Multisubunit RNA polymerases

Patrick Cramer

Transcription of the genetic information in all cells is carried out by multisubunit RNA polymerases (RNAPs). Comparison of the crystal structures of a bacterial and a eukaryotic RNAP reveals a conserved core that comprises the active site and a multifunctional clamp. Together with a further structure of eukaryotic RNAP bound to DNA and RNA, these results elucidate many aspects of the transcription mechanism, including initiation, elongation, nucleotide addition, processivity and proofreading.

Addresses
Institute of Biochemistry, Gene Center, University of Munich, Feodor-Lynen-Strasse 25, 81377 Munich, Germany;
e-mail: cramer@lmb.uni-muenchen.de

Current Opinion in Structural Biology 2002, 12:89–97

0959-440X/02/$ – see front matter

Abbreviations

bRNAP	bacterial RNA polymerase
NTP	nucleoside triphosphate
Pol I, II, III	RNA polymerase I, II, III
RNAP	RNA polymerase
yRNAP	yeast RNA polymerase II
yRNAP-EC	yeast RNA polymerase II elongation complex

Introduction

Multisubunit RNA polymerases (RNAPs) synthesize RNA from a DNA template in the course of gene transcription. Bacteria and archaea have one RNAP, whereas eukaryotes have three RNAPs, responsible mainly for the synthesis of ribosomal RNA (RNA polymerase I [Pol I]), pre-messenger RNA (Pol II) and small RNAs including transfer RNAs (Pol III). The complexity and large size of RNAPs (5–15 subunits, up to 0.6 MDa) posed formidable technical challenges to structural biologists. Over the past few years, however, X-ray crystallographic structures have been determined of a bacterial RNAP from *Thermus aquaticus* (bRNAP) at 3.3 Å resolution ([1]; reviewed in [2]), of bRNAP bound to the inhibitor rifampicin at 3.3 Å resolution [3••], of Pol II from the yeast *Saccharomyces cerevisiae* (yRNAP) in two crystal forms at 2.8 Å and 3.1 Å resolution [4••,5••], and of a yeast Pol II elongation complex (yRNAP-EC) at 3.3 Å resolution [6••]. These structures form a basis for understanding the function of all RNAPs and for dissecting the transcription mechanism by site-directed mutagenesis and further structural studies. In this review, I describe the known RNAP structures and their functional implications.

A technical *tour de force*

To solve the crystal structures of the large and asymmetric bRNAP and yRNAP multiprotein complexes, several obstacles had to be overcome ([1,4••,5••,7] and references therein). Both enzymes had to be purified from large quantities of cell culture without the benefit of over-expression. In both cases, interpretation of the experimental electron density maps relied on the placement of known subunit structures and phase combination. In the case of bRNAP, X-ray diffraction was weak, radiation sensitive and anisotropic. The interpretation of maps with weak sidechain electron density relied on noncrystallographic symmetry averaging and data from selenomethionyl bRNAP. In the case of yRNAP, a mutant yeast strain had to be used to produce Pol II lacking two substoichiometric subunits. Nonisomorphous and weakly diffracting initial Pol II crystals were improved dramatically by a soaking procedure that induced crystal shrinkage. Single heavy atoms could not be detected in Patterson maps and initial phasing therefore relied on heavy atom clusters. Single heavy atom derivatives, needed for phasing to higher resolution, could only be obtained with nonstandard compounds. Model building was greatly facilitated by sequence markers, including native zinc ions, mercury-labeled cysteines and, most notably, selenomethionine, which could be partially incorporated in yeast despite its toxicity [8].

General RNA polymerase architecture

The independently determined yRNAP and bRNAP structures reveal that five 'core' subunits underlie a general RNAP architecture. The two large subunits form the central mass of the enzyme and opposite sides of a positively charged cleft (Rpb1 and Rpb2 in yRNAP; β′ and β in bRNAP; Table 1, Figure 1). The two large subunits are anchored by two small core subunits that are involved in RNAP assembly (Rpb3–Rpb11 heterodimer in yRNAP; α homodimer in bRNAP; Figure 1). A fifth core subunit (Rpb6 in yRNAP; ω in bRNAP) further buttresses and stabilizes the large subunit [9–11] (Figure 1).

One side of the cleft (the Rpb1/β′ side) is formed by a mobile 'clamp' (see below). The other side (the Rpb2/β side) is formed by two domains: the 'lobe' and 'protrusion' domains in yRNAP, and the 'β domains 2 and 3' in bRNAP (Figure 2). The active site is located at the floor of the cleft, near the center of the enzyme (Figure 1). Beyond the active site, the cleft is blocked by a 'wall' or 'flap' (Figure 2). Just before the active site, a long α helix spans the cleft ('bridge', 'β′F helix'; Figure 1). This helix and the active site line a perforation in the floor of the cleft ('pore 1', 'secondary channel'; Figure 1), which widens towards the exterior, creating an inverted funnel. The outer rim of the funnel is lined by a pair of α helices in the largest subunit (funnel region; Figure 2).

Conserved core

A total of 22 homology regions in the core subunits [11–14] cluster around the active site. Most portions of these homology regions are structurally conserved between bRNAP and

Table 1

RNAP subunits.

Eukaryotes Pol I	Pol II	Pol III	Archaea	Bacteria	Class*	Homology regions
A190	Rpb1	C160	A′+A″	β′	Core	A–H [13]
A135	Rpb2	C128	B (B′+B″)	β	Core	A–I [12]
AC40	Rpb3	AC40	D	α	Core	α motifs I, II [14]
AC19	Rpb11	AC19	L	α	Core	α motifs I, II [14]
Rpb6	Rpb6	Rpb6	K	ω	Core/common	CR1–3 [11]
Rpb5	Rpb5	Rpb5	H	-	Common	
Rpb8	Rpb8	Rpb8	-	-	Common	
Rpb10	Rpb10	Rpb10	N	-	Common	
Rpb12	Rpb12	Rpb12	P	-	Common	
A12.2	Rpb9	C11	X	-		
A14†	Rpb4‡	-	F	-		
A43†	Rpb7‡	C25	E	-		
+two others	-	+four others	+one other	-		

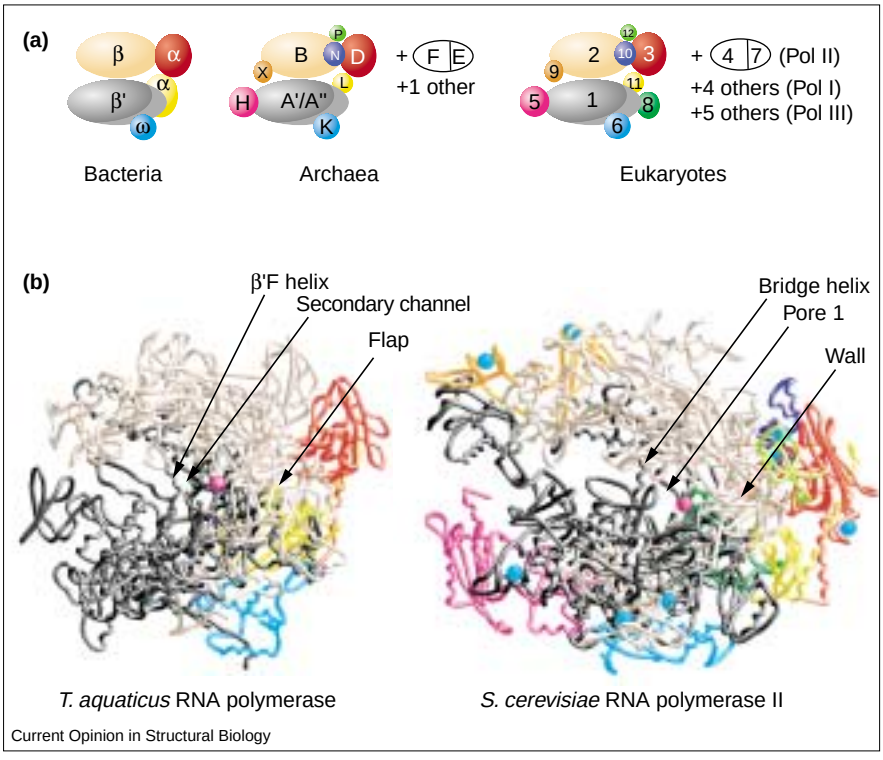
*Core, sequence partially homologous in all RNAPs; common, shared by all eukaryotic RNAPs. †Potential homologs of Pol II subunits Rpb4 and Rpb7 (S Onesti, personal communication). ‡Not included in the yRNAP structure [5••].

yRNAP (Figure 3). Many other regions also share the same structure, showing that structure is conserved better than sequence. The structurally conserved core includes the functional elements of the active center, indicating that all RNAPs share a common mechanism (see below).

Extended interaction surface of eukaryotic RNA polymerases

Whereas bacterial RNAPs consist of the five core subunits only, eukaryotic RNAPs have up to ten additional subunits (Table 1). In the yRNAP structure, subunits Rpb10 and Rpb12 appear to fill depressions on the bRNAP surface, whereas subunits Rpb5 and Rpb9 appear to extend from the bRNAP surface (Figures 1 and 3). Compared with bRNAP, yRNAP also shows additional domains in the core subunits. The largest yRNAP subunit additionally contains the 'jaw' and 'foot' domains, and has a larger 'clamp head' (Figure 3). The second largest subunit additionally contains two external domains (Figure 3). Compared with the bacterial α subunits, Rpb3 contains an exposed zinc-binding loop and

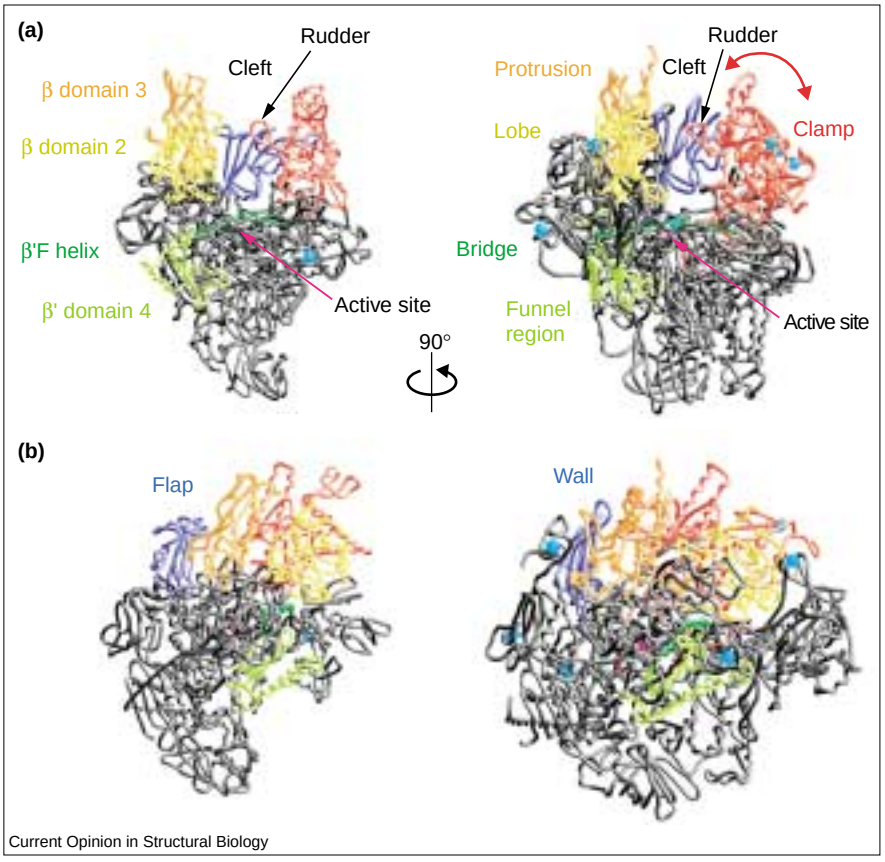
Figure 1



RNAP subunit architecture. **(a)** Schematic presentation and color code for RNAP subunits in bacterial, archaeal and eukaryotic RNAPs. Corresponding subunits have the same color. **(b)** Three-dimensional structures of *T. aquaticus* RNA polymerase (bRNAP; [1]) and *S. cerevisiae* RNA polymerase II (yRNAP; [5••]) shown in the same orientation (top view in [5••]). The subunits are colored according to (a). The active site metal ion A is shown as a pink sphere. Zinc ions are shown as blue spheres.

Figure 2

RNAP structure. Corresponding structural domains in bRNAP (left) and yRNAP (right) are depicted in the same color. The use of color is different from that in Figure 1.
(a) The view corresponds to the front view in [5••]. The direction of clamp movement is indicated by a double-headed arrow.
(b) Alternative view of the structures, related to that in (a) by a 90° rotation about the vertical axis.

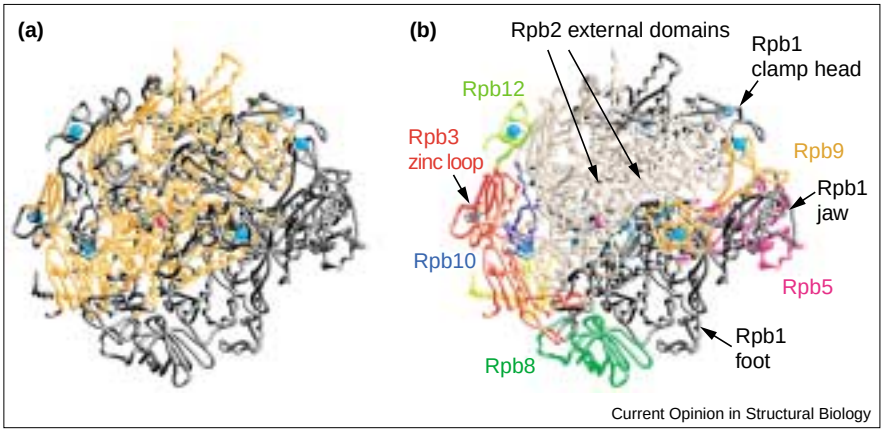


Rpb11 lacks the second α domain. At the site of this second α domain, portions of the yRNAP subunit Rpb8 are located.

Several interaction targets on the surface of bRNAP have been described (reviewed in [15]). Interactions with eukaryote-specific factors are apparently mediated by many of the additional subunits and domains of eukaryotic RNAPs. A hepatitis virus transcriptional activator binds to subunit Rpb5 [16]. Two amino acid substitutions in the Pol-II-specific zinc-binding loop of Rpb3 and adjacent to it cause a defect in activated transcription, suggesting an interaction with activator proteins [17]. bRNAP contains an activation target at a similar location [18]. The Pol-II-specific C-terminal repeat domain of the largest subunit binds mRNA processing factors [19–21]. This domain is not ordered in the yRNAP structure, but is flexibly linked to a region from which RNA exits, consistent with the role of the C-terminal repeat domain in coupling transcription to mRNA processing [5••]. Finally, the natural DNA template for bacterial and eukaryotic RNAPs is fundamentally

Figure 3

Conserved RNAP core and extended surface of eukaryotic Pol II. (a) Conserved RNAP core. Regions that are structurally conserved between the bacterial and eukaryotic RNAP structures are in orange. The view of the yRNAP structure is as in Figure 2b.
(b) Additional subunits and domains of yRNAP. This is the same view as in (a), but with the subunits colored as in Figure 1. Subunits and domains of the yRNAP structure that are not found in the bRNAP structure are located on the surface and are labeled.



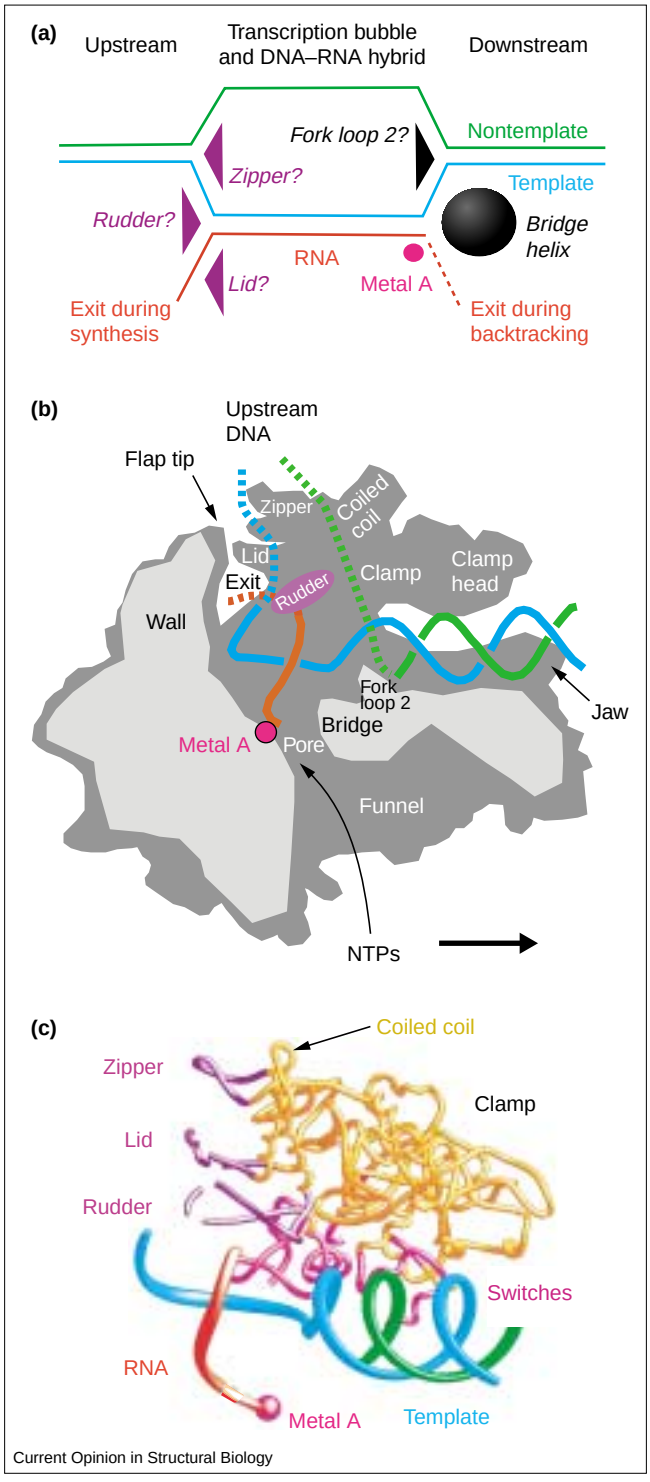


Figure 4

The RNAP elongation complex. (a) Schematic presentation of the arrangement of nucleic acids during RNA chain elongation. The DNA template and nontemplate strands are in blue and green, respectively, and the RNA is in red. The active site metal ion A is indicated by a pink sphere. Protein elements that are proposed to be involved in the maintenance of the arrangement of nucleic acids are indicated. (b) Cutaway view of the yRNAP elongation complex [6••]. The view is as in Figure 2b. Cut surfaces are lightly shaded. During transcription, DNA enters the enzyme from the right (the polymerase moves to the right). Structural features that appear to be important for function are labeled. Coloring of nucleic acids is as in (a). Exiting RNA and DNA strands are not revealed in the electron density map, but their anticipated locations are indicated by dashed lines. In this view, the clamp swings over the active center from back to front. Only one of the jaws (the lower jaw) is visible in the cutaway view. (c) Ribbon diagram of the clamp and nucleic acid backbones in the yRNAP-EC structure. The view is related to that in (b) by a 40° rotation around the vertical axis. Coloring of nucleic acids is as in (a). The switch regions at the base of the clamp are in pink and the three loops emanating from the edge of the clamp are in violet.

different. The ability of eukaryotic RNAPs to transcribe chromatin may reside in eukaryote-specific surface regions as well. These regions could act directly or indirectly through interactions with factors that facilitate chromatin transcription.

A model for archaeal RNA polymerase

In addition to the core subunits, archaeal RNAPs have between five and seven subunits, depending on the species [22,23]. As archaeal homologs have been reported for all yRNAP subunits except Rpb8 (Table 1), the overall structure of archaeal RNAPs must be very similar to the yRNAP structure. The archaeal enzymes, however, lack some external domains. For example, the domain of Rpb5 that binds the largest subunit is present in archaea, whereas the external domain of Rpb5 is not [24].

Models of the elongation complex

For RNA chain extension, RNAPs form a stable elongation complex, in which the downstream DNA duplex (the incoming DNA) is unwound to form a transcription bubble (Figure 4a). The bubble contains the DNA–RNA hybrid. At one end of the hybrid, the growing RNA end is engaged with the active site. At the other end, the hybrid strands separate and the two DNA strands combine to form the upstream duplex. A functional model of the elongation complex was derived earlier from biochemical and genetic data ([25] and references therein).

The high-resolution RNAP structures and additional data provided a more detailed picture of the elongation complex and its properties. Site-specific cross-linking of RNAPs to nucleic acids in elongation complexes allowed the modeling of DNA and RNA onto the bRNAP structure [26•] and onto a backbone model of yRNAP [27]. With electron crystallographic data on a Pol II elongation complex [28], nucleic acids were also placed onto the yRNAP backbone model [4••]. In the models, downstream DNA enters RNAP between two mobile 'jaws' (yRNAP only) and extends through the cleft towards the active site. Beyond the active site, the DNA–RNA hybrid extends upwards, at an angle of almost 90° to the incoming DNA, towards the wall/flap. The growing RNA end is located above the pore/secondary channel, allowing the entry of nucleoside triphosphates (NTPs) during RNA synthesis from below.

Structure of an elongation complex

The above models were generally confirmed with the spectacular X-ray structure of a Pol II elongation complex

(yRNAP-EC; Figure 4b) [6••]. The complex was obtained by the transcription of a DNA with a single-strand extension in the presence of only three NTPs, leading to pausing at a discrete site [6••,29]. The yRNAP-EC structure shows the exact location of the DNA and RNA phosphate groups and bases in the active center. The incoming DNA in the cleft is badly ordered, but three nucleotides before the active site, the DNA template strand becomes well ordered by binding to the bridge helix and to two 'switch' regions at the base of the clamp (Figure 4; see below). A 90° twist between subsequent nucleotides orients a DNA base towards the active site for base pairing with an RNA nucleotide. This base pair is the first of nine base pairs in the DNA–RNA hybrid emanating from the active site. The hybrid length agrees with that determined biochemically [30,31].

Discovery of the clamp

Comparison of a 6 Å X-ray map of yeast Pol II with its 16 Å electron microscopic envelope revealed a deviation, suggesting the presence of a mobile domain [7]. A subsequent backbone model of yeast Pol II suggested that the mobile domain would retain DNA in the cleft by acting as a 'clamp' [4••]. The atomic yRNAP structures in two crystal forms then trapped the clamp in two different open states [5••]. The yRNAP-EC structure revealed the clamp in a closed state, apparently induced by nucleic acid binding [6••]. A dramatic 30° rotation of the clamp occurs along with binding of the DNA template strand to three out of five 'switch' regions. These switches link the clamp to the remainder of RNAP. The switches are mobile in free yRNAP and are folded upon nucleic acid binding and clamp closure.

Clamp closure and processivity

The high stability of RNAP elongation complexes prevents the dissociation of RNAP from DNA, to allow the efficient transcription of very long genes (processivity). This stability is mainly caused by the tight binding of the DNA–RNA hybrid to RNAP [31,32]. In the yRNAP-EC structure, the hybrid is imbedded in a highly complementary binding site, created by the closure of the clamp and the folding of the switches. Because hybrid binding to the folded switches stabilizes the closed state, the switches couple clamp closure to the presence of the DNA–RNA hybrid. This coupling ensures DNA retention during transcription, when RNA is present, and DNA release after termination, when RNA is absent. Thus, the molecular basis for processivity seems to be the stabilization of the closed state of the RNAP clamp, coupled to the presence of RNA by the switches.

The clamp and switches are structurally conserved in bRNAP (except the clamp head; Figures 3 and 4) [5••]. The clamp in the bRNAP structure is almost as far closed as in the yRNAP-EC structure, and the switches are in a similar conformation, even in the absence of nucleic acids. Thus, weak crystal lattice forces apparently suffice to trap the mobile clamp in open and closed states. Easy conversion between the open and closed states seems to be important during transcription initiation, as the initial loading of promoter DNA may require clamp opening [5••], but the transition to elongation requires clamp closure [6••].

Maintenance of the transcription bubble

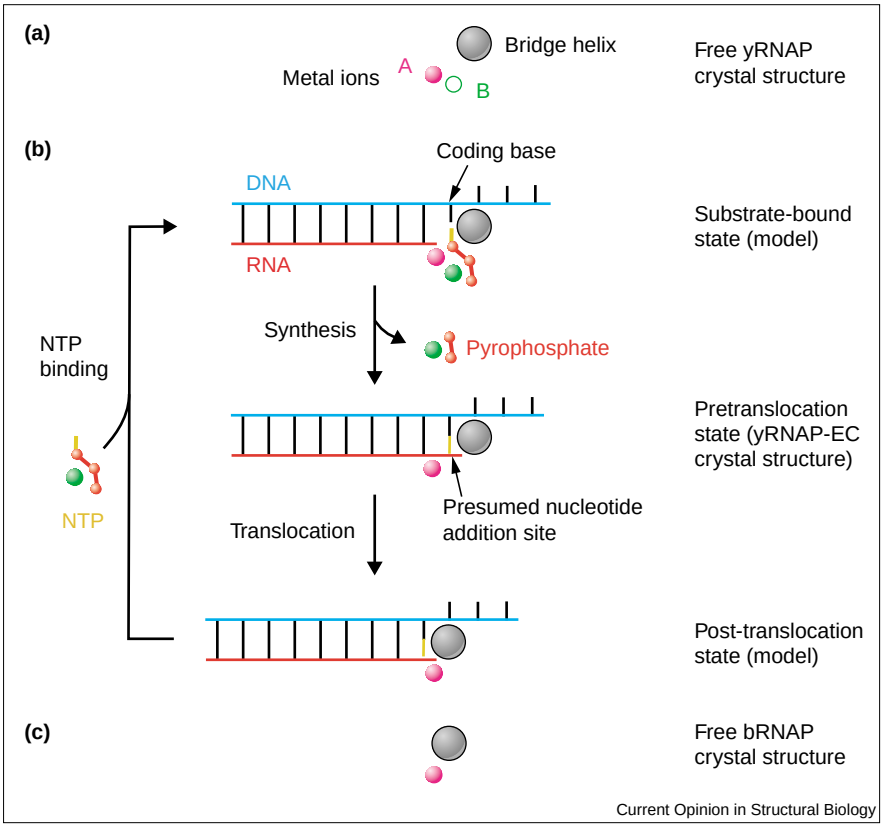
Three loops protrude from the clamp that may maintain the arrangement of nucleic acids during elongation (Figure 4). A 'rudder', first noted in the bRNAP structure, may separate RNA from DNA at the upstream end of the hybrid [1,5••,6••]. Two other loops, the 'lid' and 'zipper', could maintain the upstream end of the transcription bubble [5••]. The lid and zipper are likely to be conserved in bRNAP [5••], but are disordered in the bRNAP structure. They may cooperate in bubble maintenance with a loop on top of the wall/flap ('flap tip', 'flap loop'; Figure 4b) [5••], which binds to nascent RNA hairpins that pause or terminate bacterial transcription [33,34]. RNA hairpins generally do not affect eukaryotic transcription, maybe because the flap tip in eukaryotic RNAPs is six residues shorter. Maintenance of the downstream edge of the bubble may be attributed to the binding of the DNA template strand and to the blocking of the path of the nontemplate strand by another RNAP loop ('fork loop 2' in yRNAP, Figure 4b; 'βD loop II' in bRNAP).

The DNA nontemplate strand is disordered in the yRNAP-EC structure, maybe because the complex lacks the upstream DNA duplex and a complete bubble. It is likely that additional nucleic acid and protein elements are ordered in a natural elongation complex. The location of the nontemplate strand may change during transcription. During the initiation of bacterial transcription, a conserved coiled coil at the upper corner of the clamp (Figure 4) promotes the binding of the nontemplate strand by the bacterial RNAP cofactor σ70 [35]. A region adjacent to this coiled coil was cross-linked to the nontemplate strand in a complex with a DNA bubble [36]; however, another cross-linking study places the nontemplate strand on the other side of the cleft, into a depression between the two domains of the second largest subunit [37•], a location that was also detected in elongation complexes [26•].

DNA loading by initiation factors

To bind and open promoter DNA, RNAPs require initiation factors (the σ factor in bRNAP and general transcription factors in yRNAP). There is evidence that initiation factors help to load DNA into the RNAP cleft by restraining the RNAP clamp. Protein–DNA cross-linking showed that DNA passes through the cleft of an archaeal RNAP initiation complex [38•]. The σ factor binds to the coiled-coil region on the upper edge of the clamp [39,40] and may hold the clamp in a defined state. Given the sequence similarity of some eukaryotic general transcription factors to σ, it is possible that such a mechanism is conserved in eukaryotes. It is likely that initiation factors will also dock to other sites on the RNAP surface, to align the start site of transcription with the active site.

Figure 5



RNAP active site and RNA synthesis. (**a**) Schematic presentation of two metal ions and the bridge helix in the active site of yRNAP. The site for metal B shows low occupancy (indicated by a green circle). (**b**) Working model of the nucleotide addition cycle. The base of the added NTP substrate is in orange. Other colors are as in previous figures. The model for the post-translocation state assumes a different location of the bridge helix, as observed in the bRNAP structure (**c**).

For an initial RNAP–DNA complex ('closed complex'), two principal paths for promoter DNA through the RNAP cleft have been suggested [5••]. The DNA could pass over the wall/flap or between the wall/flap and the open clamp. The latter option appeared more appealing because the DNA could reach the floor of the cleft without distortion, whereas the other path would require a severe distortion of the DNA highly suspended above the active center. Such a distortion would, however, easily be achieved after DNA melting ('open complex') around the single-stranded regions.

Active site

Three invariant aspartate residues in a conserved amino acid motif in the largest subunit are located at the pore/secondary channel and bind a Mg^{2+} ion in the active site [1,4••,5••,41]. A second metal site, several angstroms further into the pore, is detected at the high resolution of the yRNAP structure, despite its low occupancy (Figure 5) [5••]. The second metal ion is located between the aspartate motif in the largest subunit and two conserved acidic residues in the second largest subunit. The low occupancy of the second metal site suggests that the metal ion is readily exchangeable. The observation of two metal-binding sites suggests that the two metal ion mechanism used by single subunit polymerases [42,43] may apply to RNAPs. The metal ions are accessible from one side and their location is consistent with the geometry of substrate binding observed in the yRNAP-EC structure (Figure 4) and in X-ray structures of nucleic acid complexes of single subunit DNA polymerases [44–47]. Following the nomenclature for single subunit polymerases, the first and second metal ions are referred to as metals A and B, respectively.

Nucleotide addition cycle and RNA synthesis

A working model for the nucleotide addition cycle during RNA chain elongation can be suggested based on the locations of metal ions A and B in the free yRNAP structure, DNA and RNA in the yRNAP-EC structure, and NTP substrate in the single subunit DNA polymerase structures (Figure 5). According to this model, the NTP substrate enters together with metal B and binds between the last base pair of the DNA–RNA hybrid and the bridge helix to form a base pair with the 'coding' DNA base (Figure 5b). Correct orientation of the substrates and metal ions would lead to the synthesis of a new phosphodiester bond and the release of pyrophosphate, maybe together with metal B. The resulting pretranslocation state was apparently trapped in the yRNAP-EC crystal structure, in which the RNA 3′-terminal nucleotide seems to be in the binding site for the incoming NTP ('nucleotide addition site'; Figure 5b) [6••]. The subsequent translocation of nucleic acids would align the new RNA 3′ end with metal A, ready for another cycle of RNA synthesis. In the free bRNAP structure, the bridge helix is in a different conformation, such that it comes closer to metal A (Figure 5c). This difference led to speculation that the translocation of

nucleic acids may be accompanied by a conformational change in this helix [2,5••,6••]. A corresponding 'O helix' in the active center of single subunit polymerases also stacks against the template–product nucleic acid duplex and can also undergo a conformational change [48].

Although the location of the 3′-terminal RNA nucleotide in the pretranslocated yRNAP-EC structure should reflect the approximate location of the incoming nucleotide, the chemical nature of the nucleotide addition site remains to be determined crystallographically. Incoming NTP may bind slightly differently because of the presence of two additional phosphate groups. This could allow hydrogen bond formation between the NTP 2′-OH group and the nearby conserved asparagine residue in the RNAP aspartate motif, rendering RNAPs specific for synthesizing RNA rather than DNA [6••]. Specificity for RNA may also be attributed to indirect recognition of the DNA–RNA hybrid conformation, which is intermediary between canonical A and B forms [6••].

RNA polymerase inhibitors

The RNAP structures may be used for structure-based drug design to explore species-specific inhibition of RNAPs. In particular, bacterial RNAP is a good target for antibiotics. The antibiotic rifampicin inhibits the elongation of bacterial RNAP. A co-crystal structure of bRNAP with rifampicin revealed that the inhibitor directly blocks the path of RNA in the DNA–RNA hybrid [3••]. Eukaryotic Pol II is specifically inhibited by α-amanitin, the toxin of the 'death cap' mushroom. A co-crystal structure of yRNAP with α-amanitin reveals that this inhibitor binds at a location where it can not directly interfere with nucleic acid binding or substrate access (DA Bushnell, P Cramer, RD Kornberg, unpublished data). Instead, α-amanitin is thought to act indirectly, by preventing conformational changes that underlie translocation.

Proofreading

RNAPs are able to detect the incorporation of incorrect nucleotides into the nascent RNA and to correct the error (proofreading). An error during transcription results in a mismatch base pair within the DNA–RNA hybrid. The induced distortion of the hybrid destabilizes the elongation complex and favors the reverse movement of RNAP ('backtracking') [30]. During backtracking, the RNA 3′ end disengages from the active site and a short stretch of RNA that comprises the misincorporated nucleotide is likely to be extruded through the pore/secondary channel into the funnel (Figure 4). The bridge helix will separate RNA from the DNA strand at the downstream edge of the hybrid during backtracking (Figure 4). Transcript cleavage factors (the Gre proteins in bacteria and SII/TFIIS in the Pol II system) bind to the outer rim of the funnel and may reach the RNAP active site from below, provoking the removal of the RNA stretch that comprises the misincorporated nucleotide [4••,5••,26•]. This would create a new RNA 3′ end at the active site, from which transcription can resume.

Structural dynamics and regulation

Comparison of the different RNAP structures has revealed mobile elements and conformational changes that are important for function. The most prominent mobile element, the clamp, is closed during elongation, but may open for initiation. Initiation and elongation factors may bind to the clamp and other mobile modules [5••] to regulate RNAP function. Special sequences in nascent RNA can form structures that may also bind to the clamp for RNAP regulation. The *putL* RNA from bacteriophage HK022 appears to bind to a region in the clamp [49] and may stabilize the closed state of the clamp to facilitate readthrough of termination sites. Concerted conformational changes in several RNAP elements may be required for the translocation of nucleic acids, in addition to the suggested bending of the bridge helix (see above).

Conclusions

Structural studies of RNAPs have revealed their general architecture, a conserved RNAP core and RNAP–nucleic acid interactions during transcription elongation. Specific functions for RNAP structural elements could be proposed and can now be tested by site-directed mutagenesis. The RNAP clamp plays a key role in processivity and bubble maintenance, and is a target of RNAP-binding factors and nascent RNA structures that regulate transcription. Future challenges include structural work on RNAP complexes involved in initiation, elongation and termination. An understanding of the transitions between these different phases of the transcription cycle will set the stage for mechanistic studies of transcriptional gene regulation.

Update

The only Pol II components that are not present in the yRNAP crystal structure are the two small subunits Rpb4 and Rpb7. The recent crystal structure of the archaeal homologs of these two subunits reveals a heterodimer [50]. The Rpb7 homolog, most probably present in all eukaryote RNAPs (Table 1), has an elongated two-domain structure and contains two potential RNA-binding motifs, suggesting that the Rpb4–Rpb7 complex binds near the proposed RNA exit site on the Pol II surface. The complete structure of Pol II will be needed to prove this model.

Acknowledgements

I would like to thank Roger Kornberg and his co-workers for many inspiring discussions, and Dirk Eick, Sabine Därr and Karim Armache for critical reading of the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA: Crystal structure of ***Thermus aquaticus*** core RNA polymerase at 3.3 Å resolution. *Cell* 1999, 98:811-824.

2. Darst SA: Bacterial RNA polymerase. *Curr Opin Struct Biol* 2001, 11:155-162.

3. •• Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA: Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* 2001, 104:901-912.
The co-crystal structure of bRNAP complexed with the inhibitor rifampicin explains the mechanism of action of this antibiotic.

4. •• Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD: Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 2000, 288:640-649.
The first high-resolution crystallographic backbone model of a eukaryotic RNAP reveals the subunit architecture and the location of the active site, and suggests a model for the arrangement of nucleic acids during elongation.

5. •• Cramer P, Bushnell DA, Kornberg RD: Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 2001, 292:1863-1876.
Two refined atomic structures of yeast Pol II reveal chemical details of the enzyme, two metal ions at the active site, two different open states of the clamp and functional elements that are not visible in the lower resolution structures. Structure-based sequence alignments and a detailed catalog of structural elements serve as a guide for future studies of the transcription mechanism.

6. •• Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD: Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 2001, 292:1876-1882.
The structure of an RNAP trapped 'in the act' of transcription. The structure reveals the location of phosphate groups and bases of DNA and RNA in the active center and a closed conformation of the clamp. It provides insight into many aspects of the transcription mechanism. The results also demonstrate that crystal structures of RNAP complexes can now be solved with the use of the atomic Pol II structures [5••] and molecular replacement.

7. Fu J, Gnatt AL, Bushnell DA, Jensen GJ, Thompson NE, Burgess RR, David PR, Kornberg RD: Yeast RNA polymerase II at 5 Å resolution. *Cell* 1999, 98:799-810.

8. Bushnell DA, Cramer P, Kornberg RD: Selenomethionine incorporation in ***Saccharomyces cerevisiae*** RNA polymerase II. *Structure Fold Des* 2001, 9:R3-R9.

9. Nouraini, S, Archambault J, Friesen JD: Rpo26p, a subunit common to yeast RNA polymerases, is essential for the assembly of RNA polymerases I and II and for the stability of the largest subunits of these enzymes. *Mol Cell Biol* 1996, 16:5985-5996.

10. Mukherjee K, Chatterji D: Studies on the omega subunit of ***Escherichia coli*** RNA polymerase – its role in the recovery of denatured enzyme activity. *Eur J Biochem* 1997, 247:884-889.

11. Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K: Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proc Natl Acad Sci USA* 2001, 98:892-897.

12. Sweetser D, Nonet M, Young RA: Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. *Proc Natl Acad Sci USA* 1987, 84:1192-1196.

13. Jokerst RS, Weeks JR, Zehring WA, Greenleaf AL: Analysis of the gene encoding the largest subunit of RNA polymerase II in ***Drosophila***. *Mol Gen Genet* 1989, 215:266-275.

14. Zhang G, Darst SA: Structure of the ***Escherichia coli*** RNA polymerase alpha subunit amino-terminal domain. *Science* 1998, 281:262-266.

15. Severinov K: RNA polymerase structure–function: insights into points of transcriptional regulation. *Curr Opin Microbiol* 2000, 3:118-125.

16. Cheong JH, Yi M, Lin Y, Murakami S: Human RPB5, a subunit shared by eukaryotic nuclear RNA polymerases, binds human hepatitis B virus X protein and may play a role in X transactivation. *EMBO J* 1995, 14:143-150.

17. Tan Q, Linask KL, Ebright RH, Woychik NA: Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. *Genes Dev* 2000, 14:339-348.

18. Ebright RH: RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J Mol Biol* 2000, 304:687-698.

19. Hirose Y, Manley JL: RNA polymerase II and the integration of nuclear events. *Genes Dev* 2000, 14:1415-1429.

20. Proudfoot N: Connecting transcription to messenger RNA processing. *Trends Biochem Sci* 2000, 25:290-293.

21. Steinmetz EJ: Pre-mRNA processing and the CTD of RNA polymerase II: the tail that wags the dog? *Cell* 1997, 89:491-494.

22. Langer D, Hain J, Thuriaux P, Zillig W: Transcription in archaea: similarity to that in eukarya. *Proc Natl Acad Sci USA* 1995, 92:5768-5772.

23. Darcy TJ, Hausner W, Awery DE, Edwards AM, Thomm M, Reeve JN: ***Methanobacterium thermoautotrophicum*** RNA polymerase and transcription ***in vitro***. *J Bacteriol* 1999, 181:4424-4429.

24. Todone F, Weinzierl RO, Brick P, Onesti S: Crystal structure of RPB5, a universal eukaryotic RNA polymerase subunit and transcription factor interaction target. *Proc Natl Acad Sci USA* 2000, 97:6306-6310.

25. Nudler E: Transcription elongation: structural basis and mechanisms. *J Mol Biol* 1999, 288:1-12.

26. • Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A, Darst SA: A structural model of transcription elongation. *Science* 2000, 289:619-625.
Extensive protein–nucleic acid cross-linking provides a detailed model of the bacterial RNAP transcription elongation complex and suggests structural elements with functional roles.

27. Wooddell CI, Burgess RR: Topology of yeast RNA polymerase II subunits in transcription elongation complexes studied by photoaffinity cross-linking. *Biochemistry* 2000, 39:13405-13421.

28. Poglitsch CL, Meredith GD, Gnatt AL, Jensen GJ, Chang WH, Fu J, Kornberg RD: Electron crystal structure of an RNA polymerase II transcription elongation complex. *Cell* 1999, 98:791-798.

29. Gnatt A, Fu J, Kornberg RD: Formation and crystallization of yeast RNA polymerase II elongation complexes. *J Biol Chem* 1997, 272:30799-30805.

30. Nudler E, Mustaev A, Lukhtanov E, Goldfarb A: The RNA–DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 1997, 89:38-41.

31. Kireeva ML, Komissarova N, Waugh DS, Kashlev M: The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. *J Biol Chem* 2000, 275:6530-6536.

32. Sidorenkov I, Komissarova N, Kashlev M: Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol Cell* 1998, 2:55-64.

33. Toulokhonov I, Artsimovitch I, Landick R: Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins. *Science* 2001, 292:730-733.

34. Landick R: RNA polymerase clamps down. *Cell* 2001, 105:567-570.

35. Young BA, Anthony LC, Gruber TM, Arthur TM, Heyduk E, Lu CZ, Sharp MM, Heyduk T, Burgess RR, Gross CA: A coiled-coil from the RNA polymerase beta′ subunit allosterically induces selective nontemplate strand binding by sigma(70). *Cell* 2001, 105:935-944.

36. Brodolin K, Mustaev A, Severinov K, Nikiforov V: Identification of RNA polymerase beta′ subunit segment contacting the melted region of the lacUV5 promoter. *J Biol Chem* 2000, 275:3661-3666.

37. • Naryshkin N, Revyakin A, Kim Y, Mekler V, Ebright RH: Structural organization of the RNA polymerase–promoter open complex. *Cell* 2000, 101:601-611.
DNA in an RNAP open complex is modeled based on site-specific cross-linking.

38. • Bartlett MS, Thomm M, Geiduschek EP: The orientation of DNA in an archaeal transcription initiation complex. *Nat Struct Biol* 2000, 7:782-785.
Protein–DNA cross-linking places DNA in the central cleft of RNAP.

39. Arthur TM, Burgess RR: Localization of a sigma70 binding site on the N terminus of the ***Escherichia coli*** RNA polymerase beta′ subunit. *J Biol Chem* 1998, 273:31381-31387.

40. Arthur TM, Anthony LC, Burgess RR: Mutational analysis of beta′ 260-309, a sigma 70 binding site located on ***Escherichia coli*** core RNA polymerase. *J Biol Chem* 2000, 275:23113-23119.

41. Zaychikov E, Martin E, Denissova L, Kozlov M, Markovtsov V, Kashlev M, Heumann H, Nikiforov V, Goldfarb A, Mustaev A: Mapping of catalytic residues in the RNA polymerase active center. *Science* 1996, 273:107-109.

42. Beese LS, Steitz TA: Structural basis for the 3′-5′ exonuclease activity of ***Escherichia coli*** DNA polymerase I: a two metal ion mechanism. *EMBO J* 1991, 10:25-33.

43. Steitz TA: A mechanism for all polymerases. *Nature* 1998, 391:231-232.

44. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science* 1994, 264:1891-1903.

45. Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H: Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* 1997, 36:11205-11215.

46. Doublie S, Tabor S, Long AM, Richardson CC, Ellenberger T: Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* 1998, 391:251-258.

47. Franklin MC, Wang J, Steitz TA: Structure of the replicating complex of a pol alpha family DNA polymerase. *Cell* 2001, 105:657-667.

48. Li Y, Korolev S, Waksman G: Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of ***Thermus aquaticus*** DNA polymerase I: structural basis for nucleotide incorporation. *EMBO J* 1998, 17:7514-7525.

49. Sen R, King RA, Weisberg RA: Modification of the properties of elongating RNA polymerase by persistent association with nascent antiterminator RNA. *Mol Cell* 2001, 7:993-1001.

50. Todone F, Brick P, Werner F, Weinzierl ROJ, Onesti S: Structure of an archaeal homolog of the eukaryotic RNA polymerase II Rpb4/Rpb7 complex. *Mol Cell* 2001, 8:1137-1143.