

# Electron Crystal Structure of an RNA Polymerase II Transcription Elongation Complex

Claudia L. Poglitsch,<sup>††</sup> Gavin D. Meredith,<sup>†§</sup>  
Averell L. Gnat, Grant J. Jensen, Wei-hau Chang,  
Jianhua Fu, and Roger D. Kornberg\*  
Department of Structural Biology  
Stanford University School of Medicine  
Stanford, California 94305

## Summary

The structure of an actively transcribing complex, containing yeast RNA polymerase II with associated template DNA and product RNA, was determined by electron crystallography. Nucleic acid, in all likelihood the “transcription bubble” at the active center of the enzyme, occupies a previously noted 25 Å channel in the protein structure. Details are indicative of a roughly 90° bend of the DNA between upstream and downstream regions. The DNA apparently lies entirely on one face of the polymerase, rather than passing through a hole to the opposite side, as previously suggested.

## Introduction

RNA polymerase II, responsible for all mRNA synthesis in eukaryotes, is a complex of over half a million daltons composed of 12 different subunits, conserved across species from yeast to humans (Young, 1991). The two largest subunits have counterparts in the four-subunit bacterial RNA polymerase as well. Despite the size and complexity of RNA polymerase II, additional protein factors are required for initiation and for aspects of RNA chain elongation. This constellation of transcription proteins presents a formidable challenge for structural analysis.

Structures of yeast RNA polymerase II and of *E. coli* RNA polymerase have been determined from two-dimensional (2D) crystals by electron microscopy and image processing at 16–24 Å resolution (Darst et al., 1989, 1991a; Schultz et al., 1993; Polyakov et al., 1995; Jensen et al., 1997). A notable feature of the enzymes at this resolution is a 25 Å channel, appropriate in size for binding duplex DNA. A similar feature in the X-ray structures of single-subunit DNA and RNA polymerases harbors the active center of the enzymes. Biochemical studies have indicated, however, that the active center of the multisubunit RNA polymerases is not closely related to those of the single-subunit enzymes (Treich et al., 1992; Markovtsov et al., 1996; Mustaev et al., 1997),

so the role of the 25 Å channel in the multisubunit polymerases remains to be established.

An arm of protein density surrounding the 25 Å channel in yeast RNA polymerase II was suggested to retain DNA and thereby enhance the processivity of a transcribing complex (Darst et al., 1991a). Consistent with this idea, electron crystallography revealed two alternative conformations of the arm in *E. coli* RNA polymerase: an open conformation allowing access to the channel in the polymerase holoenzyme (Darst et al., 1989), the form responsible for initiation; and a closed conformation, with the arm around the channel, in core polymerase (Polyakov et al., 1995), the form of the enzyme involved in RNA chain elongation. The occurrence of an open form of yeast RNA polymerase II under conditions conducive to initiation has been demonstrated as well (Asturias et al., 1997; Asturias and Kornberg, 1999). The two conformations of the yeast and *E. coli* RNA polymerases have so far been seen, however, only in the absence of DNA. The proposal that a transcribing polymerase adopts the closed conformation remains to be tested.

Additional features of yeast RNA polymerase II have been identified by difference electron crystallography. Two small subunits, Rpb4 and Rpb7, were located in a niche in the floor of a 25 Å groove, prompting speculation as to their role in polymerase–DNA interaction (Jensen et al., 1997). Binding sites for two transcription initiation factors, TFIIB and TFIIE, were identified at a distance from the 25 Å channel and in a position relative to the arm around the channel with important implications for DNA binding (Leuther et al., 1996). The question of whether DNA does occupy the channel during transcription is crucial for understanding the roles of these and other accessory factors.

DNase I protection mapping of RNA polymerase II transcription elongation complexes has shown the association of 40–50 base pairs of DNA with the enzyme, roughly centered on the active site (Rice et al., 1993; Selby et al., 1997). This length of DNA, sufficient to wrap nearly all the way around the polymerase, far exceeds the size of the 25 Å channel. There is also evidence from nuclease protection and direct binding studies for interaction of polymerase II with about 20 residues of RNA (Rice et al., 1991; Johnson and Chamberlin, 1994; Gu et al., 1996). The problem of polymerase–nucleic acid interaction therefore goes well beyond the question of whether DNA binds in the channel.

We report here on difference electron crystallography between RNA polymerase II–nucleic acid complexes and the polymerase alone. Initial studies, performed with polymerase–DNA template complexes, were hampered by a low yield and poor quality of 2D crystals, attributed to interference by free DNA. The work was therefore extended to paused elongation complexes in which the nucleic acids are more tightly bound. The results shed light on a number of issues, including the role of the 25 Å channel, the surrounding arm, and the paths of DNA and RNA across the surface of the enzyme.

\* To whom correspondence should be addressed (e-mail: kornberg@stanford.edu).

<sup>†</sup> These authors contributed equally to this work.

<sup>‡</sup> Present address: Caliper Technologies Corp., 605 Fairchild Drive, Mountain View, California 94043-2234.

<sup>§</sup> Present address: Department of Physiology and Biophysics, University of California–Irvine, Irvine, California 92697-4560.

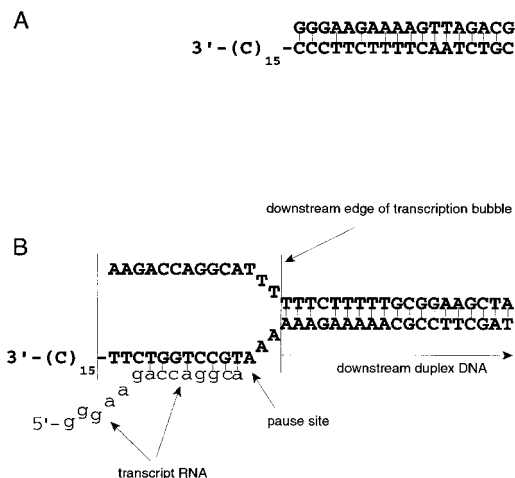


Figure 1. Transcription Templates for Structural Studies with RNA Polymerase II

(A) Tailed template used to form cocrystals with RNA polymerase II. In the presence of ribonucleotides, RNA polymerase II initiates on the 3'-tail of such templates 3–5 bases before the junction with duplex DNA (Kadesch and Chamberlin, 1982).

(B) Tailed template depicted as a paused elongation complex. The left and right vertical lines mark the upstream and downstream edges of the transcription bubble, in which template and nontemplate DNA strands are separated for about 15 bases (Uptain et al., 1997; Nudler, 1999). The pause site is indicated at the first A residue in the template strand. The transcription bubble extends 3–4 base pairs downstream of the pause site. Further downstream lie 18–19 base pairs of duplex DNA. The diagram also shows 14 residues of transcript RNA (lowercase), 8–9 residues of which are hybridized to the template strand (Uptain et al., 1997; Sidorenkov et al., 1998; Nudler, 1999). In part of this work, biotin was attached to an additional deoxycytidylate residue at the 3'-end of the nontemplate (upper) strand.

## Results

### RNA Polymerase II-Tailed Template Cocrystal

Our studies of polymerase–nucleic acid complexes began with a “tailed template,” consisting of 18 base pairs of duplex DNA with a 15-residue deoxycytidine extension or “tail” at one 3'-end (Figure 1A). Multisubunit polymerases have been shown to initiate transcription on such templates within the single-stranded region, 3–5 residues from the junction with duplex DNA, similar to the location of the polymerization site in a transcribing complex (Kadesch and Chamberlin, 1982; Gnatt et al., 1997). The tailed template is evidently recognized as half of the “transcription bubble” that normally occurs at the active center of the transcribing enzyme. Three-dimensional (3D) crystals of a yeast RNA polymerase II-tailed template complex have been reported and await X-ray structure determination (Fu et al., 1998). In order to reveal the location of the template DNA more rapidly, we formed 2D cocrystals of the polymerase-tailed template complex on positively charged lipid layers for electron microscopy. A well-characterized form of the enzyme that lacks subunits Rpb4 and Rpb7 and is capable of elongation was used because of its homogeneity and its propensity to form 2D crystals (Darst et al., 1991b). A DNA concentration of 570 nM was used to assure >90% occupancy of polymerase (190 nM),

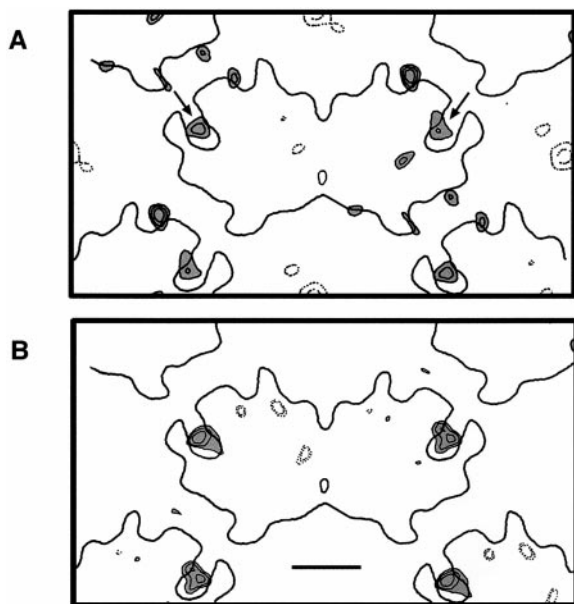
based on a  $K_D$  of 30 nM as determined from electrophoretic mobility shift assays (data not shown). The cocrystals were smaller and less abundant than crystals obtained with the polymerase alone. This was apparently due to interference by free DNA, since increasing the concentration of DNA inhibited crystal formation. The negatively charged DNA may bind to the positively charged lipid layer and prevent polymerase binding.

Despite these difficulties, six images of negatively stained, untilted crystals, taken at relatively high defocus values, were obtained, from which structure factors complete to 20 Å resolution and the corresponding projection map were derived. The crystals were evidently isomorphous with those previously obtained of the polymerase alone in the closed conformation, based on similarity of the unit cell parameters and projection maps, permitting difference analysis. A projected difference map showed peaks greater than two standard deviations ( $\sigma$ ) above the mean, located in the 25 Å channel of both molecules in the center of the unit cell (arrows in Figure 2A). Another pair of symmetrically positioned difference peaks on the periphery of the molecules near the channel was probably spurious, due to the high noise level in the map, since no corresponding density was found in the map obtained subsequently from paused elongation complexes.

### RNA Polymerase II Elongation Complex Crystal

The formation of paused elongation complexes on a minimal DNA template, with subsequent purification and 3D crystal growth has been described (Gnatt et al., 1997). The DNA consisted of 33 base pairs of duplex DNA with a 15-residue deoxycytidylate tail at one 3'-end (Figure 1B). Transcription initiated on the tail, 3–5 residues from the junction with the duplex region. In the absence of UTP, transcription paused 11 base pairs into the duplex region, at the position of the first deoxyadenylate residue in the template strand. Purified, paused complexes used for structural studies were analyzed for transcript length, by native gel electrophoresis, and for the capacity to resume RNA chain elongation upon addition of all four ribonucleoside triphosphates (data not shown). The results of all analyses were in good agreement with those reported previously (see Experimental Procedures and Gnatt et al., 1997).

Purified, paused complexes formed large (up to 2  $\mu\text{m}^2$ ) 2D crystals on positively charged lipid layers. When these crystals were negatively stained and imaged in the electron microscope, Fourier transforms of the images showed diffraction spots to about 16 Å resolution. The crystals were isomorphous with those of polymerase alone in the closed conformation, permitting difference analysis. Untilted, projected difference maps computed from subsets of images matched in defocus (nominally within 200 Å) showed a prominent peak, 2.5  $\sigma$  above the mean, in the 25 Å channel of both molecules in the unit cell (Figure 2B). In the absence of other significant peaks, the 2.5  $\sigma$  difference may be attributed to nucleic acid in the channel. The absence of any other significant differences further indicates that the structure of the polymerase in the elongation complex is the same as that of polymerase alone in the closed conformation, to the resolution of the analysis.

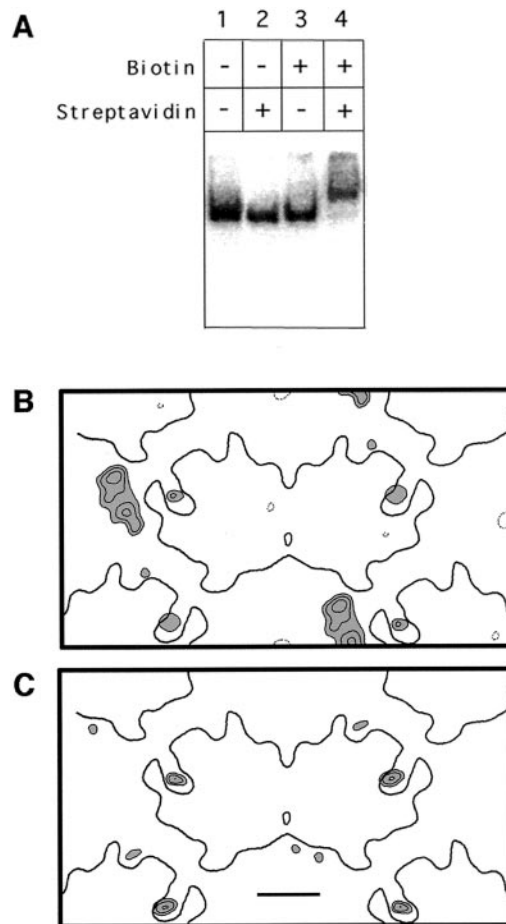


**Figure 2.** Projected Difference Maps Showing Location of Nucleic Acids in RNA Polymerase II-Tailed Template Complex and Paused Elongation Complex

(A) Projected difference map between RNA polymerase II-tailed template DNA complex and polymerase II alone. Positive (solid and gray-filled) and negative (dashed) differences are contoured in  $0.5 \sigma$  increments starting at  $\pm 2 \sigma$  from the mean and are superimposed on the outlines of the molecules in a unit cell of the RNA polymerase II crystal. The arrows point to difference peaks in the 25 Å channel. The pair of polymerase molecules in the center of the unit cell are related by a two-fold rotational symmetry axis. As the crystallographic analysis was performed without the imposition of this symmetry, differences located on either molecule arise independently. Each projection map was derived from 12 different crystalline areas. The average phase error for each map was less than  $20^\circ$ . Maps were calculated at 20 Å resolution. The scale is the same as in (B). (B) Projected difference map between paused RNA polymerase II transcription elongation complex and polymerase II alone. Differences are displayed as in (A) starting at  $\pm 2.5 \sigma$  from the mean. Each projection map was derived from 21 different crystalline areas. The average phase error for each map was less than  $20^\circ$ . Maps were calculated at 20 Å resolution. Bar, 50 Å.

#### RNA Polymerase II Elongation Complex–Streptavidin Cocrystal

To further identify the location of nucleic acids in the paused elongation complex, a DNA template was prepared with an additional deoxycytidylate residue bearing a biotin moiety at the 3'-end opposite from that bearing the 15-residue deoxycytidylate tail. The biotin was thus located at the downstream end of the template, since transcription initiates on the tail. Biotin at this location had no effect on the generation of paused elongation complexes. The biotinylated complexes bound streptavidin tightly and specifically, as shown by electrophoretic mobility shift analysis (Figure 3A). Biotinylated complexes preincubated with saturating amounts of streptavidin crystallized on lipid layers indistinguishably from nonbiotinylated complexes. Untilted, projected difference maps showed the same peak in the 25 Å channel as previously obtained for the nonbiotinylated elongation complex and, in addition, a much larger peak adjacent to the arm of polymerase density surrounding the



**Figure 3.** Streptavidin-Labeled, Paused RNA Polymerase II Elongation Complexes

(A) Gel electrophoretic mobility shift analysis of streptavidin-labeled and unlabeled paused elongation complexes. Paused complexes formed with nonbiotinylated (lanes 1 and 2) and biotinylated (lanes 3 and 4) template DNA were incubated with a two-fold molar excess of streptavidin (lane 2), stoichiometric streptavidin (lane 4), or in the absence of streptavidin (lanes 1 and 3), followed by gel electrophoresis. Complexes were radioactively labeled with the use of  $\alpha$ - $^{32}\text{P}$ -ATP and revealed by autoradiography.

(B) Projected difference map between streptavidin-labeled, paused elongation complexes and polymerase alone. Differences are displayed as in Figure 2A starting at  $\pm 2 \sigma$  from the mean. In addition to the large difference peak alongside the polymerase molecule on the left, a similar difference was seen alongside the polymerase on the right at lower contour levels (data not shown). Each projection map was derived from 22 different crystalline areas. The average phase error for each map was less than  $20^\circ$ . Maps were calculated at 20 Å resolution. The scale is the same as in (C).

(C) Control projected difference map between nonbiotinylated, paused elongation complexes crystallized in the presence of streptavidin and polymerase alone. Differences are displayed as in Figure 2A starting at  $\pm 2.5 \sigma$  from the mean. Each projection map was derived from 21 different crystalline areas. The average phase error for each projection map was less than  $20^\circ$ . Maps were calculated at 20 Å resolution. Bar, 50 Å.

channel (Figure 3B). The larger peak appeared bilobate, with a longest dimension of about 60 Å, comparable to that of tetrameric streptavidin (Darst et al., 1991c). This peak appeared alongside only one of the two polymerase molecules in the unit cell at this contour level, due

in all likelihood to preferential staining of one face of the crystal (see below). As a control, nonbiotinylated complexes were crystallized after preincubation with the same amount of streptavidin, and at similar significance levels no additional difference peak was observed (Figure 3C). This argues against nonspecific binding or other perturbation by streptavidin.

In order to reveal as much as possible of the trajectories of DNA and RNA in the elongation complex crystals, we undertook 3D reconstruction of the biotinylated complex with associated streptavidin. The difference between the elongation complex and polymerase alone was expected to be small, based on the difference density seen in projection and on the relatively small mass of the DNA and RNA component (about 5% that of the polymerase). We therefore took particular care to avoid systematic errors between the two data sets, redetermining the structure of the polymerase alongside that of the elongation complex to assure that the same microscope and imaging conditions were used and that images were assessed, digitized, and processed identically. Structure factors were trimmed in resolution and in apparent signal:noise ratio and were merged in the same manner. The polymerase data set comprised 32 untilted images and 148 tilted images, 83 of which were tilted to 45° or more. The elongation complex data set comprised 31 untilted images and 118 tilted images, 53 of which were tilted to 45° or more. Following application of a 20 Å resolution cutoff, the two sets of structure factors contained almost entirely common reflections, with a missing cone of 40°. Both data sets contained more than 85% of all possible structure factors within the 25–20 Å resolution shell, and more than 70% of all possible structure factors within the 45°–50° conical wedge. Overall, the polymerase structure factors were 98.4% complete and those for elongation complex were 92.0% complete to 20 Å resolution and 50° tilt angle (Figure 4). Image processing was performed assuming *c*12 symmetry, to average the signal from each polymerase in the unit cell, since we could not know in the event of any discrepancies between the two molecular density maps which was most likely to be valid. The average phase errors of individual measurements along lattice lines within 0.0025 Å<sup>-1</sup> in *z*\* (the direction perpendicular to the plane of the crystal) were 16.9° and 16.8° for the polymerase and the elongation complex, respectively.

The reconstruction obtained for the polymerase alone was essentially the same as that reported previously (Darst et al., 1991a), except for the presence of some density in the 25 Å channel (arrows in Figure 5A), partially blocking what was previously a 20 Å-wide hole, passing from one side of the molecule to the other. The new reconstruction is likely to be a more faithful representation of the enzyme and to show more detail than the previous one, because the data were more complete and because crystal quality was improved, due primarily to the use of the "loop transfer" procedure and a detergent treatment to remove lipid, which gave better staining (Asturias and Kornberg, 1995). Moreover, density of similar size and position was seen in the X-ray structure of polymerase alone at 6 Å resolution (Fu et al., 1999 [this issue of *Cell*]).

The polymerase and elongation complex *c*12 structure factors were aligned and scaled to each other, and

a difference map was obtained (Figure 5B) showing positive differences at 2.5  $\sigma$  above the mean. Difference density was largely confined to one face of the polymerase molecule, the face visible for the molecule shown on the left in Figure 5. A polygonal part of the difference density (labeled "d" in Figure 5B) almost entirely fills the channel. It is connected (by the density labeled "a" in Figure 5B) to the streptavidin density ("s" in Figure 5B). Two additional protruding densities ("b" and "c" in Figure 5B) emanate from the polygonal density in the directions of the previously noted "25 Å groove" and "narrow channel" (Darst et al., 1991a).

## Discussion

The difference density map obtained for the paused elongation complex is readily interpreted in terms of the configuration of nucleic acids known from biochemical studies. The active center of bacterial and eukaryotic RNA polymerases is asymmetrically located within a 15-residue single-stranded region or transcription bubble, about 3 residues from the downstream edge of the bubble (Uptain et al., 1997). In our paused complex, all 11 base pairs upstream of the pause site are melted. This is followed by another 3–4 base pairs of melted DNA to the downstream edge of the bubble, and then 18–19 base pairs of duplex DNA (Figure 1B). Of the 14–16 bases of transcript RNA, 8–9 bases will be hybridized to the template strand within the transcription bubble (Figure 1B) (Uptain et al., 1997; Sidorenkov et al., 1998; Nudler, 1999). In our difference map, streptavidin marks the downstream end of the duplex DNA. A molecular model of streptavidin and DNA downstream of the pause site (Figure 6) fits well into the densities due to streptavidin and nucleic acid ("s" and "a" in Figure 5B). The transcription bubble, at the upstream end of the 18–19 base pair duplex segment, may therefore be identified with the polygonal density in the channel ("d" in Figure 5B). The surface of the enzyme beneath the transcription bubble must include the active center cleft.

Further interpretation of nucleic acid paths on the surface of the enzyme is more speculative. A protruding density ("b" in Figure 5B) passes from the active center into the 25 Å groove and points toward the binding site of TFIIB and TATA box DNA (Leuther et al., 1996). The protruding density may be provisionally identified on this basis with upstream DNA. It is positioned above the location of subunits Rpb4 and Rpb7 in the surface of wild-type yeast RNA polymerase II (Jensen et al., 1997). Another protruding density, which passes from the active center into the narrow channel ("c" in Figure 5B) could represent the remaining component that emanates from the transcription bubble, nascent RNA. On the basis of these ideas, the path of DNA across the polymerase surface would run up the 25 Å groove to the 25 Å channel, bend roughly 90° in the channel, and then go straight for 15–20 base pairs before emerging into solution. The flexibility of single-stranded DNA within the transcription bubble would enable bending in the channel. A 90° bend is consistent with scanning force microscopy studies showing that DNA is more severely bent after the transition to elongation, with a resultant bend ranging from 90° to 100° (Rees et al., 1993;

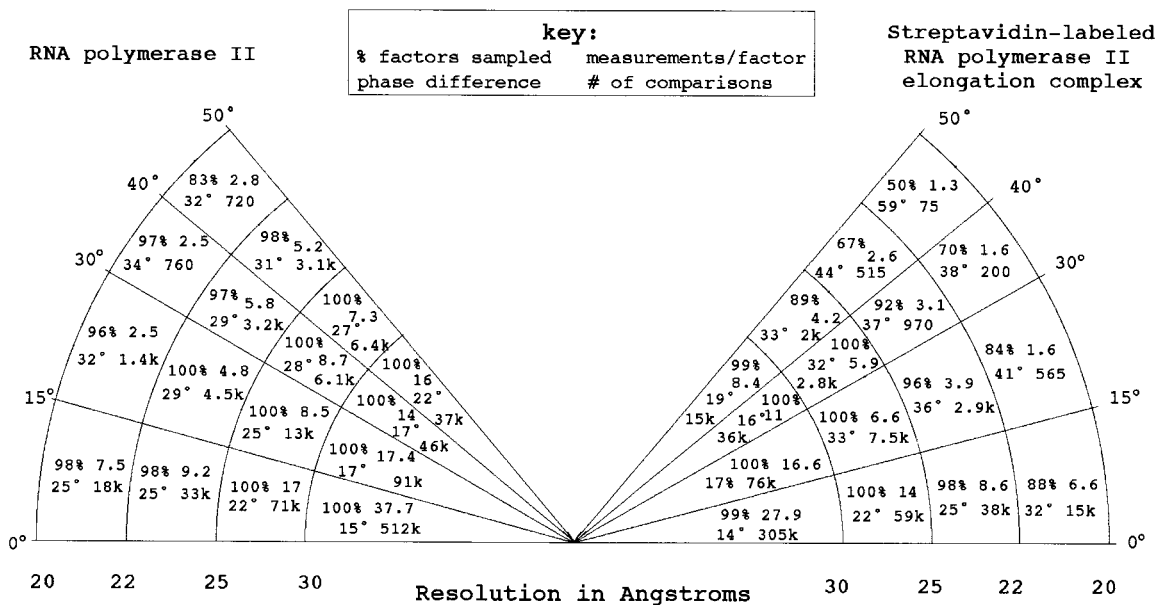


Figure 4. Reciprocal Space Coverage in 3D Reconstructions of RNA Polymerase II and Streptavidin-Labeled, Paused Elongation Complex

For reconstructions of both RNA polymerase II (left side) and paused elongation complex (right side), four statistics are presented in each of 16 regions of reciprocal space. The 16 regions are delimited by resolution and tilt-angular boundaries. In each region, the number in the upper left corner is the percentage of possible structure factors that were measured adequately enough to be sampled. Because measurements are dispersed nonuniformly along lattice lines, whether or not a structure factor is measured adequately to justify sampling can be somewhat subjective. Thus, a statistic called the redundancy of measurement is reported in the upper right corner, which is defined as the average number of measurements recorded that lie closer than half the sampling interval ( $\pm 0.00167 \text{ \AA}^{-1}$  in this case) to each sampled structure factor. Note that only measurements in sampled regions of space are included. The number in the bottom left corner is the average phase error, defined as the average phase difference between measurements within  $1/(4T)$  in  $z^*$ , where  $T$  is an estimate of the thickness of the molecule in the direction perpendicular to the plane of the crystal, and  $z^*$  is that direction in reciprocal space (here,  $T = 100 \text{ \AA}$ ). Again, only measurements in sampled regions of space are included. The number in the bottom right corner is the number of phase comparisons used to determine the average phase error.

Rippe et al., 1997). The DNA would not pass through a hole in the channel as previously imagined, but rather would be associated with the polymerase entirely on one face. This would be consistent with the evidence presented here for protein density blocking passage through the channel (Figure 5A).

Our results confirm the proposal that the closed conformation of RNA polymerase, with an arm of protein density surrounding the 25 Å channel, is the form engaged in transcription (Polyakov et al., 1995). The suggestion that the arm encircles the DNA, however, would appear inconsistent with our evidence for a DNA path confined to one face of the enzyme. The arm may nonetheless serve as a clamp, through multiple contacts with the downstream duplex DNA (labeled "a" in Figure 5B; see also Fu et al., 1999). Such a clamp might well correspond with the downstream DNA-binding site identified in *E. coli* RNA polymerase on the basis of biochemical studies (Nudler et al., 1996, 1998). In these studies, approximately 10 base pairs of downstream DNA were tightly associated with polymerase and were vital for elongation complex stability.

Our analysis provides a point of departure for interpretation of a 6 Å electron density map of RNA polymerase II from X-ray diffraction (Fu et al., 1999). The combination of the DNA path deduced from electron crystallography and the protein structure at higher resolution fits well with results of cross-linking studies, mutagenesis, and

other approaches. Detailed structure-function correlations await high resolution X-ray analysis of crystals of a paused elongation complex (Gnatt et al., 1997) and those of RNA polymerase II alone.

#### Experimental Procedures

##### Elongation Complex Preparation and Characterization

RNA polymerase II, lacking subunits Rpb4 and Rpb7, was purified from *Saccharomyces cerevisiae* by immunoaffinity chromatography and stored in 500 µg aliquots as described (Edwards et al., 1991; Gnatt et al., 1997). All tailed template DNA (Figure 1) was prepared from synthetic oligonucleotides (Ana-Gen, Palo Alto, CA) that were purified by denaturing polyacrylamide gel electrophoresis as described (Gnatt et al., 1997). The biotin group, on the biotinylated version of the template used to form paused elongation complexes (Figure 1B), was attached to an additional nucleotide on the 3'-end of the nontemplate strand by a 10-carbon linker. Oligonucleotides were further purified and desalted with a Sep-Pak C18 column (Waters, Milford, MA), dried, and resuspended in water. Oligonucleotide concentrations were determined by optical density measurements. To determine precise relative concentrations,  $\gamma$ -[ $^{32}\text{P}$ ] end-labeled long strands were titrated against unlabeled shorter strands and monitored by nondenaturing 12.5% polyacrylamide gel electrophoresis. Hybridization was performed in 50 mM Tris-Cl (pH 8.0), 1 mM EDTA, and 30 mM ammonium sulfate by heating to 75°C for 10 min followed by cooling to 20°C. Streptavidin (Boehringer Mannheim, Indianapolis, IN) was resuspended to 15 mg/ml and stored in 50 mM Tris-Cl (pH 7.5), 60 mM ammonium sulfate, 10 mM DTT, 6 mM  $\text{MgSO}_4$ , 10% (v/v) glycerol at  $-70^\circ\text{C}$ . Small aliquots were thawed at 4°C and diluted immediately before use.

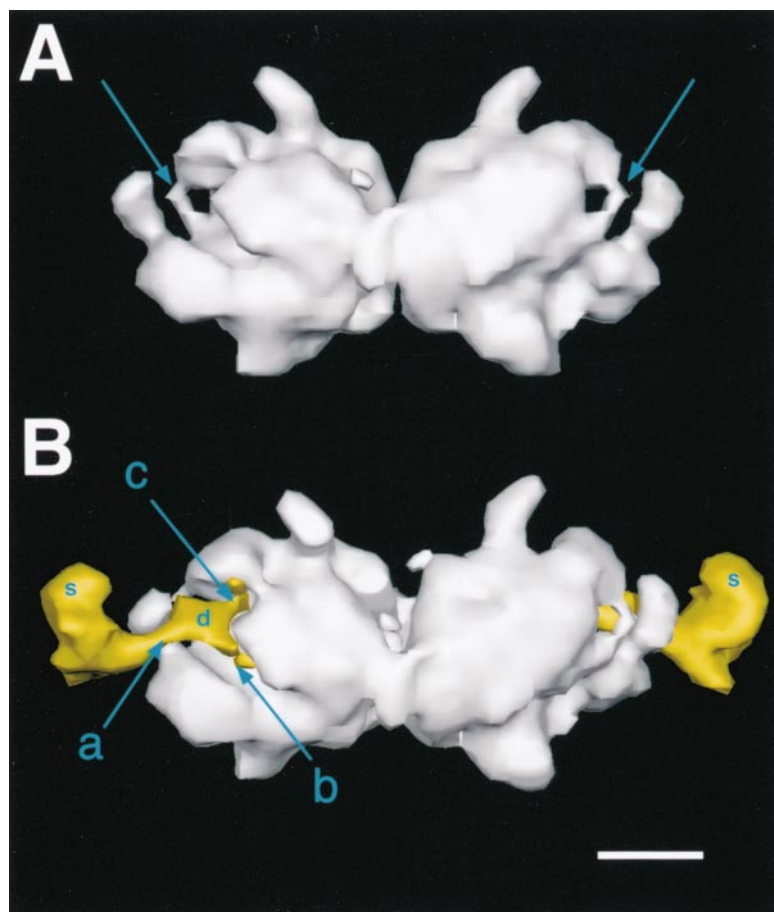


Figure 5. Three-Dimensional Reconstructions of RNA Polymerase II and of Streptavidin-Labeled, Paused RNA Polymerase II Elongation Complexes

(A) The surface envelope of a newly determined 3D reconstruction of yeast RNA polymerase II at 20 Å contoured at 0.9  $\sigma$  above the mean. A two-fold axis in the plane of the figure relates the two molecules shown, which therefore present two opposite faces of RNA polymerase II. The same channels and grooves observed in earlier 3D reconstructions of RNA polymerase II (Darst et al., 1991a; Jensen et al., 1997) are seen here. The turquoise arrows point to an additional density, not seen in other reconstructions, which partially blocks what was previously thought to be a hole. Bar, 50 Å.

(B) Three-dimensional positive differences between streptavidin-labeled, paused elongation complexes and RNA polymerase II, superimposed on the 3D reconstruction of RNA polymerase II. Differences 2.5  $\sigma$  above the mean are shown in yellow, and the surface of RNA polymerase II is rendered as in (A). Isolated, small difference densities less than about 15 Å in largest dimension could only be due to noise and were omitted for clarity. The molecular surfaces are slightly rotated with respect to the view shown in (A) to show the differences more clearly. Note that the main, continuous difference occurs entirely on one face (the face shown on the left) of the molecule. A polygonal portion of this density, labeled "d," appears in the 25 Å cleft, directly connected by density "a" to that corresponding to streptavidin, labeled "s." Two other protrusions of density, labeled "b" and "c," emanate from the cleft in the direction of the "25 Å groove" and the "narrow channel" identified earlier (Darst et al., 1991a).

For each of three preparations of elongation complexes, 0.5 mg of yeast RNA polymerase II in 50% saturated ammonium sulfate was thawed on ice and centrifuged at 13,000 rpm at 4°C for 30 min. The supernatant was removed with a glass pipette and the protein was resuspended in cold  $T_0$  buffer (50 mM Tris-Cl [pH 7.5], 10 mM DTT, 6 mM  $MgSO_4$ , 10 mM  $ZnSO_4$ , 10% glycerol) to a conductivity that matched that of  $T_{60}$  ( $T_0$  containing 60 mM ammonium sulfate). The protein concentration was determined (Bradford, 1976) with bovine serum albumin (Sigma; St Louis, MO) as the standard. Template DNA, in 30 mM ammonium sulfate, was added in 1.5-fold molar excess over RNA polymerase II. After 5 min at 20°C, HPLC-purified ATP, GTP, and CTP (Amersham Pharmacia, Piscataway, NJ) were added to final concentrations of 60  $\mu M$ , 1 mM, and 1 mM respectively. In order to radioactively label the nascent transcript, 50–100  $\mu Ci$  of  $\alpha$ -[ $^{32}P$ ]-ATP was added from a 25 mM, 10  $\mu Ci/\mu l$  stock solution (Amersham Pharmacia). The reaction was allowed to proceed for 20–30 min at 28°C. Reaction volumes were between 0.5 and 1 ml. Following incubation, the reaction mixture was loaded on a 1 ml disposable FPLC heparin-Sepharose column (Biotech HiTrap-Sepharose; Amersham Pharmacia) previously washed with 10 column volumes of  $T_{500}$  ( $T_0$  containing 500 mM ammonium sulfate) and subsequently equilibrated with 10 column volumes of  $T_{60}$ . The elongation complexes were passed through the column in  $T_{60}$  buffer under manual pressure applied with a 10 ml syringe. Fractions of 300–500  $\mu l$  were collected. The elongation complexes began to elute in the second or third fraction and were recovered in a total volume of about 1.5 ml. The fractions were pooled and subjected to 5–8 rounds of concentration in a microcon-100 (Amicon, Beverly, MA) at 4500 rpm in a table top centrifuge at 4°C under argon for 20 min and dilution in 50 mM Tris-Cl (pH 7.5) at 4°C, 60 mM ammonium sulfate,

10 mM DTT, 6 mM  $MgSO_4$ , 10% (v/v) glycerol. The final round of centrifugation served to concentrate the complexes to 0.5–2 mg/ml. Final yields ranged from 25%–50% of total polymerase recovered in elongation complexes.

Elongation complex transcript lengths were determined by extraction and denaturing gel electrophoresis of RNA as described (Gnatt et al., 1997). Elongation complexes were tested for the capacity to resume transcription by incubation with all four ribonucleotides (1 mM) for 30 min at 28°C, followed by extraction of RNA and denaturing gel electrophoresis as described (Gnatt et al., 1997). Transcript lengths, whether from paused complexes or from those allowed to resume transcription, and regardless of the presence or absence of a biotin end label, were in agreement with the lengths determined previously (Gnatt et al., 1997). The efficiency of  $\alpha$ -[ $^{32}P$ ]-adenylate incorporation ranged from 45%–100% as determined by scintillation counting. Elongation complexes were >95% pure and homogeneous as judged by their characteristic mobility in a nondenaturing gel (Gnatt et al., 1997). To monitor streptavidin binding by gel electrophoretic mobility shift, elongation complexes prepared with biotinylated and nonbiotinylated template DNA were incubated with an equimolar concentration of streptavidin for 5 min at 20°C. The complexes were then adjusted to 7% (v/v) glycerol and 0.03% bromophenol blue and electrophoresed in a 16 cm vertical 5% polyacrylamide gel in 90 mM Tris-borate (pH 8.3), 2 mM EDTA at 20°C until the bromophenol blue reached the bottom of the gel.

#### Lipid Layer Crystallography

Two-dimensional crystals of RNA polymerase II alone, in binary complexes with DNA, and in elongation complexes were formed on positively charged lipids by the lipid layer crystallization method. In

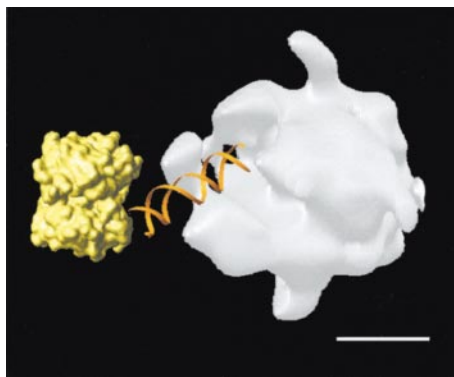


Figure 6. Modeling of Streptavidin and Downstream DNA in the Streptavidin-Labeled, Paused Elongation Complex

A 22-base pair B-DNA model (orange) was positioned on the surface of RNA polymerase II in the location of difference density attributed to downstream duplex DNA ("a" in Figure 5B). This amount of duplex DNA was chosen to approximate the length downstream from the pause site (3–4 unpaired bases followed by 18–19 base pairs of duplex DNA). The atomic coordinates of a streptavidin–biotin complex (PDB ID code 1stp, shown in yellow) were attached to the DNA model through an extended, 10-carbon chain, corresponding to the linker present in our biotinylated templates. The biotin group and 10-carbon linker are buried within the streptavidin molecule and are not visible in this view. The density due to streptavidin in Figure 5B differs in appearance from the streptavidin model shown here because the former density represents an average of two unequally imaged, symmetry-related molecules (Figure 3B). Bar, 50 Å.

all cases, the concentration of RNA polymerase II was 100  $\mu\text{g/ml}$  ( $\sim 190$  nM) in 50 mM Tris-Cl (pH 7.5), 60 mM ammonium sulfate, 6 mM  $\text{MgSO}_4$ , 10 mM DTT, and no more than 2% glycerol. Following dilution to this final concentration, the protein was chilled at 4°C for 45 min. For binary complexes, the polymerase was incubated with a 3-fold molar excess of DNA for 30 min at 20°C, prior to chilling. For experiments with streptavidin, an 8-fold molar excess was added to elongation complexes, followed by incubation for 45 min at 4°C. Droplets of protein solution, 16–18  $\mu\text{l}$ , were placed in 3.5 mm diameter, 0.5 mm deep nylon wells at 4°C (Asturias and Kornberg, 1995). The surface of the protein solution in the well was overlaid with a lipid mixture containing 0.45 mg/ml L- $\alpha$ -phosphatidylcholine and 0.05 mg/ml stearylamine (both from Avanti Polar Lipids, Alabaster, AL) in 1:1 (v/v) chloroform/hexane. After incubation for 9–12 hr at 4°C in a water-saturated argon atmosphere, crystals were transferred with a Pt/Pd loop to glow-discharged (45 s in water vapor at 0.5–1 mbar) carbon-coated 400 mesh copper-rhodium electron microscope grids (Ernest E. Fullam, Latham, NY) (Asturias and Kornberg, 1995). Thirty seconds after deposition on the grids, crystals were washed with 3  $\mu\text{l}$  of 0.05% (v/v) Tween-20 (Sigma), lightly blotted from the side with Whatman #1 filter paper (Whatman, Kent, England), washed with 3  $\mu\text{l}$  of water, blotted again, and finally stained with 3  $\mu\text{l}$  of 1% (w/v) uranyl acetate for 45 s before blotting to dryness. Care was taken never to allow the samples to dry completely until after staining.

#### Electron Microscopy, Image Processing, and Structure Calculation

Electron micrographs were recorded on SO163 film (Eastman Kodak, Rochester, NY) at a calibrated magnification of 34,300 $\times$  at defocus values of 300–700 nm (binary complex images at 640–830 nm) in low-dose mode using a CM12 transmission EM (Philips Electron Optics, Eindhoven, The Netherlands) operating at 100 kV with a LaB<sub>6</sub> filament. For images of untilted crystals, the defocus value was estimated from measurements of the positions of contrast transfer function minima in 2D Fourier transforms calculated with the program Digitalmicrograph (Gatan, Pleasanton, CA). The uncertainty in the determination of defocus values was about  $\pm 50$  nm.

For three-dimensional reconstructions, individual crystals were imaged at 4–6 (including 0° tilt) different tilt angles (up to 60°, the limit of the goniometer). Suitable micrographs were selected by optical diffractometry and digitized with a Leaf Scanner (Leaf Systems, Southborough, MA) with a 10  $\mu\text{m}$  square pixel size. Images were processed in the two-sided plane group p1 ( $a = 220 \pm 2$  Å,  $b = 220 \pm 2$  Å,  $\gamma = 121^\circ \pm 1^\circ$ ) or c2 ( $a = 384 \pm 4$  Å,  $b = 216 \pm 2$  Å,  $\gamma = 90^\circ \pm 1^\circ$ ) as described (Darst et al., 1991b; Meredith et al., 1996) by standard methods (Amos et al., 1982; Henderson et al., 1986; Crowther et al., 1996). The same unit cell parameters were determined from micrographs of frozen hydrated 2D crystals of yeast RNA polymerase II (data not shown) obtained on a Philips CM12 with a LaB<sub>6</sub> filament but otherwise as described (Asturias et al., 1998). (It proved impractical to collect sufficient data from frozen, hydrated samples for the analyses presented here). After processing, structure factors derived from individual images were trimmed to a resolution within the first zero of the contrast transfer function, or to a resolution where less than 25% of the expected structure factors were recovered with an apparent signal:noise ratio  $>1.4$  ( $\text{IQ} < 6$ ) (Henderson et al., 1986). Analysis of data from untilted crystals showed that structure factors meeting these criteria consistently merge with overall phase residuals of less than 60°, whereas below the level set by these criteria, structure factors tend to merge with greater phase errors. Smooth curves were fit to the accumulated structure factors to produce reciprocal space lattice lines. These curves were sampled at a spacing corresponding to a unit cell 300 Å thick to obtain 3D Fourier terms. Reconstructions were visualized with the XtalView software package (McRee, 1999), and surface envelopes were generated with the SYNU software package (Figure 5) (Hessler et al., 1992) or the program RIBBONS (Figure 6) (Carson, 1997). Rapid image processing was facilitated in part by automated procedures (Jensen et al., 1997). Manipulation of the DNA and streptavidin–biotin complex coordinates in relation to the RNA polymerase II map to generate Figure 6 was performed with the program O (Jones et al., 1991).

#### Difference Calculations

All crystals obtained were isomorphous, allowing direct difference calculation after appropriate trimming, scaling, and alignment. Two-dimensional difference maps were calculated between data sets where pairs of images between sets were matched for mean defocus to within 20 nm. Projection maps were scaled in real space, minimizing the cumulative difference between corresponding pixels by linear regression. Control difference projection maps were calculated by splitting the data for each average map into two groups and calculating the difference. The resultant control difference maps were featureless with few, small and asymmetrically distributed differences at  $\pm 2.5 \sigma$  (data not shown). For the 3D difference calculation, each set of structure factors was almost complete to 20 Å with a missing cone of data of 40° (Figure 4). All structure factors beyond these limits were discarded. Amplitude scaling was done in the following resolution bins:  $\infty$ –50 Å, 50–40 Å, 40–34 Å, 34–30 Å, 30–27 Å, and 1 Å bins thereafter. Reconstructions were aligned as described (Jensen et al., 1997). Three-dimensional differences were calculated directly in real space or equivalently in reciprocal space by vector subtraction.

#### Acknowledgments

We thank Lynne Mercer for advice on electron microscopy, Yang Li for polymerase used in a portion of the work, and members of our research group for helpful advice and discussions, especially Kerstin Leuther, David Bushnell, and Francisco Asturias. C. L. P. was a Damon Runyon-Walter Winchell Postdoctoral Fellow (DRG-1209) and a Katherine McCormick Scholar. G. D. M. was a National Science Foundation Fellow and a National Institutes of Health (NIH) Molecular Biophysics Training Program grant (2T32GM082494) trainee. A. G. was sponsored by the USAMRMC, DOD, Breast Cancer Initiative, DAMD17-97-7099. G. J. J. was supported by a Medical Scientist Training Program grant (GM07365) provided by the National Institute of General Medical Sciences at the NIH. This research was supported by NIH grant AI21144 to R. D. K.

Received July 7, 1999; revised August 10, 1999.

## References

- Amos, L.A., Henderson, R., and Unwin, P.N.T. (1982). Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Prog. Biophys. Mol. Biol.* **39**, 183–231.
- Asturias, F.J., and Kornberg, R.D. (1995). A novel method for transfer of two-dimensional crystals from the air/water interface to specimen grids. *J. Struct. Biol.* **114**, 60–66.
- Asturias, F.J., and Kornberg, R.D. (1999). Protein crystallization on lipid layers and structure determination of the RNA polymerase II transcription initiation complex. *J. Biol. Chem.* **274**, 6813–6816.
- Asturias, F.J., Meredith, G.D., Poglitsch, C.L., and Kornberg, R.D. (1997). Two conformations of RNA polymerase II revealed by electron crystallography. *J. Mol. Biol.* **272**, 536–540.
- Asturias, F.J., Chang, W., Li, Y., and Kornberg, R.D. (1998). Electron crystallography of yeast RNA polymerase II preserved in vitreous ice. *Ultramicroscopy* **70**, 133–143.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Chem.* **72**, 248–254.
- Carson, M. (1997). RIBBONS. *Methods Enzymol.* **277**, 493–505.
- Crowther, R.A., Henderson, R., and Smith, J.R. (1996). MRC image processing programs. *J. Struct. Biol.* **116**, 9–16.
- Darst, S.A., Kubalek, E.W., and Kornberg, R.D. (1989). Three-dimensional structure of *Escherichia coli* RNA polymerase holoenzyme determined by electron crystallography. *Nature* **340**, 730–732.
- Darst, S.A., Edwards, A.M., Kubalek, E.W., and Kornberg, R.D. (1991a). Three-dimensional structure of yeast RNA polymerase II at 16 angstroms resolution. *Cell* **66**, 121–128.
- Darst, S.A., Kubalek, E.W., Edwards, A.M., and Kornberg, R.D. (1991b). Two-dimensional and epitaxial crystallization of a mutant form of yeast RNA polymerase II. *J. Mol. Biol.* **221**, 347–357.
- Darst, S.A., Ahlers, M., Meller, P.H., Kubalek, E.W., Blankenburg, R., Ribi, H.O., Ringsdorf, H., and Kornberg, R.D. (1991c). Two-dimensional crystals of streptavidin on biotinylated lipid layers and their interactions with biotinylated macromolecules. *Biophys. J.* **59**, 387–396.
- 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.
- Fu, J., Gerstein, M., David, P.R., Gnat, A.L., Bushnell, D.A., Edwards, A.M., and Kornberg, R.D. (1998). Repeated tertiary fold of RNA polymerase II and implications for DNA binding. *J. Mol. Biol.* **280**, 317–322.
- Fu, J., Gnat, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R., and Kornberg, R.D. (1999). Yeast RNA polymerase II at 6 Å resolution. *Cell* **98**, this issue, 799–810.
- Gnat, A., Fu, J., and Kornberg, R.D. (1997). Formation and crystallization of yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* **272**, 30799–30805.
- Gu, W., Wind, M., and Reines, D. (1996). Increased accommodation of nascent RNA in a product site on RNA polymerase II during arrest. *Proc. Natl. Acad. Sci. USA* **93**, 6935–6940.
- Henderson, R., Baldwin, J.M., Downing, K.H., Lepault, J., and Zemlin, F. (1986). Structure of purple membrane from Halobacterium halobium: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy* **19**, 147–178.
- Hessler, D., Young, S.J., Carragher, B.O., Martone, M., Lamont, S., Whittaker, M., Milligan, R.A., Masliah, E., Hinshaw, J.E., and Elismann, M.H. (1992). Programs for visualization in 3-dimensional microscopy. *Neuroimage* **1**, 55–67.
- Jensen, G.J., Meredith, G., Bushnell, D.A., and Kornberg, R.D. (1997). Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.* **17**, 2353–2358.
- Johnson, T.L., and Chamberlin, M.J. (1994). Complexes of yeast RNA polymerase II and RNA are substrates for TFIIIS-induced RNA cleavage. *Cell* **77**, 217–224.
- 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.
- Kadesch, T.R., and Chamberlin, M.J. (1982). Studies of in vitro transcription by calf thymus RNA polymerase II using a novel duplex DNA template. *J. Biol. Chem.* **257**, 5286–5295.
- Leuther, K.K., Bushnell, D.A., and Kornberg, R.D. (1996). Two-dimensional crystallography of TFIIIB- and TFIIE-RNA polymerase II complexes: implications for start site selection and initiation complex formation. *Cell* **85**, 773–779.
- Markovtsov, V., Mustaev, A., and Goldfarb, A. (1996). Protein-RNA interactions in the active center of the transcription elongation complex. *Proc. Natl. Acad. Sci. USA* **93**, 3221–3226.
- McRee, D.E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
- Meredith, G.D., Chang, W.-H., Li, Y., Bushnell, D.A., Darst, S.A., and Kornberg, R.D. (1996). The C-terminal domain revealed in the structure of RNA polymerase II. *J. Mol. Biol.* **258**, 413–419.
- Mustaev, A., Kozlov, M., Markovtsov, V., Zaychikov, E., Denissova, L., and Goldfarb, A. (1997). Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. USA* **94**, 6641–6645.
- Nudler, E. (1999). Transcription elongation: structural basis and mechanisms. *J. Mol. Biol.* **288**, 1–12.
- Nudler, E., Avetisova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: protein-DNA interactions holding together the elongation complex. *Science* **273**, 211–217.
- Nudler, E., Gusarov, I., Avetisova, E., Kozlov, M., and Goldfarb, A. (1998). Spatial organization of transcription elongation complex in *Escherichia coli*. *Science* **281**, 424–428.
- Polyakov, A., Severinova, E., and Darst, S.A. (1995). Three-dimensional structure of *E. coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* **83**, 365–373.
- Rees, W.A., Keller, R.W., Vesenska, J.P., Yang, G., and Bustamante, C. (1993). Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science* **260**, 1646–1649.
- Rice, G.A., Kane, C.M., and Chamberlin, M.J. (1991). Footprinting analysis of mammalian RNA polymerase II along its transcript: an alternative view of transcription elongation. *Proc. Natl. Acad. Sci. USA* **88**, 4245–4249.
- Rice, G.A., Chamberlin, M.J., and Kane, C.M. (1993). Contacts between mammalian RNA polymerase II and the template DNA in a ternary elongation complex. *Nucleic Acids Res.* **21**, 113–118.
- Rippe, K., Guthold, M., von Hippel, P.H., and Bustamante, C. (1997). Transcriptional activation via DNA-looping: visualization of intermediates in the activation pathway of *E. coli* RNA polymerase  $\sigma^{54}$  holoenzyme by scanning force microscopy. *J. Mol. Biol.* **272**, 125–138.
- Schultz, P., Celia, H., Riva, M., Sentenac, A., and Oudet, P. (1993). Three-dimensional model of yeast RNA polymerase I determined by electron microscopy of two-dimensional crystals. *EMBO J.* **12**, 2601–2607.
- Selby, C.P., Drapkin, R., Reinberg, D., and Sancar, A. (1997). RNA polymerase II stalled at a thymine dimer: footprint and effect on excision repair. *Nucleic Acids Res.* **25**, 787–793.
- Sidorenkov, I., Komissarova, N., and Kashlev, M. (1998). Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell* **2**, 55–64.
- Treich, I., Carles, C., Sentenac, A., and Riva, M. (1992). Determination of lysine residues affinity labeled in the active site of yeast RNA polymerase II (B) by mutagenesis. *Nucleic Acids Res.* **20**, 4721–4725.
- Uptain, S.M., Kane, C.M., and Chamberlin, M.J. (1997). Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* **66**, 117–172.
- Young, R.A. (1991). RNA polymerase II. *Annu. Rev. Biochem.* **60**, 689–715.