, that coincide with the 109 residue region conserved between the *Drosophila* and human proteins. A sequence alignment alongside the predicted secondary structure of this putative domain (Fig. 1C) details the extent of its evolutionary conservation and reveals that the C-terminal helix is absent in the maize polypeptide while the N-terminal three helices are clearly conserved. As both human and *Drosophila* proteins exist with alternate N termini we indicate amino acid residues with numbers starting from the C terminus (Fig. 1C), to avoid confusion in denoting TAF20/15-derived constructs used in the following mutagenesis analysis of the globular domain. Given our sequence analysis and evolutionary comparison, the available data does not, we believe, support an alternative second-

ary structure prediction published for dTAF30 $\alpha$  (28).

**TAF20/15 Interactions within TFIID**—While pairwise interaction studies with recombinant *Drosophila* homologs of hTAF20/15, dTAF28/22 (36) or dTAF30 $\alpha$  (28), demonstrated *in vitro* associations with dTBP and the *Drosophila* hTAF135 homolog, dTAF110, we undertook to develop experimental approaches to reveal TAF-TAF interactions within the native TFIID complex that may give physiological justification for subsequent studies with recombinant polypeptides. First, we exploited the heat labile nature of TBP in anti-TAF20/15 immunoprecipitation experiments in order to identify TFIID constituents that might be in direct contact with hTAF20/15 (Fig. 3, A–C). Mild heat-treatment leads to functional inactivation of TFIID (8) via denaturation of TBP, such that TBP-associated factors are not present in heat-treated anti-TBP immunoprecipitates (37). Conversely, when TAF20/15 is immunoprecipitated from a heat-treated P11<sub>0.85</sub> M KCl fraction, TBP is not co-precipitated as revealed by silver-staining (Fig. 3A) or immunoblotting (Fig. 3B). The resulting immunoprecipitate does, however, contain associated polypeptides that stand out, upon silver staining, above the background (Fig. 3A, lane 3) and appear to represent a subset of TAFs. Western analysis with anti-TAF sera reveals that TAF55 and TAF135 polypeptides are still associated, whereas TAF250, TAF95, TAF80, and TAF31 are not (Fig. 3C, lane 3). Similar results were obtained when intact TFIID purified with anti-TAF20-Sepharose was gently washed with buffer at 47 °C for 30 min (data not shown), thereby minimizing the possibility of nonspecific retention of dissociated proteins. These data are highly suggestive that either TAF135 or TAF55, or both, are in direct contact with TAF20/15 within native TFIID.

As a second approach, we sought to partially disrupt the TFIID complex to identify direct protein interactions within this multiprotein complex. While the entire complex appeared stable to high salt concentrations, many detergents as well as low concentrations of urea or guanidinium led to the apparent denaturation of human TBP and thus complete disassembly of



**FIG. 2. The cloned cDNAs encode two bona fide subunits of human TFIID, hTAF20 and hTAF15.** *A*, Western blot analysis of TAF20/15 in human TFIID. Antibodies against recombinant TAF20 (N-terminal 96 residues) were used to probe proteins present in a phosphocellulose fraction ( $P11_{0.85}$ ) and in an immunopurified native TFIID complex ( $f:TBP-IIID$ ) following SDS-PAGE and electrotransfer onto a nitrocellulose membrane. Arrows indicate specific bands whose mobilities are indistinguishable from bacterially expressed portions of the TAF20 ORF; the upper arrow indicates a polypeptide produced from the first in-frame ATG (161 amino acids), the lower arrow a polypeptide from the second ATG (131 amino acids), and the arrowhead a polypeptide containing the C-terminal 109 residues of the TAF20 ORF. *B*, immunoprecipitation of TFIID with  $\alpha$ -TAF20 antibodies. Antigen affinity-purified antibodies against TAF20 (N-terminal 96 residues) cross-linked to Sepharose resin were used to immunopurify TAF20-containing complexes from the  $P11_{0.85}$  fraction ( $\alpha$ -TAF20) and were compared to polypeptides contained in an immunopurified TFIID preparation ( $f:TBP-IIID$ ). A silver-stained SDS-PAGE (7–17% gradient) is shown with molecular mass markers on the right and positions of previously identified class II TAFs (32) indicated on the left.

the complex, akin to heat-treatment (37). However, we found that stripping of anti-TBP-immunoprecipitates with buffers containing the detergent Sarkosyl resulted in partially disrupted TFIID complexes whose composition is directly related to the concentration of this detergent (Fig. 3D). Interestingly, TBP immunoprecipitates washed with buffers containing 0.02 and 0.05% Sarkosyl have lost most immunologically detectable TAFs but retain substantial amounts of TAF20/15 (lanes 3 and 4). Furthermore, retention of TAF20/15 and TAF135, and possibly TAF80 and TAF31, at 0.02% Sarkosyl (lane 3) does not appear to require the presence of TAF250. While we cannot exclude a nonspecific affinity of TFIID components to the immunoaffinity resin or to molecular contaminants, the result is indicative of a highly stable, direct association between TBP and TAF20/15 within native TFIID that is only fully disrupted with 0.1% Sarkosyl-containing wash buffers (lane 5). *In vitro* interaction studies with purified recombinant human TAF polypeptides has allowed examination of these interactions in more detail.

**Mapping Intra-TFIID Interactions onto hTAF20/15**—Detailed mapping of intra-TFIID protein interaction sites on TAF20/15 could lead to insights into the architectural arrangement of TFIID constituents within the complex. To this end we constructed a systematic series of TAF20/15 mutants that would allow identification of the smallest fragment of TAF20/15 capable of mediating a specific interaction in an *in vitro* interaction assay. All mutants were designed according to

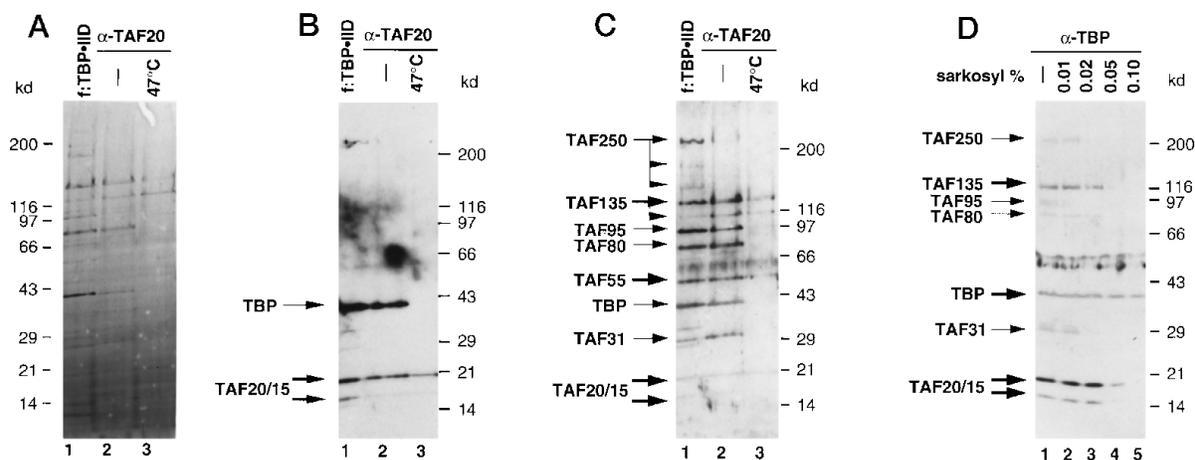
secondary structure predictions proposed in Fig. 1C, except for the constructs beginning or ending in the residue (C65) which falls within a previously proposed loop region linking two long  $\alpha$ -helices (28).

TAF20/15-derived ORFs were expressed in bacteria as fusion proteins with glutathione *S*-transferase, and amounts were normalized using Coomassie blue-stained SDS-gels (Fig. 4A) following affinity purification over glutathione-Sepharose to allow for a semiquantitative analysis. It seems noteworthy that all recombinant fusion proteins were resistant to rapid proteolysis with the exception of mutant protein C65-19. Such instability of a recombinant protein is often interpreted to be indicative of a lack of secondary structure. The set of TAF20/15 deletion mutants is thus more consistent with our structure prediction in Fig. 1C than with a previously published secondary structure proposal for dTAF30 $\alpha$  (28). In addition, we note that owing to the design of the deletion mutants, and possibly the presence of the attached GST moiety, most of the TAF20/15-derived fusion proteins are likely to attain some degree of secondary structure. However, the possibility that the residual TAF20/15 segments do not generate a completely natural fold remains a caveat commonly associated with mutagenesis studies.

Equal amounts of recombinant GST fusion proteins were used in an *in vitro* binding assay with a 1000-fold excess of bacterial proteins as nonspecific competitor to characterize interactions with previously identified TFIID constituents. Precipitates were washed in buffers containing both high salt (0.5 M) and the detergents Nonidet P-40 (0.2%) and Sarkosyl (0.02%), before probing them by Western analysis for the presence of co-precipitating proteins (results presented in panels B–F of Fig. 4 and summarized in panels G and H). Bacterially expressed full-length human TBP and the N-terminally truncated TBP core domain are thus shown, as expected from our observations with native TFIID preparations (Fig. 3D), to interact directly with hTAF20/15 and derived fragments (Fig. 4B). As the N terminus of full-length human TBP (upper arrow) causes considerable background in this assay, the TBP core domain (lower arrow), which is capable of assembly into a complete and fully functional TFIID complex (38) and mediates this specific interaction, proved a more reliable indicator. Two small fragments were each found to be sufficient for a relatively efficient interaction with the conserved TBP core domain: the long central helix B in C89-51 (lane 24) and the most C-terminal loop present in C36 (lane 21), which together may form a TBP interaction surface in the histone-fold. In similar experiments bacterially expressed TBPs from the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* and the plant *Arabidopsis thaliana* were also shown to interact specifically with hTAF20/15, demonstrating that this interaction is apparently not “species-specific” but biochemically conserved in evolution (37).

Other TFIID components were tested for association with hTAF20/15 within the *in vitro* GST interaction assay. No specific direct interaction could be demonstrated with hTAF95 or hTAF31 (data not shown) (39), whereas baculovirus-expressed hTAF80 was retained on columns bearing GST-TAF20 fusion protein (Fig. 4C). The specificity of this interaction is demonstrated by mutant studies that localize this activity to the C-terminal helices of the globular domain (C36, C53-19, and C89-51 in lanes 21, 22, and 24). The *in vitro* binding assays presented in Fig. 4D demonstrate that hTAF20/15 can interact with itself through its central helices B and C (C89-36 in lane 18), consistent with previously observed self-association of dTAF30 $\alpha$  (28).

We next investigated possible direct interactions with hu-



**FIG. 3. TAF20/15-protein interactions within native TFIID.** A, detection of TBP-independent protein interactions by TAF20/15. Phosphocellulose 0.85 M KCl (crude TFIID) fractions derived from f:TBP cell-line (lane 1) or HeLa cells (lanes 2 and 3) were used in immunoprecipitation experiments with M2-agarose (lane 1) or antigen-purified anti-TAF20/15 antibody coupled to protein A-Sepharose (lanes 2 and 3). The starting material for the sample in lane 3 was incubated at 47 °C for 15 min. Immunoprecipitation samples were separated on gradient SDS-PAGE (7–17%) and stained with silver. B, Western blot for TBP and TAF20/15 of immunoprecipitates described in A. Antisera specific for TBP and for TAF20/15 were mixed and employed in this immunoblot. The thin arrow indicates TBP, the thick arrows TAF20 and TAF15. C, Western blot for TFIID subunits of immunoprecipitates described in A. Antisera specific for individual TFIID subunits were mixed and employed in this immunoblot. The position of human TAFs is indicated. Triple arrows for TAF250 and TAF135 indicate multiple immunologically related bands whose physiological significance is currently unclear. Thick arrows are used for TAFs also present in lane 3. D, Western blot for TFIID subunits of detergent-disrupted TBP immunoprecipitates. Phosphocellulose, 0.85 M KCl (crude TFIID) fractions were immunoprecipitated with anti-TBP-Sepharose and washed in buffers containing indicated concentrations of Sarkosyl. Retained polypeptides were separated by gradient (7%–17%) SDS-PAGE and probed with a mixture of anti-TAF sera as in C. Arrows indicate respective human TAFs on the left and molecular mass markers are shown on the right.

man TAF135 and TAF55, which were suggested by the anti-TAF20/15 immunoprecipitation experiments of disrupted TFIID complexes (Fig. 3, A–C). An N-terminally truncated TAF135 derivative<sup>2</sup> that spans the entire domain conserved between the human and the *Drosophila* (dTAF110) proteins was produced in reticulocyte lysates and employed in binding assays with TAF20/15 derivatives. Direct interactions were indeed detected at moderate salt (Fig. 4E) and specifically localized to proposed helix C (C53–36 in lane 28). Similarly, bacterially expressed TAF55 (40) was bound specifically by full-length TAF20/15 and fragments thereof (Fig. 4F), indicating that its interaction surface is formed by helix C and the C-terminal adjacent loop; these domains are present jointly in C53–19 (lane 23) and individually in C36 (lane 21) and C53–36 (lane 28). Thus helix C and its adjacent loops represent a major interaction interface to other components of the TFIID complex as summarized in Fig. 4H.

#### DISCUSSION

Concerted efforts over the last few years aimed at a molecular characterization of the general initiation factor TFIID have led to the identification of more than a dozen human polypeptide components and their preliminary biochemical characterization (16). Classical studies implicating TFIID in activation mechanisms (7, 9, 10, 23, 24) have been confirmed and extended with the identification of individual subunits whose presence and capacity to interact with activators is correlated with transcriptional activation (22, 26, 40–42). However, to elucidate the biochemical mechanism underlying transcriptional regulation and TFIID's role in this process, a complete characterization of the TFIID complex and its conformational states is imperative and mandates the identification of all of its components and characterization of their three-dimensional arrangement.

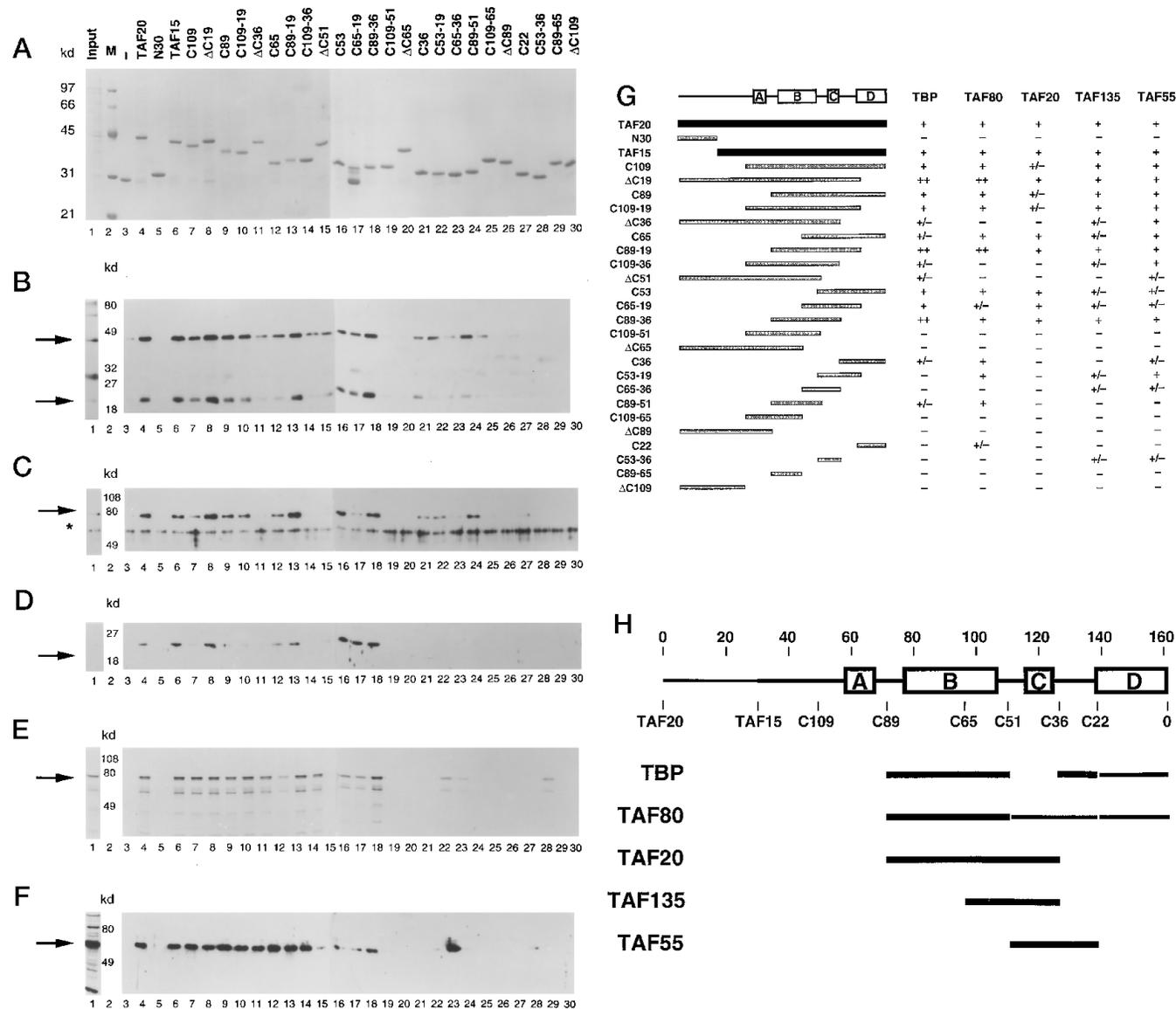
Here we report the cloning of hTAF20, the identification of

the related hTAF15 and an examination of their role(s) in the structural organization of the TFIID complex. Through sequential TFIID complex disruption experiments, we have presented the first experimental data on protein-protein interactions within native TFIID preparations. Our results indicate that TAF20/15 is associated with the TATA binding subunit, TBP, and the potential co-activator components TAF135 and TAF55. *In vitro* interaction assays with recombinant polypeptides confirmed these results and also demonstrated specific interactions of TAF20/15 with human TAF80, as well as a self-association. By means of a detailed mutagenesis analysis these interactions were mapped to discrete, though overlapping portions of a predicted secondary structure for TAF20/15 that may be used in a first attempt to model the architecture of the TFIID complex.

Elsewhere we recently presented evidence for a histone octamer-like structure within TFIID that consists of TAF80, TAF31, and TAF20/15 and was based on weak, but structurally relevant sequence homologies of the TAFs to histones H4, H3, and H2B, respectively (39). Crystallographic studies have indeed demonstrated a histone-like heterotetrameric structure for the *Drosophila* TFIID homologs, dTAF62 and dTAF42, of histones H4 and H3 (43). The results of our present studies on hTAF20/15 attain particular significance when discussed in the context of this structural motif within the TFIID complex.

First, mutagenesis studies of TAF20/15 indicate that all protein interaction sites identified map to the C-terminal 109 residues which constitute the histone H2B homologous domain. These observations reflect the demonstration that this portion of TAF20/15 is sufficient for incorporation into a functional TFIID complex (39). Furthermore, considering the lack of sequence conservation of the C-terminal 20 residues in the putative maize homolog (Fig. 1B) and in archaeobacterial H2B-like proteins (39), it appears that a minimal H2B-like fold is defined by helices A, B, and C. Our structure-function studies demonstrate that derivative proteins containing these helices are at least minimally capable of mediating all interactions tested in

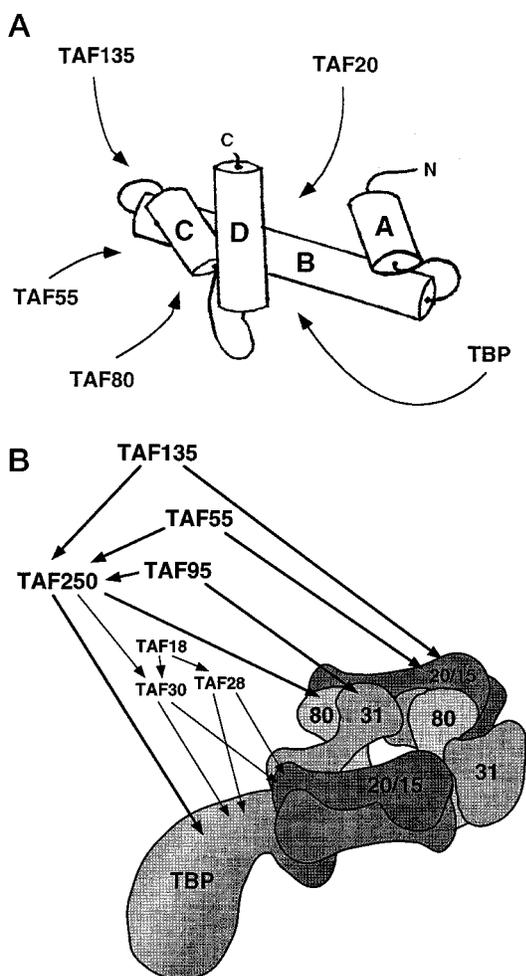
<sup>2</sup> S. Hasegawa, M. Horikoshi, S. Stevens, and R. G. Roeder, unpublished results.



**FIG. 4. Mapping protein interactions within TFIID onto TAF20/15.** *A*, Coomassie-stained gel of purified GST-TAF20/15 derivatives. TAF20/15 and fragments thereof were expressed in bacteria as fusion proteins with GST and purified from the soluble fraction with glutathione-Sepharose following dilution with control lysate to yield approximately equal amounts of purified fusion proteins from an equal input volume. One fifth of eluates (lanes 3 to 30) and one twentieth of input (lane 1) were analyzed by SDS-PAGE and stained with Coomassie Blue for visualization. TAF20 (lane 4), TAF15 (lane 6), and fragments thereof are labeled on top indicating the residues (numbered from the C terminus as indicated in Fig. 1C) included or lacking ( $\Delta$ ) with respect to TAF20, as diagrammed in *G*. Samples were loaded in the same sequence on the gel in *B-F*. *B*, interaction with TBP and the TBP core domain. Bacterially expressed hTBP and hTBPcore were incubated with normalized lysates (*A*) at 30 °C for 30 min, and affinity-purified samples were subjected to Western analysis with anti-TBP antiserum raised against hTBPcore protein. The upper arrow indicates hTBP, the lower arrow hTBP core. *C*, interaction with TAF80. Baculovirus-produced FLAG-tagged TAF80 was employed in the interaction assay as in *B*. The Western blot was developed with M2 antibody to reveal TAF80 (arrow) and a cross-reacting bacterial protein (\*). *D*, interaction with TAF20. Bacterially expressed FLAG-tagged TAF20 was used in the interaction assay as in *C*. The arrow indicates TAF20. *E*, interaction with TAF135. N-terminally truncated TAF135 containing the entire domain conserved in dTAF110 was produced in reticulocyte-lysate in the presence of [<sup>35</sup>S]methionine and employed in the interaction assay. The SDS-PAGE sample was dried and exposed to film. The arrow indicates the full-length translation product. *F*, interaction with TAF55. Bacterially expressed FLAG-tagged TAF55 was used in the interaction assay as in *C*. The arrow indicates TAF55. *G*, summary of the interaction data. TAF20, TAF15, and fragments thereof are represented by bars drawn to scale to summarize interaction data in a semiquantitative manner in four categories (++, +, +/-, -) for each interacting protein (columns labeled accordingly). The predicted secondary structure for TAF20/15 is shown above. *H*, portions of TAF20/15 involved in intra-TFIID interactions. Interaction data is summarized along the TAF20/15 polypeptide. The position and thickness of the line corresponding to a particular TAF-TAF interaction indicates the relative affinity of a particular portion of TAF20/15.

our *in vitro* assay (Fig. 4*G*). Portions of TAF20/15 shown to be involved in binding a particular TFIID component make up an interaction surface that, given the demonstrated histone homology, can be described by the H2B-like globular domain of TAF20/15 (Fig. 5*A*). Intra-TFIID interactions project on seemingly distinct surfaces in such a model for TAF20/15. Importantly, a self-association requiring the two central helices of the polypeptide is consistent with indications from crystallographic

studies that dimerization of histones H2A and H2B involves corresponding domains (44). Similarly, interaction of TAF20 with TAF80 involves the H4-homologous domain of TAF80 (15) and portions of the three C-terminal helices of TAF20/15 as observed for the H4-H2B interactions in the histone octamer structure (44). By characterizing intra-TFIID protein interactions in detail the current results support the histone-octamer model for TAFs 80, 31, and 20/15.



**FIG. 5. Modeling TAF20/15 within TFIID.** *A*, predicted interaction surfaces on an H2B-like TAF20/15. According to a structural homology between TAF20/15 and histone H2B proposed elsewhere (39), interactions impinging on TAF20/15 can be mapped onto the predicted fold based on crystallographic structures available for H2B (44, 47). Arrows indicate the predicted site of interaction with respective TFIID subunit. *B*, multiple protein interactions impinging on a hypothetical TBP-TAF core complex. Protein interactions between TFIID components characterized in this report and elsewhere (29, 31, 36)<sup>5</sup> are represented schematically as they may impinge on a TFIID core complex consisting of TBP and histone-like TAFs (39).

Second, it is noteworthy that multiple copies of TAF20/15 found within native TFIID, presumably two dimers (39), may help to accommodate the large number of observed interactions that impinge on a relatively small globular domain (Fig. 5*A*). These include associations of TAF20/15 with TBP, TAF80, TAF55, and TAF135 (this study), as well as interactions with the smaller TAF30, TAF28, and TAF18 components (29).

Finally, a strong interaction between hTAF20/15 and TBP, which we have observed both in partially disrupted native TFIID preparations (Fig. 3*D*) and between recombinant polypeptides, may represent a primary link between the proposed histone octamer-like TAF complex and the TATA box binding subunit within TFIID. This view is supported by the apparent stability of a complex consisting of recombinant hTBP, hTAF20, hTAF31, and hTAF80.<sup>3</sup> Previously reported interactions between TBP, TAF250 (30, 31), and TAF80 (15) may increase the stability of this core complex, consistent with the ability to form complexes containing TBP, dTAF230/

hTAF250, hTAF80/dTAF60, and hTAF31/dTAF40 (45).<sup>4</sup> High affinity interactions between TBP and TAF20/15 and the involvement of the latter in a presumptive histone octamer-like structure suggest an important qualification of previous models ascribing to dTAF230/hTAF250 the central role in directing assembly of the TFIID complex (30, 31). Thus, we may consider TBP bound to the histone octamer-like complex (consisting of TAF20/15, TAF31 and TAF80) as a minimal TFIID core complex with which other TAFs and possibly other co-activator proteins associate (Fig. 5*B*). However, based on previous interaction studies, the association of TAF55 (40), TAF95,<sup>5</sup> and TAF135 (29, 31, 46) with the TFIID complex may involve interaction surfaces on TAF250 as well as those in TAF20/15 proteins.

The data presented in this report represent a first attempt to characterize the three-dimensional arrangement of the TFIID multiprotein complex. Utilizing detailed mutagenesis analyses of intracomplex interactions a recently proposed structural model (39) has been confirmed and extended to include other components of TFIID. Fig. 5*B* represents a diagrammatic attempt to give a comprehensive account of all currently available structural data of the human TFIID complex (16). Future biophysical studies of TAF interactions and subcomplexes promise to further our understanding of TFIID function in transcriptional regulation by way of insights into the structure and conformation of this multiprotein complex.

**Acknowledgments**—We are grateful to the following for antisera, recombinant proteins, and plasmids: M. Horikoshi and R. Takada for  $\alpha$ -TAF250 and  $\alpha$ -TAF80, M. Guermah for  $\alpha$ -TAF31, C.-M. Chiang for TAF55 and  $\alpha$ -TAF95, T. Oelgeschläger for  $\alpha$ -TAF95, K. Hisatake and T. Ohta for TAF31 and TAF80 proteins, S. Stevens for  $\alpha$ -TAF135 as well as the partial TAF135 cDNA, and C. Baysdorfer for the *Zea mays* cDNA clone zEST00038-5. Peptide sequencing was performed by the Protein Sequencing Facility of the Rockefeller University.

#### REFERENCES

- Weil, P. A., Luse, D. S., Segall, J., and Roeder, R. G. (1979) *Cell* **18**, 469–484
- Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489
- Matsui, T., Segall, J., Weil, P. A., and Roeder, R. G. (1980) *J. Biol. Chem.* **255**, 11992–11996
- Dignam, J. D., Martin, P. L., Shastri, B. S., and Roeder, R. G. (1983) *Methods Enzymol.* **101**, 582–598
- Zawel, Z., and Reinberg, D. (1993) *Prog. Nucleic Acids Res. Mol. Biol.* **44**, 67–108
- Davison, B. L., Egly, J.-M., Mulvihill, E. R., and Chambon, P. (1983) *Nature* **301**, 680–686
- Sawadogo, M., and Roeder, R. G. (1985) *Cell* **43**, 165–175
- Nakajima, N., Horikoshi, M., and Roeder, R. G. (1988) *Mol. Cell. Biol.* **8**, 4028–4040
- Hai, T., Horikoshi, M., Roeder, R. G., and Green, M. (1988) *Cell* **54**, 1043–1051
- Horikoshi, M., Hai, T., Lin, Y.-S., Green, M. R., and Roeder, R. G. (1988) *Cell* **54**, 1033–1042
- Horikoshi, M., Carey, M. F., Kakidani, H., and Roeder, R. G. (1988) *Cell* **54**, 665–669
- Lieberman, P. M., and Berk, A. J. (1994) *Genes Dev.* **8**, 995–1006
- Hernandez, N. (1993) *Genes Dev.* **7**, 1291–1308
- Goodrich, J. A., and Tjian, R. (1994) *Curr. Opin. Cell Biol.* **6**, 403–409
- Hisatake, K., Ohta, T., Takada, R., Guermah, M., Horikoshi, M., Nakatani, Y., and Roeder, R. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 8195–8199
- Burley, S. K., and Roeder, R. G. (1996) *Annu. Rev. Biochem.* **65**, 769–799
- Hoffmann, A., Sinn, E., Yamamoto, T., Wang, J., Roy, A., Horikoshi, M., and Roeder, R. G. (1990) *Nature* **346**, 387–390
- Pugh, B. F., and Tjian, R. (1990) *Cell* **61**, 1187–1197
- Dynlacht, B. D., Hoey, T., and Tjian, R. (1991) *Cell* **66**, 563–576
- Tanese, N., Pugh, B. F., and Tjian, R. (1991) *Genes Dev.* **5**, 2212–2224
- Martinez, E., Chiang, C.-M., Ge, H., and Roeder, R. G. (1994) *EMBO J.* **13**, 3115–3126
- Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K., and Tjian, R. (1994) *Cell* **79**, 93–105
- Abmayr, S. M., Workman, J. L., and Roeder, R. G. (1988) *Genes Dev.* **2**, 542–553
- Workman, J. L., Abmayr, S. M., Cromlish, W. A., and Roeder, R. G. (1988) *Cell* **55**, 211–219
- Goodrich, J. A., Hoey, T., Thut, C. J., Admon, A., and Tjian, R. (1993) *Cell* **75**, 519–530

<sup>3</sup> T. Ohta and R. G. Roeder, unpublished observations.

<sup>4</sup> M. Guermah and R. G. Roeder, unpublished results.

<sup>5</sup> M. Guermah, Y. Tao, and R. G. Roeder, unpublished results.

26. Klemm, R. D., Goodrich, J. A., Zhou, S., and Tjian, R. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 5788–5792
27. Kokubo, T., Gong, D.-W., Yamashita, S., Takada, R., Roeder, R. G., Horikoshi, M., and Nakatani, Y. (1993) *Mol. Cell. Biol.* **13**, 7859–7863
28. Yokomori, K., Chen, J.-L., Admon, A., Zhou, S., and Tjian, R. (1993) *Genes Dev.* **7**, 2587–2597
29. Mengus, G., May, M., Jacq, X., Staub, A., Tora, L., Chambon, P., and Davidson, I. (1995) *EMBO J.* **14**, 1520–1531
30. Takada, R., Nakatani, Y., Hoffmann, A., Kokubo, T., Hasegawa, S., Roeder, R. G., and Horikoshi, M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 11809–11813
31. Weinzierl, R. O. J., Dynlacht, B. D., and Tjian, R. (1993) *Nature* **362**, 511–517
32. Chiang, C.-M., Ge, H., Wang, Z., Hoffmann, A., and Roeder, R. G. (1993) *EMBO J.* **12**, 2749–2762
33. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185**, 60–89
34. Hoffmann, A., and Roeder, R. G. (1991) *Nucleic Acids Res.* **19**, 6337
35. Kozak, M. (1986) *Cell* **44**, 283–292
36. Kokubo, T., Gong, D.-W., Wootton, J. C., Horikoshi, M., Roeder, R. G., and Nakatani, Y. (1994) *Nature* **367**, 484–487
37. Hoffmann, A. (1994) *A Molecular Characterization of Human Transcription Factor IID*, The Rockefeller University, New York
38. Zhou, Q., Boyer, T. G., and Berk, A. J. (1993) *Genes Dev.* **7**, 180–187
39. Hoffmann, A., Chiang, C.-M., Oelgeschläger, T., Xie, X., Burley, S. K., Nakatani, Y., and Roeder, R. G. (1996) *Nature* **380**, 356–359
40. Chiang, C.-M., and Roeder, R. G. (1995) *Science* **267**, 531–536
41. Jacq, X., Brou, C., Lutz, Y., Davidson, I., Chambon, P., and Tora, L. (1994) *Cell* **79**, 107–117
42. Sauer, F., Hansen, S. K., and Tjian, R. (1995) *Science* **270**, 1783–1788
43. Xie, X., Kokubo, T., Cohen, S. L., Mirza, U. A., Hoffmann, A., Chait, B. T., Roeder, R. G., Nakatani, Y., and Burley, S. K. (1996) *Nature* **380**, 316–322
44. Arents, G., Burlingame, R. W., Wang, B.-C., Love, W. E., and Moudrianakis, E. N. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10148–10152
45. Thut, C. J., Chen, J.-L., Klemm, R., and Tjian, R. (1995) *Science* **267**, 100–104
46. Kokubo, T., Gong, D.-W., Yamashita, S., Horikoshi, M., Roeder, R. G., and Nakatani, Y. (1993) *Genes Dev.* **7**, 1033–1046
47. Pruss, D., Hayes, J. J., and Wolffe, A. P. (1995) *BioEssays* **17**, 1–10

# Cloning and Characterization of Human TAF20/15: MULTIPLE INTERACTIONS SUGGEST A CENTRAL ROLE IN TFIID COMPLEX FORMATION

Alexander Hoffmann and Robert G. Roeder

*J. Biol. Chem.* 1996, 271:18194-18202.

doi: 10.1074/jbc.271.30.18194

---

Access the most updated version of this article at <http://www.jbc.org/content/271/30/18194>

#### 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

This article cites 46 references, 17 of which can be accessed free at <http://www.jbc.org/content/271/30/18194.full.html#ref-list-1>