Cell PRESS
Structure
Article

Cryo-EM Reveals Promoter DNA Binding and Conformational Flexibility of the General Transcription Factor TFIID

Hans Elmlund,[1,2,*] Vera Baraznenok,[3] Tomas Linder,[3] Zsolt Szilagyi,[2,3] Reza Rofougaran,[4] Anders Hofer,[4] Hans Hebert,[5] Martin Lindahl,[5,6] and Claes M. Gustafsson[2,3]
[1]Department of Structural Biology, Fairchild Building, Stanford University School of Medicine, Stanford, CA 94305, USA
[2]Department of Medical Biochemistry and Cell Biology, Göteborg University, SE-405 30 Göteborg, Sweden
[3]Division of Metabolic Diseases, Karolinska Institutet, Novum, SE-141 86 Huddinge, Sweden
[4]Department of Medical Biochemistry and Biophysics, Umeå University, SE-901 87 Umeå, Sweden
[5]Department of Biosciences and Nutrition, Karolinska Institutet and School of Technology and Health, Royal Institute of Technology, Novum, SE-141 87 Huddinge, Sweden
[6]Department of Molecular Biophysics, Lund University, SE-221 00 Lund, Sweden
*Correspondence: hael@stanford.edu
DOI 10.1016/j.str.2009.09.007

SUMMARY

The general transcription factor IID (TFIID) is required for initiation of RNA polymerase II-dependent transcription at many eukaryotic promoters. TFIID comprises the TATA-binding protein (TBP) and several conserved TBP-associated factors (TAFs). Recognition of the core promoter by TFIID assists assembly of the preinitiation complex. Using cryo-electron microscopy in combination with methods for ab initio single-particle reconstruction and heterogeneity analysis, we have produced density maps of two conformational states of *Schizosaccharomyces pombe* TFIID, containing and lacking TBP. We report that TBP-binding is coupled to a massive histone-fold domain rearrangement. Moreover, docking of the TBP-TAF1$_{N\text{-terminus}}$ atomic structure to the TFIID map and reconstruction of a TAF-promoter DNA complex helps to account for TAF-dependent regulation of promoter-TBP and promoter-TAF interactions.

INTRODUCTION

Transcription factor IID (TFIID) is composed of the TATA-binding protein (TBP) and several evolutionary conserved TBP-associated factors (TAFs) (Burley and Roeder, 1996), of which 13 are essential for cell viability in yeast (Yatherajam et al., 2003). Many TAFs also exists as components of other multiprotein complexes involved in the regulation of transcription, such as the SAGA complex (Grant et al., 1998) and the mammalian PCAF complex (Ogryzko et al., 1998). Most genes in yeast are dependent on TFIID (Lee et al., 2000) for the assembly of a preinitiation complex (PIC) (Hahn et al., 1989; Horikoshi et al., 1989; Peterson et al., 1990), and there is mounting evidence pointing to the importance of TAFs (Chalkley and Verrijzer, 1999; Oelgeschläger et al., 1996; Sawadogo and Roeder, 1985; Shao et al., 2005) and TBP (Kim et al., 1993; Nikolov et al., 1995) in promoter recognition. Direct TAF interactions to TBP have been reported for only TAF7 (Yatherajam et al., 2003) and the N terminus of TAF1 (Liu et al., 1998). TBP does not seem to play a crucial role in the integrity of the TFIID complex. Accordingly, TBP dynamically associates with the *Saccharomyces cerevisiae* holo-TAF complex (Sanders et al., 2002), and human TAF complexes lacking TBP have been identified (Wieczorek et al., 1998).

Atomic resolution structural information is only available for parts of the TFIID complex (Bhattacharya et al., 2007; Liu et al., 1998; Romier et al., 2007; Werten et al., 2002; Xie et al., 1996). The human histone-fold domain (HFD) subunits TAF4 and TAF12 form crystals in a (TAF4-TAF12)$_2$ arrangement (Werten et al., 2002). and the TBP-TAF1$_{11\text{-}77}$ NMR-structure (Liu et al., 1998) shows that the TAF1 N terminus is a protein mimic of the bent TATA box, which inhibits TBP-promoter interactions. Structural analysis of the complete TFIID assembly has so far been limited to low-resolution electron microscopy studies (20–30 Å) of yeast and human TFIID (Andel et al., 1999; Brand et al., 1999; Grob et al., 2006; Leurent et al., 2002, 2004; Papai et al., 2009) reporting a horseshoe-shaped structure with two approximately equally sized quaternary structure domains protruding from a central core, forming a pore through the structure. Immunolabeling of yeast TFIID (Leurent et al., 2002; Leurent et al., 2004) has localized all dimerical HFD TAFs in lobes separated at distances approaching hundreds of Ångströms, which argues against the existence of a coherent HFD arrangement within TFIID. However, because of the limited resolution of previous electron microscopic data, it has not been possible to recognize and dock the HFD crystal structures (Werten et al., 2002; Xie et al., 1996) or the TBP-TAF1$_{11\text{-}77}$ NMR-structure (Liu et al., 1998). On the basis of immunolabeling of TBP in a reconstruction of human TFIID (Andel et al., 1999), promoter DNA was suggested to bind in the central pore, but structural evidence for this interaction has not been provided.

Elucidation of the structure and dynamics of TFIID is necessary for understanding transcriptional regulation. We have taken steps in this direction and have determined single-particle cryo-EM reconstructions of two conformational states of

Schizosaccharomyces pombe TFIID, containing (TFIID) and lacking TBP (TFIIDΔTBP). We have also reconstructed TFIIDΔTBP in complex with promoter DNA. We report that TBP binding to and release from the TAFs is coupled to a massive rearrangement of the quaternary structure region comprising TAF4. Moreover, we describe the structural dynamics underlying formation of a TBP-binding site within TFIID, which indicates that promoter-TBP and promoter-TAF interactions are inhibited when TBP is bound to the TAFs.

RESULTS

Purification and GEMMA Analysis of Fission Yeast TFIID

To isolate TFIID from *S. pombe,* we used tandem affinity purification (TAP) with a TAP-tag on TAF4. After purification over IgG-Sepharose and MiniS columns, 12 distinct TAF subunits (TAF1–TAF12) were identified using SDS-PAGE (see Figures S1A and S1B available online) and MALDI-TOF mass fingerprinting of excised bands. Among the identified proteins were two closely related TAF5 paralogs, TAF5a and TAF5b, which are encoded for by two separate genes. An additional protein band of about 10 kDa could not be identified with mass fingerprinting, but the size agrees with what would be expected for TAF13. Immunoblotting showed that purified TFIID contained TBP (Figure S1C). We used Gas-phase-Electrophoretic-Mobility Macromolecule Analysis (GEMMA) to estimate the molecular weight of TFIID (Figure S1D) to 874 ± 52 kDa. GEMMA is a relatively new method for separating protein molecules in diluted protein samples according to their diameter by analyzing their different mobility in gas phase induced by a charged reduced electrospray process (Bacher et al., 2001; Kaufman et al., 1996). The method runs at atmospheric pressure and allows for determination of the mass of protein molecules with an error of ±5.6% (Bacher et al., 2001). Determination of the molecular weight of *S. pombe* TFIID to 874 ± 52 kDa agrees with the predicted molecular weight of 876 kDa. The prediction accounts for the existence of two copies each of the TAF4, TAF6, TAF9, and TAF12 subunits (Selleck et al., 2001).

3D Reconstruction of TFIID and TFIIDΔTBP

We calculated the TFIID reconstructions from 62,300 particle images, collected from 22 micrographs of a single TFIID specimen, preserved in vitrified ice, and imaged under low-dose conditions (10–15 $e^{-}/Å^{2}$) using a transmission electron microscope (EM) equipped with a field-emission gun. We used an ab initio reconstruction method based on reference-free common lines (Elmlund et al., 2008) to generate an initial TFIID 3D reconstruction. Refinement of the entire TFIID cryo-data set was performed without using sorting procedures to exclude images. The refinement halted initially at a resolution of ~25 Å, which was significantly lower than that expected for the given size of the data set. We interpreted this finding as a sign of the data set being heterogeneous and used multivariate statistical analysis to investigate the number of conformational states present in the population (see Figure S2). Two predominant conformational states were identified. We developed a method to divide the data set into homogeneous groups (see Supplemental Data). The two resulting reconstructions were identified as containing (TFIID) and lacking TBP (TFIIDΔTBP) (see below), which is in agreement with the known ability of TBP to dynamically associate with TAFs (Sanders et al., 2002). We compared projections in five evenly distributed orientations of each of the two states with corresponding projection averages. Projection averages showed an excellent agreement with projections (Figure 1B). The resolution of the two final density maps was assessed by Fourier shell correlation (FSC) to 8–10 Å according to the FSC = 0.5 criterion (Figure 1D), and the maps were low-pass filtered to 8 Å. Local symmetry analysis detected the two-fold symmetry and defined the direction of the symmetry axis for a module present in the back end of the core, the TBP-TAF region, and for two segments of the back, of which one corresponded to the $(TAF4–TAF12)_2$ assembly (see Supplemental Data and Figure S6).

Determination of the Absolute Hand

The absolute hand of a reconstruction cannot be resolved for independent projections alone. One of the human TFIID reconstructions (Grob et al., 2006) agreed well with our TFIIDΔTBP reconstruction at low resolution. The fact that the human TFIID reconstructions originate from a random conical tilt reconstruction therefore led us to attempt to determine the absolute hand by using the human TFIID reconstruction. We projected the human and *S. pombe* TFIID reconstructions in 116 even directions and aligned the *S. pombe* TFIID images in 3D by using the human TFIID projections as a reference set. We calculated a reconstruction from the aligned set of *S. pombe* TFIID projections and compared it to both hands of our reconstruction, which revealed a significant difference in correlation (Figure S5b; 0.6523 for hand1 vs. 0.4389 for hand2). Visual inspection of the results from the correlation search confirmed the assignment of absolute hand.

Docking of TBP into the TFIID Map

To simplify the interpretation of our TFIID and TFIIDΔTBP maps, we divided the reconstructions into five regions (Figure 2), which we called the core-, arm-, back, deco1-, and deco2-regions ("deco" stands for "decorative protein module"). In one of the two states (TFIID), we detected a sixth module that corresponded in size and shape to what would be expected for TBP in complex with the N-terminal stirrup of TAF1 (Figure 2A: right panel, gray). Docking of the most probable conformer of the *Drosophila* TBP-$TAF1_{11-77}$ NMR-ensemble (PDBid: 1TBA) into the TFIID map was aided by local symmetry analysis (see Figure S6) and the wealth of distinct fine structural details present in the TFIID reconstruction. Positioning of the $TAF1_{11-77}$ NMR-structure was determined by automatic real space rigid body docking in Chimera (Pettersen et al., 2004). To validate the docking, we performed common line correlation search using projections of the density map, calculated from the most probable conformer of the NMR-ensemble and the TFIID reconstruction as references and targets, respectively. The search was performed over the entire projection direction space and resulted in three correlation peaks. The first peak represented the presented docking (highest correlation peak), and the other two correlation maxima were explained by the pseudo-two-fold symmetry of the TBP-TAF complex, equivalent to the behavior of the local symmetry search (see Figure S6). We

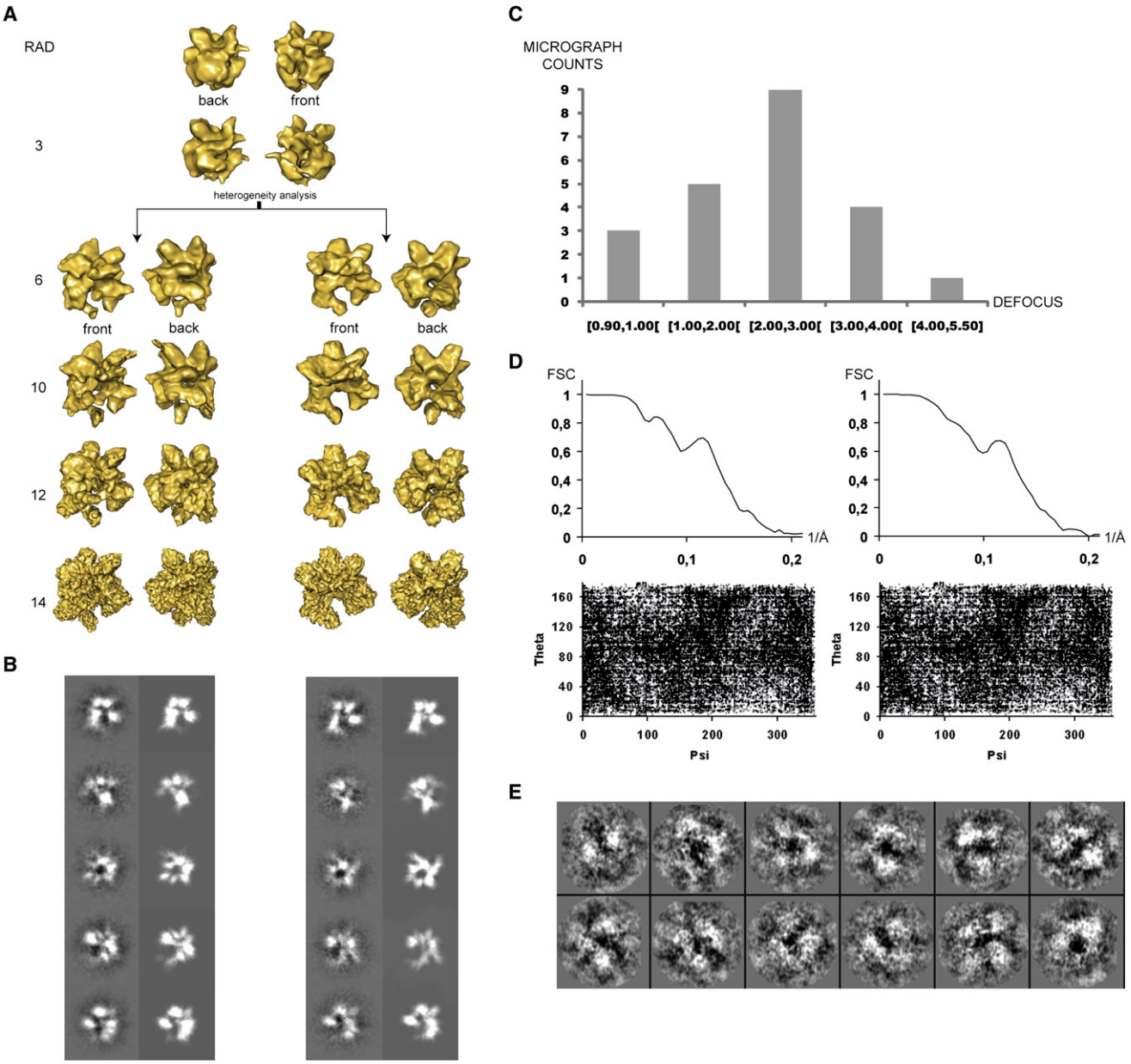

Figure 1. Validation Data for the Single-Particle 3D Reconstructions

(A) Refinement cycle for the *S. pombe* TFIIDΔTBP and TFIID reconstructions. RAD (Reference-free Alignment in a Discrete angular space) is the ab initio reconstruction (Elmlund et al., 2008). Numbering denotes round of model-based refinement. Heterogeneity analysis and supervised classification were performed as described above and in the Supplemental Data.

(B) Projections of the reconstructions (right column) and corresponding projection averages (left column) for the two states, selected according to five evenly distributed projection directions.

(C) Histogram of defocus values for the 22 micrographs of the data set.

(D) Fourier shell correlation plots (upper panel) and diagrams over the distribution of the two nonazimuthal Euler angles (lower panel). The resolution of the two TFIID reconstructions was determined to 8–10 Å according to the FSC = 0.5 criterion.

(E) Gallery of windowed particles from a micrograph acquired at 3.8 μm defocus, band-pass filtered using a [12,160] Å resolution window.

observed density features that matched well with the skewed β sheet and the C-terminal helix of TBP (Figures 3B–3D). Although our fit of TBP-TAF1$_{11-77}$ was excellent, density for the N-terminal TBP helix was hard to interpret, which may be explained by pronounced TAF-TBP interactions. The peripheral localization of TBP explains why it does not play a crucial role for the integrity of the TFIID complex and is in agreement with findings that direct TAF interactions to TBP occur only for TAF1 (Liu et al., 1998) and TAF7 (Yatherajam et al., 2003). The central position of TBP, just above the central pore, agrees

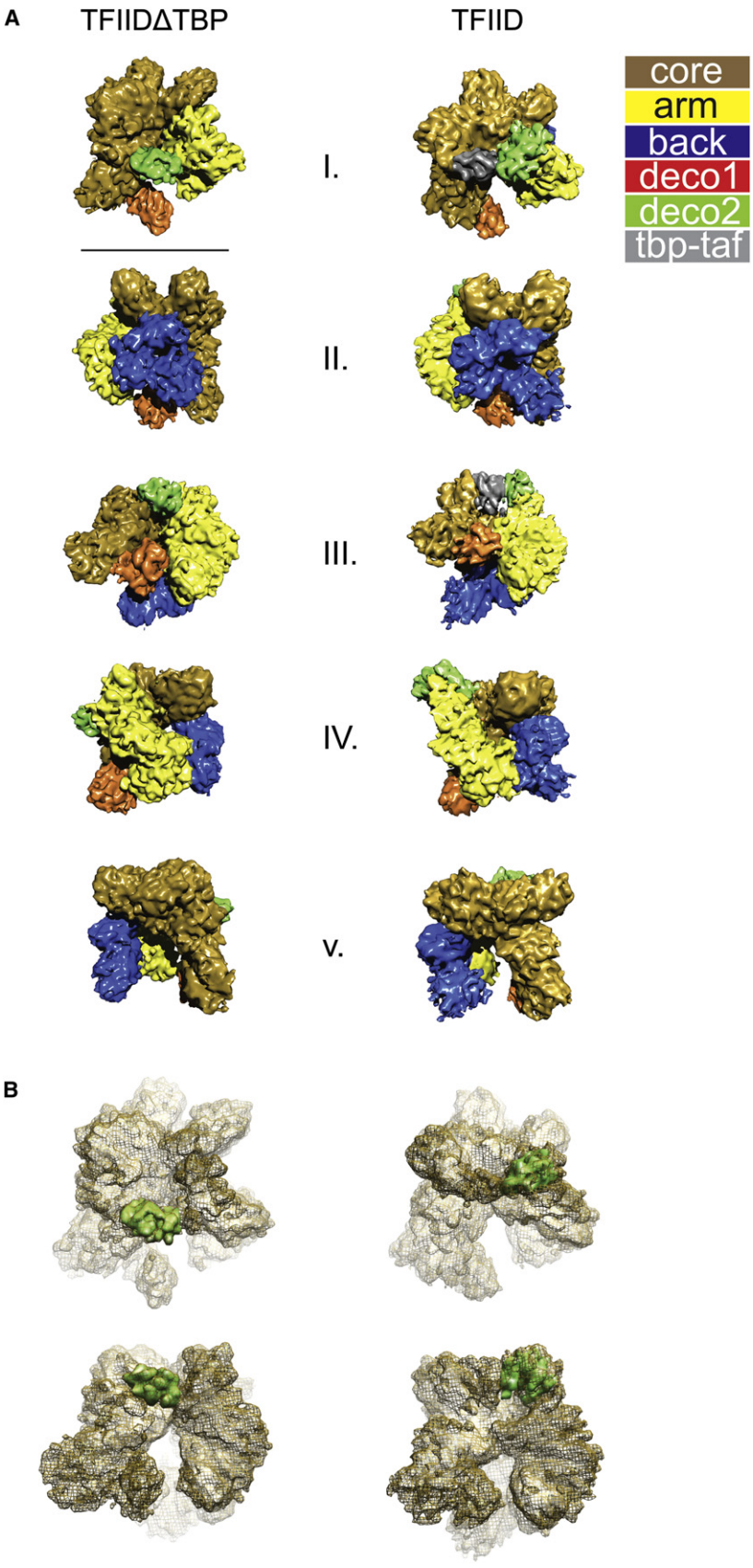

Figure 2. Quaternary Structure Arrangement of TFIID

Rendering volume was selected according to the molecular weight of TFIID.

(A) Structurally distinct and coherent quaternary regions were masked out in the TBP-lacking state (left panel) and the TBP-containing state (right panel). We describe our TFIID reconstructions as divided into five common quaternary structure regions: the core-region (brown), arm-region (yellow), back-region (blue), deco1-region (red), and deco2-region (green) ("deco" stands for "decorative protein module"), sorted in order of descending molecular weight. The tbp-taf region (gray) was present exclusively in one of the states (right panel). Viewing directions are defined as follows: I, front; II, back; III, bottom; IV, arm-side; and V, core-side. Scale bar is 150 Å.

(B) Formation of a TBP-binding site within TFIID, with the deco2 segment (green) within TFIIDΔTBP (left) and the fit of the segment to the TBP-containing reconstruction (right).

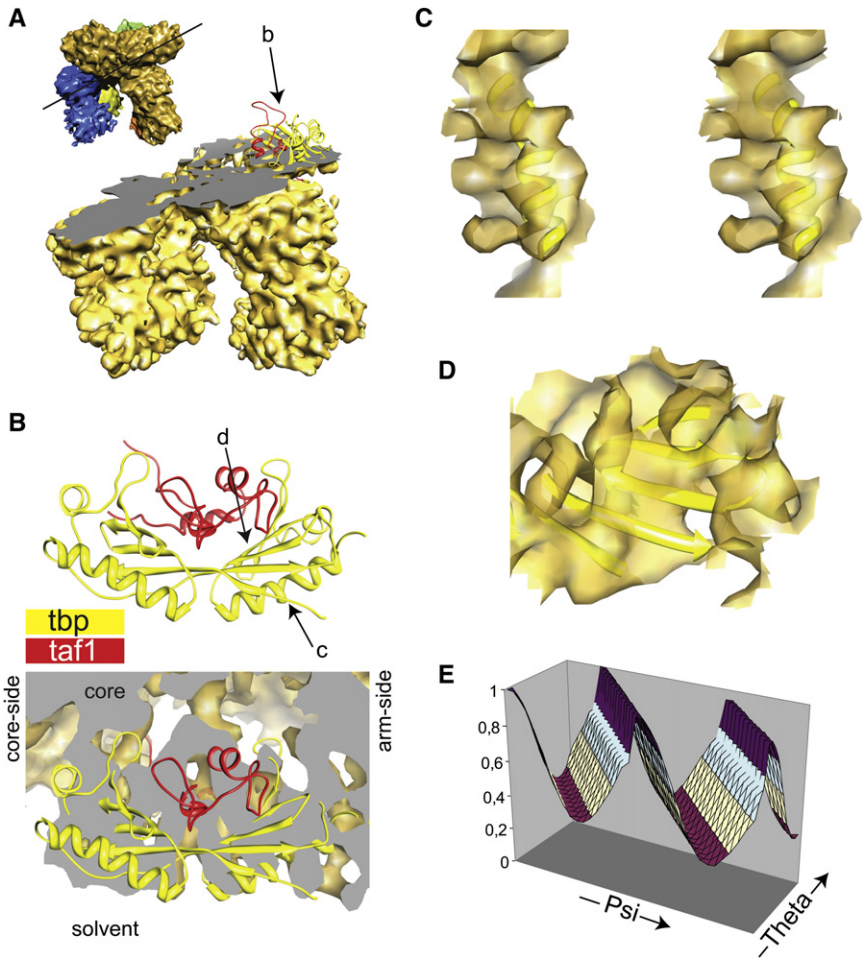

Figure 3. Docking of the TBP-TAF1$_{11-77}$ NMR-Structure

Rendering volume was selected to maximize the helical densities.

(A) Cut through the TFIID density and definition of the viewing direction in (b).

(B) View orthogonal to the cut demonstrates that TAFs occupy the concave DNA-binding surface of TBP. Viewing directions in (c) and (d) are indicated.

(C) Fit of the C-terminal TBP helix in stereo.

(D) View facing the DNA-binding pocket of TBP shows the fit of the C-terminal end of the skewed TBP β sheet.

(E) Validation of the fit of the TBP-TAF1$_{11-77}$ NMR structure by common line correlation-based search over the entire projection direction space. The analysis revealed three correlation peaks. The first (leftmost) correlation peak represents the presented docking (highest correlation peak), and the second peak is an artificial peak induced by the pseudo-two-fold symmetry axis present perpendicular to the real two-fold symmetry axis of TBP. All two-fold symmetrical objects have a pseudo-two-fold symmetry axis perpendicular to the real axis. First in the D_2 point-group this pseudo-axis becomes a real two-fold axis. The third peak represents the symmetry-induced correlation peak.

with the earlier immunolabeling of TBP in human TFIID (Andel et al., 1999). Moreover, the high degree of structural similarity between the TBP-TAF1$_{11-77}$ NMR data and the corresponding region of our TFIID reconstruction supports the previously reported inhibitory role of the TAFs (Liu et al., 1998; Mal et al., 2004). We could not identify the TBP-TAF$_{11-77}$ NMR structure within the other conformer (TFIIDΔTBP), which evidently lacked the TBP density. In agreement with Sanders et al. (2002), we therefore identify the two states as containing and lacking TBP.

The formation of a TBP-binding site involved one of the small decorative protein modules (Figure 2; deco2). To characterize the rearrangement of deco2 in the TFIIDΔTBP to TFIID transition, we fitted the segment of the TBP-lacking reconstruction into the TBP-containing map. Figure 2B shows the deco2 segment within TFIIDΔTBP (left) and the fit of the segment to the TBP-containing reconstruction (right). To confirm the ability of TBP to dynamically associate with the TAFs, we added purified TBP in seven-fold molar excess to the original heterogeneous preparation and collected 14,991 single-particle cryo-EM images. To investigate any preference of the population to any of the two resolved TFIID states, we applied the principle of supervised classification (Gao et al., 2004), using our final reconstructions as templates for the common line correlation-based matching. Addition of TBP significantly affected the equilibrium between the two states. Although the original preparation showed an almost equal distribution between the states (TFIIDΔTBP, 31,974 particles; TFIID, 30,326 particles), the preparation with added TBP showed a preference for the TBP-containing conformation (TFIIDΔTBP, 5510 particles; TFIID, 9481 particles).

Docking of (hTAF4-hTAF12)$_2$ into the TFIID Map

The crystal structure of human TAF4 and in complex with human TAF12 has been solved. The two hTAF4-hTAF12 heterodimers interact to form a heterotetramer (hTAF4-hTAF12)$_2$ (Werten et al., 2002). Multiple sequence alignment between the hTAF4 and hTAF12 crystal structure sequences (PDBid: 1H3O) and the corresponding complete human, *Drosophila*, *S. pombe*, and *S. cerevisiae* TAF sequences, using ClustalW (Thompson et al., 1994), did not reveal any inserts or deletions within the structurally characterized regions (data not shown). The (hTAF4-hTAF12)$_2$ crystal structure should therefore be a good template for modeling the *S. pombe* (TAF4-TAF12)$_2$ complex. Local symmetry analysis identified a two-fold symmetrical module in the back region of the TFIID map (see Figure S6C), which fitted the size and shape of (hTAF4-hTAF12)$_2$. Docking was performed as described above. As expected, the correlation search used for validation resulted in three peaks. The first peak represented the presented docking (highest correlation peak), and the other two maxima were explained by the two-fold symmetry of the (TAF4-TAF12)$_2$ assembly (Figure 4). In the fitting of (TAF4-TAF12)$_2$, rodlike density features in the TFIID

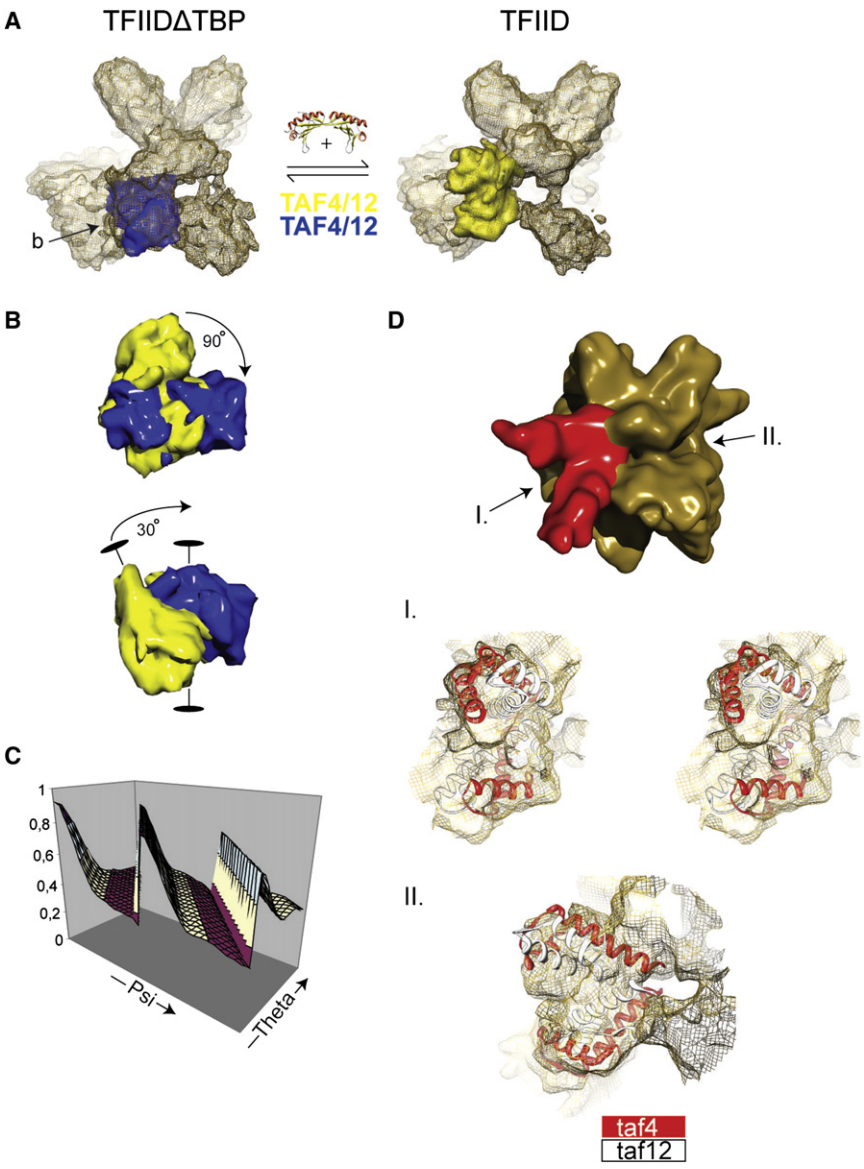

Figure 4. Identification, Docking, and Structural Dynamics Analysis of the $(TAF4\text{-}TAF12)_2$ Histone-Fold Tetramer Within TFIID

(A) The $(TAF4\text{-}TAF12)_2$ segment (yellow) within the TBP-containing state (right) and the fit of the segment (blue) to the TBP-lacking state (left).

(B) Views parallel and perpendicular to the two-fold symmetry axis of the $(TAF4\text{-}TAF12)_2$ segment within the TBP-lacking (blue) and TBP-containing (yellow) state, with the relative rotation and tilt around the two-fold symmetry axis indicated.

(C) Validation of the fit of the $(hTAF4\text{-}hTAF12)_2$ crystal structure by common line correlation-based search over the entire projection direction space. The analysis revealed three correlation peaks. The first (leftmost) correlation peak represents the presented docking (highest correlation peak), and the second peak is an artificial peak induced by the pseudo-two-fold symmetry axis present perpendicular to the real two-fold symmetry axis of $(TAF4\text{-}TAF12)_2$. All two-fold symmetrical objects have a pseudo-two-fold symmetry axis perpendicular to the real axis. First in the D_2 point-group this pseudo-axis becomes a real two-fold axis. The third peak represents the symmetry-induced correlation peak.

(D) Immunolabeling of the C terminus of TAF4 with the density of the ordered antibody regions in red according to the difference calculated between the labeled and the native structure. I, Stereo view along the two-fold symmetry axis of the docked $(hTAF4\text{-}hTAF12)_2$ crystal structure. II, View perpendicular to the two-fold symmetry axis of the docked $(hTAF4\text{-}hTAF12)_2$ crystal structure.

reconstruction coincided with the C-terminal hTAF4 helices. Our finding that the $(TAF4\text{-}TAF12)_2$ crystallographic unit also corresponded to a biological unit was puzzling, because previous immunolabeling of yeast TFIID identified TAF4 in lobes separated at distances approaching hundreds of Ångströms (Leurent et al., 2002, 2004). We therefore used an antibody recognizing the calmodulin-binding peptide of the C-terminal TAP-tag on TAF4 to specifically label the TAF4 C terminus. Low-resolution reconstructions were calculated from 22,000 cryo-EM images, with 11,201 and 10,799 particles assigned to the TFIIDΔTBP and TFIID states, respectively. No labeling was observed in the TFIIDΔTBP state, whereas binding to two closely associated regions was observed in the TBP-containing state, located in the center of the back (Figure 4D). It was obvious that the labeled protein module was present in the TFIIDΔTBP state, and the absence of structured antibody regions in the TBP-lacking reconstruction indicated that the conformational rearrangement had masked the epitope-binding surfaces. It appeared as if the $(TAF4\text{-}TAF12)_2$ units in the TFIIDΔTBP and TFIID reconstructions were related by a composite rotation. To further investigate this hypothesis, we masked out the region corresponding to $(TAF4\text{-}TAF12)_2$ in the TBP-containing state. Positioning of the $(TAF4\text{-}TAF12)_2$ density within the TFIIDΔTBP reconstruction revealed a coordinate transformation involving a ~30° tilt of and a ~90° rotation around the two-fold symmetry axis of the docked $(TAF4\text{-}TAF12)_2$ structure (Figures 4A and 4B). We concluded that TBP binding to and release from the TAFs was coupled to a massive conformational rearrangement of the back region, likely involving a composite rotation of the $(TAF4\text{-}TAF12)_2$ tetramer.

3D Reconstruction of a TFIIDΔTBP-Promoter DNA Complex

Finally, we wanted to monitor interactions between TFIID and promoter DNA. A DNA fragment composed of 101 base pairs of the *S. pombe* NMT1 promoter (50 and 51 base pairs upstream and downstream of the transcription start site, respectively) was added in seven-fold molar excess to the original TFIID preparation, and low-resolution reconstructions were calculated from 14,300 cryo-EM images, with 7280 and 7020 particles assigned to the TFIIDΔTBP and TFIID states, respectively. In previous

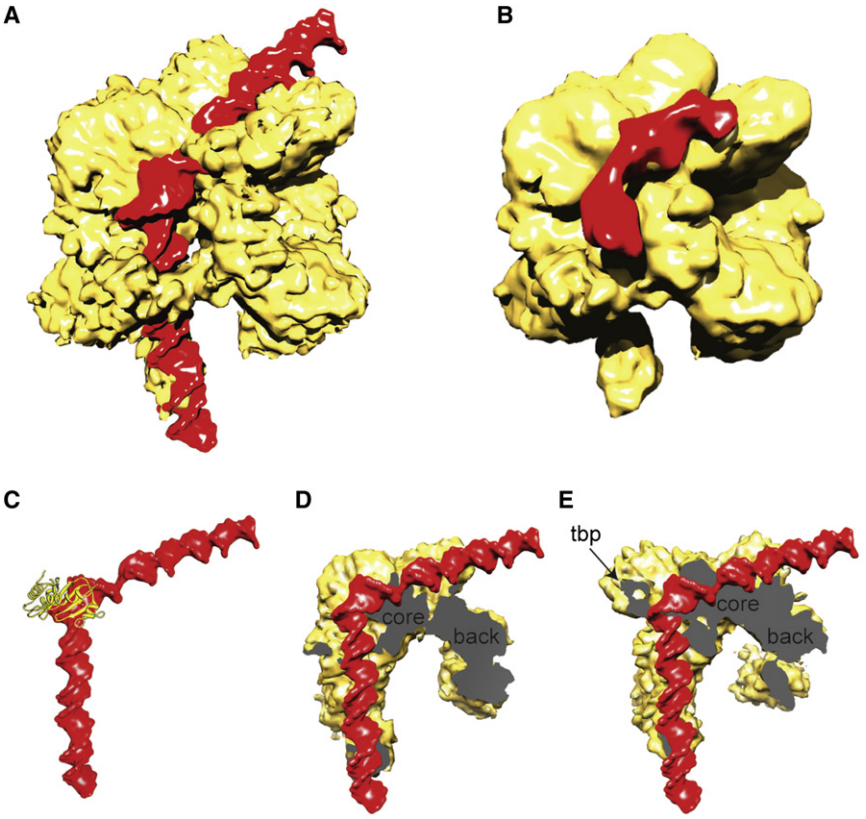

Figure 5. Promoter DNA Binding to TFIIDΔTBP

(A) An elongated promoter was modeled on the basis of the TBP-promoter-DNA crystal structure (shown in C) and docked to the TBP-lacking reconstruction according to the position of TBP.

(B) Reconstruction of a TFIIDΔTBP-promoter complex with the DNA colored red on the basis of difference mapping between the native TFIIDΔTBP reconstruction and the map calculated from the DNA-bound population.

(D and E) Cut along the DNA binding groove serves to illustrate the position of the promoter DNA kink and TBP in relation the TBP-lacking (D) and TBP-containing (E) reconstruction. Quaternary structure regions are indicated.

gel-shift experiments, we had observed direct binding between this DNA fragment and the TFIID complex (data not shown). The bent DNA structure appeared in the reconstruction calculated from particles mapped to the TFIIDΔTBP state in the position exactly expected from docking of TBP (Figure 5). No DNA binding was observed to the TBP-containing state. Our inability to visualize DNA binding to the TBP-containing complex does not imply that DNA associates only with TFIIDΔTBP. Rather, if promoter DNA binds with less specificity or exhibits a higher degree of flexibility when associated to the TBP-containing complex, this may prevent us from visualizing it as a result of averaging over heterogeneous states. The relative spatial arrangement of the localized protein modules with respect to the promoter DNA is illustrated in Figure 6.

DISCUSSION

In this work, we used single-particle cryo-EM to study the structure and dynamics of TFIID. First, we reported that purified TFIID exists in two predominant conformational states, which does not exclude the possibility that additional intermediate states exist with low occupancy in the data set. However, our method for heterogeneity analysis relies on the separation of images that cluster in factor space, and because we failed to detect more than two clusters, separation of additional states would be random. Second, we confirmed the ability of TBP to dynamically associate with the TAFs, and we characterized the structural dynamics underlying formation of a TBP-binding site within TFIID. Third, we reported that TBP binding to the TAFs is coupled to a massive rearrangement of a quaternary structure region comprising TAF4. Finally, our data demonstrate direct interactions between promoter DNA and a TBP-less TAF complex. Our localization of TBP agrees with its previously reported position, just above the central pore (Andel et al., 1999). Furthermore, docking of the NMR structure of TBP-TAF1$_{11-77}$ reveals a TBP configuration with the concave DNA-binding surface pointing toward the center of mass of the TFIID complex. This configuration indicates that TBP interactions with TFIID must be disrupted

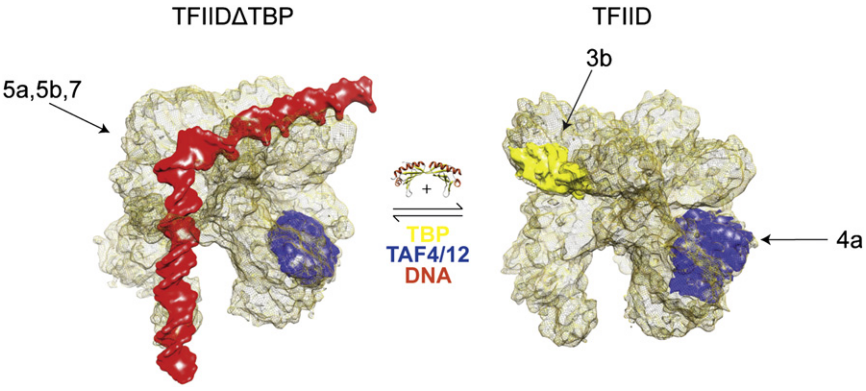

Figure 6. Relative Spatial Arrangement of DNA and Localized Protein Modules

Mesh representation of the TBP-lacking (left) and TBP-containing (right) state, with DNA (red), TBP (yellow), and (TAF4-TAF12)$_2$ (blue) indicated. Arrows indicate viewing directions and corresponding figure number.

before a complex between TBP and promoter DNA can be formed (Kim et al., 1993).

Structural Architecture of TFIID

The HFD TAF4 structure is present in two closely associated copies located in the center of the back. Our localization of TAF4 is based on vitrified specimens and new methods for 3D reconstruction, and it receives strong support from the (hTAF4-hTAF12)$_2$ crystal structure (Werten et al., 2002) in which two copies of each of the TAF4 and TAF12 HFD structures form a closely entangled two-fold symmetric tetramer. The two-fold symmetry of the (TAF4-TAF12)$_2$ structure is confirmed by local symmetry analysis of the back region. TAF4 appears to be present in two copies, and our positioning of the C-terminal TAF4 helix within the EM-density corresponds exactly with the extra density attributed to the structured regions of the anti-TAF4 antibodies. Our data therefore suggest that the (TAF4-TAF12)$_2$ crystallographic unit also corresponds to a biological unit. Certain regions of the crystal structure are not well defined in the TFIID map. Residues 17–26 of hTAF4 and 47–57 of hTAF12 in the crystallized sequence segments are highly conserved loop and helix end regions that are fully accessible to the solvent. The potentially high flexibility of these regions explains why they have poor occupancy in the TFIID reconstruction. The conformational transition coupled to binding of TBP leads to a composite rotation of (TAF4-TAF12)$_2$ such that exactly these conserved and potentially flexible regions become involved in interactions with the arm segment. Docking of (TAF4-TAF12)$_2$ places the terminal regions where the sequence continues either toward the arm segment or toward the additional domains of the back. A likely explanation is that the remaining TAF4 and TAF12 sequence parts entangle with the arm segment or the additional domains of the back, which moves in concert with the HFD tetramer and appears to be the "glue" that holds the arm and core segments together in the conformational transition of the TFIID structure. This observation agrees with the finding that TAF4, rather than TBP or TAF1, plays the most critical role in maintaining the stability of TFIID (Wright et al., 2006). Similar to what has been described for human TFIID (Grob et al., 2006), a conformational breathing of the molecule occurs, with the two main quaternary domains changing their relative position. The HFD rearrangement is directly related to TBP binding to and releasing from the TAFs. How this rearrangement may relate to other aspects of TFIID biology remains to be established.

DNA-TBP-, DNA-TAF, and TAF-TBP-Interactions and Regulation of Transcription

Loading of TBP onto the promoter is a crucial step in gene activation. Transcriptional activity in yeast strongly correlates with promoter occupancy by general factors such as TBP, TFIIA, and TFIIB, but not with occupancy of TAFs (Kuras et al., 2000). Our docking of the TBP-TAF1$_{11-77}$ NMR-structure into the TFIID map and reconstruction of TFIIDΔTBP bound to promoter DNA give a number of important structural insights, which may help to explain why we observe promoter DNA binding to TFIID only in the absence of TBP. We observe DNA binding to a two-fold symmetric groove that protrudes from the TBP-binding site and along the surface of the core. TBP thus appears to form a molecular lid over the groove, with its DNA-binding surface pointing toward the center of mass of the TFIID structure. This configuration of TBP within TFIID may prevent access of promoter DNA to the concave DNA-binding pocket of TBP by steric hindrance. It is tempting to speculate that TBP-TAF interactions must be disrupted before promoter-DNA interactions can take place with either TBP or TAFs. The TFIIDΔTBP to TFIID transition resolved here may describe the structural dynamics underlying this process, involving a reversible conformational change that may affect binding of TBP to the TAFs and allow for DNA entrance. It should be noted that this model does not exclude the existence of a TBP-containing and promoter-bound TAF complex, which may represent the last intermediate in the TFIID-dependent catalytic cycle of loading TBP onto the promoter.

In stark contrast to binding of TBP, promoter DNA binding to the TAFs does not introduce any conformational rearrangements of note. The DNA configuration within the promoter-bound TAF complex would result in major steric clashes between TAFs and pol II according to the current model for the positioning of pol II with respect to promoter DNA in the PIC (see Bushnell et al., 2004 and Chen et al., 2007 and references therein). This observation could indicate that the TAFs need to either partly dissociate from the promoter or change their positioning before entry of the polymerase into the PIC. In support of this notion, immobilized template assays have demonstrated that TFIID and TFIIA can assemble at promoters and block further PIC assembly (Ranish et al., 1999). How this block is relieved is not well understood, but it could perhaps be a regulatory target for transcriptional activators and repressors.

The TATA consensus does not appear to be a major determinant of TBP binding in yeast (Kim and Iyer, 2004). Only approximately 20% of the genes in yeast contain a TATA box, and these genes are associated with responses to stress, they are highly regulated, and they preferentially utilize SAGA before TFIID to load TBP (Basehoar et al., 2004). One possibility is therefore that TFIIDΔTBP plays an especially important role in DNA bending and TBP loading at promoters lacking a TATA consensus (illustrated in Figure 7). Another possibility is that of transcription without TBP, which would require stabilization of the bent promoter structure by general factors other than TBP to accomplish initiation. The latter possibility agrees with what has been observed for the TBP-lacking human TAF complex, TFTC (Wieczorek et al., 1998). We see no reason to exclude any of the two pathways. The TAFs are involved in both activation and repression of transcription. They inhibit TBP-mediated basal transcription in the absence of activator and they stimulate activated transcription in synergy with the Mediator complex (Burley and Roeder, 1998; Guermah et al., 1998, 2001). The steric hindrance of promoter-TAF and promoter-TBP interactions exerted by TBP when bound to the TAFs may provide one part of the structural basis for the dual role of TAFs in transcription regulation.

EXPERIMENTAL PROCEDURES

Purification of TBP and TFIID

Recombinant spTBP with 6xHis tag in the PET21b(+) vector was expressed in *Escherichia coli* (Cod+) cells for 15 hr at room temperature and purified over Ni column followed by heparin column. For purification of TAP-tagged TFIID, 15 l

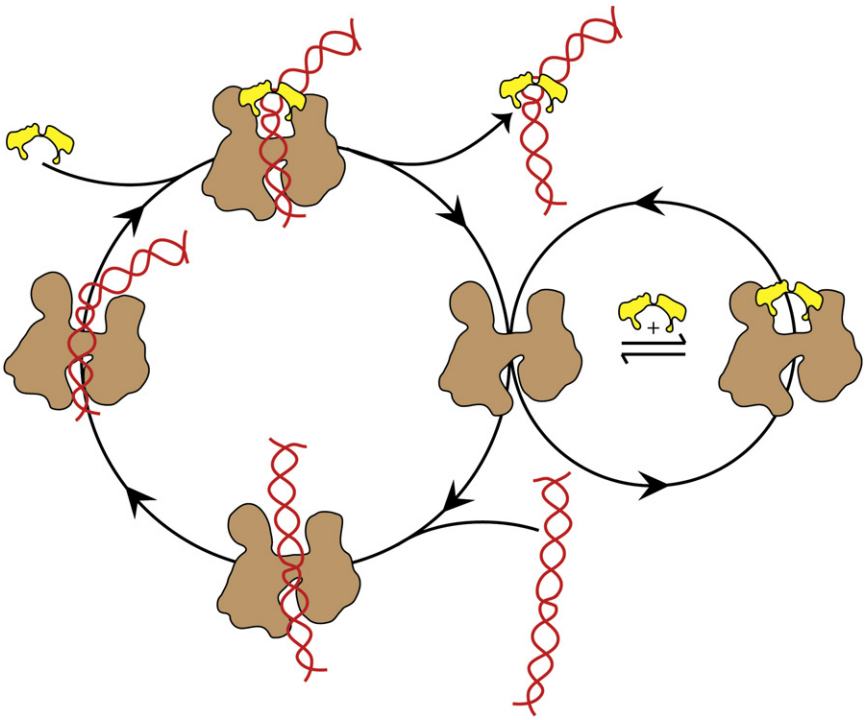

Figure 7. A Model for TFIID Function Based on Our Current Structural Knowledge about TFIID

TFIID (brown) exists in a dynamic equilibrium between TBP-containing and TBP-lacking states. Promoter DNA (red) binding occurs primarily to the TBP-lacking state. TFIIDΔTBP introduces the bent promoter structure required for initiation and in TBP-dependent transcription it may function by presenting a protein-DNA surface for TBP to interact with. At promoters that are not dependent on TBP, general factors other than TBP must stabilize the promoter kink.

of *S. pombe* cells was grown to OD_{600} 3.0–4.5 in YES medium supplemented with 0.2 g/l adenine. Cells were collected by centrifugation (JA-10, Beckman Coulter, at 2,500 rpm, for 7 min at +4°C), washed once with ice-cold water, and frozen in liquid nitrogen. Cells were broken in a Freezer/Mill 6850 (SPEX CertiPrep, NJ) using the following program: 10 min precooling, 5 cycles with 2 min beating, and 2 min rest at stringency 14. Broken cells were suspended in 0.5 ml of 3*TAP buffer (200 mM KOH-HEPES [pH 7.8], 15 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM EDTA, and 15% glycerol) per gram of cell pellet. DTT (0.5 mM) and protease inhibitors were added last, and the amount was adjusted to fit the total volume. After clearing of the supernatant by centrifugation (JA-10, at 9000 rpm, for 15 min, at +4°C), 1/9 volume of 2 M KCl was added followed by stirring for 15 min. After ultracentrifugation (Ti45, Beckman Coulter, at 42,000 rpm, for 30 min, at +4°C), the supernatant was frozen. IgG beads (300 μl; slurry; Amersham Biosciences) was added per each of 30–45 ml of extract and was incubated for 1 hr at +4°C. IgG beads were collected by centrifugation (JA-17, Beckman Coulter, at 1,000 rpm, for 2 min, at +4°C) and were washed with IgG buffer (10 mM Tris-HCl and 150 mM KOAc [pH 8.0]) and then by tobacco etch virus (TEV) protease cleavage buffer (25 mM HEPES-KOH [pH 7.6], 150 mM KOAc, 1 mM DTT, 0.5 mM EDTA, and 0.05% NP-40). TFIID was eluted by incubation for 1.5–2 hr at +16°C with 200 units of TEV protease in 1.5–2 ml of TEV protease cleavage buffer. Eluted TFIID was loaded onto a MiniS column (PE 4.6/50) and was subjected to 10 column volumes of gradient 0.15–1.5 M KOAc in the buffer (25 mM HEPES-KOH [pH 7.6], 10% glycerol, 0,05% NP-40, 1 mM DTT, protease inhibitors, and 2 mM MgOAc). TFIID eluted around 0.5 M KOAc. Fractions containing 0.03–0.08 g/l of TFIID were used for the following experiments. TFIID was concentrated by TCA precipitation, and TAF subunits were resolved on 12% SDS PAGE. Protein bands were excised, were in-gel digested by sequencing grade-modified trypsin from Promega, and were analyzed by MALDI-TOF MS on an Ultraflex TOF/TOF instrument (Bruker). TBP was detected by Western blot, developed with anti-*S. cerevisiae* TBP (Santa Cruz TBP (y-240): sc-33736).

GEMMA

TFIID was exchanged into a buffer consisting of 150 mM ammonium acetate (pH 7.8) using Sephadex-25 chromatography and was diluted into a protein concentration of 0.005 mg/ml in ammonium acetate (final concentration, 95 mM). GEMMA analysis was performed as described elsewhere (Rofougaran et al., 2006) using a pressure drop of 1.8 psi to direct the flow-rate. The results from the GEMMA analysis are shown in Figure S1.

Generation of Complex TFIID with TAP-TAF4 Antibody

Affinity-purified rabbit polyclonal anti-TAP antibody, recognizing C terminus of the TAP construct after TEV cleavage, came from OpenBiosystems (CAB1001). To prepare complexes, approximately five times molar excess of antibodies were added to the TFIID preparation. After mixing, TFIID-antibody complexes were dialyzed against binding buffer (10 mM Tris-HCl [pH 8.0], 150 mM NaCl, 0.05% NP-40, and 1 mM DTT).

Cryo-EM and Single-Particle Processing

Vitrified specimen was prepared, and image acquisition and digitization were performed as described elsewhere (Elmlund et al., 2008), resulting in a sampling size of 2.33 Å/pixel at the specimen level. Defocus of the micrographs varied from 1.0 to 5.5 μm to avoid systematic loss of information due to the contrast transfer function (CTF). Selection of particles, CTF-correction, 2D alignment, image classification, and initial model generation were performed as described elsewhere (Elmlund et al., 2008). We used the method described in Supplemental Data to do heterogeneity analysis. Refinement and supervised classification was performed in Strul (Lindahl, 2001). UCSF chimera (Pettersen et al., 2004) was used for segmentation, docking and visualization.

ACCESSION NUMBERS

Maps have been deposited in the EMDataBank with accession codes EMD-5134 and EMD 5135.

SUPPLEMENTAL DATA

Supplemental data include six figures and may be found with this article online at http://www.cell.com/structure/supplemental/S0969-2126(09)00375-X.

ACKNOWLEDGMENTS

We thank Roger D. Kornberg, Anders Liljas, David Bushnell, Stefan Björklund, Per Elias, and Philip Koeck for critical reading of the manuscript and helpful suggestions; the LUNARC centre for distributed computing; and Pasi Purhonen for microscope assistance. This work was supported by grants from the Swedish Research Council to H.E, C.M.G., and H.H, from the Swedish Cancer Society to C.M.G., and from the Swedish Foundation for Strategic Research to C.M.G. The remaining authors have no financial interest related to this work. H.E. and C.M.G. designed research; H.E., V.B., T.L., and R.R. performed research; H.E. developed the method for heterogeneity analysis; and H.E. and C.M.G. analyzed data. All authors contributed to writing of the article.

Received: March 27, 2009
Revised: September 7, 2009
Accepted: September 12, 2009
Published: November 10, 2009

REFERENCES

Andel, F., 3rd, Ladurner, A.G., Inouye, C., Tjian, R., and Nogales, E. (1999). Three-dimensional structure of the human TFIID-IIA-IIB complex. Science *286*, 2153–2156.

Bacher, G., Szymanski, W.W., Kaufman, S.L., Zollner, P., Blaas, D., and Allmaier, G. (2001). Charge-reduced nano electrospray ionization combined with differential mobility analysis of peptides, proteins, glycoproteins, noncovalent protein complexes and viruses. J. Mass Spectrom. *36*, 1038–1052.

Basehoar, A.D., Zanton, S.J., and Pugh, B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. Cell *116*, 699–709.

Bhattacharya, S., Takada, S., and Jacobson, R.H. (2007). Structural analysis and dimerization potential of the human TAF5 subunit of TFIID. Proc. Natl. Acad. Sci. USA *104*, 1189–1194.

Brand, M., Leurent, C., Mallouh, V., Tora, L., and Schultz, P. (1999). Three-dimensional structures of the TAF(II)-containing complexes TFIID and TFTC. Science *286*, 2151–2153.

Burley, S.K., and Roeder, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu. Rev. Biochem. *65*, 769–799.

Burley, S.K., and Roeder, R.G. (1998). TATA box mimicry by TFIID: Autoinhibition of pol II transcription. Cell *94*, 551–553.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: An RNA polymerase II-TFIIB cocrystal at 4.5 angstroms. Science *303*, 983–988.

Chalkley, G.E., and Verrijzer, C.P. (1999). DNA binding site selection by RNA polymerase II TAFs: a TAF(II)250-TAF(II)150 complex recognizes the Initiator. EMBO J. *18*, 4835–4845.

Chen, H.T., Warfield, L., and Hahn, S. (2007). The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat. Struct. Mol. Biol. *14*, 696–703.

Elmlund, H., Lundqvist, J., Al-Karadaghi, S., Hansson, M., Hebert, H., and Lindahl, M. (2008). A new cryo-EM single-particle *ab initio* reconstruction method visualizes secondary structure elements in an ATP-fuelled AAA+ motor. J. Mol. Biol. *375*, 934–947.

Gao, H.X., Valle, M., Ehrenberg, M., and Frank, J. (2004). Dynamics of EF-G interaction with the ribosome explored by classification of a heterogeneous cryo-EM dataset. J. Struct. Biol. *147*, 283–290.

Grant, P.A., Schieltz, D., Pray-Grant, M.G., Steger, D.J., Reese, J.C., Yates, J.R., and Workman, J.L. (1998). A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. Cell *94*, 45–53.

Grob, P., Cruse, M.J., Inouye, C., Peris, M., Penczek, P.A., Tjian, R., and Nogales, E. (2006). Cryo-electron microscopy studies of human TFIID: Conformational breathing in the integration of gene regulatory cues. Structure *14*, 511–520.

Guermah, M., Malik, S., and Roeder, R.G. (1998). Involvement of TFIID and USA components in transcriptional activation of the human immunodeficiency virus promoter by NF-kappa B and Sp1. Mol. Cell. Biol. *18*, 3234–3244.

Guermah, M., Tao, Y., and Roeder, R.G. (2001). Positive and negative TAF(II) functions that suggest a dynamic TFIID structure and elicit synergy with TRAPs in activator-induced transcription. Mol. Cell. Biol. *21*, 6882–6894.

Hahn, S., Buratowski, S., Sharp, P.A., and Guarente, L. (1989). Isolation of the gene encoding the yeast Tata binding-protein TFIID: a gene identical to the Spt15 suppressor of Ty element insertions. Cell *58*, 1173–1181.

Horikoshi, M., Wang, C.K., Fujii, H., Cromlish, J.A., Weil, P.A., and Roeder, R.G. (1989). Cloning and structure of a yeast gene encoding a general transcription initiation-factor TFIID that binds to the Tata box. Nature *341*, 299–303.

Kaufman, S.L., Skogen, J.W., Dorman, F.D., Zarrin, F., and Lewis, K.C. (1996). Macromolecule analysis based on electrophoretic mobility in air: Globular proteins. Anal. Chem. *68*, 1895–1904.

Kim, J., and Iyer, V.R. (2004). Global role of TATA box-binding protein recruitment to promoters in mediating gene expression profiles. Mol. Cell. Biol. *24*, 8104–8112.

Kim, J.L., Nikolov, D.B., and Burley, S.K. (1993). Co-crystal structure of Tbp recognizing the minor-groove of a Tata element. Nature *365*, 520–527.

Kuras, L., Kosa, P., Mencia, M., and Struhl, K. (2000). TAF-containing and TAF-independent forms of transcriptionally active TBP in vivo. Science *288*, 1244–1248.

Lee, T.I., Causton, H.C., Holstege, F.C.P., Shen, W.C., Hannett, N., Jennings, E.G., Winston, F., Green, N.R., and Young, R.A. (2000). Redundant roles for the TFIID and SAGA complexes in global transcription. Nature *405*, 701–704.

Leurent, C., Sanders, S., Ruhlmann, C., Mallouh, V., Weil, P.A., Kirschner, D.B., Tora, L., and Schultz, P. (2002). Mapping histone fold TAFs within yeast TFIID. EMBO J. *21*, 3424–3433.

Leurent, C., Sanders, S.L., Demeny, M.A., Garbett, K.A., Ruhlmann, C., Weil, P.A., Tora, L., and Schultz, P. (2004). Mapping key functional sites within yeast TFIID. EMBO J. *23*, 719–727.

Lindahl, M. (2001). Strul: a method for 3D alignment of single-particle projections based on common line correlation in Fourier space. Ultramicroscopy *87*, 165–175.

Liu, D., Ishima, R., Tong, K.I., Bagby, S., Kokubo, T., Muhandiram, D.R., Kay, L.E., Nakatani, Y., and Ikura, M. (1998). Solution structure of a TBP-TAF(II)230 complex: Protein mimicry of the minor groove surface of the TATA box unwound by TBP. Cell *94*, 573–583.

Mal, T.K., Masutomi, Y., Zheng, L., Nakata, Y., Ohta, H., Nakatani, Y., Kokubo, T., and Ikura, M. (2004). Structural and functional characterization on the interaction of yeast TFIID subunit TAF1 with TATA-binding protein. J. Mol. Biol. *339*, 681–693.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal-structure of a Tfiib-Tbp-Tata-element ternary complex. Nature *377*, 119–128.

Oelgeschläger, T., Chiang, C.M., and Roeder, R.G. (1996). Topology and reorganization of a human TFIID-promoter complex. Nature *382*, 735–738.

Ogryzko, V.V., Kotani, T., Zhang, X.L., Schiltz, R.L., Howard, T., Yang, X.J., Howard, B.H., Qin, J., and Nakatani, Y. (1998). Histone-like TAFs within the PCAF histone acetylase complex. Cell *94*, 35–44.

Papai, G., Tripathi, M.K., Ruhlmann, C., Werten, S., Crucifix, C., Weil, P.A., and Schultz, P. (2009). Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. Structure *17*, 363–373.

Peterson, M.G., Tanese, N., Pugh, B.F., and Tjian, R. (1990). Functional domains and upstream activation properties of cloned human TATA binding protein. Science *248*, 1625–1630.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF chimera: A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Ranish, J.A., Yudkovsky, N., and Hahn, S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. Genes Dev. *13*, 49–63.

Rofougaran, R., Vodnala, M., and Hofer, A. (2006). Enzymatically active mammalian ribonucleotide reductase exists primarily as an alpha(6)beta(2) octamer. J. Biol. Chem. *281*, 27705–27711.

Romier, C., James, N., Birck, C., Cavarelli, J., Vivares, C., Collart, M.A., and Moras, D. (2007). Crystal structure, biochemical and genetic characterization of yeast and E-cuniculi TAF(II)5 N-terminal domain: Implications for TRID assembly. J. Mol. Biol. *368*, 1292–1306.

Sanders, S.L., Garbett, K.A., and Weil, P.A. (2002). Molecular characterization of *Saccharomyces cerevisiae* TFIID. Mol. Cell. Biol. *22*, 6000–6013.

Sawadogo, M., and Roeder, R.G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the Tata box region. Cell *43*, 165–175.

Selleck, W., Howley, R., Fang, Q.J., Podolny, V., Fried, M.G., Buratowski, S., and Tan, S. (2001). A histone fold TAF octamer within the yeast TFIID transcriptional coactivator. Nat. Struct. Biol. *8*, 695–700.

Shao, H., Revach, M., Moshonov, S., Tzuman, Y., Gazit, K., Albeck, S., Unger, T., and Dikstein, R. (2005). Core promoter binding by histone-like TAF complexes. Mol. Cell. Biol. *25*, 206–219.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). Clustal-W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.

Werten, S., Mitschler, A., Romier, C., Gangloff, Y.G., Thuault, S., Davidson, I., and Moras, D. (2002). Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). J. Biol. Chem. *277*, 45502–45509.

Wieczorek, E., Brand, M., Jacq, X., and Tora, L. (1998). Function of TAF(II)-containing complex without TBP in transcription by RNA polymerase II. Nature *393*, 187–191.

Wright, K.J., Ii, M.T.M., and Tjian, R. (2006). TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. Proc. Natl. Acad. Sci. USA *103*, 12347–12352.

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). Structural similarity between TAFs and the heterotetrameric core of the histone octamer. Nature *380*, 316–322.

Yatherajam, G., Zhang, L., Kraemer, S.M., and Stargell, L.A. (2003). Protein-protein interaction map for yeast TFIID. Nucleic Acids Res. *31*, 1252–1260.

