

RNA Polymerase II/TFIIF Structure and Conserved Organization of the Initiation Complex

Wen-Hsiang Chung,^{1,3} John L. Craighead,^{1,3}
Wei-Hau Chang,² Chukwudi Ezeokonkwo,¹
Avital Bareket-Samish,² Roger D. Kornberg,²
and Francisco J. Asturias^{1,*}

¹Department of Cell Biology
The Scripps Research Institute
10550 North Torrey Pines Road
La Jolla, California 92037

²Department of Structural Biology
Stanford University School of Medicine
Fairchild Research Building
Stanford, California 94305

Summary

The structure of an RNA polymerase II/general transcription factor TFIIF complex was determined by cryo-electron microscopy and single particle analysis. Density due to TFIIF was not concentrated in one area but rather was widely distributed across the surface of the polymerase. The largest subunit of TFIIF interacted with the dissociable Rpb4/Rpb7 polymerase subunit complex and with the mobile “clamp.” The distribution of the second largest subunit of TFIIF was very similar to that previously reported for the σ subunit in the bacterial RNA polymerase holoenzyme, consisting of a series of globular domains extending along the polymerase active site cleft. This result indicates that the second TFIIF subunit is a true structural homolog of the bacterial σ factor and reveals an important similarity of the transcription initiation mechanism between bacteria and eukaryotes. The structure of the RNAPII/TFIIF complex suggests a model for the organization of a minimal transcription initiation complex.

Introduction

X-ray structures of bacterial and eukaryotic RNA polymerases show highly similar protein folds in the active center region, indicative of a conserved mechanism of RNA chain elongation (Cramer et al., 2001; Zhang et al., 1999). On the other hand, it has long been thought that the mechanisms of initiation of RNA chains by these enzymes are fundamentally different. Most bacterial promoters consist of conserved sequences about 10 and 35 bp upstream of the transcription start site (deHaeseht et al., 1998), whereas promoters for the principal eukaryotic enzyme, RNA polymerase II (RNAPII), most often contain a TATA box about 25 bp upstream. Bacterial promoters are recognized by a single RNA polymerase subunit, known as σ factor, whereas initiation by RNAPII requires five general transcription factors, termed TFIIB, -IID, -IIE, -IIF, and -IIH, comprising, minimally, 15 additional polypeptides (Kornberg, 1999). We report here that despite the far greater complexity of

the RNAPII initiation machinery, the overall organization of a minimal eukaryotic initiation complex resembles closely that of the bacterial holoenzyme.

The bacterial σ factor serves two distinct roles: recognizing the -10 and -35 regions of the promoter and facilitating the formation of an unwound region of DNA surrounding the transcription start site (Gross et al., 1998). The role in promoter recognition entails not only promoter binding but also the capacity of σ to disrupt a nonspecific RNA polymerase-DNA complex. One of the general transcription factors, TFIIF (IIF), can disrupt a nonspecific RNAPII-DNA complex as well (Conaway and Conaway, 1990; Killeen and Greenblatt, 1992; Tan et al., 1995). Moreover, a subunit of IIF, termed Tfg2 in yeast and RAP30 in mammalian cells, is a clear amino acid sequence homolog of σ (Henry et al., 1994; Sopta et al., 1989). Information about the organization of the two largest yeast IIF subunits (yeast Tfg1/mammalian Rap74 and yeast Tfg2/mammalian Rap30) derived from sequence analysis is summarized in Figure 1.

Recent X-ray studies of bacterial RNA polymerase holoenzymes, which include the σ factor, have revealed an unusual structure of σ arrayed in a series of discrete domains across the surface of the largest polymerase subunits, connected by extended linkers (Murakami et al., 2002b; Vassylyev et al., 2002). These structural domains correspond to regions of the σ factor amino acid sequence previously defined on the basis of evolutionary conservation and genetic analysis. The X-ray structure of a bacterial polymerase-promoter DNA complex has shown how the multiple σ factor domains serve to guide the DNA toward the active center cleft and participate in its unwinding. We have used cryo-electron microscopy (cryo-EM) and image analysis to determine the structure of a yeast RNAPII/IIF complex. Regions of IIF density were localized by comparison with a reconstruction of RNAPII previously determined by cryo-EM (Craighead et al., 2002).

Results and Discussion

RNAPII/IIF Complex Purification and Characterization

Transcription factor IIF was first identified in mammalian cells based on its tight interaction with RNAPII (Sopta et al., 1985), and the human RNAPII/IIF complex can be immunoaffinity purified using antibodies against the largest (Rap74) IIF subunit (Cho et al., 1997). Affinity purification of the yeast RNAPII/IIF complex was attempted by introducing a TAP tag (Puig et al., 2001) at the C terminus of Tfg2, the middle yeast IIF subunit. The resulting Tfg2-TAP yeast strain was analyzed by PCR and immunoblotting to ensure correct integration. The cells showed a normal growth curve (doubling time of 120 min). Furthermore, the strain showed neither heat nor cold sensitivity phenotype (data not shown), indicating that the presence of the TAP-tag did not interfere significantly with the normal *in vivo* function of IIF. Affinity purification of the TAP-tagged yeast IIF resulted in

*Correspondence: asturias@scripps.edu

³These authors contributed equally to this work.

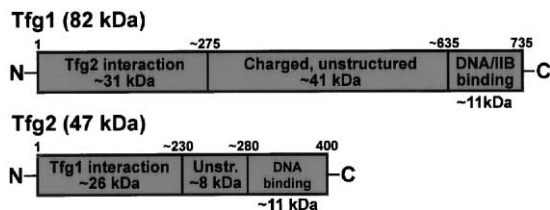


Figure 1. Domain Organization of the Two Largest Yeast IIF Subunits

Tfg1 and Tfg2, the two largest yeast IIF subunits (homologs of human IIF subunits Rap74 and Rap30, respectively), are similarly organized, with N- and C-terminal globular domains and a disordered, charged central region. Sequence analysis and alignment of the yeast and human IIF subunits suggest the approximate position of domain boundaries.

isolation of an RNAPII/IIF complex (Figure 2). The purity and homogeneity of the RNAPII/IIF complex were important for the electron microscopy work, and were tested by quantitative Coomassie blue and Sypro orange staining. The complex was found to be stoichiometric within the precision of the SDS-PAGE analysis. To support the biological relevance of the cryo-EM results, transcriptional activity of pure IIF obtained by high-salt displacement from the RNAPII/IIF complex was tested in an nonspecific transcription assay, and the purified factor was found to be active in promoter-dependent transcription from a G-less cassette (data not shown).

Calculation and Initial Interpretation of the RNAPII/IIF Structure

Random conical tilt and back projection methods (Rademacher, 1988) were used to calculate a 30 Å resolution reconstruction from images of RNAPII/IIF particles preserved in stain. This initial reconstruction was used as

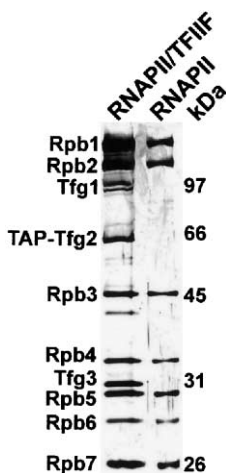


Figure 2. Purification of the RNAPII/IIF Complex

Immunopurification of the RNAPII/IIF complex was carried out using a TAP tag introduced at the C terminus of the Tfg2 IIF subunit (Puig et al., 2001). The eluate from the calmodulin column used in the last step of the purification procedure was fractionated by SDS-10% PAGE and silver stained. Copurification of IIF and RNAPII provides evidence for the strong interaction between the transcription factor and the polymerase.

the reference volume for refinement of unstained RNAPII/IIF particle images from frozen-hydrated specimens using a previously developed procedure (Craighead et al., 2002). A data set including ~30,000 RNAPII/IIF images generated a structure with a resolution of ~18 Å, estimated using the Fourier Shell correlation method (van Heel, 1987). The resulting RNAPII/IIF structure is shown in Figure 3A (blue). The overall structure of the RNAPII/IIF complex is similar to that of RNAPII alone. However, significant differences are immediately revealed by comparison with the structure of RNAPII alone, which is shown in Figure 3B (orange). The RNAPII/IIF structure shows additional density around the clamp module, in association with polymerase subunits Rpb4/Rpb7, above the “wall” that defines the upstream end of the active site cleft, between the clamp and shelf polymerase modules, and between the “feet” of the polymerase. Comparison between the RNAPII/IIF and RNAPII structures also suggests a rearrangement of the polymerase induced by binding of factor IIF.

In order to better interpret these results, the RNAPII/IIF and RNAPII structures were carefully aligned, and a difference map (RNAPII/IIF – RNAPII) was calculated. The results of the difference analysis are shown in Figure 4. Panels in Figure 4A show overlaid slabs of the RNAPII/IIF (blue) and RNAPII (orange) maps that illustrate how differences between the two structures are due not only to the presence of IIF density (marked by white cross marks) in the RNAPII/IIF complex but also to changes in the relative position of previously identified RNAPII domains (Cramer et al., 2001) that move as rigid modules. Two of the major rearrangements are illustrated in the middle panel of Figure 4A, where the arrows indicate changes in the position of the jaw-lobe module and the Rpb1 foot between the positions marked by the dotted white lines. Examining the aligned RNAPII/IIF and RNAPII structures as 3D maps and analyzing overlaid slabs like the ones shown in Figure 4A makes it possible to unequivocally distinguish between features in the (RNAPII/IIF – RNAPII) difference map related to polymerase conformational changes and those related to IIF density. The way in which movement of specific polymerase subunits gives rise to features in the (RNAPII/IIF – RNAPII) difference map is illustrated in Figure 4B, where the appropriate differences are shown as yellow patches overlaid on the X-ray structure of the 12 subunit RNAPII (Armache et al., 2003; Bushnell and Kornberg, 2003). Binding of IIF to polymerase causes a rotation of the shelf and jaw-lobe modules and pushes apart the polymerase Rpb1 and Rpb8 feet. This analysis confirms that the overall changes in RNAPII conformation are those indicated by the arrows in the bottom panel of Figure 3B.

Features in the (RNAPII/IIF – RNAPII) difference map that arise from the presence of IIF density are shown in Figure 4C as blue densities on the surface of the RNAPII X-ray structure. Interpretation of the (RNAPII/IIF – RNAPII) difference map shown in Figure 4C identifies more precisely the location of IIF density first revealed by visual comparison of the RNAPII/IIF and RNAPII structures shown in Figure 3. A significant portion of IIF density is closely associated with polymerase subunits Rpb4/Rpb7 (labeled 3 in Figure 4C). Additional IIF density is located along the length of the active site cleft of

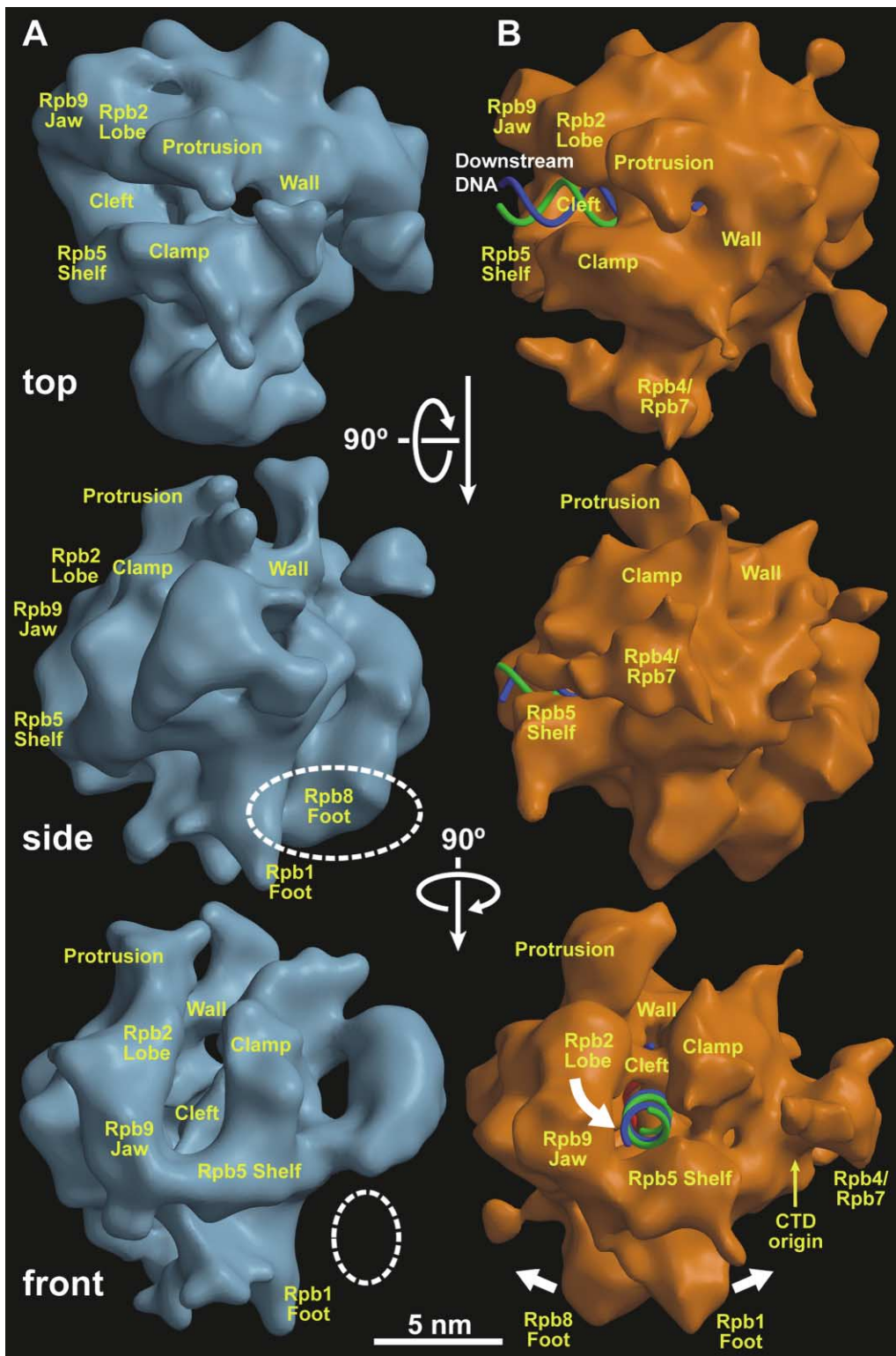


Figure 3. Cryo-EM Reconstruction of the Yeast RNAPII/IIF Complex

(A) Top, side, and front views of the ~ 18 Å resolution reconstruction of RNAPII/IIF generated from about 30,000 images of unstained, frozen-hydrated particles. The threshold for rendering was chosen based on a molecular weight of ~ 700 kDa for the RNAPII/IIF complex. Additional density is visible at a slightly lower threshold in the position encircled by the dashed lines, suggesting the presence of a disordered IIF domain. (B) An ~ 18 Å resolution cryo-EM reconstruction of RNAPII (Craighead et al., 2002) is shown for comparison, with promoter DNA modeled into the active site cleft to help orient the viewer. Comparison of the RNAPII/IIF and RNAPII structures reveals the presence of extra density around the active site cleft, polymerase feet, and polymerase subunits Rpb4/Rpb7, as well as changes in polymerase conformation, which are indicated by the arrows at the bottom.

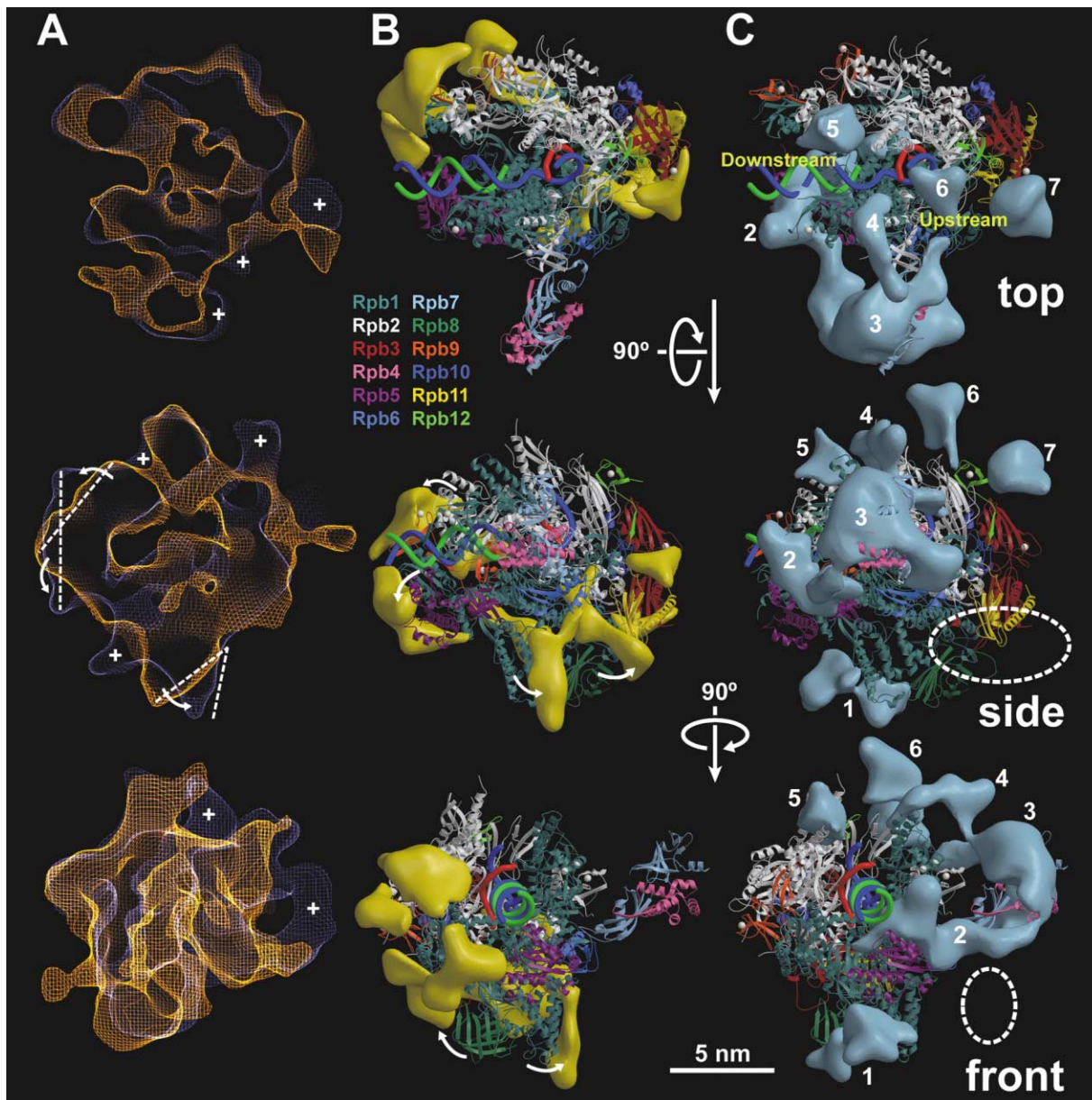


Figure 4. Analysis of the RNAPII/IIF Cryo-EM Reconstruction by Difference Mapping and Comparison with RNAPII

The RNAPII/IIF volume was amplitude corrected (amplitudes were normalized to X-ray amplitudes from an RNAPII/DNA/RNA complex (Gnatt et al., 2001) and aligned to a previously calculated RNAPII cryo-EM reconstruction (Craighead et al., 2002).

(A) Thin slabs of the aligned RNAPII/IIF (blue) and RNAPII (orange) reconstructions show the close correspondence between most portions of the two structures but also highlight the presence of some of the additional density due to the presence of IIF (marked by the white cross marks) and of changes in RNAPII conformation induced by IIF binding (marked by the change in position of white dotted lines; movement in the direction indicated by the arrows). Large portions of the RNAPII structure move as rigid modules, as previously reported from comparison of different 3D crystal forms of the enzyme (Cramer et al., 2001).

(B) Features in the (RNAPII/IIF – RNAPII) difference map related to polymerase conformational changes are shown in yellow, overlaid on the X-ray structure of the enzyme (promoter DNA was modeled into the polymerase active site cleft to help orient the viewer). The panels show specific RNAPII subunits involved in the IIF-induced changes in polymerase conformation. While the overlaid slabs shown in (A) are very helpful to differentiate between conformational changes and IIF density, the conformational changes can only be fully appreciated by looking at the entire structure. Note, for example, the change in the position of the Rpb8 foot, which is not apparent in the slab representation shown in the middle panel of (A). This analysis provides a detailed understanding of the overall changes in the structure of RNAPII summarized in the bottom panel of Figure 3B.

(C) Density due to IIF is shown in blue, overlaid on the X-ray structure of RNAPII. IIF density is apparent near the feet (1), at the downstream end of the active site cleft, filling the gap between the clamp and polymerase subunit Rpb5 (2), in close association with polymerase subunits Rpb4/Rpb7 (3), at the upstream end of the active site cleft (4, 5, and 6), and behind the wall that forms the end of the active site cleft (7). The position of a partially ordered IIF domain visible at a lower threshold is indicated as in Figure 3A.

polymerase, starting behind the wall (Figure 4C, 7), near subunit Rpb12 (Figure 4C, 6), next to the protrusion (Figure 4C, 5), along the top of the clamp (Figure 4C, 4), past the downstream end of the active site cleft near Rpb5 (Figure 4C, 2), and into the Rpb1/Rpb8 feet (Figure 4C, 1). Consistent with the sequence homology between IIF and the bacterial σ factor, comparison with the X-ray structure of the bacterial holoenzyme (Murakami et al., 2002b; Vassilyev et al., 2002) reveals a remarkable correspondence between the distribution of IIF density around the active site cleft and that of σ factor density in the bacterial complex. Further interpretation of the RNAPII/IIF structure required additional information about the location of individual IIF subunits, especially Tfg2, predicted by sequence comparison analysis to be the direct counterpart of the bacterial σ factor.

Tfg2 Subunit Localization

The TAP tag introduced in factor IIF for immunopurification of the RNAPII/IIF complex left a calmodulin binding peptide at the C terminus of the Tfg2 subunit. A cryo-EM reconstruction after calmodulin labeling of the RNAPII/IIF complex at this site indicated that the IIF-related density located behind polymerase subunit Rpb12 (Figure 4C, 7) corresponds to the C terminus of Tfg2 (data not shown). To complete the mapping of the Tfg2 subunit, we determined the structure of the complex formed by RNAPII and Tfg2. It is important to note that, although stable assemblies are formed between RNAPII and individual IIF subunits, several observations indicate that their structure may only partially correspond to that of the full RNAPII/IIF complex. For example, in the human RNAPII/IIF complex the interaction of Rap30 with RNAPII appears to be influenced by the presence of the largest IIF subunit, Rap74 (Fang and Burton, 1996; Wei et al., 2001). Biochemical characterization of RNAPII/IIF complexes formed using recombinant IIF subunits (Tan et al., 1994) suggests that formation of a fully functional IIF complex requires cofolding of IIF subunits (Wei et al., 2001). This observation is explained by the structure of the Rap74/Rap30 dimerization domain (Gaiser et al., 2000), which reveals a complex interaction. Therefore, we did not expect to rely on the RNAPII/Tfg2 reconstruction as a source of definitive structural information but planned to use it as a guide for identifying Tfg2 density in the reconstruction of the entire RNAPII/IIF complex and in the (RNAPII/IIF – RNAPII) difference map.

The RNAPII/Tfg2 complex was formed by incubating RNAPII with purified, recombinant Tfg2, and the structure of the resulting complex was determined. The RNAPII/Tfg2 image set was refined against two distinct references: the RNAPII structure (the refined RNAPII/Tfg2 volume should have some additional features due to the presence of Tfg2, and possibly an altered RNAPII conformation) and the RNAPII/IIF structure (the refined RNAPII/Tfg2 volume should lose some features, as only one of three IIF subunits was present). The RNAPII/Tfg2 volume obtained using the RNAPII/IIF structure as a reference is shown in Figure 5A. All of the IIF-related densities arranged along the active site cleft in the RNAPII/IIF structure (labeled 1, 2, 4, 5, 6, and 7 in Figure 4C) were present in the RNAPII/Tfg2 reconstruction. No-

tably absent was the IIF density associated with polymerase subunits Rpb4/Rpb7 (Figure 4C, 3). The arrangement of Tfg2 density around the RNAPII active site cleft corresponds closely to that of σ in the bacterial holoenzyme structure (Murakami et al., 2002b; Vassilyev et al., 2002), demonstrating that Tfg2 is the true structural homolog of the bacterial σ factor. This observation is consistent with the well-documented sequence homology between Tfg2 and the bacterial σ factor (see Figure 1), and with the functional homology between Tfg2/Rap30 and the σ factor (Conaway and Conaway, 1990; Groft et al., 1998; Killeen and Greenblatt, 1992; McCracken and Greenblatt, 1991). The distribution of Tfg2 density along the active site cleft of polymerase thus explains the biochemical function of Tfg2/Rap30 in modulating RNAPII/DNA interactions and the observation that Rap30 can bind to *E. coli* RNA polymerase but is displaced by σ (McCracken and Greenblatt, 1991).

The extended arrangement of Tfg2 density and the similarity with σ factor structure suggest that the Tfg2 N terminus is most likely located near the downstream end of the RNAPII active site cleft (densities labeled 1 and 2 in Figure 4C). The N and C termini of Tfg2 would be connected by a linker traveling through the active site cleft, as observed for σ in the bacterial holoenzyme structure (Murakami et al., 2002b; Vassilyev et al., 2002). That connection between different globular Tfg2 domains would not be apparent at the resolution of the cryo-EM reconstruction, but its existence is supported by analysis of the Tfg2 sequence, and by NMR structural studies of Rap30, which indicate that the central portions of both proteins (including amino acids 220–280 in Tfg2, and 90–160 in Rap30) possess no regular secondary structure in the absence of interaction with RNAPII (Groft et al., 1998; Henry et al., 1994). It has been reported that part of the middle portion of Rap30 (amino acids 101–170) interacts with RNAPII subunit Rpb5 both in vitro and in vivo (Wei et al., 2001). Tfg2 density (Figure 4C, 2) is clearly observed filling the gap between the downstream end of the polymerase clamp and subunit Rpb5. The location of this portion of Tfg2 corresponds to the location determined by fluorescence resonance energy transfer studies for the negatively charged σ 1.1 region in the bacterial RNAP/promoter open complex (Mekler et al., 2002) (σ 1.1 was not localized in the X-ray structure of the bacterial holoenzyme). It should be noted that the position of the Tfg2 density located next to the protrusion (labeled 5 in Figure 4C) does not correspond to the position of any of the σ domains identified in the X-ray structure of the bacterial holoenzyme, but it closely matches the location determined by fluorescence energy transfer studies for the σ 1.1 region when DNA is not present (Mekler et al., 2002).

Additional IIF Density

Decoration of RNAPII with recombinant Tfg1 could in principle provide information regarding the location of the largest IIF subunit, but cloning and expression of Tfg1 in bacteria could not be carried out because of apparent toxicity of the subunit to *E. coli* (Henry et al., 1994). However, information about IIF density not related to Tfg2 came from analysis of variability in the RNAPII/IIF particle images. Inspection of the RNAPII/IIF recon-

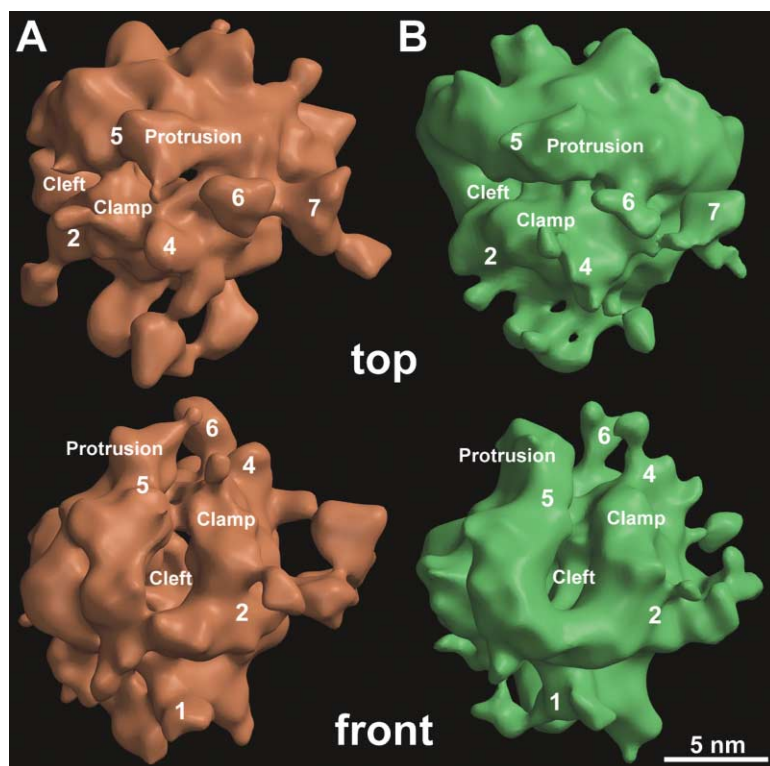


Figure 5. Cryo-EM Reconstructions of the RNAPII/Tfg2 Complex and of a Partial RNAPII/IIF Complex

(A) In order to determine the location of IIF subunit Tfg2, RNAPII was incubated with purified, recombinant Tfg2, and the resulting complex was imaged. About 10,000 images of unstained, frozen hydrated particles were refined using as a reference the cryo-EM reconstruction of the RNAPII/IIF complex. Examination of the RNAPII/Tfg2 reconstruction and comparison with the structure of the entire RNAPII/IIF complex reveals the location of Tfg2. Different domains of Tfg2 are labeled as in Figure 4C.

(B) Statistical analysis of a subset of the RNAPII/IIF particle images indicated that in ~25% of the particles in the data set the density beside the clamp (due to polymerase subunits Rpb4/Rpb7 and associated IIF density) were missing. A supervised image classification scheme was used to identify that subset of particles, which were then used to generate the reconstruction shown in the figure. This analysis, along with other considerations detailed in the text, indicates that the IIF density associated with polymerase subunits Rpb4/Rpb7 most likely corresponds to IIF subunit Tfg1. The Tfg1/Tfg2 interface would be located near the Tfg2 density labeled 2. Different IIF domains (all connected to Tfg2) are labeled as in Figure 4C.

struction indicated that a significant amount of IIF density (Figure 4C, 3) not related to Tfg2 is closely associated with the Rpb4/Rpb7 polymerase subunit heterodimer, which has a well-established tendency to dissociate from the rest of the enzyme. Therefore, we tried to establish if Rpb4/Rpb7 and associated IIF density could be missing from a fraction of the RNAPII/IIF particles. Images of particles in the “top” orientation (looking from above the active site cleft), where the presence or absence of the Rpb4/Rpb7 and associated IIF density would be most easily detectable, were subject to multivariate statistical analysis and hierarchical ascendant classification (Frank, 1996). About 30% of images of particles in the top orientation were grouped in a class where the Rpb4/Rpb7 and associated IIF density was clearly missing. Competitive supervised classification analysis (Craighead et al., 2002) was used to extend the variability determination to include all of the RNAPII/IIF particle images (multivariate statistical analysis and hierarchical ascendant classification would have performed poorly on particle projections where Rpb4/Rpb7 density overlapped significantly with density from other portions of the RNAPII structure), using as references projections of the RNAPII/IIF structure and projections of the same volume after the portion corresponding to Rpb4/Rpb7 and associated IIF density had been computationally removed.

In agreement with the statistical analysis of images of molecules in the top orientation, about 75% of the particles in the complete data set (~22,500 particle images) were identified as projections of the intact RNAPII/IIF reference, while the other 25% (~7500 particle images) lacked Rpb4/Rpb7 and the associated IIF density, and generated the RNAPII/IIF reconstruction shown in

Figure 5B. While Rpb4/Rpb7 and the IIF density associated with them are entirely absent in this reconstruction, all densities identified as corresponding to subunit Tfg2 (labeled 1, 2, 4, 5, 6, and 7 in Figures 4C and 5A) are present. The Tfg2 density filling the gap between the downstream end of the clamp and polymerase subunits Rpb5 (Figure 4C, 2) is seen to protrude toward the area occupied in the full RNAPII/IIF reconstruction by polymerase subunits Rpb4/Rpb7 and the associated IIF density.

The IIF density associated with subunits Rpb4/Rpb7 could correspond to either Tfg1 or Tfg3, and definitive experimental evidence regarding its identity was not obtained. However, several lines of evidence suggest that this density (Figure 4C, 3) most likely corresponds to Tfg1. Comparison of low-resolution structures of human RNAPII and RNAPII/IIF complexes calculated by cryo-EM (our unpublished data) reveals the presence of significant IIF density associated with the human homologs of yeast polymerase subunits Rpb4/Rpb7. This density can only correspond to Rap74, the largest human IIF subunit (homologous to Tfg1), as there is no mammalian homolog for yeast IIF subunit Tfg3. Sequence analysis indicates that Tfg1 and its mammalian counterpart, Rap74, have a similar organization. In both proteins, the N terminus mediates interaction with the second IIF subunit, Tfg2/Rap30 (Gaiser et al., 2000; Tan et al., 1995), and an unstructured central region is followed by a globular domain at the C terminus, which interacts with factor IIB and with DNA (Fang and Burton, 1996; Kamada et al., 2001). Analysis of the Tfg1 sequence (see Figure 1) reveals the presence of a large, highly charged middle section (Henry et al., 1994), which is hypersensitive to trypsin digestion and highly exposed in the RNAPII/IIF

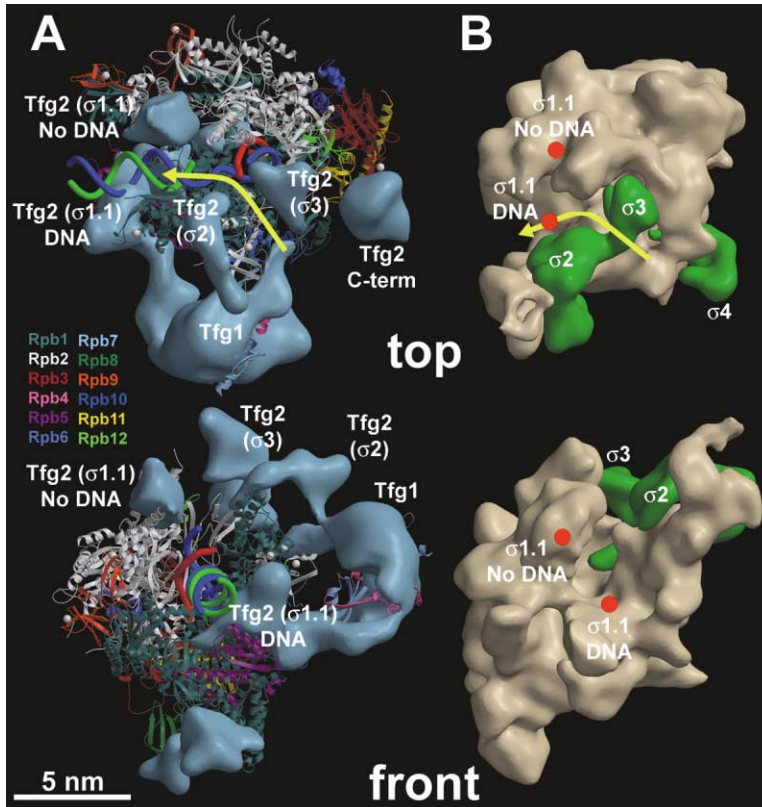


Figure 6. Location of IIF Subunits in the RNAPII/IIF Complex and Comparison with the Structure of the Bacterial Holoenzyme

(A) In the yeast RNAPII/IIF complex, subunit Tfg1 (the yeast homolog of human Rap74) is located beside the polymerase clamp, in close association with polymerase subunits Rpb4/Rpb7. Subunit Tfg2, the yeast homolog of human Rap30 and the bacterial σ factor, has an extended structure. The Tfg2 N terminus is most likely located near the polymerase feet, where it might contact Tfg1. From there, Tfg2 extends as a series of globular domains that run along the active site cleft (the connectivity between them is not apparent at the resolution of the reconstruction). The Tfg2 C terminus is located behind the polymerase wall, where it is perhaps involved in interactions with promoter DNA near the TATA box. The location of the boundary between yeast IIF subunits Tfg1 and Tfg2 near Rpb5 is suggested by the structure of the partial RNAPII/IIF complex shown in Figure 5B, but it is not indicated, as we suspect that Tfg1 may actually extend further to interact with the Tfg2 N terminus.

(B) The arrangement of the σ factor in the bacterial holoenzyme structure closely resembles the distribution of Tfg2 density in the eukaryotic RNAPII/IIF complex, with a number of globular σ domains arranged along the active site cleft of the polymerase. The arrows in the top views of the yeast and bacterial complexes indicate the position at which DNA enters the upstream end of the active

site cleft. In the bacterial complex DNA enters the cleft by passing through σ_2 and σ_3 domains. In the eukaryotic complex, two globular Tfg2 domains are placed to direct DNA to the active site. The σ and Tfg2 N termini are also similarly positioned at the downstream end of the active site cleft and seem to respond in the same manner to the presence or absence of DNA in the cleft (see text for details). The remarkable correspondence revealed by this comparison of the RNAPII/IIF and bacterial holoenzyme structures suggests a conserved mode of interaction with promoter DNA at initiation.

complex (Yong et al., 1998). The size of this disordered middle section of Tfg1 (roughly comprising amino acids 280–630 with a mass of ~ 40 kDa) matches closely the size of the IIF density associated with Rpb4/Rpb7. We suggest that the IIF density beside the polymerase clamp most likely corresponds to the central portion of subunit Tfg1. Mass spectrometry characterization of yeast multiprotein complexes detected extensive interaction of Tfg1 with Rpb7, and also with Rpb9 and Rpb3 (Gavin et al., 2002). Since the N termini of Tfg1 and Tfg2 mediate interaction of the two subunits, it is possible that part of the density clearly visible between the RNAPII feet could correspond to the N terminus of Tfg1. The preceding arguments would imply that the partially disordered IIF density located behind and below the Rpb4/Rpb7-associated IIF density, near the expected position for interaction of IIB with RNAPII (Cramer et al., 2001), might correspond to the Tfg1 C-terminal domain, which would remain disordered until interaction with DNA, IIB, and TBP constrains its position, contributing to the stabilization of the preinitiation complex (Parvin and Sharp, 1993).

We did not obtain any direct evidence regarding the location of Tfg3, the smallest yeast IIF subunit. About 30% of the mass of factor IIF was not localized in the RNAPII/IIF reconstruction, and portions of Tfg1 appear to be only partially ordered (see Figures 3 and 4). It

is possible that Tfg3 may only be ordered when other components of the transcription preinitiation complex are present and was not detected in our map of the RNAPII/IIF complex.

Structure of RNAPII/IIF and Organization of the Initiation Complex

The distribution of IIF density in the RNAPII/IIF complex, as well as the close correspondence between the arrangement of the middle IIF subunit Tfg2 and that of the σ factor in the bacterial holoenzyme structure, is illustrated in Figure 6. The structure of the RNAPII/IIF complex clearly reflects a 1:1 stoichiometry, with a single copy of IIF interacting with polymerase. Reported crosslinking between human IIF and widely separated positions on promoter DNA might be explained by the extended distribution of IIF density and not result from the presence of multiple copies of the transcription factor in the preinitiation complex as previously suggested (Robert et al., 1998). By considering the structure of the RNAPII/IIF complex and published structures of prokaryotic and eukaryotic polymerase-nucleic acid complexes, it is possible to predict the way in which promoter DNA might interact with RNAPII. The RNAPII/IIF structure can also be used as the basis for deriving a model for organization of the catalytic core of the eukaryotic transcription machinery (the minimal com-

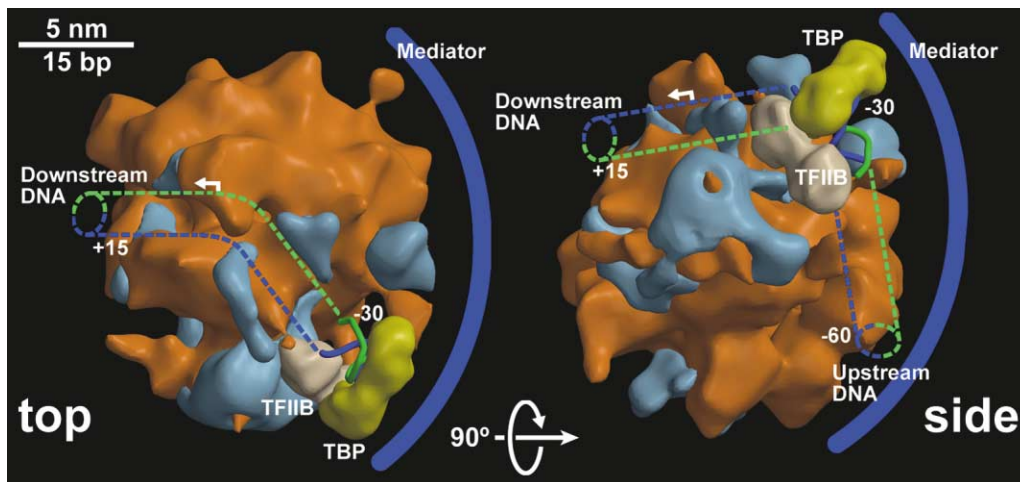


Figure 7. Model for the Structure of an RNAPII/IIF/IIB/TBP/DNA Minimal Transcription Complex and Its Interaction with Mediator

The information regarding distribution of IIF density in the RNAPII/IIF complex, along with biochemical and structural information from other studies, leads to a model for organization of the catalytic core of the eukaryotic transcription machinery, including RNAPII/IIF/IIB/TBP and DNA. The X-ray structure of the TBP/IIB/DNA complex (Tsai and Sigler, 2000) was filtered to 15 Å resolution for inclusion in the model. Information about interaction with the Mediator complex that conveys regulatory information to the basal transcription machinery comes from the structure of the yeast RNAPII/Mediator holoenzyme complex (Davis et al., 2002). In the figure, Mediator is represented by the blue shape to the right of polymerase. In the preinitiation complex promoter DNA would be threaded between RNAPII and Mediator, implying that the two must be separately recruited to a promoter. See text for a more detailed discussion.

plex capable of promoter-directed initiation from a supercoiled template) including RNAPII and factors IIF, IIB, and TBP (Parvin and Sharp, 1993; Tyree et al., 1993).

The structure of a bacterial holoenzyme/DNA complex indicates that in the prokaryotic initiation complex DNA enters the upstream end of the RNAP active site cleft passing between σ factor domains σ_2 and σ_3 (Korzheva et al., 2000; Mekler et al., 2002; Murakami et al., 2002a). We anticipate that in the eukaryotic initiation complex DNA will pass between the two corresponding Tfg2 domains at the upstream end of the active site cleft in the RNAPII/IIF complex (see Figure 6) and that, prior to formation of the transcription bubble, DNA will travel along the top of the clamp module. Cryo-EM and X-ray analysis of the structure of the complete, 12 subunit yeast RNAPII (Armache et al., 2003; Bushnell and Kornberg, 2003; Craighead et al., 2002), suggests that RNAPII does not normally adopt a conformation where the clamp domain creates an active site cleft wide enough for interaction with double-stranded DNA as previously suggested (Cramer et al., 2001; Murakami et al., 2002a). Instead, it seems most likely that DNA strand separation prompted by formation of the transcription bubble allows the template strand to melt and reach (through a narrow active site cleft [Craighead et al., 2002; Vassylyev et al., 2002]) its final position as revealed by the X-ray structure of an elongating yeast RNAPII/DNA/RNA complex (Gnatt et al., 2001).

The trajectory of upstream promoter DNA can be anticipated based on the direction from which DNA enters the active site (see Figure 6), and on the structures of the bacterial RNAP holoenzyme (Murakami et al., 2002b; Vassylyev et al., 2002) and of a σ_4 /DNA complex (Campbell et al., 2002). Upstream DNA is expected to extend along the back face of RNAPII. The topology of eukaryotic promoters and of the RNAPII active site cleft then

determine the approximate location of the TATA box and consequently of the TBP/IIB/DNA complex. This assignment agrees with the location of IIB around the RNAPII “dock” determined by preliminary structural characterization of the RNAPII/IIB complex (Cramer et al., 2001). The previous considerations allowed us to position the X-ray structure of the TBP/IIB/DNA complex (Tsai and Sigler, 2000) with respect to the cryo-EM reconstruction of the RNAPII/IIF complex.

Finally, the structure of the yeast Mediator/RNAPII holoenzyme (Davis et al., 2002) provides information about the interaction of Mediator, the multisubunit complex required to convey regulatory information to the basal transcription machinery (Bjorklund, 1996; Flanagan et al., 1991; Malik and Roeder, 2000) with the minimal transcription complex. The interaction of Mediator with RNAPII involves an area in the back face of the polymerase structure, centered on polymerase subunits Rpb3/Rpb11 (Davis et al., 2002). The corresponding area (α/α' homodimer) on the surface of the bacterial RNAP plays a critical role in bacterial regulation (Busby and Ebright, 1999). According to the model of the eukaryotic preinitiation complex proposed above, upstream promoter DNA would travel down the back face of RNAPII, along the interface between polymerase and the Mediator complex. This interface appears quite open in the low resolution structure of the yeast Mediator/RNAPII holoenzyme (Davis et al., 2002), perhaps because in the assembled initiation complex it must accommodate promoter DNA and factors IIB and TBP. The model of the initiation complex derived from these considerations, including the minimal set of factors required for promoter-directed initiation and illustrating the mode of interaction with the Mediator complex, is shown in Figure 7. A clear implication of the model (with promoter DNA traveling between RNAPII and Mediator) is that Mediator

and polymerase cannot be recruited as a holoenzyme complex but must be recruited separately. It has been observed that Mediator can be recruited to a promoter ahead of RNAPII (Cosma et al., 2001). This suggests that Mediator, along with factors IID and IIB, may provide a docking platform for the RNAPII/IIF complex. Consistent with this proposition, it has been reported that Mediator and factors IIA, IID, IIE, and IIH provide a platform for reinitiation that is stabilized by activators (Yudkovsky et al., 2000).

Experimental Procedures

TFIIF TAP Tagging and Purification of RNAPII/IIF Complex

Yeast factor IIF was affinity tagged on the C terminus of Tfg2 as described (Borggreve et al., 2001). In brief, PCR was performed with forward (5'-GCGGAGGCTGACTTGGAGATGAAATAGAAATGGAA GATGTCGTTCCATGGAAAGAGAAG) and reverse (5'-TAGGGCTC AAGAACTGCGTAAATATAAAATTAATGAAGAAAATCTACGACTCA CTATAGGG) primers using pBS1479 (Rigaut et al., 1999) as a template. The resulting PCR product was used to transform yeast strain CB010 (Mata, pep4::His3/prb1::LEU2,prc1::HISG,can1,ade2,trp1,ura3,his3,leu2-3,112), as described (Ito et al., 1983). Transformants were screened by PCR and immunoblotting of cell extracts as described (Borggreve et al., 2001), resulting in isolation of strain Tfg2-TAP (Mata, pep4::His3/prb1::LEU2,prc1::HISG,can1,ade2,trp1,ura3,his3,leu2-3,112,Tfg2::Tfg2-TAP-Trp).

Cells were grown in a fermentor to an A_{600} of 2.4 and harvested. A whole-cell lysate was generated by bead beating as described in Sayre et al. (1992), and purification of the TAP-tagged RNAPII/IIF complex was performed following a published protocol (Puig et al., 2001). In brief, the lysate was rotated with 3 ml of IgG beads (Sigma) for 2 hr at 4°C. The beads were washed with 100 ml of 25 mM Tris-acetate (pH 7.9), 200 mM potassium acetate, 5 mM EDTA, 0.1% NP40, 10% glycerol, 1 mM DTT, and protease inhibitors. This was followed by a wash with 100 ml of 25 mM Tris-acetate (pH 7.9), 500 mM ammonium sulfate, 5 mM EDTA, 0.1% NP40, 10% glycerol, 1 mM DTT, and protease inhibitors. An additional wash with 100 ml of TEV cleavage buffer (25 mM Tris-acetate [pH 7.9], 200 mM potassium acetate, 5 mM EDTA, 0.1% NP40, 1 mM DTT) was performed. After addition of 10 ml of TEV cleavage buffer and 1000 units of TEV protease (Gibco) to the washed beads, the cleavage reaction was incubated at 16°C for 2 hr. The complex was then eluted with 3 column volumes of 25 mM Tris-acetate (pH 7.9), 200 mM potassium acetate, 0.1% NP40, 10% glycerol, 1 mM magnesium acetate, 1 mM imidazole, 2 mM CaCl₂, 1 mM DTT, and protease inhibitors. The IgG eluate was rotated with 3 ml of calmodulin beads (Stratagene) for 1 hr at 4°C, poured into a column, and drained by gravity. A pure TFIIF/RNAPII was eluted with 10 ml of 25 mM Tris-acetate (pH 7.9), 200 mM potassium acetate, 0.1% NP40, 10% glycerol, 1 mM magnesium acetate, 1 mM imidazole, 2 mM EGTA, 1 mM DTT, and protease inhibitors. The complex eluted immediately after the void volume. The purification was monitored by SDS-PAGE and immunoblotting. The stoichiometry of the resulting complex was tested by quantitative SDS-PAGE using Coomassie blue and Sypro orange staining.

Tfg2 Cloning, Expression, and Purification

Tfg2 was cloned using a modified version of a previously published protocol (Henry et al., 1994). PCR was used to amplify the 1200 base pair sequence of tfg2 from yeast genomic DNA using a forward primer with sequence 5'-ATATCCATGGCTCACCATCACCACCAT-CATAGCAGTGGTTTCAGCAGG-GGCA-3' and a reverse primer 5'-ATATGGATCCCTAAACGACATCTTCCA-TTT-3'. The forward primer contained an NcoI site (underlined) and six histidine codons, while the reverse primer contained a BamHI site (underlined). The purified PCR fragment was cloned into a pET11d system (Novagen).

The pET11d plasmid containing tfg2 was transformed into *E. coli* BL21 (DE3) and grown at 30°C with shaking in 3 liters of Luria Broth (LB) containing 300 µg/ml of ampicillin (Sigma). The cells were grown to an OD of 0.6 and induced with 0.4 mM IPTG (Fisher Biotech). Two hours after the induction the cells were harvested by centrifugation at 4000 rpm for 10 min in a Beckman Allegra 25R centrifuge.

The cells were resuspended in lysis buffer (0.05 M Tris at pH 8.0, 0.3 M NaCl, 5 mM β-mercaptoethanol, 0.02 M imidazole, 10% glycerol, 0.1% Triton-X 100, and protease inhibitors). After a brief sonication, cellular debris was removed by centrifugation at 10,000 rpm for 20 min. The soluble fraction contained Tfg2.

Batch purification using Ni²⁺-nitrilotriacetic acid-agarose (Ni-NTA) resin (Qiagen) was used to purify Tfg2. In brief, 1 ml of Ni-NTA resin was added to the soluble fraction containing Tfg2 (20 ml) and incubated overnight at 4°C with gentle shaking. After incubation, the slurry was put in a column and allowed to drain by gravity. The beads were washed four times using a total of 40 ml of wash buffer (0.02 M imidazole at pH 8.0, 0.3 M NaCl, 10% glycerol, and protease inhibitors). For elution, 20 ml of elution buffer (0.2 M imidazole at pH 8.0, 0.3 M NaCl, 10% glycerol, and protease inhibitors) was used to elute the column four times. A second batch purification using essentially the same steps was also performed. The eluates were pooled (20 ml), dialyzed into the lysis buffer, and incubated for 4 hr with 400 µl of the Ni-NTA resin. After incubation, the slurry was washed four times in the column as before with 40 ml of wash buffer. After the wash step, 10 ml of elution buffer was incubated with the resin for 5 min, and collected. This process was repeated five times. Most of the pure Tfg2 used for electron microscopy analysis was obtained in the third elution (10 ml, 0.97 mg/ml) (the first two elutions contained full-length Tfg2 and a 31 kDa C-terminal degradation product and were not used for structural work).

Electron Microscopy Sample Preparation and Imaging

RNAPII/IIF aliquots (~0.5 mg protein/ml) were dialyzed against 50 mM Tris-HCl (pH 7.5) at 4°C, 150 mM KOAc, 5 mM MgCl₂, 1 mM EDTA, 5 mM DTT, and then diluted with the same buffer to a final concentration of about 50 µg of protein/ml. About 3 µl of protein solution was applied to freshly glow-discharged (in the presence of amylamine), carbon-coated Maxtaform, 300 mesh Cu/Rh grids (Ted Pella, Inc., Redding, CA). The grids were then either negatively stained with a 1% uranyl acetate solution using the "carbon layer sandwich" technique to improve particle staining (Stoffler and Stoffler-Meilicke, 1983; Tischendorf et al., 1974), or flash-frozen and preserved in amorphous ice (Dubochet et al., 1988). Samples of RNAPII/Tfg2 complex were prepared similarly, using a mixture with a 1:10 RNAPII/Tfg2 molar ratio, which was incubated at room temperature for 30 min prior to preservation of the specimen in negative stain or ice.

A CM120 transmission electron microscope (Philips/FEI) fitted with a LaB₆ filament and operating at an accelerating voltage of 100 kV was used to image specimens preserved in negative stain. Areas of the specimen were imaged at 0° and 55° on Kodak SO-163 film, under low-dose conditions and at ~0.3 µm underfocus and 60,000× ± 1% magnification to provide the data necessary to calculate a reconstruction using the random conical tilt method (Frank, 1996). Images of molecules preserved in amorphous ice were recorded only at 0°, using a Philips CM200 (FEI/Philips) microscope with a field emission source, operating at 120 kV, with underfocus values in the range of 1.37–3.32 µm and at a magnification of 66,000× ± 1%.

All micrographs were digitized on a Zeiss SCAI flat-bed scanning densitometer (ZI/Carl Zeiss) using a 7 µm sampling step size. Digitized images of particles preserved in negative stain were 3-fold pixel averaged, which resulted in a pixel size of 3.312 Å on the object scale. Digitized images of particles preserved in amorphous ice were 2-fold pixel averaged for a final pixel size of 2.055 Å on the object scale (calibrated using images of catalase crystals). All image processing was carried out using the SPIDER software package (Frank et al., 1996). Particle images were interactively selected and windowed using WEB.

Calculation of the RNAPII/IIF Reconstruction

Approximately 3800 images of RNAPII/IIF particle preserved in negative stain were selected from 0° micrographs and subject to reference-free alignment, multivariate statistical analysis (MVSA), and hierarchical ascendant classification (HAC) (Frank, 1996). A subset of approximately 800 images showing RNAPII/IIF particles in approximately the same orientation was identified, as required for application of the random conical tilt reconstruction method. The

corresponding particles from the 55° micrographs were used to calculate an initial reconstruction of the RNAPII/IIF complex. After appropriate scaling to account for the difference in pixel size between images from stained and unstained specimens, this reconstruction was used as the initial reference volume for calculation of an improved volume from images of unstained particles.

Approximately 30,000 images of unstained RNAPII/IIF particles were divided into ten groups, according to image defocus. Defocus value and contrast transfer function (CTF) parameters were estimated from the digitized micrographs. Each defocus group had a spread of roughly 200 nm, which resulted in CTF coherence to approximately 0.1 Å⁻¹. A final volume was generated from this data set by refinement through iterative reference projection matching (Penczek et al., 1994).

Calculation of the RNAPII/Tfg2 Reconstruction

Initial analysis of the RNAPII/Tfg2 data was carried out like that of the RNAPII/IIF data. A reconstruction was generated from images of negatively stained particles using the random conical tilt method. The resulting volume was similar to the RNAPII/IIF reconstruction in negative stain, except that the density beside the polymerase clamp appeared smaller. In view of the similarity to the RNAPII/IIF volume, images of unstained RNAPII/Tfg2 particles were refined using as a reference the RNAPII/IIF cryo-EM reconstruction. We reasoned that the volume reconstructed from the RNAPII/Tfg2 images should show some but not all of the features in the RNAPII/IIF structure, and we expected some of the density apparently due to IIF to disappear as the refinement progressed. As a check, a second refinement was set up using the RNAPII structure as the initial reference, expecting that some extra features would develop in the refined volume as the number of iterations increased. Comparing the progression and final result of these two refinements would make possible identification of Tfg2 in the RNAPII/IIF structure.

Multivariate Statistical Analysis and Reconstruction of a Partial RNAPII/IIF Complex

Because a significant amount of IIF density is associated with the Rpb4/Rpb7 heterodimer, which has a well-established tendency to dissociate from the rest of the polymerase, individual RNAPII/IIF particle images were checked for variability in that area. A supervised classification approach with a mix reference data set was implemented (Craighead et al., 2002), using as references the RNAPII/IIF reconstruction and the same reconstruction minus the density in the location corresponding to Rpb4/Rpb7.

Volume Rendering and Analysis

Each final reconstruction was rendered using a threshold that resulted in a volume consistent with the molecular weight of the complex in question, assuming an average protein density of 0.83 Da Å⁻³. All maps were normalized to a 0–3σ density range and manually aligned to one another. A difference between the RNAPII/IIF and RNAPII structures was calculated to facilitate interpretation of the RNAPII/IIF structure, while extra density in the RNAPII/Tfg2 volume was identified by visual inspection and comparison with the structure of the RNAPII/IIF complex.

Acknowledgments

This research was supported by NIH grants R01-60607 to F.J.A., and GM-36659 to R.D.K. F.J.A. is a Scholar of the Leukemia and Lymphoma Society of America. J.L.C. is supported by an NSF predoctoral fellowship. We thank David Bushnell for critical reading of the manuscript and insightful comments about purification of the RNAPII/IIF complex, and Richard Ebright for helpful comments regarding interpretation of some details of the RNAPII/IIF structure.

Received: March 3, 2003

Revised: September 16, 2003

Accepted: September 25, 2003

Published: October 23, 2003

References

- Armache, K.J., Kettenberger, H., and Cramer, P. (2003). Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. USA* *100*, 6964–6968.
- Bjorklund, S., Kim, Y.-J. (1996). Mediator of transcriptional regulation. *Trends Biochem. Sci.* *21*, 335–337.
- Borggreve, T., Davis, R., Bareket-Samish, A., and Kornberg, R.D. (2001). Quantitation of the RNA polymerase II transcription machinery in yeast. *J. Biol. Chem.* *276*, 47150–47153.
- Busby, S., and Ebright, R.H. (1999). Transcription activation by catabolite activator protein (CAP). *J. Mol. Biol.* *293*, 199–213.
- Bushnell, D.A., and Kornberg, R.D. (2003). Complete, 12-subunit RNA polymerase II at 4.1 Å resolution: implications for the initiation of transcription. *Proc. Natl. Acad. Sci. USA* *100*, 6969–6973.
- Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity σ subunit. *Mol. Cell* *9*, 527–539.
- Cho, H., Maldonado, E., and Reinberg, D. (1997). Affinity purification of a human RNA polymerase II complex using monoclonal antibodies against transcription factor IIF. *J. Biol. Chem.* *272*, 11495–11502.
- Conaway, J.W., and Conaway, R.C. (1990). An RNA polymerase II transcription factor shares functional properties with *Escherichia coli* sigma 70. *Science* *248*, 1550–1553.
- Cosma, M.P., Panizza, S., and Nasmyth, K. (2001). Cdk1 triggers association of RNA polymerase to cell cycle promoters only after recruitment of the mediator by SBF. *Mol. Cell* *7*, 1213–1220.
- Craighead, J., Chang, W., and Asturias, F. (2002). Structure of yeast RNA polymerase II in solution. Implications for enzyme regulation and interaction with promoter DNA. *Structure* *10*, 1117–1125.
- Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* *292*, 1863–1876.
- Davis, J.A., Takagi, Y., Kornberg, R.D., and Asturias, F.A. (2002). Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol. Cell* *10*, 409–415.
- deHaseth, P.L., Zupancic, M.L., and Record, M.T., Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.* *180*, 3019–3025.
- Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowell, A.W., and Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* *21*, 129–228.
- Fang, S.M., and Burton, Z.F. (1996). RNA polymerase II-associated protein (RAP) 74 binds transcription factor (TF) IIB and blocks TFIIB-RAP30 binding. *J. Biol. Chem.* *271*, 11703–11709.
- Flanagan, P.M., Kelleher, R.J., III, Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1991). A mediator required for activation of RNA polymerase II transcription *in vitro*. *Nature* *350*, 436–438.
- Frank, J. (1996). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies* (San Diego: Academic Press).
- Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M., and Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* *116*, 190–199.
- Gaiser, F., Tan, S., and Richmond, T.J. (2000). Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. *J. Mol. Biol.* *302*, 1119–1127.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* *415*, 141–147.
- Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* *292*, 1876–1882.
- Groft, C.M., Uljon, S.N., Wang, R., and Werner, M.H. (1998). Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. *Proc. Natl. Acad. Sci. USA* *95*, 9117–9122.

- Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* 63, 141–155.
- Henry, N.L., Campbell, A.M., Feaver, W.J., Poon, D., Weil, P.A., and Kornberg, R.D. (1994). TFIIF-TAF-RNA polymerase II connection. *Genes Dev.*, 2868–2878.
- Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153, 163–168.
- Kamada, K., De Angelis, J., Roeder, R.G., and Burley, S.K. (2001). Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc. Natl. Acad. Sci. USA* 98, 3115–3120.
- Killeen, M.T., and Greenblatt, J.F. (1992). The general transcription factor RAP30 binds to RNA polymerase II and prevents it from binding nonspecifically to DNA. *Mol. Cell. Biol.* 12, 30–37.
- Kornberg, R.D. (1999). Eukaryotic transcriptional control. *Trends Cell Biol.* 9, M46–M49.
- Korzheva, N., Mustae, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). A structural model of transcription elongation. *Science* 289, 619–625.
- Malik, S., and Roeder, R.G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* 25, 277–283.
- McCracken, S., and Greenblatt, J. (1991). Related RNA polymerase-binding regions in human RAP30/74 and *Escherichia coli* σ 70. *Science* 253, 900–902.
- Mekler, V., Kortkhonjia, E., Mukhopadhyay, J., Knight, J., Revyakin, A., Kapanidis, A.N., Niu, W., Ebright, Y.W., Levy, R., and Ebright, R.H. (2002). Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* 108, 599–614.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002a). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290.
- Murakami, K.S., Masuda, S., and Darst, S.A. (2002b). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296, 1280–1284.
- Parvin, J.D., and Sharp, P.A. (1993). DNA topology and a minimal set of basal factors for transcription by RNA polymerase II. *Cell* 73, 533–540.
- Penczek, P., Grassucci, R.A., and Frank, J. (1994). The ribosome at improved resolution: new techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy* 53, 251–270.
- Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24, 218–229.
- Radermacher, M. (1988). The three-dimensional reconstruction of single particles from random and non-random tilt series. *J. Electron Microsc. Tech.* 9, 359–394.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032.
- Robert, F., Douziech, M., Forget, D., Egly, J.M., Greenblatt, J., Burton, Z.F., and Coulombe, B. (1998). Wrapping of promoter DNA around the RNA polymerase II initiation complex induced by TFIIF. *Mol. Cell* 2, 341–351.
- Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1992). Purification and properties of *Saccharomyces cerevisiae* RNA polymerase II general initiation factor α . *J. Biol. Chem.* 267, 23383–23387.
- Sopta, M., Carthew, R.W., and Greenblatt, J. (1985). Isolation of three proteins that bind to mammalian RNA polymerase II. *J. Biol. Chem.* 260, 10353–10360.
- Sopta, M., Burton, Z.F., and Greenblatt, J. (1989). Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II. *Nature* 341, 410–414.
- Stoffler, G., and Stoffler-Meilicke, M. (1983). The ultrastructure of macromolecular complexes studied with antibodies. In *Modern Methods in Protein Chemistry*, H. Tesche, ed. (Berlin: De Gruyter), pp. 409–455.
- Tan, S., Aso, T., Conaway, R.C., and Conaway, J.W. (1994). Roles for both the RAP30 and RAP74 subunits of transcription factor IIF in transcription initiation and elongation by RNA polymerase II. *J. Biol. Chem.* 269, 25684–25691.
- Tan, S., Conaway, R.C., and Conaway, J.W. (1995). Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc. Natl. Acad. Sci. USA* 92, 6042–6046.
- Tischendorf, G.W., Zeichhardt, H., and Stoffler, G. (1974). Determination of the location of proteins L14, L17, L18, L19, L22, L23 on the surface of the 50S ribosomal subunit of *Escherichia coli* by immune electron microscopy. *Mol. Gen. Genet.* 134, 187–208.
- Tsai, F.T., and Sigler, P.B. (2000). Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.* 19, 25–36.
- Tyree, C.M., George, C.P., Lira-DeVito, L.M., Wampler, S.L., Dahmus, M.E., Zawel, L., and Kadonaga, J.T. (1993). Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II. *Genes Dev.* 7, 1254–1265.
- van Heel, M. (1987). Similarity measures between images. *Ultramicroscopy* 21, 95–100.
- Vassilyev, D.G., Sekine, S., Laptchenko, O., Lee, J., Vassilyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- Wei, W., Dorjsuren, D., Lin, Y., Qin, W., Nomura, T., Hayashi, N., and Murakami, S. (2001). Direct interaction between the subunit RAP30 of transcription factor IIF (TFIIF) and RNA polymerase subunit 5, which contributes to the association between TFIIF and RNA polymerase II. *J. Biol. Chem.* 276, 12266–12273.
- Yong, C., Mitsuyasu, H., Chun, Z., Oshiro, S., Hamasaki, N., and Kitajima, S. (1998). Structure of the human transcription factor TFIIF revealed by limited proteolysis with trypsin. *FEBS Lett.* 435, 191–194.
- Yudkovsky, N., Ranish, J.A., and Hahn, S. (2000). A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408, 225–229.
- Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811–824.