

Structure of Yeast RNA Polymerase II in Solution: Implications for Enzyme Regulation and Interaction with Promoter DNA

John L. Craighead,¹ Wei-hau Chang,^{2,4} and Francisco J. Asturias^{1,3}

¹Department of Cell Biology
The Scripps Research Institute
La Jolla, California 92037

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

Summary

An 18 Å resolution structure of the 12-subunit yeast RNA polymerase II (RNAPII) calculated from electron microscope images of single particles preserved in amorphous ice reveals the conformation of the enzyme in solution. The Rpb4/Rpb7 polymerase subunit complex was localized and found to be ideally positioned to determine the path of the nascent RNA transcript. The RNAPII structure suggests a revised mode of interaction with promoter DNA and demonstrates that regulation of RNAPII must involve structural changes that render the enzyme competent for initiation.

Introduction

RNA polymerase II (RNAPII) is the enzyme responsible for synthesis of all mRNAs in eukaryotic cells and forms the core of the transcription machinery [1, 2]. Transcription by RNAPII mediates differentiation during development and serves as the focal point for regulation in processes such as cellular metabolism and circadian rhythms [3, 4]. RNAPII along with up to six general transcription factors and promoter DNA forms the basal transcription machinery [5, 6], which is capable of promoter-directed initiation of transcription [7–9], but is unresponsive to regulation [10]. Regulatory signals are conveyed to the basal transcription machinery during initiation by Mediator [11] and other coactivator complexes [12]. General transcription factors [13], regulatory proteins [14, 15], and RNA processing enzymes [16] are also associated with RNAPII over the course of transcription.

While viral and bacterial RNA polymerases include only a few (<6) subunits [17, 18], RNAPII from *Saccharomyces cerevisiae* is composed of 12 subunits (denoted in order of decreasing molecular weight, Rpb1 to Rpb12), most likely indicative of the increased complexity required for transcription regulation in eukaryotic organisms [19, 20]. Yeast subunits Rpb1, 2, 3, 11, and 6 are homologous to the bacterial β' , β , α' , α'' , and ω subunits, respectively, and Rpb5, 6, 8, 10, and 12 are shared between the three yeast RNA polymerases [1, 2, 21], reflecting the high degree of homology between all RNA

polymerases. In higher eukaryotes RNAPII is highly conserved across species; for example, the yeast and human enzymes exhibit 53% overall sequence identity [22].

A great deal of information regarding the structure of eukaryotic RNAPII has come from studies of a 10-subunit form of yeast RNAPII ($\Delta 4/7$ RNAPII). Following biochemical purification from cells in exponential growth phase, substoichiometric amounts (<20%) of subunits Rpb4 and Rpb7 are found, while in early stationary phase the majority (>90%) of RNAPII molecules contain all 12 subunits [23]. Rpb4 and Rpb7 form a heterodimeric complex that dissociates from the 12-subunit enzyme under mild denaturing, native gel electrophoretic, or DEAE-Sephadex anion exchange chromatographic conditions [24]. RNAPII purified from an *RPB4* deletion strain also lacks Rpb7 (suggesting that Rpb4 mediates interaction of Rpb7 with the 10-subunit RNAPII [25]) and constitutes a source of functional, homogenous $\Delta 4/7$ RNAPII that is unimpaired in transcription elongation, fully active in transcription initiation when supplemented with the missing subunits [26], and suitable for crystallographic structural analysis [27]. However, structural analysis of the 12-subunit RNAPII is essential because the 10- and 12-subunit enzymes are not functionally equivalent. Unlike RNAPII, $\Delta 4/7$ RNAPII is incapable of promoter-directed initiation in a reconstituted transcription system [26]. Cells lacking the nonessential Rpb4 subunit (221 amino acids, 25 kDa) grow at moderate temperatures and show only modest changes in transcriptional activity affecting expression from a subset of inducible genes [28, 29, 25], but at high temperatures the deletion of Rpb4 results in global downregulation of transcription [30]. To date, no explanation for the essential role of Rpb7 (171 amino acids, 19 kDa) has been put forward, but the presence of two putative single-stranded nucleic acids binding domains in the structure, perhaps related to its function, has been reported [31].

Perhaps the most significant question still unresolved is the relationship between the conformation and function of RNAPII. High-resolution X-ray structures of two crystallographic forms of $\Delta 4/7$ RNAPII [32, 22] and of a transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex [33] have been reported. Comparison of the two $\Delta 4/7$ RNAPII crystallographic forms led to a description of the enzyme as being composed of four distinct modules [22]. A core module that contains the active center accounts for roughly half of the molecular weight of the polymerase. Three additional mobile modules: the jaw-lobe, shelf, and clamp, form the binding cleft that surrounds the RNAPII active site. The hinged clamp module plays a critical role in controlling access to the active site, which determines how RNAPII will interact with its promoter DNA substrate. However, the clamp conformation relevant for transcription initiation has not been identified. All $\Delta 4/7$ RNAPII structures derived from X-ray analysis of 3D crystals, including that of a transcribing $\Delta 4/7$ RNAPII/

³Correspondence: asturias@scripps.edu

⁴Present address: Department of Physics, Stanford University, Stanford, California 94305.

Key words: transcription regulation, transcription initiation, Rpb4/Rpb7, cryo-electron microscopy, RNA polymerase conformation

DNA/RNA complex, show the clamp in different “competent” conformations, in which the active site cleft is accessible. A previous $\Delta 4/7$ RNAPII reconstruction derived from electron crystallographic analysis of 2D crystals shows the clamp in a “collapsed” conformation, where access to the active site is blocked [34].

To address these unresolved questions cryo-electron microscopy (cryo-EM) was used to determine the structure of RNAPII molecules with a full complement of subunits under physiologically relevant conditions.

Results

Images of negatively stained specimens (used for development and testing of the data analysis strategy) and of flash-frozen RNAPII specimens were recorded on film and digitized. Particles were selected from the micrographs using an automated particle selection routine [35]. Following manual screening, 4,287 RNAPII particles preserved in amorphous ice, 9,349 RNAPII negatively stained particles and 7,687 negatively stained $\Delta 4/7$ RNAPII particles were selected for further processing. All data were analyzed using the SPIDER image processing software package [36]. In order to determine alignment parameters for the RNAPII EM images, a set of 83 reference projections separated by 15° was generated from the X-ray crystal structure of $\Delta 4/7$ RNAPII [32]. For both frozen-hydrated and negatively stained particles, refinement of an RNAPII reconstruction obtained by back-projection from the aligned images indicated variability in the position of the clamp module in the input image set. Therefore, an additional reference projection set was generated from the X-ray crystal structure of $\Delta 4/7$ RNAPII identical to the first one, except that prior to calculating reference projections the clamp module was computationally moved to the “collapsed” position observed in the $\Delta 4/7$ RNAPII structure derived from electron crystallographic analysis of 2D crystals [34]. Visual inspection of corresponding projections in the two reference sets suggested that the position of the clamp module was more apparent for certain molecular orientations (Figure 1A). This observation was quantified by calculating crosscorrelation coefficients between projections corresponding to the same orientation in the competent and collapsed clamp reference sets. Crosscorrelation values followed a bimodal distribution (Figure 1B). A total of 27 orientations for which crosscorrelation coefficients were above the 0.976 minimum were classified as relatively insensitive to clamp conformation. The remaining 56 orientations were classified as sensitive to clamp position. Thus, a complete reference set was generated, including all 139 (27 clamp-insensitive + 56 clamp-sensitive, competent + 56 clamp-sensitive, collapsed) projections needed to segregate particles according to their orientation and clamp conformation.

Using the complete reference projection set, particles from each RNAPII data set were classified into groups. Following classification, each group was checked to insure that it was homogeneous (same particle orientation and clamp conformation), and its members were subjected to reference-free alignment [37] in order to avoid introducing any reference-related bias. For each RNAPII data set, reference-free class averages were

separated into two subsets, depending on clamp conformation (competent versus collapsed). About 20% of RNAPII particle images (negatively stained or preserved in amorphous ice) occurred in orientations that were classified as insensitive to the position of the clamp module. The distribution of clamp conformations for the remaining particles (80%) was as follows: 75% were in a competent clamp conformation and 25% had a collapsed clamp. For $\Delta 4/7$ RNAPII particles preserved in negative stain, a higher proportion (40%) had a collapsed clamp, providing evidence that the presence of the Rpb4/Rpb7 complex favors a competent clamp conformation. Images of particles in orientations for which the clamp position could not be distinguished do not contribute significantly to determining the position of the clamp module reflected in the final 3D reconstructions and were not included in our quantitation of the clamp conformation distribution. Now that the particle classification strategy worked equally well for negatively stained and frozen specimens, further analysis of the flash-frozen RNAPII images was pursued.

Images in 0° tilt micrographs provided all projections necessary to calculate a RNAPII 3D reconstruction, as the distribution of particle orientations was nearly isotropic (Figure 1C). Back projection was used to generate an initial 3D reconstruction by combining class averages using relative angles determined by common lines [38], and candidate volumes were screened against the known structure of $\Delta 4/7$ RNAPII. The initial reconstruction was then subjected to refinement by projection matching [39], using a strategy that included contrast transfer function (CTF) corrections and simultaneously optimized translational and rotational alignment parameters for each image. Following refinement and CTF correction, amplitudes were normalized based on RNAPII X-ray amplitudes [40, 22, 41], and the final reconstruction was filtered to 18 Å resolution, as estimated by the Fourier shell correlation method (Figure 1D) with a cutoff value between the 0.5 (~ 20 Å) and 3σ (~ 16 Å) criteria [42–44]. The weighted sum of the power spectra for all the micrographs included in the RNAPII data set indicates that a narrow frequency range centered at $1/26$ Å was under sampled (Figure 1D), but reduced sampling of such a narrow frequency range has a negligible effect on a reconstruction, as previously demonstrated [45]. The final structure of RNAPII in solution is shown in Figure 2 (see animated movies of Figures 2 and 3, available as Supplementary Material with this article online).

The threshold for the cryo-EM map was carefully chosen to reflect the molecular weight of RNAPII, and the reconstruction was compared to the published $\Delta 4/7$ RNAPII X-ray structures. The conformation of RNAPII in our cryo-EM reconstruction resembles most closely the conformation of polymerase in the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex [33], as the position of the clamp module observed in solution does not match the clamp conformation detected in either form 1 or 2 of the $\Delta 4/7$ RNAPII X-ray structures [32, 22]. The X-ray structures also differ from the cryo-EM reconstruction in that they appear slightly more “compact” along the a unit cell axis, as if compressed by $\sim 5\%$ in this direction (Figure 2A). Features on the face of the RNAPII X-ray structures opposite to the Rpb4/Rpb7 subunit complex

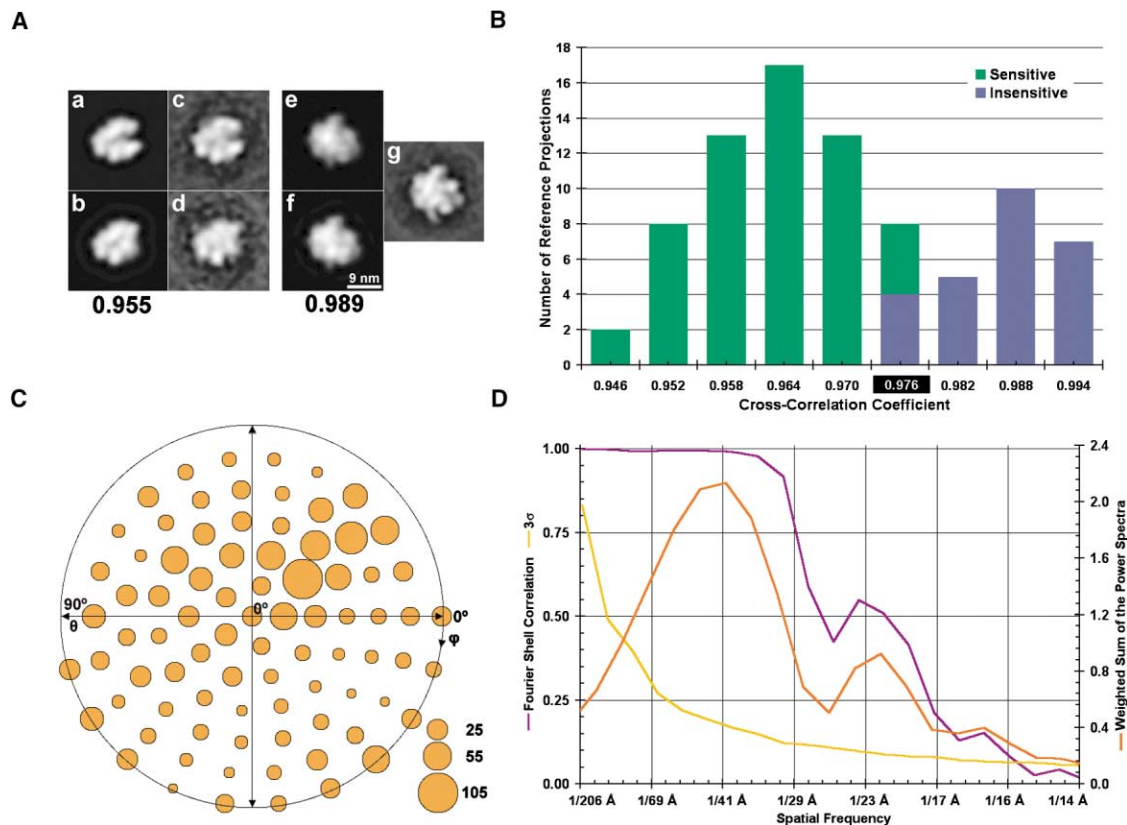


Figure 1. Data Analysis Strategy for Detection of the Clamp Conformation in RNAPII and Characterization of the Image Data Set

(A) Example of reference projections from competent and collapsed RNAPII clamp conformations (a and b) and the corresponding reference-free aligned class averages for images of $\Delta 4/7$ RNAPII preserved in negative stain (c and d). A typical crosscorrelation coefficient between corresponding views in the competent versus collapsed sets is comparatively low (0.955). For particle orientations in which the position of the clamp is not distinguishable, the competent and collapsed reference projections (e and f) are more similar (crosscorrelation coefficient 0.989), and all particles in each of such orientations were used to generate a single reference-free aligned class average (g).

(B) Bimodal distribution of crosscorrelation coefficients between corresponding projections in the competent and collapsed reference data sets. If the crosscorrelation coefficient was less than 0.976 (green), the corresponding orientation was classified as sensitive to clamp conformation. For crosscorrelation coefficients greater than 0.976 (blue), the orientations were considered insensitive to clamp conformation. This agrees well with the conclusion reached after visual inspections of the competent and collapsed reference projection sets.

(C) Distribution of particle orientations for RNAPII samples preserved in amorphous ice. A similar, isotropic distribution was observed for RNAPII and $\Delta 4/7$ RNAPII particles preserved in negative stain.

(D) Sampling of spatial frequencies in the flash-frozen RNAPII data set, and resolution measurement of the corresponding 3D reconstruction. Shown in orange is the sum of power spectra for all micrographs in the data set, weighted by the number of particles included from each micrograph. The resolution of the RNAPII reconstruction was estimated by calculating the degree of correlation between two partial reconstructions (Fourier shell correlation method). The Fourier shell correlation function (purple line) drops below the commonly accepted 0.5 threshold at ~ 20 Å and crosses the 3σ plot (yellow line) at ~ 16 Å. Therefore, the best estimate of the resolution of the reconstruction is taken to be 18 Å.

(which is roughly perpendicular to the direction of the *a* unit cell axis) correspond well with surface features of the 20 Å cryo-EM map, but appear displaced by ~ 3 –7 Å toward the center of the enzyme. At the same time, regions beneath the DNA binding cleft and active site of RNAPII that do not contain density in the cryo-EM map and are occupied with protein density in the crystal structure.

Several densities absent in the $\Delta 4/7$ RNAPII X-ray structures are observed in the cryo-EM reconstruction. They include an extension of the Rpb2 N terminus and surrounding region (34 amino acids, ~ 4 kDa, Figure 2Aa) and the Rpb2 protrusion domain (56 amino acids, ~ 6 kDa, Figure 2Ac). Residues from Rpb1 and Rpb2 (23 amino acids, ~ 3 kDa, Figure 2Ad) form a “catwalk” above the bridge helix that spans the active site cleft.

Of particular interest is the connection between the top, back portion of the clamp module and the rest of the RNAPII structure. A number (13 amino acids, ~ 2 kDa, Figure 2Ab) of presumably mobile Rpb2 residues in that area were not identified in the $\Delta 4/7$ RNAPII X-ray structures. In our reconstruction, the Rpb2 wall domain and the clamp module are fused and block the back end of the active site cleft above the saddle domain. The origin of other features in the cryo-EM map (Figure 2Af, 2Ag, and 2Ah) cannot be unequivocally determined, but there are a number of (presumably disordered) N and C termini in the corresponding areas of the RNAPII X-ray structure.

The most prominent difference between the $\Delta 4/7$ RNAPII X-ray structures and the cryo-EM RNAPII reconstruction is the large density (highlighted in white in the

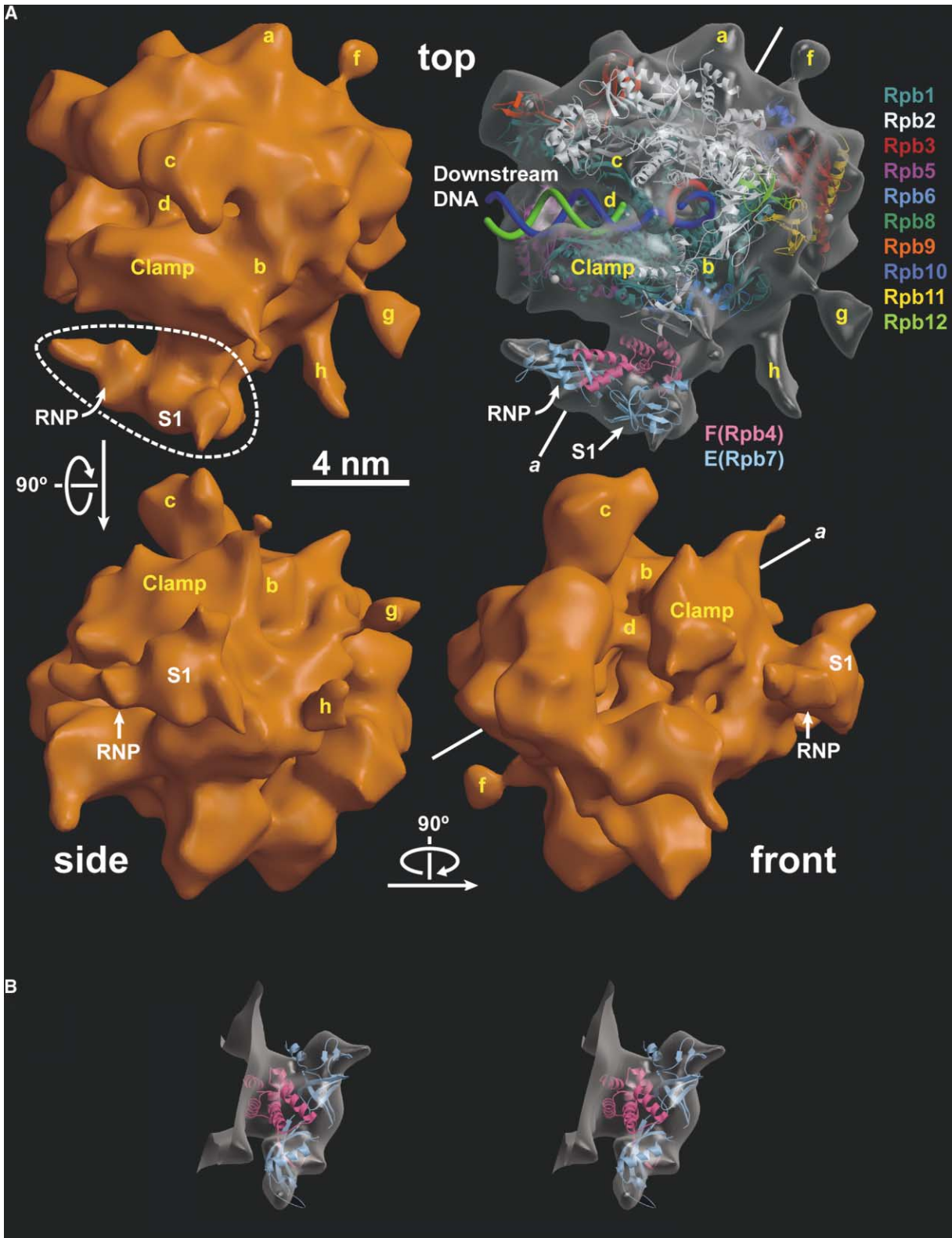


Figure 2. Surface Representation of the RNAPII Reconstruction and X-Ray Structures of the Transcribing $\Delta 4/7$ RNAPII/DNA/RNA and *M. jannaschii* E/F Complexes

(A) Three different views of the RNAPII reconstruction: from the top (looking from above the active site cleft, upper left), from the side (looking at the Rpb4/Rpb7 complex and the side of the clamp, lower left), and the front (looking into the active site cleft from the direction of the downstream DNA, lower right). The *a* unit cell axis from the X-ray structure of the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex (upper and lower right) describes the direction along which the cryo-EM reconstruction appears slightly compressed. Density not attributable to residues

upper left panel in Figure 2A) attached to the outside of the structure at the hinge between the clamp and core modules (Figure 2A). A broad connection links an extended structure reaching forward along the hinge region to the main body of RNAPII. That this density corresponds to RNAPII subunits Rpb4 and Rpb7 is confirmed by rigid-body fitting of the X-ray structure of the complex formed by *Methanococcus jannaschii* subunits F (107 amino acids, 13.4 kDa) and E (187 amino acids, 20.7 kDa), the Archaeal homolog of the eukaryotic Rpb4/Rpb7 complex, into the corresponding portion of the RNAPII map (Figure 2B). Based on this fit and the homology between the Archaeal and eukaryotic complexes, we conclude that Rpb4 connects Rpb7 to the 10-subunit polymerase in the manner apparent in the structure.

Discussion

The work reported here illustrates the power of a model-based method for deconvoluting orientational and conformational variability in the analysis of EM images of a molecule that exists in a small number of discrete conformational states. A similar concept has been applied to classify images of viral particles [46]. By separating images according to particle conformation, the cryo-EM analysis provides a description of the conformations adopted by individual RNAPII particles under physiologically relevant conditions. Although reference projections were used to separate images into homogeneous groups in the first step of our analysis, the resulting groups were checked to confirm the homogeneity of their members, and the images in each group were aligned independently of any reference, generating an accurate, unbiased representation of the structure of RNAPII in solution.

Our results indicate that the conformational variability of RNAPII in solution is primarily related to changes in the position of the mobile clamp module. In a minority (~25%) of RNAPII particles, the clamp is in a collapsed position, but in the majority of particles the position of the clamp does not impede access to the polymerase active site. At the resolution of our RNAPII structure, the back end of the active site cleft appears fully blocked (Figure 2Ab), which is in agreement with what was observed in the structure of the transcribing polymerase [33], where the top/back of the clamp module is in close proximity to the Rpb2 wall domain. If the distribution of clamp conformations in the RNAPII molecules is described using Boltzman statistics, only a small amount of energy (1.1 $k_B T$) is required to switch the conformation of the clamp between the competent and collapsed positions. However, molecules in which the position of the clamp would allow a straight, double-stranded promoter

DNA to be placed into the enzyme active site cleft were not detected. Although such conformation of RNAPII has been observed in 3D crystals [22], it must not be energetically favorable under physiologically relevant conditions. The observed compression of RNAPII by ~5% along a particular direction in the X-ray structures of the $\Delta 4/7$ form of the enzyme results in slight relative movement of several domains. This compression most likely resulted from shrinking of the 3D crystal unit cell along that same direction (that of the a unit cell axis) as a consequence of partial dehydration of the crystals required to improve diffraction quality [32, 22]). Discrepancies between the cryo-EM reconstruction and the X-ray structures related to this effect can be resolved by expanding the X-ray structure (or compressing the EM map) along the a unit cell axis.

As the conformation of RNAPII in solution resembles closely the transcribing conformation of the enzyme revealed by X-ray crystallography [33], it is informative to consider possible interactions between nucleic acid and protein density in the active site cleft. We fitted the RNAPII portion of the X-ray structure of the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex into the RNAPII reconstruction and examined the placement of the nucleic acids. The result is illustrated in Figure 3. The top and side views of the enzyme were sliced to show the region around the active site cleft (upper and lower left panels in Figure 3). The template DNA strand from the $\Delta 4/7$ RNAPII/DNA/RNA transcribing complex threads through the cleft to reach the active site, going under the thin connection ("catwalk," labeled d in Figures 2A and 3) that spans the active site cleft. The nontemplate DNA strand appears to be forced over the catwalk, which may maintain the transcription bubble during elongation. The separation of the template and nontemplate strands at this point explains why, in the absence of general transcription factors, RNAPII can only initiate transcription from a tailed DNA template. Given the conformation of RNAPII in solution, only a single-stranded overhanging tail would be capable of reaching the enzyme active site. Upstream of the active site, the DNA/RNA hybrid turns sharply upward, thereby avoiding the blockage at the back end of the active site cleft. If the path of the nascent RNA transcript is extended from the last ribose nucleotide observed in the DNA/RNA hybrid, it appears that RNA must exit the active site cleft near the top of the clamp.

Regarding the location of the Rpb4/Rpb7 complex, the match between the asymmetric shapes of the *M. jannaschii* E/F complex and the Rpb4/Rpb7 density in the RNAPII reconstruction leaves little doubt about the position and orientation of the eukaryotic complex (Figures 2A and 2B). It is clear from the structure of RNAPII that the Rpb4/Rpb7 complex interacts with the 10-sub-

in the X-ray structure is noticeable in several areas (a and f-h), but, most importantly, in and around the active site cleft (b-d). The mobile clamp module is also labeled. Density due to the Rpb4/Rpb7 complex is also marked (dotted white line, upper left), as are the locations of two single-stranded nucleic acid binding domains (S1 and RNP) on Archaeal polymerase subunit E (Rpb7 homolog). The upper right panel shows a top view of the RNAPII rendered as a transparent surface. The X-ray structures of the transcribing $\Delta 4/7$ RNAPII/DNA/RNA [33] and *M. jannaschii* E/F complexes [60] (PDB accession codes 1I6H and 1G03, respectively) have been fit interactively using the program O [61]. (B) Stereo view of the Rpb4/Rpb7 complex region of the RNAPII reconstruction and the fit of the Archaeal E/F complex X-ray structure. This figure and others were prepared with BOBSCRIPT [62] and Raster3D [63] at a threshold consistent with the molecular weight of RNAPII, calculated using $0.833 \text{ Da}/\text{\AA}^3$ as the average density of a protein.

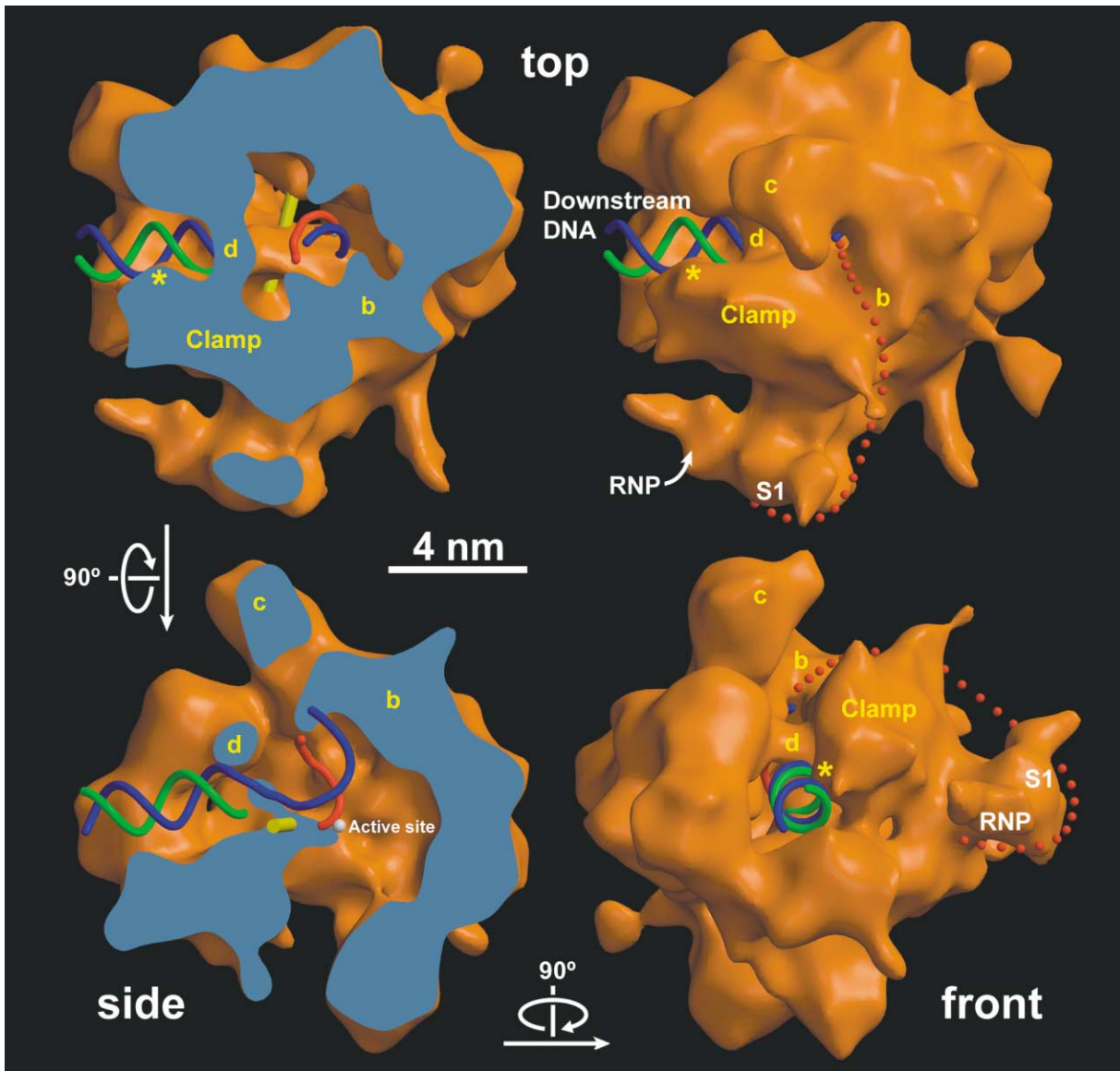


Figure 3. Structure of RNAPII around the Active Site Cleft and Its Relation to Substrate DNA and the RNA Transcript

Top (upper left) and side (lower left) views of RNAPII were “sliced” to show the conformation around the active site cleft. The position of template (blue) and nontemplate (green) DNA strands and of the nascent RNA transcript (red) are based exclusively on the docking of the $\Delta 4/7$ RNAPII portion of the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex X-ray structure to the RNAPII reconstruction. The active site metal ion (white sphere) and the Rpb1 bridge helix (yellow cylinder) proposed to be involved in RNAPII translocation [33] are shown to provide a frame of reference. Protein density in the active site cleft area is tightly packed around the nucleic acid. The downstream end of the inside face of the clamp (labeled with an asterisk) appears to be shaped to fit within the major groove of the downstream DNA. The template DNA strand continues under a thin connection that spans the active site cleft (d) and may maintain the transcription bubble during elongation. The sharp turn of the DNA/RNA hybrid upstream of the active site causes it to change direction and avoid the blockage at the back end of the active site cleft (b). The red dots in the upper and lower right panels describe a path for the nascent RNA transcript after exiting the RNAPII active site.

unit RNAPII primarily through Rpb4, as suggested by extensive genetic and biochemical evidence gathered from studies of yeast, human, and plant model systems [26, 47, 48]. Density is observed accounting for the entire mass of Rpb7, given the similarity in size between Rpb7 and its Archaeal homolog, subunit E (19 and 21 kDa, respectively). However, subunit F, the Archaeal homolog of Rpb4, is only about half the size of the eukaryotic subunit (13.4 kDa versus 25 kDa), due to a large insertion

loop and a long N-terminal extension [49]. Therefore, about 12 kDa of Rpb4 density has not been identified in our cryo-EM reconstruction. This density may contribute to some of the features identified in the active site cleft area, or may have been only partially detected because of high mobility of the corresponding residues.

The position of the Rpb4/Rpb7 complex, adjacent to the location where the nascent RNA transcript exits the active site cleft of RNAPII, suggests that two putative

single-stranded nucleic acid binding domains in Rpb7 [31] are likely involved in determining the path of the newly synthesized RNA as it is transcribed. This is illustrated in the right panels of Figure 3. The red dots mark a path from the last RNA nucleotide identified in the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex, to the location of the two single-stranded nucleic acid binding domains in Rpb7. In agreement with biochemical evidence, if the RNA were to “turn back” in the manner suggested in Figure 3, after the addition of about 30 bases to the newly synthesized RNA chain, its 5' end would be positioned near the base of the Rpb1 C-terminal domain, which has been reported to interact with capping and other RNA-processing enzymes [50–52]. However, binding of RNA to Rpb7 does not explain completely the essential role that the Rpb4/Rpb7 complex plays in initiation, as RNA transcripts prepared during the abortive phase of initiation would not be long enough to reach the single-stranded nucleic acid binding sites in Rpb7. The possibility that changes in polymerase conformation induced by the Rpb4/Rpb7 complex may be essential for initiation cannot be ruled out, but analysis of a lower-resolution structure of $\Delta 4/7$ RNAPII in solution (calculated from particles preserved in negative stain) does not support that possibility. It is conceivable that the presence of the Rpb4/Rpb7 complex may be required for appropriate interaction of RNAPII with additional components of the transcription machinery required for initiation.

Biological Implications

The structure of RNAPII in solution prevents the enzyme from interacting with promoter DNA, suggesting that control of the enzyme conformation must play an essential role in transcription regulation. Loading of promoter DNA onto RNAPII must require conformational changes, likely triggered by interaction with additional factors involved in the recruitment of polymerase to its DNA substrate. The general transcription factors TFIIB and TFIIF could induce such a conformational change. Alternatively, interaction of RNAPII with a coactivator complex could render the polymerase proficient for transcription initiation. Consistent with this idea, it has been recently reported that the Mediator coactivator complex is recruited to a promoter before RNAPII [53]. Whatever the case, as the structure of RNAPII in solution is close to the conformation of the enzyme during elongation, it would be most economical to minimize the changes that the enzyme must undergo to interact with promoter DNA. A promoter may initially interact with RNAPII by binding diagonally across the active site cleft, contacting the entrance to the cleft (Figure 4i) and the blockage at its back end (Figure 4b). This is illustrated in Figure 4, which also shows the current model, where promoter DNA is positioned in the fully open active site cleft that could only be generated by significant (and energetically unfavorable) rearrangement of RNAPII density. In agreement with previous reports [54], formation of the transcription bubble by action of TFIIF at a distance (no foreseeable change in the conformation of the active site cleft area would generate enough room

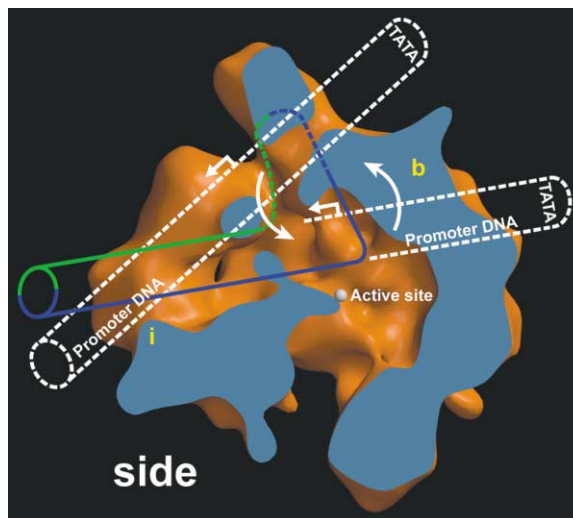


Figure 4. Two Possible Modes of Interaction between RNAPII and Promoter DNA

Positioning straight, double-stranded DNA along the base of the RNAPII active site cleft would require significant conformational changes. The changes required for interaction with DNA would be minimized if the promoter bound diagonally across the polymerase, making contact with the downstream end of the cleft (i) and the density that defines its upstream end (b). Upon interaction with TFIIF and formation of a transcription bubble, the DNA would melt, the transcription initiation site would reach the active site, and the template DNA strand would adopt the conformation detected in the transcribing $\Delta 4/7$ RNAPII/DNA/RNA complex.

to accommodate the $70 \times 125 \text{ \AA}$ IIF complex [55, 56]) would then cause the DNA to “melt” into the active site cleft.

Experimental Procedures

RNAPII was purified following a published protocol [27, 58], which included as a last step immunoaffinity chromatography using a monoclonal antibody (8WG16) directed against the Rpb1 C-terminal domain (CTD) [32]. Cells were harvested in the early stationary phase of growth, where the majority (>90%) of RNAPII molecules contain subunits Rpb4 and Rpb7. For control experiments, $\Delta 4/7$ RNAPII was purified similarly from an *RPB4* deletion yeast strain. Samples for electron microscopy were either negatively stained with a 2% uranyl acetate solution (RNAPII and $\Delta 4/7$ RNAPII) for methodological development, or flash-frozen and preserved in amorphous ice (RNAPII) for final structure determination [59].

Prior to sample preparation, RNAPII aliquots (1.26 mg protein/mL, in 60mM $(\text{NH}_4)_2\text{SO}_4$, 50 mM Tris-HCl [pH 7.5], 5 mM DTT, 10% glycerol) were dialyzed against 50 mM Tris-HCl (pH 7.5) at 4°C, 200 mM KOAc, 5 mM MgCl_2 , 1 mM EDTA, 5 mM DTT, and then diluted to a final enzyme concentration of about 50 nM. About 3 μL of protein solution were applied to freshly glow-discharged (in the presence of amylamine), carbon-coated Maxtaform, 300-mesh Cu/Rh grids (Ted Pella).

All images were collected under low-dose conditions on Kodak SO-163 film. For negatively stained specimens, images were recorded at 0.3 μm underdefocus and $60,000 \times \pm 1\%$ magnification, using a Philips CM120 (FEI/Philips) microscope, equipped with a conventional filament, operating at an accelerating voltage of 100 kV. Images of molecules preserved in amorphous ice were recorded similarly, with underfocus values in the range of 1.8–3.3 μm and at a magnification of $66,000 \times \pm 1\%$, but using a Philips CM200 (FEI/Philips) with a field emission source, operating at 120 kV. All micrographs were digitized on a Zeiss SCAI flat-bed scanning densitome-

ter (ZI/Carl Zeiss) with a step size of 7 μm . Digitized images of particles preserved in negative stain were 3-fold pixel averaged, which resulted in a pixel size of 3.312 \AA 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 \AA on the object scale.

Particles were divided into seven 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 \AA^{-1} . To determine the quality of sampling as a function of frequency, a weighted sum of the power spectra of all individual micrographs was calculated.

Supplementary Material

Supplementary Material includes animated movies of Figures 2 and 3 and can be found online at <http://images.cellpress.com/supmat/supmatin.htm>.

Acknowledgments

We thank Gavin Meredith and Yang Li for preparing RNAPII and $\Delta 4/7$ RNAPII. We thank William Young at TSRI Research Computing for access to computing resources, and Roger D. Kornberg and David A. Bushnell for discussions regarding the $\Delta 4/7$ RNAPII X-ray structures. We gratefully acknowledge members of the Asturias laboratory for discussion and critical reading of the manuscript. J.L.C. is supported by an NSF Graduate Research Fellowship. F.J.A. is a Leukemia and Lymphoma Society of America Scholar. This work was supported by NIH grant GM60607 (F.J.A.).

Received: May 15, 2002

Accepted: June 27, 2002

References

1. Young, R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.* 60, 689–715.
2. Sentenac, A., Riva, M., Thuriaux, P., Buhler, J.-M., Treich, I., Carles, C., Werner, M., Ruet, A., Huet, J., Mann, C., et al. (1992). Yeast RNA polymerase subunits and genes. In *Transcriptional Regulation*, vol. 2., S.L. McKnight and K.R. Yamamoto, eds. (Plainview, NY: Cold Spring Harbor Laboratory Press), pp. 27–54.
3. Gurney, A.L., Park, E.A., Liu, J., Giralt, M., McGrane, M.M., Patel, Y.M., Crawford, D.R., Nizielski, S.E., Savon, S., and Hanson, R.W. (1994). Metabolic regulation of gene transcription. *J. Nutr.* 124, 1533S–1539S.
4. Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. *Nat. Rev. Genet.* 2, 702–715.
5. Roeder, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci.* 21, 327–335.
6. Conaway, R.C., and Conaway, J.W. (1997). General transcription factors for RNA polymerase II. *Prog. Nucleic Acid Res. Mol. Biol.* 56, 327–346.
7. Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell* 18, 469–484.
8. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11, 1475–1489.
9. Lue, N.F., and Kornberg, R.D. (1987). Accurate initiation at RNA polymerase II promoters in extracts from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 84, 8839–8843.
10. Kornberg, R.D. (1999). Eukaryotic transcriptional control. *Trends Cell Biol.* 9, M46–M49.
11. Myers, L.C., and Kornberg, R.D. (2000). Mediator of transcriptional regulation. *Annu. Rev. Biochem.* 69, 729–749.
12. Malik, S., and Roeder, R.G. (2000). Transcriptional regulation through mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* 25, 277–283.
13. Conaway, J.W., Shilatfard, A., Dvir, A., and Conaway, R.C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* 25, 375–380.
14. Shilatfard, A. (1998). Factors regulating the transcriptional elongation activity of RNA polymerase II. *FASEB J.* 12, 1437–1446.
15. Conaway, J.W., and Conaway, R.C. (1999). Transcription elongation and human disease. *Annu. Rev. Biochem.* 68, 301–319.
16. Hirose, Y., and Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev.* 14, 1415–1429.
17. Yura, T., and Ishihama, A. (1979). Genetics of bacterial RNA-polymerases. *Annu. Rev. Genet.* 13, 59–97.
18. Ishihama, A., and Nagata, K. (1988). Viral-RNA polymerases. *CRC Crit. Rev. Biochem.* 23, 27–76.
19. Archambault, J., and Friesen, J.D. (1993). Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiol. Rev.* 57, 703–724.
20. Woychik, N.A. (1998). Fractions to functions: RNA polymerase II thirty years later. *Cold Spring Harb. Symp. Quant. Biol.* 63, 311–317.
21. Minakhin, L. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc. Natl. Acad. Sci. USA* 98, 892–897.
22. Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 \AA resolution. *Science* 292, 1863–1876.
23. Jensen, G.J., Meredith, G., Bushnell, D.A., and Kornberg, R.D. (1998). Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.* 17, 2353–2358.
24. McKune, K., Richards, K.L., Edwards, A.M., Young, R.A., and Woychik, N.A. (1993). RPB7, one of two dissociable subunits of yeast RNA polymerase II, is essential for cell viability. *Yeast* 9, 295–299.
25. Pillai, B., Sampath, V., Sharma, N., and Sadhale, P. (2001). Rpb4, a nonessential subunit of core RNA polymerase II of *Saccharomyces cerevisiae* is important for activated transcription of a subset of genes. *J. Biol. Chem.* 276, 30641–30647.
26. Edwards, A.M., Kane, C.M., Young, R.A., and Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J. Biol. Chem.* 266, 71–75.
27. Edwards, A.M., Darst, S.A., Feaver, W.J., Thompson, N.E., Burgess, R.R., and Kornberg, R.D. (1990). Purification and lipid-layer crystallization of yeast RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 87, 2122–2126.
28. Woychik, N.A., and Young, R.A. (1989). RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. *Mol. Cell. Biol.* 9, 2854–2859.
29. Rosenheck, S., and Choder, M. (1998). Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes in vitro. *J. Bacteriol.* 180, 6187–6192.
30. Miyao, T. (2001). Deletion of the RNA polymerase subunit RPB4 acts as a global, not stress-specific, shut-off switch for RNA polymerase II transcription at high temperatures. *J. Biol. Chem.* 276, 46408–46413.
31. Orlicky, S.M., Tran, P.T., Sayre, M.H., and Edwards, A.M. (2001). Dissociable Rpb4-Rpb7 subassembly of rna polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation. *J. Biol. Chem.* 276, 10097–10102.
32. Cramer, P., Bushnell, D.A., Fu, J., Gnat, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640–649.
33. Gnat, 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 \AA resolution. *Science* 292, 1876–1882.
34. Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991). Three-dimensional structure of yeast RNA polymerase II at 16 \AA resolution. *Cell* 66, 121–128.
35. Lata, K.R., Penczek, P., and Frank, J. (1995). Automatic particle picking from electron-micrographs. *Ultramicroscopy* 58, 381–391.
36. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjaj, 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.
37. Penczek, P., Radermacher, M., and Frank, J. (1992). Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy* **40**, 33–53.
 38. Penczek, P.A., Zhu, J., and Frank, J. (1996). A common-lines based method for determining orientations for $N > 3$ particle projections simultaneously. *Ultramicroscopy* **63**, 205–218.
 39. Frank, J., Penczek, P., Agrawal, R.K., Grassucci, R.A., and Heagle, A.B. (2000). Three-dimensional cryoelectron microscopy of ribosomes. *Methods Enzymol.* **317**, 276–291.
 40. Gabashvili, I.S., Agrawal, R.K., Spahn, C.M.T., Grassucci, R.A., Svergun, D.I., Frank, J., and Penczek, P. (2000). Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. *Cell* **100**, 537–549.
 41. Saad, A., Ludtke, S.J., Jakana, J., Rixon, F.J., Tsuruta, H., and Chiu, W. (2001). Fourier amplitude decay of electron cryomicroscopic images of single particles and effects on structure determination. *J. Struct. Biol.* **133**, 32–42.
 42. Saxton, W.O., and Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial-cell envelope protein. *J. Microsc.* **127**, 127–138.
 43. Unser, M., Trus, B.L., and Steven, A.C. (1987). A new resolution criterion based on spectral signal-to-noise ratios. *Ultramicroscopy* **23**, 39–51.
 44. Vanheel, M. (1987). Similarity measures between images. *Ultramicroscopy* **21**, 95–99.
 45. Boisset, N., and Mouche, F. (2000). *Sepia officinalis* hemocyanin: a refined 3D structure from field emission gun cryoelectron microscopy. *J. Mol. Biol.* **296**, 459–472.
 46. Baker, T.S., and Cheng, R.H. (1996). A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.* **116**, 120–130.
 47. Khazak, V., Estojak, J., Cho, H., Majors, J., Sonoda, G., Testa, J.R., and Golemis, E.A. (1998). Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRPB4 with its partner hsRPB7 and with pol II. *Mol. Cell. Biol.* **18**, 1935–1945.
 48. Larkin, R.M., and Guilfoyle, T.J. (1998). Two small subunits in *Arabidopsis* RNA polymerase II are related to yeast RPB4 and RPB7 and interact with one another. *J. Biol. Chem.* **273**, 5631–5637.
 49. Werner, F., Eloranta, J.J., and Weinzierl, R.O. (2000). Archaeal RNA polymerase subunits F and P are bona fide homologs of eukaryotic RPB4 and RPB12. *Nucleic Acids Res.* **28**, 4299–4305.
 50. Furuichi, Y., and Shatkin, A.J. (2000). Viral and cellular mRNA capping: past and prospects. *Adv. Virus Res.* **55**, 135–184.
 51. Shuman, S. (2001). Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog. Nucleic Acid Res. Mol. Biol.* **66**, 1–40.
 52. Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell* **108**, 501–512.
 53. 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.
 54. Kim, T.K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**, 1418–1421.
 55. Chang, W.H., and Kornberg, R.D. (2000). Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell* **102**, 609–613.
 56. Schultz, P., Fribourg, S., Poterszman, A., Mallouh, V., Moras, D., and Egly, J.M. (2000). Molecular structure of human TFIIH. *Cell* **102**, 599–607.
 57. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285–1290.
 58. Thompson, N.E., Aronson, D.B., and Burgess, R.R. (1990). Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. *J. Biol. Chem.* **265**, 7069–7077.
 59. 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.
 60. Todone, F., Brick, P., Werner, F., Weinzierl, R.O., and Onesti, S. (2001). Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol. Cell* **8**, 1137–1143.
 61. Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119.
 62. Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132–134.
 63. Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524.

Note Added in Proof

The structure of a prokaryotic RNA polymerase holoenzyme-DNA complex published after our manuscript was submitted for publication [57] fully substantiates the mode of interaction between eukaryotic RNA polymerase and DNA put forward here on the basis of the structure of yeast RNA polymerase II in solution.