Molecular Cell, Vol. 16, 955–965, December 22, 2004, Copyright ©2004 by Cell Press

Complete RNA Polymerase II Elongation Complex Structure and Its Interactions with NTP and TFIIS

Hubert Kettenberger,[1] Karim-Jean Armache,[1] and Patrick Cramer*
Gene Center
Department of Chemistry and Biochemistry
Ludwig-Maximilians-University of Munich
Feodor-Lynen-Str. 25
81377 Munich
Germany

Summary

The crystal structure of the complete 12 subunit RNA polymerase (pol) II bound to a transcription bubble and product RNA reveals incoming template and nontemplate DNA, a seven base pair DNA/RNA hybrid, and three nucleotides each of separating DNA and RNA. The complex adopts the posttranslocation state and accommodates a cocrystallized nucleoside triphosphate (NTP) substrate. The NTP binds in the active site pore at a position to interact with a DNA template base. Residues surrounding the NTP are conserved in all cellular RNA polymerases, suggesting a universal mechanism of NTP selection and incorporation. DNA-DNA and DNA-RNA strand separation may be explained by pol II-induced duplex distortions. Four protein loops partition the active center cleft, contribute to embedding the hybrid, prevent strand reassociation, and create an RNA exit tunnel. Binding of the elongation factor TFIIS realigns RNA in the active center, possibly converting the elongation complex to an alternative state less prone to stalling.

Introduction

During transcription of protein-coding genes, RNA pol II forms a stable elongation complex with DNA and RNA (Erie, 2002; Shilatifard et al., 2003). In the elongation complex, downstream DNA is unwound before entering the polymerase active site and is rewound when it exits the complex as an upstream DNA duplex. In the unwound bubble region, the DNA template strand forms a hybrid duplex with product RNA that emerges from the active site. Pol II can select NTP substrates in a template-directed manner, synthesize RNA, translocate along DNA, maintain the transcription bubble, and separate RNA from DNA at the upstream end of the hybrid. In addition, pol II has a weak RNA nuclease activity that is stimulated by the elongation factor TFIIS, which also promotes readthrough of blocks to elongation (Fish and Kane, 2002; Wind and Reines, 2000). Understanding pol II elongation requires three-dimensional (3D) structures of elongation complexes in various functional states, with all nucleic acids present, and in the absence and presence of TFIIS.

Electron microscopy studies of the pol II elongation complex allowed initial placement of downstream DNA onto a crystallographic model of the ten subunit pol II core (Cramer et al., 2000; Poglitsch et al., 1999). Subsequent structures of the pol II core and a minimal elongation complex, formed by transcription of a "tailed" DNA template, revealed the DNA-RNA hybrid with three DNA nucleotides downstream and the interacting pol II elements (Cramer et al., 2001; Gnatt et al., 1997, 2001). A recent structure of the pol II core bound to a hybrid with unpaired upstream strands showed additional nucleotides of separating DNA and RNA beyond the hybrid (Westover et al., 2004a). A molecular view of the elongation complex has emerged from these studies (Cramer, 2004b; Gnatt, 2002; Shilatifard et al., 2003). Downstream DNA enters at the pol II jaws and extends through the cleft toward the active site. The hybrid emerges from the active site toward the wall at about a right angle with downstream DNA. The growing RNA 3′ end lies above a pore through which nucleotides may enter. Prominent loops were proposed to separate RNA from DNA at the end of the hybrid (Cramer et al., 2001; Gnatt et al., 2001; Westover et al., 2004a).

In the two structures of pol II core-nucleic acid complexes, upstream DNA was absent, the binding site for NTP was not determined, and downstream DNA was either absent or largely disordered and involved in crystal packing (Gnatt et al., 2001; Westover et al., 2004a). The model of the pol II-TFIIS complex lacked nucleic acids (Cramer, 2004a; Kettenberger et al., 2003). Thus, previous work has left several questions about the elongation mechanism unanswered. What is the location of downstream and upstream DNA, and which pol II elements interact with them? How are the strands of downstream DNA separated, and how does this mechanism compare to DNA-RNA strand separation? Is the presumed location of the NTP substrate correct? How does the elongation factor TFIIS influence DNA and RNA in the elongation complex? Finally, what aspects of the elongation mechanism are conserved in the two other nuclear RNA polymerases, pol I and pol III?

Answers to these questions are provided by structural data presented here. A refined atomic X-ray crystallographic model of the complete pol II in complex with a DNA bubble and product RNA, together with additional electron density maps of this elongation complex in the presence of an NTP substrate analog or TFIIS, provides unbiased 3D views of the natural transcribing enzyme.

Results and Discussion

Complete Pol II-Bubble-RNA Complex

We reconstituted complete pol II elongation complexes from endogenous yeast pol II core, recombinant Rpb4/7 subcomplex, a 41-mer DNA duplex with an 11 nucleotide mismatched bubble region, and an RNA 20-mer with eight 3′-terminal nucleotides complementary to the DNA template strand in the bubble (Figure 2A, Experimental Procedures). After removal of excess Rpb4/7 and nucleic acids by gel filtration, the complex formed crystals

*Correspondence: cramer@lmb.uni-muenchen.de
[1]These authors contributed equally to this work.

Table 1. X-ray Diffraction Data and Refinement Statistics

Crystal Composition	Pol II-Bubble-RNA[a]	Pol II-Bubble-RNA (with Br-dU)[a,b]	Pol II-Bubble-RNA + GMPCPP[a,c]	Pol II + TFIIS[a,d]	Pol II-Bubble-RNA + TFIIS[a,i]
Unit cell axes (Å)[e]	221.4, 392.5, 283.2	220.6, 393.1, 282.0	222.5, 393.1, 283.7	218.9, 395.3, 281.0	220.2, 395.7, 282.1
Wavelength (Å)	0.9200	0.9195 (Br peak)	0.97960	0.9790	0.9919
Resolution range (Å)[f]	50–4.0 (4.14–4.0)	50–4.2 (4.35–4.20)	50–4.5 (4.66–4.50)	50–3.8 (3.9–3.8)	50–4.0 (4.14–4.0)
Unique reflections[f]	102,020 (10,148)	69,541 (6,559)	58,178 (5,134)	115,508 (9,618)	103,460 (10,312)
Completeness (%)[f]	99.3 (99.5)	77.5 (73.8)	78.4 (69.8)	96.2 (81.0)	98.7 (99.2)
Redundancy [f]	4.0 (4.0)	4.9 (4.3)	2.8 (2.5)	2.9 (1.2)	4.3 (4.3)
Mosaicity (°)	0.35–1.23[g]	0.51	0.20–0.54[g]	0.21–0.53[h]	0.29–0.37 [g]
R_{sym} (%)[f]	9.0 (27.9)	7.6 (18.6)	8.1 (23.1)	8.9 (35.5)	7.0 (27.9)
$I/\sigma(I)$[f]	13.7 (4.8)	16.6 (7.2)	7.2 (3.0)	9.3 (1.9)	20.3 (5.4)
Peaks in anomalous Fourier	–	8 Zn in pol II (15.1–8.1 σ) 1 Br in DNA (5.4 σ)	–	8 Zn in pol II (10.9–5.1 σ) 1 Zn in TFIIS (9.1 σ)	–
R_{free} after rigid body refinement (%)		28.6	29.7		28.1
Refinement					
Number of residues	3926 amino acid residues, 36 nucleotides			4001 amino acid residues	
RMSD bonds (Å)	0.010			0.011	
RMSD angles (°)	1.55			1.62	
R_{cryst} (%)	25.3			28.2	
R_{free} (%)	27.6			29.4	

[a] Diffraction data were collected at beamline X06SA at the Swiss Light Source.
[b] The complex contained a DNA template strand in which thymine at position −4 was replaced by the isosteric bromouracil, and the C at position +1 was exchanged to G.
[c] The complex was cocrystallized with the nonreactive NTP analogue GMPCPP.
[d] Data taken from Kettenberger et al. (2003).
[e] Crystals belong to the space group $C222_1$.
[f] Values in parentheses correspond to the highest resolution shell.
[g] Due to radiation damage, mosaicity increased during data collection and was refined in segments of five frames.
[h] Mosaicity was refined in segments of 40 frames.
[i] The complex crystals were soaked with the inactive TFIIS variant D290A/E291A.

of a maximum size of 0.3 × 0.15 × 0.1 mm. The reconstituted complex could elongate RNA upon NTP addition (data not shown). After extensive screens of crystals and cryocooling conditions, data to 4.0 Å resolution were obtained with a careful data collection strategy.

Phasing of the data with the core pol II structure (Cramer et al., 2001; Gnatt et al., 2001) resulted only in noisy additional electron density for nucleic acids in the polymerase cleft (Figure 2E). We could, however, build and refine an atomic model for the complete pol II at 3.8 Å resolution with the help of a new crystal structure of the free polymerase subcomplex Rpb4/7. These structures will be described in detail elsewhere (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data). When the refined complete pol II structure was used for phasing, the nucleic acid density improved dramatically (Figure 2D). Backbone phosphate groups and individual nucleotides were resolved, but the identity of the nucleotides was not revealed at this resolution. We therefore introduced 5-bromouracil in the DNA template strand and located the bromine atom at position −4 in an anomalous difference Fourier map, defining the register of nucleic acids (Figure 2B, Table 1). Nucleic acids were built into unbiased electron density. Restrained refinement was feasible at 4 Å (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data) and lowered the free R-factor to 27.6%.

Overview of Nucleic Acid Structure

The final model comprises all 12 pol II subunits and 19, 7, and 10 residues of template DNA, nontemplate DNA, and RNA, respectively (Figure 1). Strong, continuous electron density is observed for two-thirds of a turn of downstream DNA before the active site, for the DNA-RNA hybrid, for the template strand connecting between the downstream duplex and the hybrid, and for three nucleotides of both DNA and RNA upstream of the hybrid (Figure 2). The complex adopts the posttranslocation state, with no density in the presumed NTP binding site (position +1) (Figures 2A and 2D). The nucleic acids are not involved in crystal packing, and their location is thus determined by interactions with the polymerase, providing an unbiased view of the elongation complex.

Although all nucleic acids of a natural elongation complex are present in our crystals, only selected regions are bound by the polymerase and thus visible in the electron density. DNA entering the cleft is highly mobile but is bound for strand separation closer to the active site (Figure 2D). Downstream DNA runs along the Rpb1-face of the cleft (Figure 1). Upstream of the point of DNA strand separation, the DNA template strand positions a base in the active site that directs RNA synthesis. The DNA-RNA hybrid forms numerous polymerase contacts, which are generally in accordance with earlier observations (Gnatt et al., 2001) and confer stability and thus processivity to the elongation complex (Kireeva et al., 2000b). Upstream of the hybrid, the polymerase continues to contact the RNA and the DNA template strand, and these contacts may be mainly responsible for strand separation and proper RNA exit. The upstream DNA duplex and the nontemplate strand in the bubble are

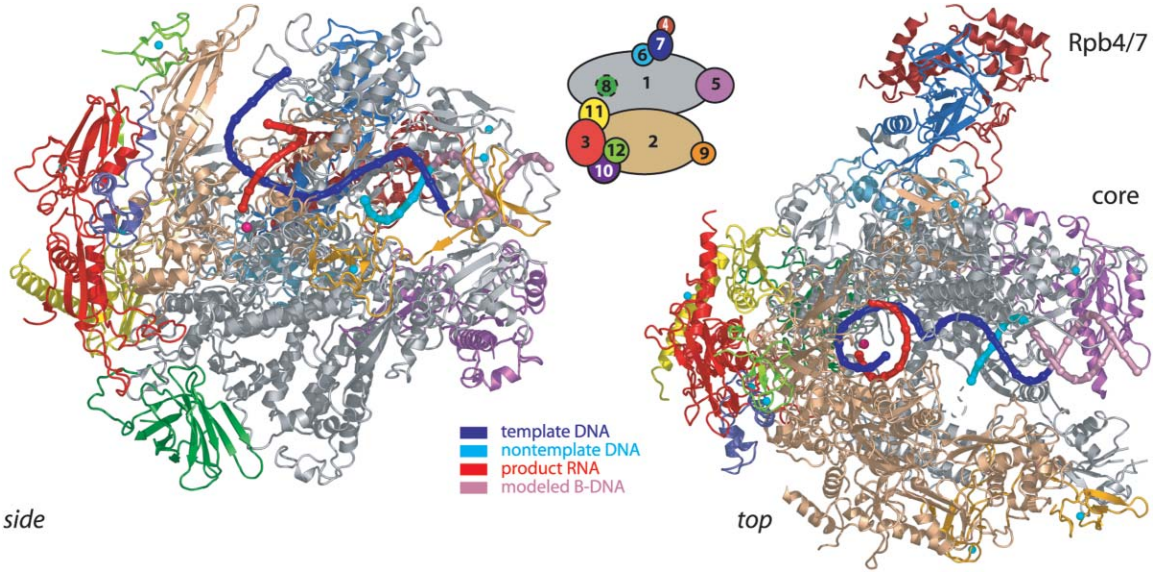


Figure 1. Structure of the Complete Pol II Elongation Complex

Two views of a ribbon model of the protein subunits and nucleic acids. The polymerase subunits Rpb1–Rpb12 are colored according to the key between the views. Template DNA, nontemplate DNA, and product RNA are shown in blue, cyan, and red, respectively. Phosphorous atoms are indicated as spheres and extrapolated B-form downstream DNA is colored in light pink. Eight zinc ions and the active site magnesium ion are depicted as cyan spheres and a magenta sphere, respectively. This color code is used throughout. Secondary structure assignments for pol II are according to Cramer et al. (2001) and K.-J.A., S. Mitterweger, A. Meinhart, and P.C. (unpublished data).

not observed, and they do not contribute significantly to elongation complex stability.

DNA Unwinding

In contrast to previous structures of pol II-nucleic acid complexes, seven base pairs of downstream DNA are observed from positions +3 to +9 (Figures 1 and 2). Our observation of DNA template-nontemplate base pairs from positions +3 to +5 precludes a model of several NTP molecules lining up before the active site by binding to complementary DNA template bases (Zhang and Burton, 2004). Position +2 in the template strand, however, is unpaired and corresponds to the point of strand separation in our complex. The downstream DNA forms some direct contacts with pol II and is surrounded by several polymerase elements (Figures 2C and 3). The ordered part of the downstream DNA apparently occupies a DNA binding site identified previously in bacterial polymerase (Nudler, 1999; Nudler et al., 1996, 1998; Naryshkin et al., 2000; Korzheva et al., 2000), and its sequence can influence polymerase pausing (Palangat et al., 1998, 2004). The overall position of downstream DNA agrees with those inferred from electron microscopy (Cramer et al., 2000; Poglitsch et al., 1999) and from scattered electron density features in a crystallographic map (Gnatt et al., 2001).

The observed conformation of downstream DNA suggests how the template strand is removed from the nontemplate strand before the active site. DNA from registers +5 to +9 is essentially in canonical B form, but the template strand deviates from this conformation from positions +2 to +4 (Figure 3A). This deviation is apparently achieved by three positively charged residues in switch 2 (R326, K330, and R337), which "pull" the template strand away from the duplex axis, and by three negatively charged residues in switch 1 that repel the DNA strand (E1403, E1404, and E1407) (Figure 3A). Thus, an enzyme-induced distortion and destabilization of the incoming DNA duplex may drive strand unwinding. In addition, the polymerase fork loop 2 sterically blocks duplex binding, interferes with the nontemplate strand upstream of position +3, and prevents reassociation of separated strands (Figure 3A).

RNA Displacement and Exit

The hybrid comprises seven base pairs in the post-translocation state and would be eight base pairs long upon nucleotide incorporation. At register −8, the Watson-Crick edges of the bases are 4 Å apart, indicating that the RNA begins to separate from the DNA template strand. At position −9 and −10, the RNA deviates from an extrapolated course beyond the point of strand separation (Figure 3A), suggesting that pol II induces a distortion of the stable heteroduplex structure. The separating strands form backbone contacts with pol II residues that are generally conserved among nuclear RNA polymerases, suggesting that these enzymes make use of a common mechanism (Figures 2A, 3A, and 4). Contacts of pol II from RNA positions −8 to −10 may constitute a previously characterized RNA binding site that contributes to elongation complex stability (Nudler, 1999). The contacts to the separating strands were not observed in the tailed-template structure (Gnatt et al., 2001), apparently due to formation of an overextended hybrid. Lack of these contacts explains the negative effect of an overextended hybrid on elongation complex stability (Kireeva et al., 2000a).

In principle, it is possible that the design of the nucleic

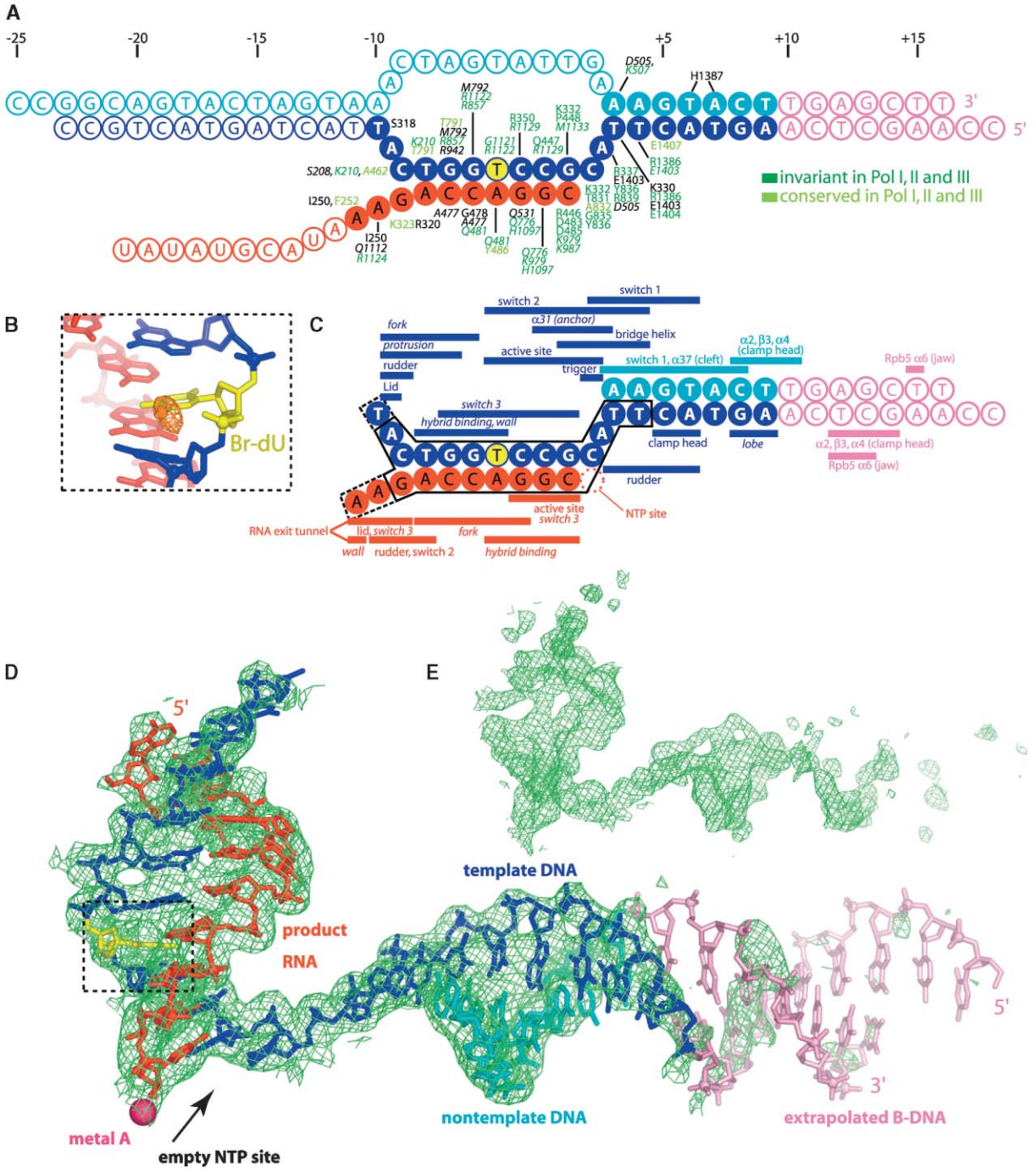


Figure 2. Nucleic Acids in the Elongation Complex

(A) Nucleic acid constructs and pol II interactions. The template DNA strand, nontemplate DNA strand, and product RNA are colored blue, cyan, and red, respectively, and this color code is used throughout. Extrapolated canonical B-DNA is in light magenta. Filled and open circles denote nucleotides that are ordered and disordered, respectively, in the electron density of the cocrystals. Pol II residues that are within 4 Å distance of nucleic acids are depicted. Residues that are invariant, conserved, and differing among the three yeast nuclear RNA polymerases are in dark green, light green, and black, respectively. Italics are used to distinguish Rpb2 from Rpb1 residues. The yellow circle marks the thymine substituted by 5-bromo uracil and located by anomalous diffraction (B). Nucleotide positions are numbered with respect to the nucleotide addition site (position +1). Upstream and downstream nucleotides/positions have negative and positive numbers, respectively.

(B) Bromine anomalous difference Fourier map. The map was calculated from data set 2 (Table 1) taken at the bromine absorption maximum and from phases from the complete pol II structure without nucleic acids. It is contoured at 4.0 σ. The observed peak has a height of 5.4 σ and coincides with the location of the bromine atom in a 5-bromouracil residue incorporated at position −4 of the DNA template strand, thereby defining the register of nucleic acids. The view is related to that in (A) by a 180° rotation about a vertical axis. This and other figures were prepared with PYMOL (DeLano Scientific).

(C) Long-range pol II–nucleic acid interactions. Bars indicate pol II domains and elements within a distance of 8 Å from the nucleic acids

acids used in our studies, and in particular the A-A mismatch at position −9, result in an artificial situation and that the hybrid in a natural elongation complex may be one base pair longer. Indeed, we can not exclude that a different nucleic acid sequence at the point of DNA-RNA strand separation would allow for Watson-Crick base pair formation at position −8, because only a 1 Å decrease in the distance between the bases would be required. Overall, however, the location of the separating upstream strands detected here is very similar to that seen in the pol II core complex with a synthetic hybrid (Westover et al., 2004a), although the sequence of nucleic acids differs, strongly arguing for the sequence-independent nature of the nucleic acid locations and the polymerase-nucleic acid interactions observed here.

Maintenance of the upstream end of the hybrid also involves the lid, a prominent loop that protrudes from the edge of the clamp (Figures 3 and 4). Modeling shows that an extrapolated template strand clashes with the lid (Figure 3A). Except phenylalanine 252, the lid residues do not contact separating nucleic acids and are weakly conserved between the three nuclear RNA polymerases (Figure 3B). Hence, the lid apparently plays a steric role in DNA-RNA strand separation. The lid and the saddle between the wall and the clamp (Cramer et al., 2001) form a narrow RNA exit tunnel that only allows passage of a single strand (Figure 4C). A similar ordering of the loops occurs in the pol II core bound to a DNA-RNA hybrid (Westover et al., 2004a). The RNA 5′ end is located at the entrance to this tunnel, and its extension through the tunnel leads to two potential RNA binding grooves that flank the pol II dock domain (Cramer et al., 2000, 2001) (Figure 4B). There is no electron density in the grooves, although the RNA in our complex is long enough to reach them, indicating that RNA does not interact stably with the grooves in the absence of other factors.

Initiation-Elongation Transition

After transcription initiation, RNA growing along the identified path may trigger structural transitions that lead to a stable elongation complex. Growing RNA would displace the initiation factor TFIIB, which binds to the dock domain (Chen and Hahn, 2003) and extends over the saddle into the hybrid site (Bushnell et al., 2004; Chen and Hahn, 2004). The transition from the closed preinitiation complex to an open complex includes movement of downstream DNA, which lies inside the cleft during elongation but is apparently suspended above the cleft in the initiation complex (Armache et al., 2003; Bushnell and Kornberg, 2003; Craighead et al., 2002). Upstream DNA is mobile during elongation, but during initiation it is bound by several factors and may lie above the cleft (Chen and Hahn, 2004) as in a bacterial RNA polymerase–DNA complex (Murakami et al., 2002).

The transition also goes along with an almost complete embedding of the DNA-RNA hybrid, the hallmark of the elongation complex (Figure 4). Two loops that extend from opposite sides of the cleft, the rudder and fork loop 1, are mobile in free pol II (Cramer et al., 2001) but are ordered and contact each other in the elongation complex, creating two compartments in the cleft for holding downstream DNA and the hybrid (Figure 4B) (Westover et al., 2004a). Given the apparent importance of the lid, fork loops, and rudder for separation of the nucleic acids, it is surprising that these loops are not well conserved in pol I and pol III and even deviate in length (Figure 3B). Apparently these loops are only responsible for a passive topological partitioning of the cleft.

The initiation-elongation transition was structurally characterized for the single subunit RNA polymerase of phage T7, which does not require initiation factors. A large portion of this polymerase refolds during the transition, to create a hybrid binding site and an RNA exit tunnel (Steitz, 2004; Tahirov et al., 2002; Yin and Steitz, 2002). In pol II, the hybrid site and exit tunnel preexist in the initiation-competent form of the enzyme, but are partially filled during initiation with TFIIB (Bushnell et al., 2004) and maybe also with the general factors TFIIE (Forget et al., 2004; Kim et al., 2000; Meinhart et al., 2003) and TFIIF (Chung et al., 2003). These factors may all contribute to the stabilization of an early transcribing complex, but must liberate the hybrid site and RNA exit tunnel during formation of a stable elongation complex. Whereas the initiation complexes of the single-subunit polymerase and pol II are very different, the arrangement of nucleic acids in the elongation complexes is similar, including the length of the DNA-RNA hybrid, the observed 90° twist of template bases at position +2 and +1, and the relative position of the hybrid with respect to downstream DNA.

NTP Binding, Selection, and Incorporation

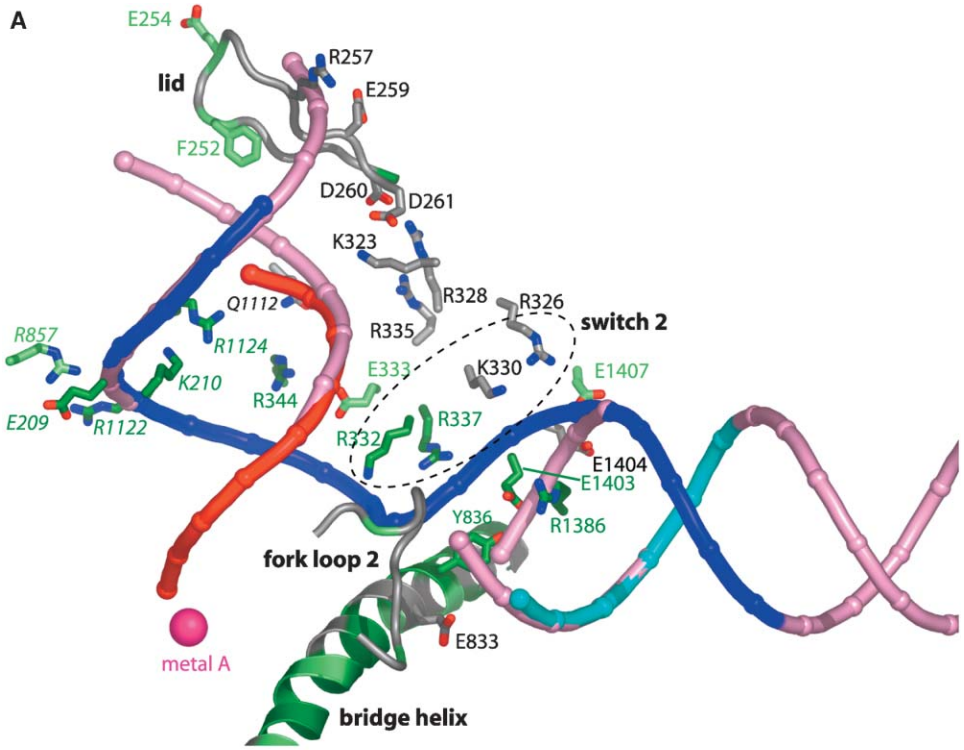
The electron density shows an open space at the 3′ end of the RNA, which corresponds to the presumed binding site for substrate NTP (Figure 2D). The DNA template cytidine at position +1 points toward this site, ready for binding a complementary GTP. This prompted us to cocrystallize the complex with the nonreactive GTP analog GMPCPP, (see Experimental Procedures). The resulting cocrystals diffracted weakly and were very radiation sensitive, but complete diffraction data to 4.5 Å resolution were obtained by merging datasets from five crystals (Table 1).

A difference Fourier map revealed electron density in

(Cramer et al., 2001). The extent and color of the bars define the DNA or RNA nucleotides contacted by pol II. Rpb2 elements are in italics. The encircled nucleotide positions have previously been observed in pol II–nucleic acid complex structures (solid black line, Gnatt et al. [2001]; dashed black line, Westover et al. [2004a]).

(D) Electron density map. A $2F_o$-F_c map was calculated from data set 1 (Table 1) by using phases from the final model with nucleic acids omitted. The map is contoured at 1.0 σ. The final model for the nucleic acids is superimposed. The region within the dashed box is magnified in 1B. Note the absence of electron density in the substrate binding (+1) site.

(E) Initial electron density map. The $2F_o$-F_c map has been calculated with the same experimental data as in (D) but was phased with the incomplete 10-subunit core pol II structure (PDB-ID 1I6H). The map is contoured at 1.0 σ.



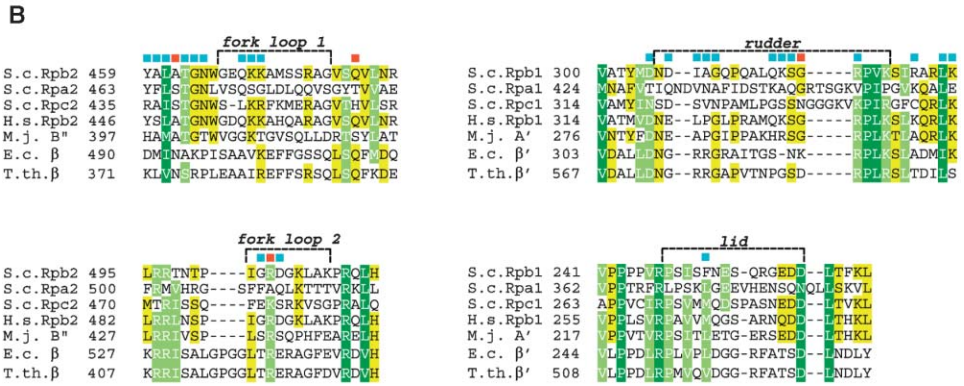


Figure 3. Nucleic Acid Strand Separation

(A) DNA unwinding and DNA-RNA strand separation. Nucleic acid backbones are colored as in Figure 2. Spheres represent the location of phosphorous atoms. Downstream DNA that has been extrapolated as a canonical B form duplex is in light magenta. The DNA/RNA hybrid at positions −1 to −5 has been used to extrapolate the hybrid structure upstream that is also shown in light magenta. The active site metal ion A is shown as a pink sphere. Pol II residues that are apparently involved in nucleic acid strand separation are colored according to their conservation in pol I, II, and III as in Figure 2A. The bridge helix, fork loop 2, and the lid are depicted. Basic amino acids in switch 2 that may pull the template strand upwards are encircled with a dashed line.

(B) Sequence alignments of loops that partition the cleft. Amino acid alignments of the two fork loops in Rpb2 and the rudder and lid in Rpb1 are colored according to conservation between eukaryotic, archaeal, and bacterial RNA polymerases. Invariant, conserved, and weakly conserved residues are colored in dark green, light green, and yellow, respectively. Red and blue squares above the alignments indicate residues within 4 Å and 8 Å distance from nucleic acids, respectively. Sequences in subunits of pol II of *S. cerevisiae* (S.c.), human (H.s.) and *Thermus thermophilus* (T.th.) are aligned with their homologs in S.c. pol I and III and *E. coli* (E.c.) and *Methanococcus janaschii* (M.j.) RNA polymerase.

the pol II pore near the presumed NTP site between the RNA 3′ end and the bridge helix (Figure 5A). Two density lobes can be ascribed to the nucleotide base and triphosphate moiety (Figure 5A). The nucleotide base could form a Watson-Crick base pair with the DNA template base at position +1, as required for template-directed RNA synthesis. The NTP ribose would be close to the side chain of Rpb1 residue N479, which could bind the ribose 2′-OH group, selecting for ribonucleotides (Gnatt et al., 2001). Indeed, NTP versus desoxy-NTP discrimination in bacterial RNA polymerase involves the residue corresponding to N479 (Svetlov et al., 2004). Polymerase residues lining the NTP site and the active site are highly conserved among species and among nuclear RNA polymerases (Figure 5B), suggesting that the mechanism of NTP binding, selection, and incorporation is the same in all these enzymes.

The NTP triphosphate moiety could be bound by four

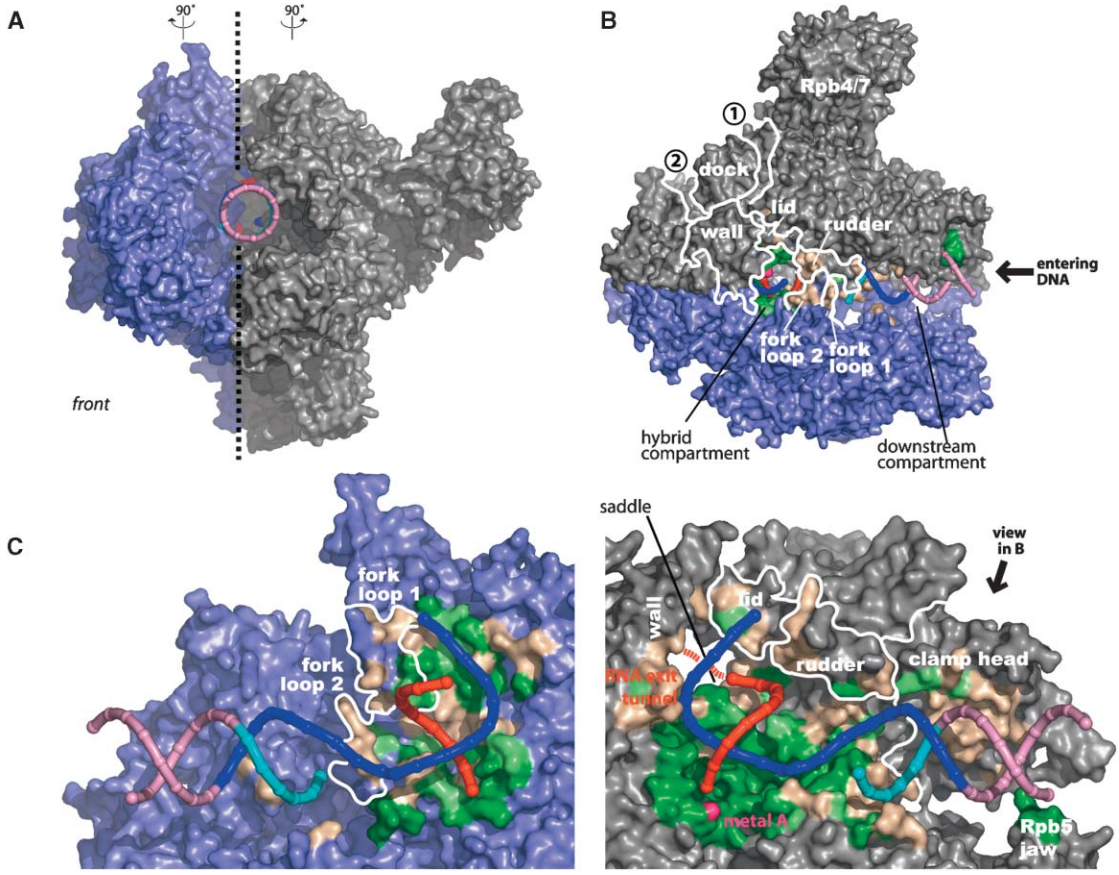


Figure 4. Conservation and Compartmentalization of the Polymerase Cleft

(A) Overall structure of the pol II-bubble-RNA complex. The complete pol II is shown as a molecular surface, and nucleic acid backbones are drawn as ribbons. The view corresponds to the front view (Cramer et al., 2000, 2001). A dashed line indicates the slice plane used to create the views in (C).

(B) Top view of the model in (A). Pol II loops are outlined that partition the enzyme cleft. During transcription elongation, DNA enters from the right. Previously proposed RNA exit grooves are labeled 1 and 2.

(C) Conservation of nucleic-acid interaction surfaces. The model in (A) was intersected along the plane indicated in (A), and the resulting halves were rotated by 90° around a vertical axis in opposite directions. The molecular surface of residues within 8 Å distance from nucleic acids is colored in beige. Residues that are invariant and conserved between pol I, II, and III are highlighted in dark and light green, respectively. Pol II elements that contact nucleic acids and partition the enzyme cleft are outlined.

conserved basic residues (Figure 5) but is not in a position that allows nucleotide incorporation into RNA. The triphosphate is too far away from the catalytic metal ion A and from Rpb2 residue D837, which may bind a second catalytic metal ion (Cramer et al., 2001; Sosunov et al., 2003), and is required for full activity of an archaeal RNA polymerase (Werner and Weinzierl, 2002). Thus, the observed NTP site is apparently not identical to the insertion site occupied during catalysis but may correspond to a preinsertion site. We predict that for nucleotide incorporation, pol II adopts a more closed conformation and creates an NTP insertion site possibly by movement of the bridge helix toward the active site. An NTP in the closed insertion site could form all contacts required for incorporation, including the metal ion interactions with the triphosphate.

Comparison with structures of functional complexes of T7 RNA polymerase supports our results and their interpretation. The overall location of the NTP with respect to the active site, DNA and RNA, is the same in recent substrate complexes of T7 RNA polymerase (Landick, 2004; Temiakov et al., 2004; Yin and Steitz, 2004). Cocrystallization of a T7 RNA polymerase elongation complex with the NTP analog AMPCPP can result in NTP binding to a preinsertion site, and in an open conformation at the active site (Temiakov et al., 2004), as observed here for a pol II elongation complex cocrystallized with GMPCPP. Analysis of the preinsertion site in T7 RNA polymerase suggested that the NTP triphosphate moiety is bound by three basic residues in bacterial RNA polymerase (Temiakov et al., 2004), which correspond exactly to the basic residues found near the difference density in the pol II complex.

TFIIS Realigns RNA

A previous backbone model of pol II bound by TFIIS showed that this elongation factor inserts a hairpin loop into the pol II pore to complement the active site with two acidic residues that may position a metal ion and a nucleophilic water molecule for RNA cleavage (Kettenberger et al., 2003). Several aspects of TFIIS function, however, remained unclear because this structural work

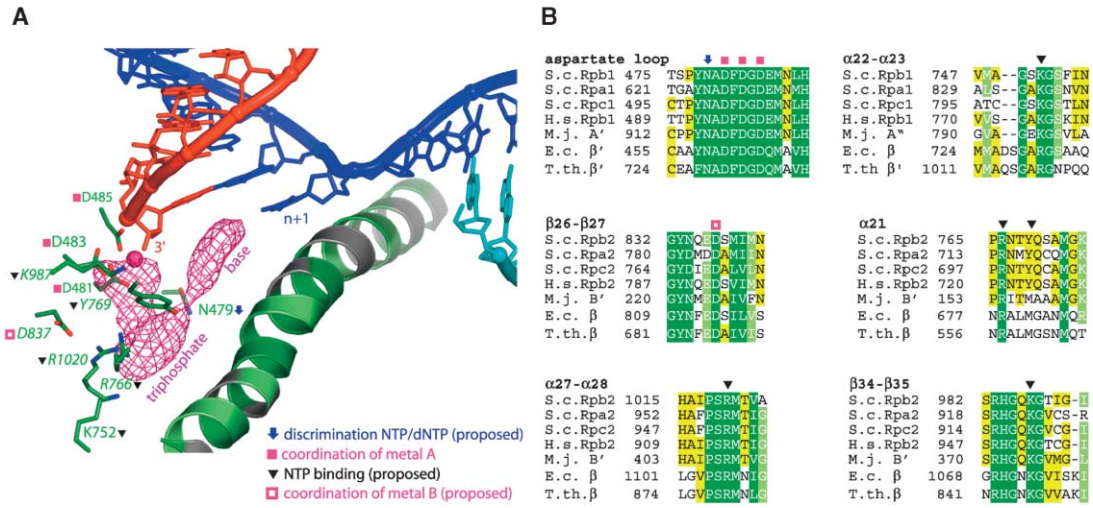


Figure 5. NTP Binding Site

(A) Fourier difference electron density for a GTP substrate analog. The F_o-F_c map was calculated from data of the pol II–bubble-RNA complex after cocrystallization with GMPCPP (Table 1), was phased with the final elongation complex structure, and is contoured at 2.8 σ. Two prominent lobes of electron density are seen in the presumed NTP binding site that is empty in the pol II-bubble-RNA structure. One lobe can be attributed to the nucleotide base, approximately positioned to form a Watson-Crick base pair with the template DNA base at position +1 (the "coding" base). The other lobe may be attributed to the triphosphate moiety of the NTP. Invariant pol II residues in the vicinity of the density peak are shown. Between the two lobes, the ribose moiety would be located, at a position where its 2′-OH group could form a hydrogen bond with Rpb1 residue N479, consistent with a role of this residue in NTP over dNTP selection.

(B) Sequence alignment of residues surrounding the NTP site. Sequences in subunits of pol II of S.c., H.s., and T.th. are aligned with their homologs in S.c. pol I and III and E.c. and M.j. RNA polymerase.

did not include nucleic acids. We therefore investigated what effect TFIIS binding has on the elongation complex structure. We soaked pol II-bubble-RNA complex crystals with a TFIIS variant that is inactive in RNA cleavage stimulation, because acidic hairpin residues D290 and E291 were mutated to alanine. For phasing, we refined an atomic model of the pol II-TFIIS complex by using the complete pol II structure and previous data to 3.8 Å resolution (Table 1, Experimental Procedures). The resulting unbiased difference density in the pol II-bubble-RNA-TFIIS complex was noisy but allowed modeling of four base pairs of the DNA-RNA hybrid (positions −1 to −4, Figure 6).

TFIIS induces an unexpected shift of the RNA strand with phosphate groups moving up to 4 Å (Figure 6). This realignment of the RNA in the active site may allow for an optimized inline attack of the scissile phosphodiester bond by a nucleophilic water molecule during RNA cleavage. Analysis of distances in the refined pol II-TFIIS structure further shows that the acidic residue D290 in the TFIIS hairpin could directly contact and activate a nucleophilic water molecule in cooperation with a metal ion. TFIIS does not simply reposition the DNA-RNA hybrid as a rigid body, but its exact conformation is not revealed at the limited resolution of the data.

Facilitated readthrough of blocks to elongation is generally thought to result from TFIIS-stimulated RNA cleavage, which creates new RNA 3′ ends at the active center, thereby allowing pol II to overcome a barrier with multiple trials. Our new structural data suggest alternative explanations for how TFIIS facilitates readthrough. First, the TFIIS-induced repositioning of RNA may result in a different elongation complex conformation that is less

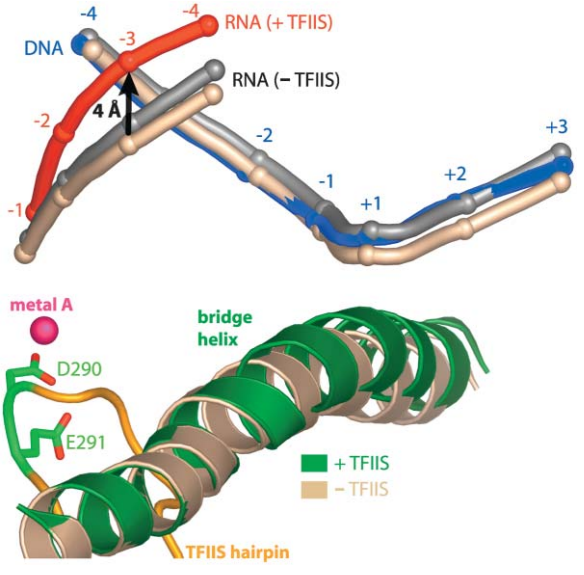


Figure 6. TFIIS-Induced RNA Realignment

Selected elements in the pol II active center that move upon TFIIS binding are shown. The bridge helix, DNA, and RNA in the pol II-bubble-RNA-TFIIS complex are in green, blue, and red, respectively. The TFIIS hairpin is in orange with the two acidic functionally essential and invariant residues in green. Nucleic acids in the pol II-bubble-RNA complex structure after superposition of residues in the active site aspartate loop or in switch 2 are shown in beige and gray, respectively. Switch 2 moves slightly upon TFIIS binding (Kettenberger et al., 2003), explaining the difference in the two superpositions.

prone to pausing. Recent kinetic data are consistent with an alternative TFIIS-induced state of the elongation complex (Zhang and Burton, 2004). Second, the acidic residues in the TFIIS hairpin could assist in NTP binding. The TFIIS hairpin overlaps the NTP site described above, but small changes in the hairpin conformation may allow for metal-mediated stabilizing interactions with the NTP triphosphate.

Conclusions

Here, we present the 3D structural organization of the complete pol II elongation complex, with entering DNA, exiting RNA, and incoming NTP substrate and show how the complex responds to binding of the elongation factor TFIIS. The results generally apply to all three nuclear RNA polymerases, because the regions relevant for elongation show a strong structural conservation in pol I and pol III (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data). In the future, we must trap and structurally resolve the many different functional states of the enzyme during the nucleotide addition cycle. It will also be important to find ways of increasing the resolution of these large and transient complexes, to clarify chemical details of the mechanism. A long-term challenge is the stabilization and structure determination of other complexes of pol II with nucleic acids and transcription factors, which represent intermediates in the mRNA transcription cycle. In the meantime, the data presented here provide exciting insights into the sophisticated conformational regulation of the dynamic RNA polymerase machinery.

While this paper was in revision, independent crystallographic studies of a pol II core elongation complex were published by Kornberg and colleagues that revealed two additional NTP binding sites in the active center (Westover et al., 2004b). Whereas we detected NTP binding to a "preinsertion site," Kornberg and coworkers observed binding to the catalytic "insertion site" proposed above, but only when an NTP was used that could base pair with the DNA template (a mismatched NTP instead bound adjacently in an inverted orientation). Differences in the NTP binding sites apparently result from differences in the experimental design. Whereas we used a nonreactive NTP analog, Kornberg and colleagues used natural NTPs and prevented their incorporation with a chain-terminating residue at the RNA 3′ end. As a consequence, the two studies provide complementary insights into nucleotide addition by pol II.

Experimental Procedures

Preparation of Pol II-Nucleic Acid Complexes

Endogenous yeast ten subunit pol II core and recombinant Rpb4/7 heterodimer were purified as described (Armache et al., 2003). DNA bubble-RNA complexes were annealed by mixing equimolar amounts of synthetic template DNA, nontemplate DNA, and RNA in TE buffer at a final concentration of 100 μM, heating the mixture to 90°C, and slow cooling to room temperature. We designed a mismatch bubble of only 11 nucleotides and assumed that the polymerase would open the bubble further if required for stable binding. The pol II-bubble-RNA complex was assembled by incubating core pol II with a 5-fold molar excess of recombinant Rpb4/7 and a 1.5-fold molar excess of the bubble-RNA construct in assembly buffer (50 mM Hepes [pH 7.5], 40 mM ammonium sulfate, 5 μM $ZnCl_2$, 5% glycerol, and 10 mM DTT) for 15 min at 20°C. The complex was purified by gel filtration (Superose 6 hr, Amersham) in pol II buffer (5 mM Hepes [pH 7.25], 40 mM ammonium sulfate, 10 μM $ZnCl_2$, and 10 mM DTT). The complex was concentrated to 3.5–4.5 mg/mL, and an additional amount of bubble-RNA complex was added to a final concentration of 2 μM prior to crystallization. The bromine-derivatized template DNA contained 5-bromo-uracile at positions −4, −7, and −10.

Crystal Growth and Treatment

Crystals were grown at 20°C with the hanging drop vapor diffusion method by mixing 2 μl of sample solution with 2 μl of reservoir solution (200 mM ammonium acetate, 150 mM magnesium acetate, 50 mM Hepes [pH 7.0], 5% PEG 6000, and 5 mM TCEP). Crystals grew to a maximum size of 0.3 × 0.15 × 0.1 mm. Crystals were transferred stepwise to mother solution containing additionally 0%–22% glycerol over 8 hr, slowly cooled down to 8°C, incubated for another 24 hr, and plunged into liquid nitrogen. Crystals additionally containing the substrate analog GMPCPP (Guanosine-5′-[(α,β)-methyleno]-triphosphate) were grown under the same conditions in the presence of 3 mM GMPCPP (Jena Biosciences). To ensure maximum occupancy, GMPCPP was also present in the last cryo solution at a concentration of 3 mM. The TFIIS double mutant D290A/E291A was obtained by site-directed mutagenesis of a previously used TFIIS variant comprising domains II and III and an N-terminal hexahistidine tag and was purified as described (Kettenberger et al., 2003). Crystals of the pol II-bubble-RNA-TFIIS complex were obtained by soaking preformed cryoprotected pol II-bubble-RNA complex crystals for 24 hr in the last cryo solution containing additionally 1 mg/mL TFIIS double mutant D290A/E291A as described (Kettenberger et al., 2003).

X-Ray Structural Analysis

Diffraction data were collected in 0.25° increments at the protein crystallography beamline X06SA of the Swiss Light Source (Table 1). Raw data were processed and scaled with HKL2000 (Otwinowski and Minor, 1996). Structures were solved by molecular replacement with the refined atomic model of the complete 12 subunit pol II (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data). After rigid body refinement with CNS (Brunger et al., 1998), a $2F_o$-F_c map showed clear continuous electron density for the DNA-RNA hybrid and several bases of downstream duplex DNA, whereas phasing with the structure of the pol II core alone produced very noisy maps (Figure 2). The DNA-RNA hybrid structure from the previous elongation complex model (Gnatt et al., 2001) was fit into the density and adjusted manually. To model downstream DNA, a canonical B-DNA was positioned, and individual nucleotides were manually adjusted to fit the unbiased $2F_o$-F_c map. The initial model for the nucleic acids was combined with the pol II structure (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data) and was subjected to refinement of atomic positions and B factors with CNS. Refinement was monitored carefully with the free R factor, calculated from the same set of excluded reflections as in refinement of free pol II (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data). To locate nucleic acids in the complex crystal after soaking with TFIIS, an atomic model of the previously reported pol II-TFIIS complex (Kettenberger et al., 2003) was built. The complete pol II structure (K.-J.A., S. Mitterweger, A. Meinhart, and P.C., unpublished data) was subjected to rigid body refinement in previously defined groups against data of the pol II-TFIIS complex extending to 3.8 Å resolution (Kettenberger et al., 2003). A model-phased electron density map was improved by solvent flipping with CNS and allowed building of an atomic model for domain III and the linker of TFIIS (residues 240–309). After careful refinement of atomic positions and B factors, the resulting pol II-TFIIS model was used to phase the pol II-bubble-RNA-TFIIS complex, revealing electron density for the hybrid at the active site.

Acknowledgments

We thank C. Schulze-Briese and the staff at beamline X06SA of the Swiss Light Source for help. We thank members of the Cramer lab and Katja Sträßer for comments on the manuscript. Supported by

the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

Received: October 7, 2004
Revised: November 19, 2004
Accepted: November 23, 2004
Published: December 21, 2004

References

Armache, K.-J., Kettenberger, H., and Cramer, P. (2003). Architecture of the initiation-competent 12-subunit RNA polymerase II. Proc. Natl. Acad. Sci. USA *100*, 6964–6968.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography &NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Bushnell, D.A., and Kornberg, R.D. (2003). Complete RNA polymerase II at 4.1 Å resolution: implications for the initiation of transcription. Proc. Natl. Acad. Sci. USA *100*, 6969–6972.

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

Chen, H.T., and Hahn, S. (2003). Binding of TFIIB to RNA polymerase II: mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. Mol. Cell *12*, 437–447.

Chen, H.T., and Hahn, S. (2004). Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. Cell *119*, 169–180.

Chung, W.H., Craighead, J.L., Chang, W.H., Ezeokonkwo, C., Bareket-Samish, A., Kornberg, R.D., and Asturias, F.J. (2003). RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. Mol. Cell *12*, 1003–1013.

Craighead, J.L., Chang, W.H., and Asturias, F.J. (2002). Structure of yeast RNA polymerase II in solution: implications for enzyme regulation and interaction with promoter DNA. Struct. Fold. Des. *10*, 1117–1125.

Cramer, P. (2004a). RNA polymerase II structure: from core to functional complexes. Curr. Opin. Genet. Dev. *14*, 218–226.

Cramer, P. (2004b). Structure and function of RNA polymerase II. Adv. Protein Chem. *67*, 1–42.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, 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.

Cramer, P., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science *292*, 1863–1876.

Erie, D.A. (2002). The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. Biochim. Biophys. Acta *1577*, 224–239.

Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. Biochim. Biophys. Acta *1577*, 287–307.

Forget, D., Langelier, M.F., Therien, C., Trinh, V., and Coulombe, B. (2004). Photo-cross-linking of a purified preinitiation complex reveals central roles for the RNA polymerase II mobile clamp and TFIIE in initiation mechanisms. Mol. Cell. Biol. *24*, 1122–1131.

Gnatt, A. (2002). Elongation by RNA polymerase II: structure-function relationship. Biochim. Biophys. Acta *1577*, 175–190.

Gnatt, A., Fu, J., and Kornberg, R.D. (1997). Formation and crystallization of yeast RNA polymerase II elongation complexes. J. Biol. Chem. *272*, 30799–30805.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876–1882.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell *114*, 347–357.

Kim, T.K., Ebright, R.H., and Reinberg, D. (2000). Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science *288*, 1418–1422.

Kireeva, M.L., Komissarova, N., and Kashlev, M. (2000a). Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. J. Mol. Biol. *299*, 325–335.

Kireeva, M.L., Komissarova, N., Waugh, D.S., and Kashlev, M. (2000b). The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J. Biol. Chem. *275*, 6530–6536.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). A structural model of transcription elongation. Science *289*, 619–625.

Landick, R. (2004). Active-site dynamics in RNA polymerases. Cell *116*, 351–353.

Meinhart, A., Blobel, J., and Cramer, P. (2003). An extended winged helix domain in general transcription factor E/IIEα. J. Biol. Chem. *278*, 48267–48274.

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.

Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., and Ebright, R.H. (2000). Structural organization of the RNA polymerase-promoter open complex. Cell *101*, 601–611.

Nudler, E. (1999). Transcription elongation: structural basis and mechanisms. J. Mol. Biol. *288*, 1–12.

Nudler, E., Avetissova, 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., Avetissova, E., Kozlov, M., and Goldfarb, A. (1998). Spatial organization of transcription elongation complex in *Escherichia coli*. Science *281*, 424–428.

Otwinowski, Z., and Minor, W. (1996). Processing of X-ray diffraction data collected in oscillation mode. Meth Enzym *276*, 307–326.

Palangat, M., Meier, T.I., Keene, R.G., and Landick, R. (1998). Transcriptional pausing at +62 of HIV-1 nascent RNA modulates formation of the TAR RNA structure. Mol. Cell *1*, 1033–1042.

Palangat, M., Hittinger, C.T., and Landick, R. (2004). Downstream DNA selectively affects a paused conformation of human RNA polymerase II. J. Mol. Biol. *341*, 429–442.

Poglitsch, C.L., Meredith, G.D., Gnatt, A.L., Jensen, G.J., Chang, W.H., Fu, J., and Kornberg, R.D. (1999). Electron crystal structure of an RNA polymerase II transcription elongation complex. Cell *98*, 791–798.

Shilatifard, A., Conaway, R.C., and Conaway, J.W. (2003). The RNA polymerase II elongation complex. Annu. Rev. Biochem. *72*, 693–715.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J. *22*, 2234–2244.

Steitz, T.A. (2004). The structural basis of the transition from initiation to elongation phases of transcription, as well as translocation and strand separation, by T7 RNA polymerase. Curr. Opin. Struct. Biol. *14*, 4–9.

Svetlov, V., Vassylyev, D.G., and Artsimovitch, I. (2004). Discrimination against deoxyribonucleotide substrates by bacterial RNA polymerase. J. Biol. Chem. *279*, 38087–38090.

Tahirov, T.H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W.T., Vassylyev, D.G., and Yokoyama, S. (2002). Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. Nature *420*, 43–50.

Temiakov, D., Patlan, V., Anikin, M., McAllister, W.T., Yokoyama, S., and Vassylyev, D.G. (2004). Structural basis for substrate selection by t7 RNA polymerase. Cell *116*, 381–391.

Werner, F., and Weinzierl, R.O. (2002). A recombinant RNA polymer-

ase II-like enzyme capable of promoter-specific transcription. Mol. Cell *10*, 635–646.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004a). Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. Science *303*, 1014–1016.

Westover, K.D., Bushnell, D.A., and Kornberg, R.D. (2004b). Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. Cell *119*, 481–489.

Wind, M., and Reines, D. (2000). Transcription elongation factor SII. Bioessays *22*, 327–336.

Yin, Y.W., and Steitz, T.A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science *298*, 1387–1395.

Yin, Y.W., and Steitz, T.A. (2004). The structural mechanism of translocation and helicase activity in t7 RNA polymerase. Cell *116*, 393–404.

Zhang, C., and Burton, Z.F. (2004). Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism. J. Mol. Biol. *342*, 1085–1099.

Accession Numbers

The Protein Data Bank accession numbers for the pol II-bubble-RNA complex, the pol II-bubble-RNA-NTP complex, and the refined pol II-TFIIS complex (replacing the previous backbone model with accession code 1NT9) are 1Y1W, 1Y77, and 1Y1V, respectively. The accession code for the backbone model of the pol II-bubble-RNA-TFIIS complex is 1Y1Y.