Bacterial RNA polymerase

Seth A Darst

The recently determined crystal structure of a bacterial core RNA polymerase (RNAP) provides the first glimpse of this family of evolutionarily conserved cellular RNAPs. Using the structure as a framework, a consistent picture of protein–nucleic acid interactions in transcription complexes has been accumulated from cross-linking experiments. The molecule can be viewed as a molecular machine, with distinct structural features hypothesized to perform specific functions. Comparison with the α-carbon backbone of a eukaryotic RNAP reveals close structural similarity.

Addresses
The Rockefeller University, Box 224, 1230 York Avenue, New York, NY 10021, USA; e-mail: darst@rockefeller.edu

Current Opinion in Structural Biology 2001, 11:155–162

0959-440X/01/$ – see front matter

Abbreviations

bp	base pairs
EM	electron microscopy
nt	nucleotides
NTD	N-terminal domain
RNAP	RNA polymerase
TEC	ternary elongation complex

Introduction

RNA in all cellular organisms is synthesized by a complex molecular machine, the DNA-dependent RNA polymerase (RNAP). In its simplest bacterial form, the enzyme comprises at least four subunits, with a total molecular mass of around 400 kDa. The eukaryotic enzymes comprise upwards of a dozen subunits, with a total molecular mass of around 500 kDa. The catalytically competent bacterial core RNAP (subunit composition $\alpha_2\beta\beta'\omega$) is evolutionarily conserved in sequence, structure and function from bacteria to man [1,2•,3,4••]. Sequence conservation points to structural and functional homologies, making the simpler bacterial RNAPs excellent model systems for understanding the cellular RNAPs in general.

The basic elements of the transcription cycle were elucidated through study of the prokaryotic system. In both prokaryotes and eukaryotes, promoter-specific initiation of transcription requires protein factors in addition to the catalytic core RNAP [5]. In bacteria, specific transcription initiation by RNAP requires a single polypeptide, the σ factor, which binds to the core RNAP to form the holoenzyme [6]. The holoenzyme locates specific DNA sequences called promoters within the double-stranded DNA, forms the open complex by melting the DNA surrounding the transcription start site and initiates the synthesis of an RNA chain. Once the RNA chain becomes 10 nucleotides (nt) or so, σ is released [7] and the core RNAP elongates the RNA processively while translocating itself and the melted transcription bubble along the DNA template. RNAP finally releases itself and the completed transcript after encountering a specific termination signal.

Since the initial isolation of the bacterial enzyme [8], a wealth of biochemical, biophysical and genetic information has accumulated on RNAP and its complexes with nucleic acids and accessory factors [9–12]. An essential step to aid incorporation of this information into a complete understanding of the transcription mechanism and its regulation is to determine three-dimensional structures of RNAP and its complexes with DNA, RNA and regulatory factors at different stages of the transcription cycle.

Low-resolution structures (in the 25–12 Å range) of cellular RNAPs, provided by electron microscopy (EM) and image processing, reveal a molecule shaped like a crab claw, with a groove or channel that is an appropriate size for accommodating double-helical DNA [13–18,19•,20]. The first steps towards high-resolution structures of cellular RNAPs have recently been taken, providing a 3.3 Å resolution structure of a bacterial core RNAP [4••,21•] (Figure 1) and an α-carbon backbone of a eukaryotic RNAP from X-ray analysis at 3 Å resolution [22••].

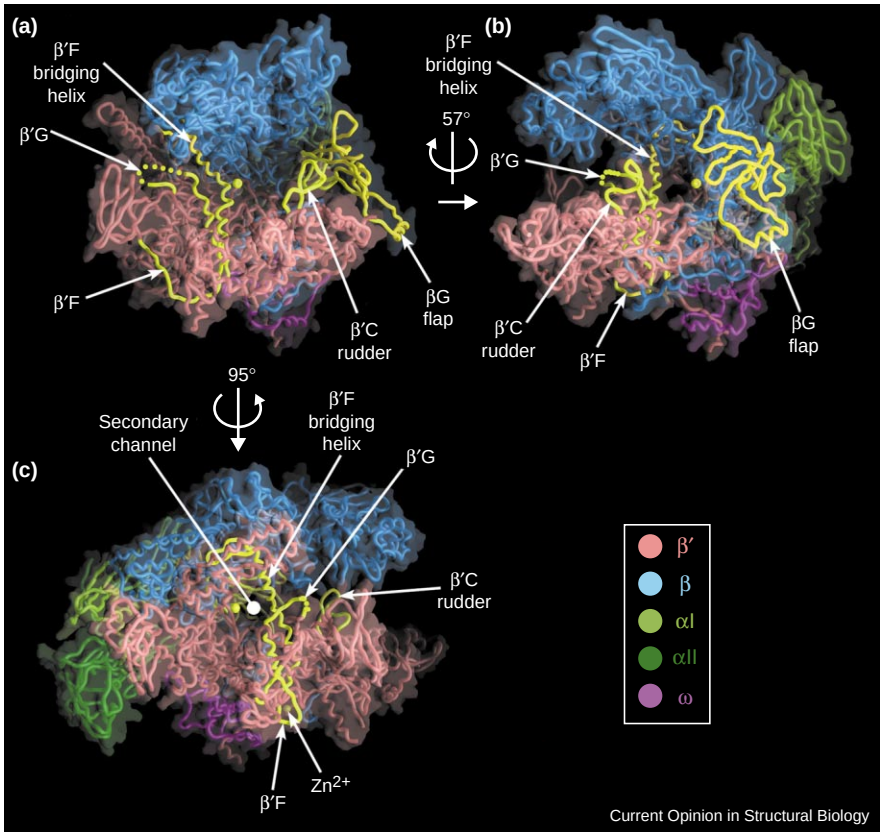
The new structural information promises to have a major impact on the transcription field, guiding functional and structural investigations at a level of unprecedented detail. For instance, cross-linking approaches have been used to generate detailed models of complexes between RNAP and nucleic acids [23••–25••]. The new structures also reveal similarities between the prokaryotic and eukaryotic polymerases that, although predicted from sequence analyses [1] and previously revealed at low resolution [16], extend to a much greater degree than expected.

Structure of the bacterial RNA polymerase

By far the best-characterized cellular RNAP is that from *Escherichia coli*. But, despite years of effort by many groups, three-dimensional crystals of *E. coli* RNAP suitable for X-ray analysis have never been obtained. We expanded our investigations to thermophilic eubacteria, focusing on *Thermus aquaticus* (*Taq*) because of the relative ease with which large amounts of biomass could be obtained.

The core RNAP isolated from *Taq* comprised four distinct polypeptides of about 170, 125, 35 and 11 kDa. There was no sequence information available for the *Taq* RNAP subunits, so the three largest subunits were cloned and sequenced, identifying them as β′, β and α [4••]. The smallest polypeptide was tentatively identified as ω [4••] and this was confirmed by subsequent cloning and sequencing [2•]. The *Taq* RNAP subunits contain all of the expected conserved regions (Figure 2a).

Figure 1



The *Taq* core RNAP structure. The α-carbon backbone is shown as worms, along with the transparent molecular surface. Various structural features discussed in the text are labeled and highlighted in yellow. The individual subunits are color-coded as indicated in the key. The Mg^{2+} ion chelated at the active center is indicated by a yellow sphere. The Zn^{2+} ion bound in β′ is indicated by a light green sphere. (a) View looking into the active site channel, perpendicular to the axis of the main channel. This view is rotated (b) 57° clockwise about the vertical axis and (c) 95° counter-clockwise about the vertical axis (this view looks directly down the axis of the secondary channel).

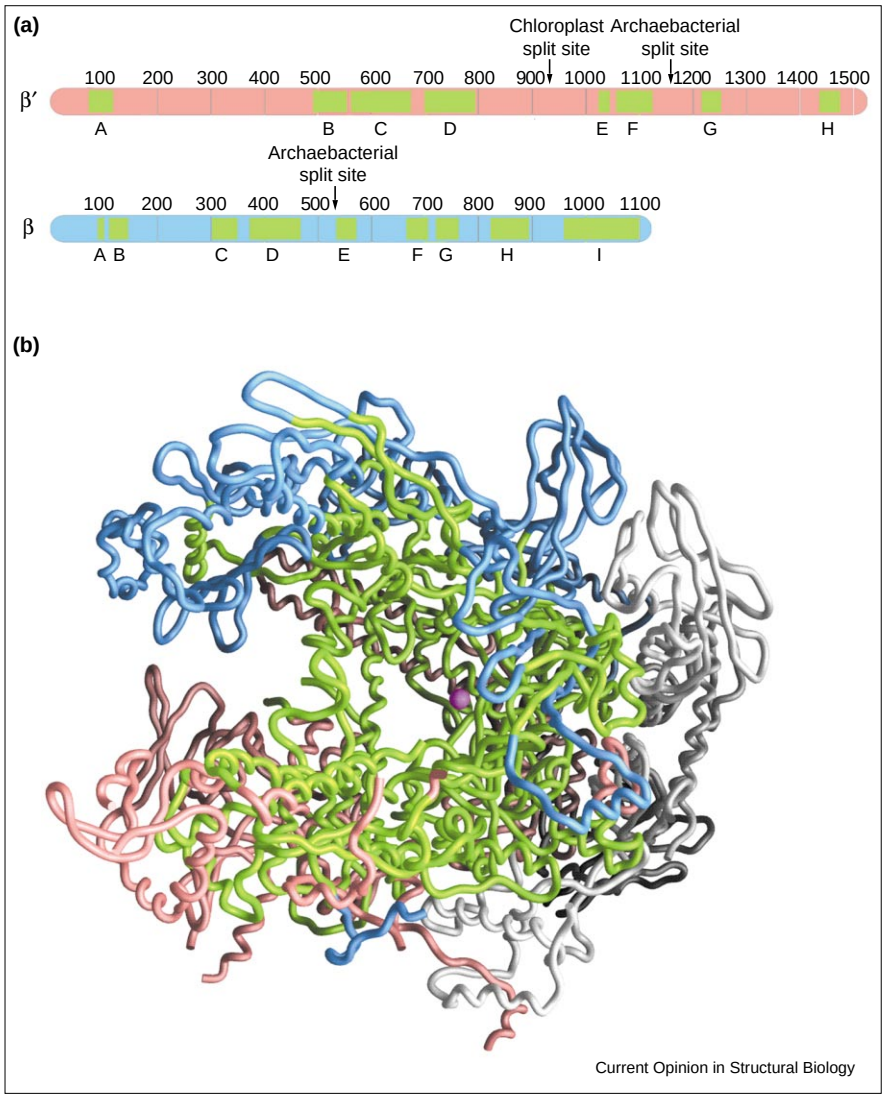
With the pure active enzyme [21•,26], crystals of *Taq* core RNAP were obtained relatively rapidly. The asymmetric unit contained one 380 kDa RNAP molecule. Even the frozen crystals were radiation sensitive and the diffraction was anisotropic. Useful diffraction typically extended to 3.2 Å × 3.4 Å at undulator beamlines. The structure was solved by the method of multiple isomorphous replacement, using a combination of metal clusters and conventional derivatives [4••]. There were four important milestones in determining the structure. First, the recognition within the experimental maps of the fold of the α subunit N-terminal domain (NTD) dimer (the structure from *E. coli* was previously solved in my laboratory [3]), which allowed the objective assessment of further attempts to improve the maps. Second, the use of SOLOMON [27] for density modification, which dramatically improved the experimental map. Third, the preparation and crystallization of selenomethionyl *Taq* core RNAP, which allowed the localization of methionine residues from the resulting Fourier difference peaks. Finally, the identification of the active site Mg^{2+} chelated by the three aspartic acid residues of an absolutely conserved NADFDGD motif in β′ conserved region D (β′D; see Figure 2a) from a single-site Pb^{2+} derivative (it was shown that Pb^{2+} binds with high affinity in the same site as Mg^{2+}; A Mustaev, personal communication).

The current model (Figure 1) [21•] contains essentially complete β and ω subunits, about 75% of the β′ subunit, the α NTD dimer, one Mg^{2+} ion (chelated at the active center) and one Zn^{2+} ion, important for the folding of β′ [28]. Regions lacking electron density that are presumably disordered in the crystal include both α C-terminal domains [29], as well as a 38-residue segment of β′ that includes a second Zn^{2+}-binding motif near β′A. In addition, a nonconserved domain of 300 residues inserted between β′A and β′B (Figure 2a) is currently not modeled. The structure brings into focus a wide range of independent biochemical, biophysical and genetic data that is available in the literature (summarized in [4••]).

The shape and size of the *Taq* core RNAP X-ray structure (Figure 1) correspond extremely well with the structure of the *E. coli* core RNAP obtained using cryo-EM [18]. The shape is reminiscent of a crab claw, with an internal channel running along its full length (between the claws). One arm of the claw is primarily β, the other primarily β′. The molecule is about 150 Å long (from the back to the tips of the claws), 115 Å tall and 110 Å wide (parallel with the channel). The channel has many internal features, but the overall width is approximately 27 Å.

Evolutionary relationships for each of the bacterial core subunits have been identified among all organisms from bacteria to man [2•,3,30–32]. These relationships are particularly strong between the large β′ and β subunits, which contain colinearly arranged segments of conserved

Figure 2



Evolutionarily conserved segments of the RNAP large subunits form the structural core. (a) The bars schematically represent the primary sequences of the *Taq* β′ (top, pink) and β (bottom, blue) subunits, with amino acid numbering directly above the bars [4••]. The split sites of the large subunits, which occur as two separate peptides in chloroplast and archaebacterial RNAPs, are also indicated. Evolutionarily conserved regions among prokaryotes, chloroplasts, archaebacteria and eukaryotes [4••], originally defined for β′ by Jokerst *et al.* [31] and for β by Sweetser *et al.* [32], are denoted as light green boxes and are labeled (A–H for β′, A–I for β) directly below. (b) Backbone worm diagram of *Taq* core RNAP (same view as Figure 1b). The two α monomers and ω are colored white. The large subunits are colored pink (β′) and blue (β), except for the evolutionarily conserved segments shown in (a), which are colored light green. The active site Mg^{2+} ion is shown as a magenta sphere.

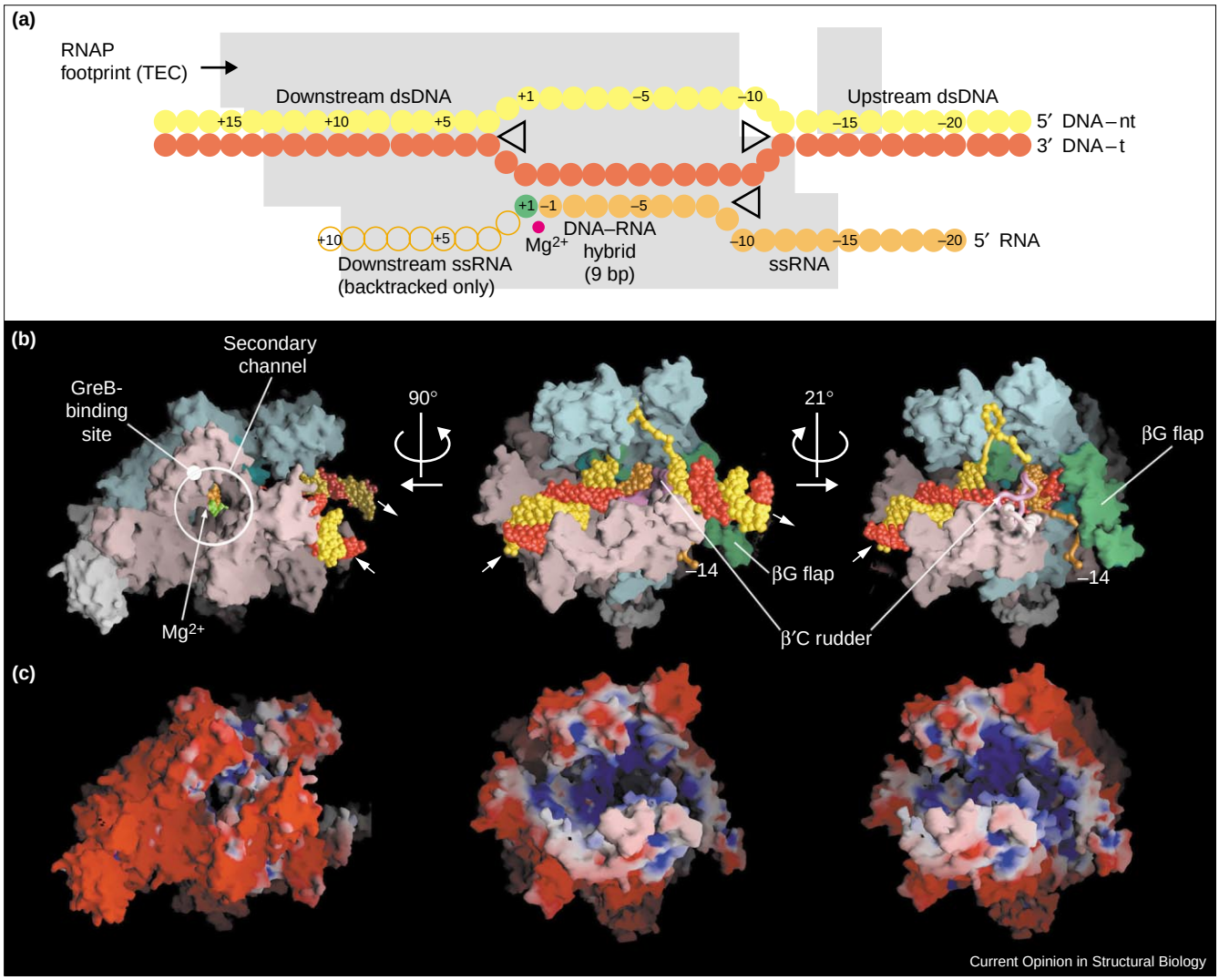
sequence (Figure 2a). Although these segments are distributed throughout the primary sequences, they all converge around the active center Mg^{2+} in the folded assembly (Figure 2b). Thus, we expect this central 'core' of the molecule (defined by a roughly 80 Å diameter sphere surrounding the active center Mg^{2+}) to have an identical fold in RNAPs from all other organisms and anticipate that the additional subunits unique to the archaeal and eukaryotic RNAPs would decorate the outer, relatively nonconserved, rim of the molecule (see below).

Considering the RNAP as a molecular machine, distinct structural features or 'parts' that appear to be important for function can be described (denoted in Figure 1). One feature, which includes βG, forms a flap-like structure that appears to be flexibly connected to the rest of the RNAP (the βG flap). Region F of β′ is most remarkable; it begins in the upper domain of β′, where it interacts with β, and then forms a helical segment that traverses the middle of the main channel, bridging the top and bottom of the main channel (the β′F bridging helix). The active center Mg^{2+} is positioned on the back wall of the main channel, directly opposite the bridging helix. The β′F bridging helix conspires with β′G, which extends into the main channel, to form a structure that bifurcates the main RNAP channel into two separate channels. The smaller of these channels (the secondary channel) is roughly 12 Å in diameter, not large enough to accommodate double-stranded nucleic acid. Finally, a coiled-coil-like structure extends from the main channel and supports a loop-like structure (comprising part of β′C) that protrudes into the main channel, reminiscent of an upside-down rudder.

Models of the ternary elongation complex

Decades of functional studies [33–36] and more recent structural evidence [4••,22••,25••,37] indicate that cellular RNAPs operate as complex molecular machines, with extensive and dynamic interactions with the template

Figure 3



The TEC model and RNAP charge distribution. **(a)** Nucleic acids within the TEC are schematically illustrated, with transcription going from right to left. Solid circles represent the DNA (red: template strand, t; yellow: nontemplate strand, nt) and RNA (orange) chains. The incoming nucleotide substrate is green. Open circles show the segment of single-stranded downstream RNA in the backtracked TEC. The positions of nucleotide residues are numbered relative to the position of the incoming nucleotide substrate, which is denoted +1. The 12–14 base transcription bubble is shown [50–53]. Regions of the DNA and RNA protected by the RNAP in chemical and enzymatic footprinting studies are indicated by gray shading [52–57]. The black triangles indicate that the RNAP itself actively controls the length of the transcription bubble and the RNA–DNA hybrid. **(b)** Three views of the TEC model. The model consists of three components: the *Taq* core RNAP crystal structure [4••], shown as color-coded molecular surfaces (β, cyan; β′, pink; α and ω, white; catalytic Mg^{2+}, magenta sphere); the DNA template (template strand, red; nontemplate strand, yellow); and the RNA transcript (gold) plus incoming nucleotide substrate (green). The nucleic acid backbones are shown as worms with CPK phosphate atoms. For base-paired sections of the nucleic acids, the bases are also shown. The directions of the entering downstream DNA and exiting upstream DNA are indicated (large arrows). The middle structure shows a view perpendicular to the main active site channel, which runs roughly horizontal. Parts of the protein structure discussed in the text are colored (βG flap in green, β′C rudder in magenta). The structure on the left shows a view down the secondary channel, showing the path for diffusion of the incoming nucleotide substrate (green) into the active site. The catalytic Mg^{2+} ion (magenta sphere) is just visible to the left of the substrate nucleotide. The structure on the right shows a view roughly along the axis of the exiting upstream duplex DNA. For clarity, parts of the β′ subunit nearest the viewer are shown as a backbone worm without the corresponding molecular surface (the β′C rudder in magenta and its coiled-coil platform in pink) and the upstream duplex DNA has been removed. Parts of the protein structure discussed in the text are colored (βG flap in green). **(c)** Same three views of the RNAP, but with the surface shaded according to the electrostatic potential (red, acidic; white, neutral; blue, basic [48]).

DNA, product RNA [25••] and regulatory molecules. Using the *Taq* core RNAP structure as a framework, several groups have used systematic protein–nucleic acid cross-linking and mapping techniques to begin defining these interactions in detail. These studies have led to a consistent picture of the topography of the prokaryotic open

promoter complex [23••], the trajectory of the DNA in an archaeal pre-initiation complex [24••] and the disposition of the DNA and RNA in the ternary elongation complex (TEC) containing DNA template and RNA transcript (Figure 3) [25••].

In the TEC model, about 12 bp of downstream duplex DNA is nestled in a trough formed by β′ and enclosed on top by β. In the main RNAP channel, a 9 bp RNA–DNA hybrid (from +1 to –8; the register of the incoming nucleotide substrate is denoted '+1'; upstream positions, where RNAP has been, are denoted –1, –2, etc.; and downstream positions, where RNAP is going, are denoted +2, +3, etc.) forms a remarkable complementary fit with the protein structure. The hybrid extends from the enzyme active site and the β′F bridging helix at +1, along a region of the β subunit harboring rifampicin-resistance mutations [21•], to the βG flap and the β′C rudder. The single-stranded upstream RNA (from –9 to –14) is then extruded through the channel underneath the βG flap domain.

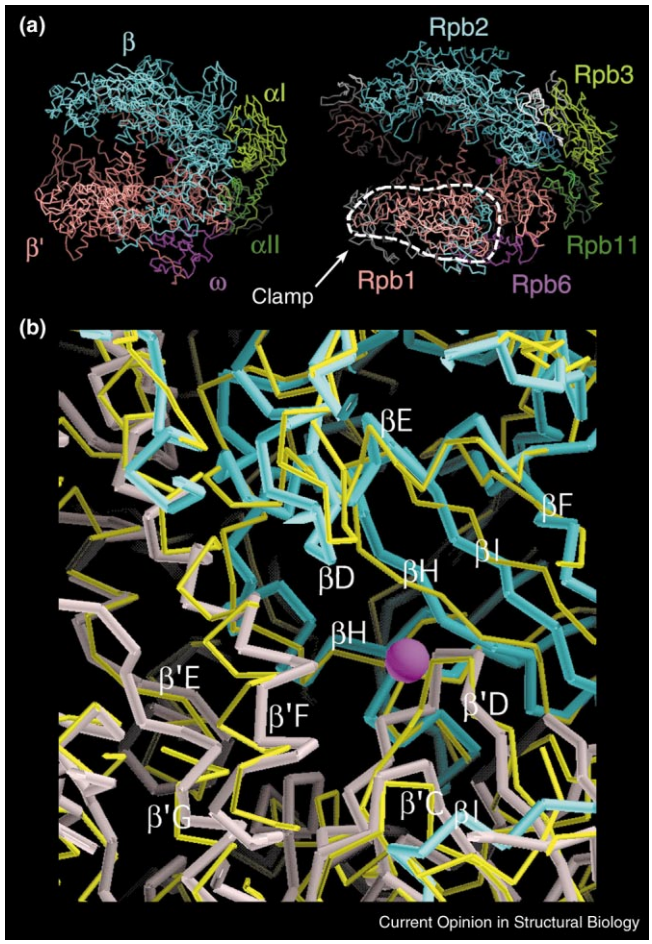
The RNA–DNA hybrid is 8–9 bp [38], but the transcription bubble spans 12–14 nt, leaving 3–6 nt of single-stranded template DNA upstream of the hybrid to reach the upstream side of the rudder and re-anneal with the nontemplate DNA. The nontemplate strand DNA within the bubble is held apart from the template DNA in a groove between two structural domains of the β subunit. The incoming downstream duplex DNA and the exiting upstream duplex DNA form a bend angle of about 90°, arising mainly from a kink between the downstream duplex and the RNA–DNA hybrid. A bend angle of 92° was experimentally determined by atomic-force microscopy of active TECs [39].

The transcription complex models are strongly supported by an examination of the electrostatic charge distribution around the RNAP molecule (Figure 3c). The cellular RNAPs tend to be very acidic overall, with pIs near 5, and the *Taq* RNAP is acidic (red) over most of its surface. The exceptions to this rule, where the surface of the molecule is neutral (white) or basic (blue), correspond exactly to the proposed locations of the nucleic acid elements within the TEC. This includes the underside of the flap domain, where the upstream single-stranded RNA is placed. These results probably rule out an alternative path proposed for the upstream RNA (path 1 of [22••]) as the protein surface of this path is very acidic.

Comparison between archaeal and eukaryotic RNA polymerases

Comparison of the *Taq* core RNAP structure [4••,21•] with the α-carbon backbone determined for yeast RNAP II [22••] reveals that the overall shape similarity known from low-resolution EM studies [16] extends to the overall architecture of the conserved subunits (Figure 4a). As predicted [4••], the six subunits unique to the crystallized yeast RNAP II (white in Figure 4a) are arranged around the outer rim of the conserved core. The highly conserved segments of the largest subunits converge around the active site of the enzymes (Figure 2); immediately surrounding the active site, the sequence homology approaches 100% [4••]. As a consequence, the folds of the two enzymes in this region are essentially identical (Figure 4b).

Figure 4



Taq RNAP versus yeast RNAP II. (a) Analogous views of the *Taq* core RNAP structure (left) [4••] and yeast RNAP II (right) [22••], but with reassignment of 40 C-terminal residues of Rpb1 based on refined atomic coordinates (P Cramer, D Bushnell, RD Kornberg, personal communication). Color coding for the *Taq* core RNAP is the same as in Figure 1. Yeast RNAP II homologs of the bacterial core subunits are color-coded similarly. Subunits unique to the crystallized yeast RNAP II (Rpb5, Rpb8, Rpb9, Rpb10 and Rpb12) are shown in white. The mobile 'clamp' domain [22••] is outlined. (b) The region around the RNAP active site, surrounding the Mg^{2+} ion (shown as a magenta sphere), is depicted. The backbone of β′ is shown in pale pink and that of β is shown in blue. The backbone of yeast RNAP II is shown in yellow and is aligned with the bacterial structure. The locations of some selected evolutionarily conserved segments are labeled.

Conclusions

From the close similarities between the bacterial core RNAP and RNAPs from higher organisms, we expect that most, if not all, of the basic functional mechanisms are

conserved. The new structural information and the resulting transcription complex models invite hypotheses as to how these functions are realized in terms of specific structural elements of the RNAP. Some of these proposed structure/function relationships are briefly described below.

The transcription bubble, the RNA–DNA hybrid and the β′C rudder

The lengths of the transcription bubble and of the RNA–DNA hybrid are controlled by the RNAP during transcript elongation [34]. The TEC model suggests that the rudder may participate in these functions, maintaining the upstream edge of the RNA–DNA hybrid by removing the RNA transcript from the DNA template strand, while also maintaining the upstream edge of the bubble.

Secondary channel function

It has been proposed that the secondary channel provides access to the active center for the NTP substrates, which cannot enter through the main channel as it is blocked by the DNA and RNA in the transcription complex [4••,*22*••,*25*••,40]. The structural basis for this hypothesis is illustrated in Figure 3b, showing a clear path for the incoming nucleotide (green) through the secondary channel.

RNAP can transiently move backward during transcription, which is important for proofreading and for traversing obstacles on the DNA. During backtracking, the RNAP and associated transcription bubble move backward along the DNA, while the RNA reverse threads through the complex. Another role for the secondary channel, to accommodate the 3′ proximal single-stranded RNA fragment generated by RNAP backtracking [38,41–43], has therefore been proposed [*22*••,*25*••,44]. The backtracked TEC is an obligate intermediate in factor-induced transcript cleavage [43] and transcript cleavage factors — GreA and GreB of prokaryotes and SII of eukaryotes, which are functional but not structural homologs — all bind at the outer entrance to the secondary channel, where the 3′ end of the RNA would emerge in a backtracked complex (Figure 3b) [*22*••,*25*••,45].

Conformational changes, the clamp and the bridging helix

Comparison of the TEC model with available footprinting data, as well as other considerations, suggested that there is a conformational change in the RNAP in the TEC compared with the crystal structure, namely more complete closure of the RNAP around the nucleic acids occupying the main channel and possible movements of the β′F bridging helix to better accommodate the RNA–DNA hybrid [25••]. Structural evidence points to conformational changes in the RNAPs and the similarities between the bacterial and eukaryotic enzymes may extend to conformational dynamics. From a comparison of the yeast RNAP II crystal structure with low-resolution maps derived from EM and image processing [14], a large conformational change in the enzyme was noted, involving the opening or closing of the active site channel through a mobile element (termed the clamp; Figure 4a) comprising an N-terminal segment of Rpb1, the most C-terminal domain of Rpb2 and Rpb6 [*22*••,40]. Similar flexibility of the bacterial RNAP (comprising the N-terminal segment of β′, the C-terminal domain of β and ω) is observed when comparing the X-ray crystal structure of *Taq* core RNAP with the low-resolution map of *E. coli* core RNAP derived from EM and image processing ([18]; SA Darst, N Opalka, A Polyakov, C Richter, G Zhang, personal communication).

Imagine a simple mechanical model consisting of a set of jaws hinged on one side, but bridged in the middle by a post attached to the top and bottom jaws. It is apparent that opening or closing of the jaws requires tilting of the bridging post, just as opening and closing of the RNAP around the main channel (via movements that include the clamp) would effect movement (such as tilting) of the bridging helix, suggesting that the conformational changes are linked. The β′F bridging helix may interact with the downstream end of the RNA–DNA hybrid [25••] and is highly conserved in sequence, especially where it passes near the active site Mg^{2+}. Conformational changes of this helix are likely to have profound consequences for the catalytic and translocation functions of the RNAP.

The βG flap and transcription termination

The 'canonical' bacterial intrinsic termination signal in DNA comprises a GC-rich dyad symmetry element followed by a run of thymine residues [46]. The resulting RNA transcript forms a hairpin, which is essential for termination. The region of the transcript that forms the hairpin (upstream of –7) is precisely the same region that interacts with the rudder and then passes under the βG flap. The formation of the bulky hairpin could initiate the breakdown of the TEC by altering the position of the βG flap and disrupting nucleic acid interactions with the rudder.

In summary, in the past year, long-awaited high-resolution structural information on cellular RNAPs has finally appeared. These structures are having a major impact on the transcription field by guiding the design and implementation of ever more incisive genetic, biochemical and biophysical experiments to probe the RNAP functional mechanism. Functional insights into transcription complexes of the RNAP and nucleic acids have come from structural models developed using extensive cross-link mapping data, leading to specific and testable predictions for structure/function relationships. Even more exciting are the prospects of solving structures of the transcription complexes at various stages of the transcription cycle [47]. In the end, a complete understanding of the transcription cycle will come from a combination of structural studies of trapped states, providing snapshots of the polymerase at various stages of its function, and functional studies using genetic, biochemical and biophysical probes informed by the accumulating structural data.

Acknowledgements

I thank past and present members of my laboratory, as well as our collaborators in the laboratories of Alex Goldfarb and Konstantin Severinov, for their many contributions. I thank Roger Kornberg and Patrick Cramer for providing the yeast RNAP II α-carbon backbone structure. The work in the author's laboratory was funded by the National Institutes of Health and the Pew Foundation. Figures 1, 2b and 3b,c were generated using GRASP [48]. Figure 4 was generated using O [49].

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. Archambault J, Friesen JD: Genetics of RNA polymerases I, II and III. *Microbiol Rev* 1993, 57:703-724.

2. • Minakhin L, Bhagat S, Brunning A, Campbell EA, Darst SA, Ebright RH, Severinov K: Bacterial RNA polymerase subunit ω 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.

This paper shows that the bacterial core RNAP subunit ω shares sequence, structural and functional homology with archaeal RNAP subunit RpoK and with an essential, shared subunit of eukaryotic RNAPs I, II and III (Rpb6). This settles the issue of whether ω is really an RNAP subunit, shows that it plays an important functional role in RNAP assembly and extends the evolutionary link between the prokaryotic and eukaryotic cellular RNAPs to all of the bacterial core RNAP subunits.

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

4. •• 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.

This paper describes the first X-ray crystal structure of a cellular RNAP – *Taq* core RNAP.

5. Burgess RR, Travers AA, Dunn JJ, Bautz EKF: Factor stimulating transcription by RNA polymerase. *Nature* 1969, 221:43-44.

6. Gross CA, Lonetto M, Losick R: Bacterial sigma factors. In *Transcriptional Regulation*. Edited by Yamamoto K, McKnight S. New York: Cold Spring Harbor Laboratory Press; 1992.

7. Travers AA, Burgess RR: Cyclic re-use of the RNA polymerase sigma factor. *Nature* 1969, 222:537-540.

8. Chamberlin M, Berg B: Deoxyribonucleic acid-directed synthesis of ribonucleic acid by an enzyme from ***Escherichia coli***. *Proc Natl Acad Sci USA* 1962, 48:81-94.

9. von Hippel PH, Bear DG, Morgan WD, McSwiggen JA: Protein-nucleic acid interactions in transcription: a molecular analysis. *Annu Rev Biochem* 1984, 53:389-446.

10. Erie DA, Yager TD, von Hippel PH: The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective. *Annu Rev Biophys Biomol Struct* 1992, 21:379-415.

11. Sentenac A, Riva M, Thuriaux P, Buhler J-M, Treich I, Carles C, Werner M, Ruet A, Huet J, Mann C *et al.*: Yeast RNA polymerase subunits and genes. In *Transcriptional Regulation*. Edited by McKnight SL, Yamamoto KR. New York: Cold Spring Harbor Laboratory Press; 1992:27-54.

12. Gross CA, Chan CL, Lonetto MA: A structure/function analysis of ***Escherichia coli*** RNA polymerase. *Philos Trans R Soc London Ser B* 1996, 351:475-482.

13. Darst SA, Kubalek EW, Kornberg RD: Three-dimensional structure of ***Escherichia coli*** RNA polymerase holoenzyme determined by electron crystallography. *Nature* 1989, 340:730-732.

14. Darst SA, Edwards AM, Kubalek EW, Kornberg RD: Three-dimensional structure of yeast RNA polymerase II at 16 Å resolution. *Cell* 1991, 66:121-128.

15. Schultz P, Celia H, Riva M, Sentenac A, Oudet P: Three-dimensional model of yeast RNA polymerase I determined by electron microscopy of two-dimensional crystals. *EMBO J* 1993, 12:2601-2607.

16. Polyakov A, Severinova E, Darst SA: Three-dimensional structure of ***Escherichia coli*** core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* 1995, 83:365-373.

17. Darst SA, Polyakov A, Richter C, Zhang G: Insights into ***Escherichia coli*** RNA polymerase structure from a combination of X-ray and electron crystallography. *J Struct Biol* 1998, 124:115-122.

18. Darst SA, Polyakov A, Richter C, Zhang G: Structural studies of ***Escherichia coli*** RNA polymerase. *Cold Spring Harbor Symp Quant Biol* 1998, 63:269-276.

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

The first experimental evidence, provided by EM and image processing of yeast RNAP II transcription complexes in two-dimensional crystals, that nucleic acids occupy the large RNAP channel, which had long been hypothesized. The authors also showed, however, that the path of the nucleic acids was not straight through the channel, as presumed.

20. Opalka N, Mooney RA, Richter C, Severinov K, Landick R, Darst SA: Direct localization of a β subunit domain on the three-dimensional structure of ***Escherichia coli*** RNA polymerase. *Proc Natl Acad Sci USA* 1999, 97:617-622.

21. • 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, in press.

The authors describe the co-crystal structure of *Taq* core RNAP complexed with the antibiotic inhibitor rifampicin. Structural and biochemical results demonstrate the inhibition mechanism of the antibiotic, along with an improved refinement of the bacterial RNAP structure.

22. •• 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.

This paper provides the first high-resolution view of a eukaryotic RNAP – yeast RNAP II.

23. •• Naryshkin N, Revyakin A, Kim Y, Mekler V, Ebright RH: Structural organization of the RNA polymerase-promoter open complex. *Cell* 2000, 101:601-611.

The results of protein–DNA cross-linking, along with the *Taq* core RNAP structure [4••], provide a structural model of the prokaryotic open promoter complex.

24. •• Bartlett MS, Thomm M, Geiduschek EP: The orientation of DNA in an archaeal transcription initiation complex. *Nat Struct Biol* 2000, 7:782-785.

The results of protein–DNA cross-linking, along with the *Taq* core RNAP structure [4••], provide a structural model for the trajectory of the DNA in an archaeal pre-initiation complex.

25. •• 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.

The results of protein–DNA and protein–RNA cross-linking, along with the *Taq* core RNAP structure [4••], provide a structural model of the TEC.

26. Minakhin L, Nechaev S, Campbell EA, Severinov K: Recombinant ***Thermus aquaticus*** RNA polymerase, a new tool for structure-based analysis of transcription. *J Bacteriol* 2001, 183:71-76.

27. Abrahams JP, Leslie AGW: Methods used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Crystallogr D* 1996, 52:30-42.

28. Markov D, Naryshkina T, Mustaev A, Severinov K: A zinc-binding site in the largest subunit of DNA-dependent RNA polymerase is involved in enzyme assembly. *Genes Dev* 1999, 13:2439-2448.

29. Ebright RH, Busby S: ***Escherichia coli*** RNA polymerase α subunit: structure and function. *Curr Opin Genet Dev* 1995, 5:197-203.

30. Allison LA, Moyle M, Shales M, Ingles CJ: Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell* 1985, 42:599-610.

31. 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.

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

33. Chamberlin MJ: New models for the mechanism of transcription elongation and its regulation. *Harvey Lectures* 1993, 88:1-21.

34. Korzheva N, Mustaev A, Nudler E, Nikiforov V, Goldfarb A: Mechanistic model of the elongation complex of ***Escherichia coli*** RNA polymerase. *Cold Spring Harbor Symp Quant Biol* 1998, 63:337-345.

35. Mustaev A, Kozlov M, Markovtsov V, Zaychikov E, Denissova L, Goldfarb A: Modular organization of the catalytic center of RNA polymerase. *Proc Natl Acad Sci USA* 1997, 94:6641-6645.

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

37. Mooney RA, Landick R: RNA polymerase unveiled. *Cell* 1999, 98:687-690.

38. 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, 89:33-41.

39. Rees WA, Keller RW, Vesenka JP, Yang G, Bustamante C: Evidence of DNA bending in transcription complexes imaged by scanning force microscopy. *Science* 1993, 260:1646-1649.

40. 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.

41. Reeder TC, Hawley DK: Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. *Cell* 1996, 87:767-777.

42. Komissarova N, Kashlev M: RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J Biol Chem* 1997, 272:15329-15338.

43. Komissarova N, Kashlev M: Transcriptional arrest: ***Escherichia coli*** RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci USA* 1997, 94:1755-1760.

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

45. Polyakov A, Richter C, Malhotra A, Koulich D, Borukhov S, Darst SA: Visualization of the binding site for the transcript cleavage factor GreB on ***Escherichia coli*** RNA polymerase. *J Mol Biol* 1998, 281:262-266.

46. d'Aubenton Carafa Y, Brody E, Thermes C: Prediction of rho-independent ***Escherichia coli*** transcription terminators. A statistical analysis of their RNA stem-loop structures. *J Mol Biol* 1990, 216:835-858.

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

48. Nicholls A, Sharp KA, Honig B: Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 1991, 11:281-296.

49. Jones TA, Zou J-Y, Cowan S, Kjeldgaard M: Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 1991, 47:110-119.

50. Siebenlist U, Simpson RB, Gilbert W: ***E. coli*** RNA polymerase interacts homologously with two different promoters. *Cell* 1980, 20:269-281.

51. Kirkegaard K, Buc H, Spassky A, Wang JC: Mapping of single-stranded regions in duplex DNA at the sequence level: single-strand-specific cytosine methylation in RNA polymerase-promoter complexes. *Proc Natl Acad Sci USA* 1983, 80:2544-2548.

52. Mecsas J, Cowing DW, Gross CA: Development of RNA polymerase-promoter contacts during open complex formation. *J Mol Biol* 1991, 220:585-597.

53. Zaychikov E, Denissova L, Heumann H: Translocation of the ***Escherichia coli*** transcription complex observed in the registers 11 to 20: 'jumping' of RNA polymerase and asymmetric expansion and contraction of the 'transcription bubble'. *Proc Natl Acad Sci USA* 1995, 92:1739-1743.

54. Metzger W, Schickor P, Heumann H: A cinematographic view of ***Escherichia coli*** RNA polymerase translocation. *EMBO J* 1989, 8:2745-2754.

55. Schickor P, Metzger W, Wladyslaw W, Lederer H, Heumann H: Topography of intermediates in transcription initiation of ***E. coli***. *EMBO J* 1990, 9:2215-2220.

56. Monforte JA, Kahn JD, Hearst JE: RNA folding during transcription by ***Escherichia coli*** RNA polymerase analyzed by RNA self-cleavage. *Biochemistry* 1990, 29:7882-7890.

57. Komissarova N, Kashlev M: Functional topography of nascent RNA in elongation intermediates of RNA polymerase. *Proc Natl Acad Sci USA* 1998, 95:14699-14704.