Three-Dimensional Structures of the TAF$_{II}$-Containing Complexes TFIID and TFTC

Marjorie Brand,* Claire Leurent,* Véronique Mallouh, Làszlò Tora, Patrick Schultz†

TBP (TATA-binding protein)–associated factors (TAF$_{II}$s) are components of large multiprotein complexes such as TFIID, TFTC, STAGA, PCAF/GCN5, and SAGA, which play a key role in the regulation of gene expression by RNA polymerase II. The structures of TFIID and TFTC have been determined at 3.5-nanometer resolution by electron microscopy and digital image analysis of single particles. Human TFIID resembles a macromolecular clamp that contains four globular domains organized around a solvent-accessible groove of a size suitable to bind DNA. TFTC is larger and contains five domains, four of which are similar to TFIID.

In eukaryotes, transcription initiation of protein-encoding genes by RNA polymerase II was thought to require transcription factor TFIID, which is composed of TBP and several TAF$_{II}$s (1–3). TFIID recognizes specific sequence elements of the promoter, such as the TATA box and the initiator, and nucleates the assembly of a preinitiation complex. This situation was complicated by the discovery in human cells of distinct TFIID complexes (4–6), and at least two different subpopulations based on the absence (TFIIDα) or presence (TFIIDβ) of the subunit TAF$_{II}$30 were identified (5, 7). In addition, a human (h) TAF$_{II}$-containing complex, hTFTC (for TBP-free TAF$_{II}$-containing complex), has been described, which does not contain TBP but is able to direct PIC formation (8). In parallel, additional human and yeast (y) TAF$_{II}$-containing complexes with slightly different subunit compositions were isolated: hPCAF/GCN5 (9), hSTAGA (SPT3-TAF$_{II}$31-GCN5 acetyl transferase) (10), and ySAGA (SPT-ADA-GCN5 acetyl transferase) (11). These complexes carry a histone acetyl transferase activity either on TAF$_{II}$250 (TFIID) or on GCN5 (hTFTC, hPCAF/GCN5, hSTAGA, ySAGA), and all GCN5-containing complexes are able to acetylate histones in both a free and a nucleosomal context (1, 9–11). Thus, the TAF$_{II}$-containing complexes are modular entities composed of a common core of seven or eight TAF$_{II}$s present in all complexes and a set of subunits specific for each complex.

To study the structural features underlying this modularity, we investigated the molecular architecture of TFIIDβ and TFTC by electron microscopy. The complexes were purified from human HeLa cells by a two-step immunopurification procedure (12). Antibodies directed against hTAF$_{II}$30, a subunit present in both complexes, were used to immunoprecipitate all TAF$_{II}$30-containing complexes from HeLa cell nuclear extracts, and the complexes were eluted under nondenaturating conditions with an excess of epitope peptide. The TAF$_{II}$30-containing complexes were separated by a second immunoprecipitation using an antibody raised against hTBP (8). TFTC complexes, in which TBP is absent, remained in the supernatant, whereas TFIIDβ complexes were eluted with an excess of the corresponding epitope peptide (13).

Purified TFIIDβ (hereafter called TFIID) and TFTC complexes were adsorbed onto carbon-coated electron microscopy grids, negatively stained with uranyl acetate and imaged in a transmission electron microscope (14). Direct observation of either TFIID or TFTC preparations showed fields of particles of similar size, indicating that these complexes exist as discrete and homogeneous entities, as previously suggested by biochemical studies (5, 8). Numerical image analysis was performed, using the IMAGIC statistical package (15) to improve the resolution and the statistical significance of these observations (16). Individual molecular images of the complexes were aligned, clustered into classes to discriminate between different orientations or conformations of the particles, and average images for each class were calculated (17).

The TFIID complexes appeared as round particles of about 15 to 20 nm in diameter (Fig. 1A). Image analysis of two independent TFIID preparations, including a total of 3160 particles, reproducibly revealed that 48% of the complexes were preferentially oriented on the support whereas about 20% of the particles adsorbed in an upside down orientation, thus yielding a mirror symmetry view. The average image (Fig. 1B) revealed a slightly elongated structure about 19 nm by 15 nm in size formed by two compact domains separated by a pronounced solvent-accessible groove 2 to 4 nm wide. Each domain was further divided into two subdomains 6 to 8 nm in size (labeled 1 to 4 in Fig. 1B). The size of the highly purified TFIID complexes and the absence of internal symmetry indicated that the observed structure corresponds to a monomer and not to a dimer of TFIID, as previously suggested (18). About 20% of the complexes showed a more extended structure of about 25 nm long and 12 nm wide (Fig. 1C). This molecular view differs from the previous view only in the angle between the two major domains (Fig. 1, B and C) which caused the width at the entry of the groove to vary between 4 and 10 nm. The size of this solvent-accessible groove suggests that it may constitute a DNA binding interface. The observed "open" and "closed" conformations raise the possibility that TFIID may act like a macromolecular clamp upon binding to DNA.

Six TFIID subunits contain characteristic histone folds (19–21), some of which have
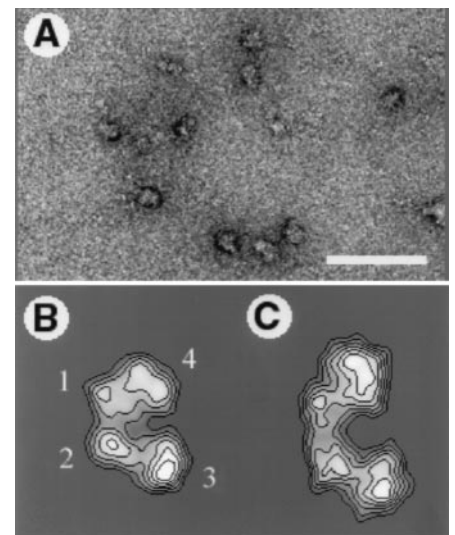

Fig. 1. Molecular views of TFIID complexes. (**A**) Electron micrograph of negatively stained TFIID complexes shows discrete globular particles. (**B** and **C**) Characteristic views of TFIID complexes obtained reproducibly upon analysis of two independent TFIID preparations (1774 and 1386 molecular images, respectively). Two views of the same complex were obtained, which differ by the angle between two protein domains indicating a flexibility in the linker region. The stain-excluding domains, represented in white as plots of equal density, are numbered from 1 to 4. The resolution of the class average was estimated to about 3.5 nm by Fourier ring correlation (34). Bar, 100 nm in (A) and 13 nm in (B) and (C).

Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/Université Louis Pasteur, Boite Postale 163, F-67404 Illkirch cedex, Communauté Urbaine de Strasbourg, France.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: pat@wotan.u-strasbg.fr

been crystallized and solved to atomic resolution (20, 22). The presence of these motifs and the finding that TFIID can introduce negative supercoils upon interaction with DNA (23) led to the hypothesis that TFIID may contain a histone octamer–like structure (24). The four domains in the TFIID complex are almost equivalent in size and none are large enough to hold all six histone motif–containing TAF$_{II}$s. Therefore, a compact histone octamer–like structure incorporating all the TAF$_{II}$s that have a histone motif is unlikely. Taking into account the fact that histone folds mainly tend to dimerize, our data favor the hypothesis that they may form interfaces for protein-protein interactions. Histone fold pairs could thus be located in each subdomain, or alternatively between subdomains, to organize the observed structure of TFIID.

In contrast to TFIID, more than 80% of the TFTC complexes had a characteristic elongated shape of 16 nm by 27 nm in size (Fig. 2A). The analysis of 458 molecular images of TFTC complexes showed a preferred adsorption of the particles and revealed mirror symmetry related views arising from upside down orientations of the particles. The overall architecture of the TFTC complexes consists of five subdomains (labeled 1 to 5 in Fig. 2B) about 6 to 8 nm in size separated by solvent-accessible grooves. The arrangement of subdomains 1 to 4 and the presence of a
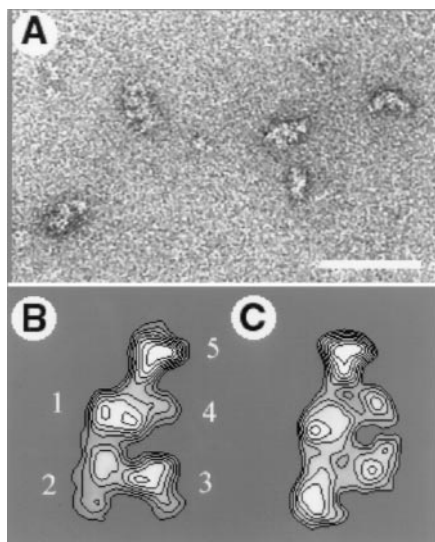
pronounced groove about 2 to 3 nm wide between domains 1,4 and 2,3 indicated that the structure of TFTC is locally similar to that of TFIID. The elongated shape of TFTC is due to domain 5, which is absent in TFIID and is attached to domain 1 or 4.

The image analysis further revealed two subpopulations of images (Fig. 2, B and C). The overall shape of the two molecules was conserved, and minor differences in domains 1, 3, 4, and 5 could be due to a slight angular variation. The density variation in domain 2 could not be explained, either by a small difference in orientation or by a conformational change, and likely represents a heterogeneity in the image population probably due to a variability in the polypeptide composition. Thus, it is conceivable that some of the polypeptides that are only found in TFTC (1), but not in the other TAF$_{II}$-GCN5 complexes (for example, TAF$_{II}$150, TAF$_{II}$135, TAF$_{II}$100, TAF$_{II}$80), may be located in domain 2 of TFTC.

To investigate the three-dimensional (3D) organization of TFIID and TFTC, a conical tilt series was generated by inclining the grid in the microscope (25, 26). The calculated 3D models of TFIID and TFTC had thresholds set to volumes of 780 and 1212 nm^3, respectively, which corresponds to about 70% of the expected mass of the complexes in each case, assuming a protein density of 1.3 g/cm^3. A lower value of the reconstructed volume is generally found for molecules preserved in negative stain and is probably due to stain penetrating the surface of the molecule or to a slight shrinkage of the structure
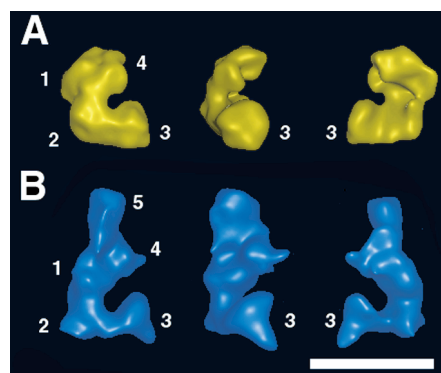
upon drying or electron exposure (27). The 3D organization of TFIID (Fig. 3A) and TFTC (Fig. 3B) show striking similarities. Both complexes share a solvent-accessible groove that allows an unambiguous alignment of the two models. The size of this groove is suitable to accommodate a double-stranded DNA molecule that would be buried inside the complex and be completely surrounded by protein domains forming a molecular clamp. The groove is bent across the surface of the molecule and continues on both sides by marked depressions corresponding to the domain interfaces. Additional experiments will be required to define the path and orientation of the promoter DNA in the TFIID complex in order to explain the results of DNA footprinting experiments, which show that TFIID protects about 73 base pairs from deoxyribonuclease I action (28), and account for the topological strain applied on the DNA upon binding to TFIID (23).

Comparison of the two 3D models further indicates that the structure of TFIID is almost included in that of TFTC. Subdomains 1 to 3, which form the walls of the groove, have comparable sizes and positions in both complexes and thus are probably shared by the two complexes. Seven TAF$_{II}$s, which add up to a mass of 58% of TFIID, are also in TFTC (1, 2). These common subunits are probably organized into a similar core structure in the two complexes. Subdomain 4 is smaller in TFIID than in TFTC where it constitutes the main interaction site for the additional domain 5. Domain 5 has a relative volume corresponding to a mass of 350 to 400 kD, which may correspond to the size of the TFTC-specific subunit TRRAP (1). The described 3D architectures of TFIID and TFTC give a structural hint to how distinct macromolecular assemblies sharing common subunits but differing by their specific subunits can be built up (29).

These 3D models will provide a useful framework to integrate structural information from various sources such as high-resolution x-ray (20, 22, 30) or nuclear magnetic resonance data (31), electron crystallographic data (27, 32), or protein-DNA cross-linking experiments (33) and will help to identify the interaction sites with components of the transcriptional apparatus.

Fig. 2. Molecular views of TFTC complexes. (**A**) Electron micrograph of negatively stained TFTC complexes shows elongated wedge-shaped particles. (**B** and **C**) Characteristic views of TFTC complexes obtained upon analysis of 458 molecular images. The stain-excluding domains, represented in white as plots of equal density, are numbered from 1 to 5. Two subpopulations of particles were identified that differed in the protein density in domain 2. The resolution of the class average was estimated to about 3.5 nm by Fourier ring correlation (34). Bar, 100 nm in (A) and 13 nm in (B) and (C).

Fig. 3. Three-dimensional models of TFTC and TFIID complexes. The models were aligned with respect to the position of the solvent-accessible groove. The three different orientations of the models correspond to rotations around a vertical axis by 80° increments. The domains are numbered from 1 to 4 (TFIID) or from 1 to 5 (TFTC), and domain 3 was labeled in all the views for sake of clarity. (**A**) A total of 115 different views of the most compact TFIID complexes (Fig. 1B) were merged to calculate the 3D model. (**B**) A total of 112 different tilted views from the most abundant form of TFTC complexes (Fig. 2B) were merged to calculate the 3D model. Bar, 20 nm.

References and Notes
1. M. Brand, K. Yamamoto, A. Staub, L. Tora, *J. Biol. Chem.* **274**, 18285 (1999).
2. B. Bell and L. Tora, *Exp. Cell Res.* **246**, 11 (1999).
3. W. P. Tansey and W. Herr, *Cell* **88**, 729 (1997).
4. C. Brou *et al.*, *EMBO J.* **12**, 489 (1993).
5. X. Jacq *et al.*, *Cell* **79**, 107 (1994).
6. A. Bertolotti, Y. Lutz, D. J. Heard, P. Chambon, L. Tora, *EMBO J.* **15**, 5022 (1996).
7. G. Mengus *et al.*, *EMBO J.* **14**, 1520 (1995).
8. E. Wieczorek, M. Brand, X. Jacq, L. Tora, *Nature* **393**, 187 (1998).
9. V. V. Ogryzko *et al.*, *Cell* **94**, 35 (1998).

10. E. Martinez, T. K. Kundu, J. Fu, R. G. Roeder, *J. Biol. Chem.* **273**, 23781 (1998).
11. P. A. Grant *et al.*, *Cell* **94**, 45 (1998).
12. For immunoprecipitation, mouse monoclonal antibodies (mAbs), anti-hTAF$_{II}$30 1H8 (*8*), and anti-hTBP 2C1 (*5*), were purified as described in (*4*). HeLa cell nuclear extract was first passed through a single-stranded (ss) DNA cellulose column to eliminate proteins that bind nonspecifically to the resin used later in immunoprecipitations. Fractions were immunoprecipitated with mAbs bound to protein G–Sepharose (Pharmacia) as described in (*8*). These complexes bound to protein G–Sepharose were washed three times with IP buffer [25 mM tris-HCl (pH 7.9), 10% (v/v) glycerol, 0.1% NP-40, 0.5 mM dithiothreitol, 5 mM MgCl$_2$] containing 500 mM KCl and two times with IP buffer containing 100 mM KCl. After washing, bound proteins were eluted with a 1000-fold excess of the corresponding epitope peptide and analyzed by SDS–polyacrylamide gel electrophoresis (SDS-PAGE). The polypeptide composition of the isolated complexes were verified either on silver-stained SDS-PAGE gels or by protein immunoblots as described (*1*, *8*).
13. The following subunits were identified in hTFIID: TBP, TAF$_{II}$250, TAF$_{II}$150, TAF$_{II}$135, TAF$_{II}$100, TAF$_{II}$80, TAF$_{II}$55, TAF$_{II}$31, TAF$_{II}$30, TAF$_{II}$28, TAF$_{II}$20, and TAF$_{II}$18. The hTFTC complex contained TAF$_{II}$150, TAF$_{II}$135, TAF$_{II}$100, TAF$_{II}$80, TAF$_{II}$55, TAF$_{II}$31, TAF$_{II}$30, TAF$_{II}$20, TAF$_{II}$18, GCN5-L, PAF65β (*9*), hADA3, hSPT3, TRRAP, and other noncharacterized subunits (*1*).
14. We cross-linked 5 μl of the purified TFTC or TFIIDβ preparations for 10 s with 0.05% glutaraldehyde and placed them on a 10-nm-thick carbon film previously treated by a glow discharge in air. After 2 min of adsorption, the grid was negatively stained with a 2% uranyl acetate solution. The images were formed on a Philips CM120 Transmission Electron Microscope operating at 100 kV and were analyzed as described (*16*).
15. M. van Heel, G. Harauz, E. V. Orlova, *J. Struct. Biol.* **116**, 17 (1996).
16. C. Klinger *et al.*, *EMBO J.* **15**, 4643 (1996).
17. M. van Heel and J. Frank, *Ultramicroscopy* **6**, 187 (1981).
18. A. K. P. Taggart and B. F. Pugh, *Science* **272**, 1331 (1996).
19. A. Hoffmann *et al.*, *Nature* **380**, 356 (1996).
20. X. Xie *et al.*, *Nature* **380**, 316 (1996).
21. Y.-G. Gangloff *et al.*, *Mol. Cell. Biol.*, in press.
22. C. Birck *et al.*, *Cell* **94**, 239 (1998).
23. T. Oelgeschlager, C. M. Chiang, R. G. Roeder, *Nature* **382**, 735 (1996).
24. A. Hoffmann, T. Oelgeschlager, R. G. Roeder, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8928 (1997).
25. M. Radermacher, T. Wagenknecht, A. Verschoor, J. Frank, *J. Microsc.* **146**, 113 (1987).
26. The 3D models were calculated from conical tilt series produced by recording the same field of molecules at 60° and 0° tilt. Molecular images were extracted from the untilted micrographs, aligned, and classified into homogeneous groups to identify subpopulations of similarly oriented particles. Tilt and in-plane rotation parameters yield the Euler angles for each tilted view, about 100 of which were merged by filtered back-projection to calculate the 3D molecular envelope.
27. S. A. Darst, A. M. Edwards, E. W. Kubalek, R. D. Kornberg, *Cell* **66**, 121 (1991).
28. B. A. Purnell, P. A. Emanuel, D. S. Gilmour, *Genes Dev.* **8**, 830 (1994).
29. R. E. Kingston, *Nature* **399**, 199 (1999); M. Hampsey and D. Reinberg, *Curr. Opin. Genet. Dev.* **9**, 132 (1999).
30. Y. Nakatani, S. Bagby, M. Ikura, *J. Biol. Chem.* **271**, 6575 (1996).
31. D. Liu *et al.*, *Cell* **94**, 573 (1998).
32. P. Schultz, H. Celia, M. Riva, A. Sentenac, P. Oudet, *EMBO J.* **12**, 2601 (1993).
33. T. K. Kim *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12268 (1997).
34. M. van Heel, *Ultramicroscopy* **21**, 95 (1987).
35. We thank Y. Lutz for antibody production, J. C. Dantonel and B. Bell for discussions and critical reading of the manuscript, P. Eberling for peptide synthesis, the

Three-Dimensional Structure of the Human TFIID-IIA-IIB Complex

Frank Andel III,[1] Andreas G. Ladurner,[3] Carla Inouye,[3] Robert Tjian,[2] Eva Nogales[1,3]*

The multisubunit transcription factor IID (TFIID) is an essential component of the eukaryotic RNA polymerase II machinery that works in concert with TFIIA (IIA) and TFIIB (IIB) to assemble initiation complexes at core eukaryotic promoters. Here the structures of human TFIID and the TFIID-IIA-IIB complex that were obtained by electron microscopy and image analysis to 35 angstrom resolution are presented. TFIID is a trilobed, horseshoe-shaped structure, with TFIIA and TFIIB bound on opposite lobes and flanking a central cavity. Antibody studies locate the TATA-binding protein (TBP) between TFIIA and TFIIB at the top of the cavity that most likely encompasses the TATA DNA binding region of the supramolecular complex.

The accurate and regulated transcription of protein coding genes in all eukaryotic organisms requires the assembly at specific promoter elements of a complex molecular machine that includes general transcription factors in association with RNA polymerase II (RNA pol II) (*1*). Recognition of core promoter DNA sequences by TFIID (*2*) is followed by the assembly of a fully activated preinitiation complex that contains TFIIA, TFIIB, TFIIE, TFIIF, TFIIH, and RNA pol II (*3*). In the absence of activators bound to enhancer elements, this core transcription complex can accurately initiate basal levels of RNA synthesis. In the presence of gene-selective enhancer and promoter binding activators, significantly elevated levels of transcription initiation can be achieved. A key step in the multistep process of gene activation is the recruitment and assembly of the TFIID-IIA-IIB complex at the TATA DNA region that is found in many core promoters of eukaryotic genes. The TBP subunit and some of the TBP-associated factors (TAF$_{II}$) subunits (TAF$_{II}$250, TAF$_{II}$150, and TAF$_{II}$70) make specific contacts with the TATA box and other core promoter elements, including initiator (INR) and downstream promoter elements (DPE) (*4*). Other TAF$_{II}$'s of the multisubunit TFIID complex such as TAF$_{II}$130, -100, -55, -32, -30, and -28 likely serve as targets of activation domains involved in the recruitment and stabilization of TFIID at core promoters by upstream enhancer binding factors (*5–7*).

The binding of TFIID to the core promoter is coordinated with the assembly of an active preinitiation complex that includes IIA and IIB. X-ray diffraction and nuclear magnetic resonance (NMR) studies have revealed the structures of various subdomains and truncated fragments of IIA and IIB and of subunits contained within TFIID. For example, the high-resolution structures of TBP bound to TATA DNA and to domains of either IIA or IIB have been determined (*8*). However, both the size of the complex and the inherent difficulties in obtaining large quantities of purified holo-TFIID, -IIA, and -IIB have precluded conventional x-ray diffraction studies of the full complex. Consequently, the overall shape and relative position of the components within the TFIID-IIA-IIB complex remain unknown. As a first step toward determining the structure of the intact native human TFIID, we used electron microscopy (EM) and single-particle image analysis to obtain the structure of TFIID and its complex with full-length IIA and IIB at 35 Å resolution.

Homogeneous preparations of TFIID suitable for EM studies were purified as described (*9*). For preparations of TFIID bound to IIB and IIA, recombinant IIB and IIA subunits were purified and reconstituted with antibody affinity-purified TFIID in vitro (*10–*

[1]Life Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [2]Howard Hughes Medical Institute, [3]Molecular and Cell Biology Department, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: enogales@lbl.gov