
This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of September 23, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/296/5571/1280.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2002/05/15/296.5571.1280.DC1.html>

This article **cites 29 articles**, 16 of which can be accessed free:

<http://www.sciencemag.org/content/296/5571/1280.full.html#ref-list-1>

This article has been **cited by** 241 article(s) on the ISI Web of Science

This article has been **cited by** 100 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/296/5571/1280.full.html#related-urls>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>

Structural Basis of Transcription Initiation: RNA Polymerase Holoenzyme at 4 Å Resolution

Katsuhiko S. Murakami, Shoko Masuda, Seth A. Darst*

The crystal structure of the initiating form of *Thermus aquaticus* RNA polymerase, containing core RNA polymerase ($\alpha_2\beta\beta'\omega$) and the promoter specificity σ subunit, has been determined at 4 angstrom resolution. Important structural features of the RNA polymerase and their roles in positioning σ within the initiation complex are delineated, as well as the role played by σ in modulating the opening of the RNA polymerase active-site channel. The two carboxyl-terminal domains of σ are separated by 45 angstroms on the surface of the RNA polymerase, but are linked by an extended loop. The loop winds near the RNA polymerase active site, where it may play a role in initiating nucleotide substrate binding, and out through the RNA exit channel. The advancing RNA transcript must displace the loop, leading to abortive initiation and ultimately to σ release.

Transcription initiation is a major control point of gene expression. RNA polymerase (RNAP), the central enzyme of transcription, contains an ~400-kD catalytic core (subunit composition $\alpha_2\beta\beta'\omega$ in bacteria) conserved in structure and function among all cellular organisms (1–3). Promoter-specific initiation requires additional proteins, ranging from an ~750-kD collection of more than a dozen basal initiation factors for messenger RNA (mRNA) synthesis in eukaryotes (4) to a single polypeptide in bacteria, the σ subunit, which binds core RNAP to form the holoenzyme (5).

The bacterial RNAP holoenzyme forms an initial closed promoter complex by recognizing two hexamers of consensus DNA sequence: the Pribnow box (–10 element), centered at about –10 with respect to the start site (+1), and the –35 element (6). A series of isomerization steps yields the transcription-competent open promoter complex, in which about 14 base pairs of DNA are melted in a region including the start site (the transcription bubble). In the presence of nucleotide substrates, RNA synthesis begins.

Before transitioning to a stable elongation complex, the enzyme initiates repeatedly, generating and releasing short transcripts, usually 2 to 12 nucleotides (nt) in length (7, 8), without dissociating from the promoter. The rate of synthesis of these abortive products can vastly exceed that of the full-length transcript. By the time the transcript reaches a length of around 12 nt, the complex acquires the properties of the elongation complex, with a stable hold on both the DNA template and the RNA transcript (9).

Structures of most components of the holoenzyme have been determined. The 3.3 Å

resolution structure of core RNAP from *T. aquaticus* (*Taq*) revealed a crab claw-shaped molecule with a 27 Å wide internal channel (3, 10). The enzyme active site is located on the back wall of the channel, where an essential Mg^{2+} ion is chelated. High-resolution structures of three domains of *Taq* σ^A have also been determined (11).

These studies revealed the structural components of the holoenzyme and give insight into their individual roles in transcription initiation. To understand how these elements function together, we need to know the structure of the holoenzyme. Here, we present the 4 Å resolution x-ray crystal structure of the *Taq* RNAP holoenzyme [Figs. 1A and 2, Table 1, fig. S1 (12)]. The structure provides insight into the multiple roles of the σ subunit in transcription initiation and provides a basis for the design of future experiments probing the initiation process.

Overall structure. *Taq* σ^A belongs to a homologous family closely related to *Escherichia coli* σ^{70} , with distinct regions of highly conserved amino acid sequence [Fig. 1B, fig. S2 (12–14)]. The σ structure comprises three, flexibly linked domains, σ_2 , σ_3 , and σ_4 , that contain conserved regions 1.2 to 2.4, 3.0 to 3.1, and 4.1 to 4.2, respectively [Fig. 1B, fig. S2 (11, 12)]. In the holoenzyme, the three σ domains lay spread out across the upstream face (15, 16) of the RNAP (Figs. 1A and 2). The COOH-terminus of σ_3 (residue 332) and the NH₂-terminus of σ_4 (residue 366) in the holoenzyme are separated by 45 Å. This distance is spanned by an extended 33-residue linker, comprising primarily σ region 3.2, that winds through the RNAP active-site channel and out through the RNA exit channel. A disordered loop (σ residues 337 to 345) may extend even further toward the active site, consistent with cross-linking results that place σ residues 333 to 386 within

5 Å of the initiating nucleotide γ -phosphate (17). The chain then continues into σ_4 , which is clamped to the α helix at the tip of the β flap (β flap-tip helix; Fig. 1A).

Conserved structural features of RNAP, the β' zipper and β' lid [using the nomenclature introduced for yeast RNAP II (1)], were disordered in the bacterial core RNAP (3, 10) but resolved in the holoenzyme (Fig. 2). Electron density for the β' NH₂-terminal Zn²⁺-binding domain (β' ZBD, also disordered in core RNAP) was present but difficult to interpret. The region was modeled (Fig. 2), but after β' residue 45, the path of the peptide backbone and the sequence register are extremely tentative.

The structure of the β' lid is very similar to that of the RNAP II Rpb1 lid (1). In the holoenzyme, the main RNAP channel is relatively closed, allowing the β' lid to interact with the inner surface of the β flap, creating a protein tunnel through which σ region 3.2 threads (Fig. 2). For this protein-protein interaction to establish itself, a profound conformational change of the core RNAP is required.

Genetic and biophysical studies implicate σ region 2.4 residues in recognition of the –10 element (5), region 2.3 residues in sequence-specific binding of the melted nontemplate strand of the –10 element in the open complex (18), and region 3.0 residues (11) in recognition of the extended –10 motif (19). The structure of *Taq* σ^A_4 bound to –35 element DNA (11) confirmed earlier genetic studies (20, 21) implicating residues within the recognition helix of the σ_4 helix-turn-helix motif in –35 element binding. In the holoenzyme, all promoter recognition determinants of σ are solvent exposed, ready to engage promoter DNA (Fig. 2), and their spacing is roughly consistent with their binding sites on the DNA.

Conformational changes of RNAP.

Large conformational changes of RNAP, mainly due to swinging motions of the clamp domain [Fig. 3, table S1 (12)] that open and close the main channel by >20 Å, have been observed for yeast RNAP II (1, 15, 22, 23) and bacterial RNAPs (24). Additional mobile modules have been defined by comparing different crystal forms of yeast RNAP II (1).

Comparison of the core RNAP structure (3, 10) with the core RNAP within the holoenzyme reveals five modules that move essentially as rigid bodies relative to each other [Fig. 3, table S1 (12)]. The bulk of the enzyme lies in a “core” module containing the two α -subunit NH₂-terminal domains, the ω subunit, and regions of β and β' around the active site. Four additional modules that move relative to the core module between the core RNAP structure and the holoenzyme structure essentially frame the main channel on three sides (Fig. 3). The overall effect of these changes is to close the width of the main channel by about 10 Å in the holoenzyme compared with core RNAP.

The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA.

*To whom correspondence should be addressed. E-mail: darst@rockefeller.edu

The clamp, $\beta 1$, and β flap domains make extensive interactions with σ_2 , σ_3 , and σ_4 , respectively [Fig. 1A; table S1 (12)], and the conformational changes of these domains are likely the result of σ binding. The $\beta 2$ domain, by contrast, does not interact with σ . The conformational change of this domain is likely to be in response to changes in the other RNAP modules.

σ -core RNAP interactions. In the holoenzyme, each σ domain, as well as the linkers connecting them, makes extensive interactions with core RNAP (Fig. 1, table S2). The total contact area of the σ -core RNAP interface (8230 Å²) is nearly twice the area of the largest protein-protein recognition interface (4660 Å²) (25) and is comparable to the largest contact areas of oligomeric proteins (~10,000 Å²) (26).

The holoenzyme structure confirms genetic studies indicating that the σ -core RNAP interface involves several regions of σ (27), as well as studies suggesting that the primary interface

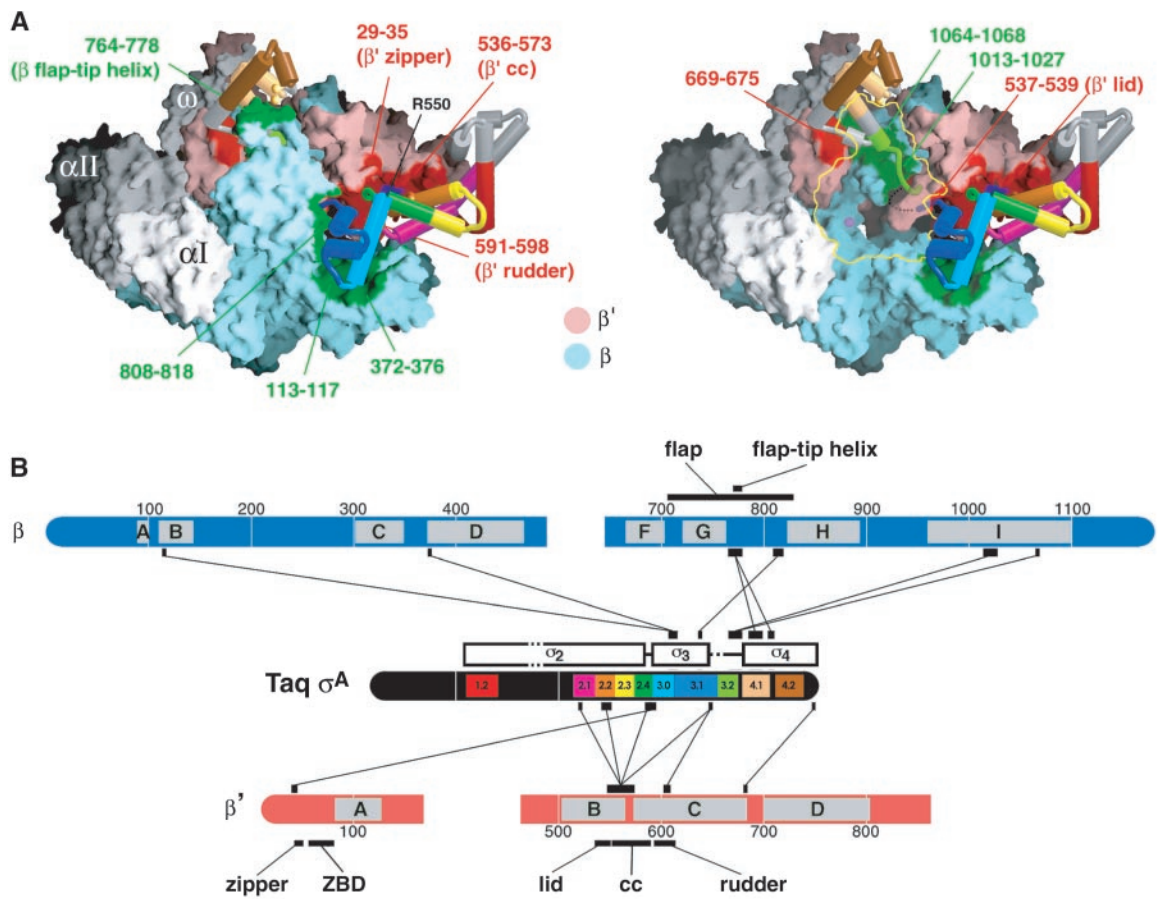
involves the exposed, polar surface of the σ region 2.2 helix and the β' coiled-coil (cc) (28, 29). The contact area of the σ_2 -core RNAP interface is the largest of any of the σ domains [table S2 (12)], and this interface mainly involves association with the β' cc (Fig. 1). Mutants that cause defects in σ -core RNAP binding have been isolated on the exposed faces of the σ region 2.2 helix (27, 30) and β' cc (28), and these residues face each other across the interface (Fig. 1).

Region 1.1 of σ . The holoenzyme crystals could be obtained only by using *Taq* σ^A with an NH₂-terminal truncation of 91 residues (Table 1). Among the σ^{70} family, the primary or group 1 σ subunits uniquely harbor an ~90-amino acid NH₂-terminal extension, region 1.1. This region is poorly conserved in sequence, but the characteristic acidity is preserved. In *Taq* σ^A residues 1 to 91, fully one-third of the residues are negatively charged, for a total charge of -20 and a calculated isoelectric point (pI) of 3.95.

Two functions have been ascribed to region 1.1. First, it autoinhibits promoter recognition by free σ (31, 32). This autoinhibition is presumed to be relieved upon σ binding to core RNAP. Second, region 1.1 can accelerate open complex formation at some promoters (33), which suggests that a step, rate limiting only at some promoters, is facilitated by region 1.1.

Although region 1.1 is missing from the crystals, the electrostatic surface of the holoenzyme provides insight into the initiation role of region 1.1. Although the overall charge of RNAP is highly negative (the pI of core RNAP is near 5), the electrostatic charge distribution over the surface is very asymmetric (Fig. 4A). The outside of the enzyme is almost uniformly acidic, whereas surfaces that interact with nucleic acids, and especially the inner walls of the main channel, are basic. The NH₂-terminal end of the σ fragment in the crystals comprises an α helix that points into the main RNAP channel, directly toward the active-site Mg²⁺ (Figs. 2 and

Fig. 1. *Taq* RNAP holoenzyme structure and σ -core RNAP interactions. (A) Structure of holoenzyme and map of the σ -core RNAP interface. (Left) The core component of the holoenzyme is shown as a molecular surface, color coded as follows: α I, α II, ω : gray; β : cyan; β' : pink. The σ subunit is shown as an α -carbon backbone worm, with α helices shown as cylinders, color coded according to the conserved regions as schematized in (B). Surfaces of RNAP within 4 Å of any σ atoms are color coded green (β) or red (β') and labeled. Positions in σ region 2.2 (orange backbone) and 4.1 (tan backbone) where substitutions cause defects in core RNAP binding are indicated by orange or tan α -carbon spheres, respectively (27, 30). The exposed surface of β' Arg⁵⁵⁰ on the β' cc important for σ interaction is colored blue and labeled (28). (Right) The same view, except that the β flap has been removed, revealing the σ_2 - σ_3 linker and its interactions with the β' lid and β region I underneath. (B) Schematic map of the σ -core RNAP interface. The bars represent the primary sequences of β (top, blue), β' (bottom, pink), and σ^A (middle, black). Portions of β and β' are omitted, as indicated by breaks in the bars. The gray boxes denote evolutionarily conserved regions among prokaryotic, chloroplast, archaeobacterial, and eukaryotic β (labeled A to I) and β' (A to H) homologs (3, 40). The conserved regions (11, 13, 14) are labeled and color



coded. The domain architecture of the σ segment in the holoenzyme crystals is indicated above the bar, with rectangles indicating the three structured domains, and lines representing the flexible linkers. Disordered segments within the holoenzyme structure are indicated by dashes. Important structural features of the β and β' subunits are indicated above (for β) and below (for β'). The lines connecting between σ and β or β' denote regions of the subunits that interact (<4 Å) in the holoenzyme. (A) was made with the program GRASP (41).

RESEARCH ARTICLES

4A). Extending from the NH₂-terminus of this helix, the acidic region 1.1 would likely locate directly inside the main channel, where it could interact with the basic surface of the channel walls.

In the holoenzyme structure, the RNAP channel is relatively closed. The distance across the open end of the channel is about 15 Å, so double-stranded DNA could not enter. We suggest that σ region 1.1 inside the channel could widen the channel to facilitate DNA entry, possibly explaining the acceleration of open complex formation at some promoters by region 1.1. At a later stage of initiation, other signals may induce closure of the claws, ejecting region 1.1 from inside the channel (Fig. 4B). The closed

conformation of the holoenzyme in the crystals may be induced by crystal packing and is unlikely to be normal for free holoenzyme, possibly explaining why crystals with full-length σ could not be obtained (in addition, the high salt concentration in the crystallization solution would likely inhibit the electrostatic interactions). Support for this hypothesis comes from hydroxyl-radical protein footprinting, showing that *E. coli* σ^{70} region 1.1 is exposed to hydroxyl radicals in free σ^{70} , strongly protected in the holoenzyme but exposed again in the binary complex of holoenzyme with promoter DNA (34). In the holoenzyme-DNA complex (35), the conformation of the enzyme channel is closed even further relative to the holoenzyme

alone, but in this case, isomorphous crystals could be obtained either with or without σ region 1.1. Two observations indicate that in these crystals, σ region 1.1, when present, was located outside the RNAP channel, flexible, and exposed to solution. First, in crystals where region 1.1 was present, it was not observed in the electron density maps; and second, the region was very sensitive to proteolysis within the crystals.

Abortive initiation and σ . A model for the nucleic acid path through RNAP in an elongation complex was determined from cross-link mapping (16). The modeled components included downstream duplex DNA, RNA-DNA hybrid extending from the RNAP

Fig. 2. *Taq* RNAP holoenzyme structure. Views of the *Taq* RNAP holoenzyme structure, shown as a molecular surface but with important features of core RNAP shown as α -carbon backbone worms without the corresponding surfaces (color coding of surfaces and worms is indicated). The molecular surface of σ is transparent, allowing the orange α -carbon backbone worm to be seen as well. (A) The same view as in Fig. 1A. The Zn²⁺ ion bound in the β' ZBD is shown as a light-green sphere. Surfaces of σ corresponding to residues important for promoter recognition and melting are color coded as follows: melting/−10 element nontemplate strand binding, yellow; −10 element recognition, green; extended −10 element recognition, blue; −35 element recognition, brown. (B) Partial, magnified view, obtained from (A) by rotation about the horizontal axis as indicated. Obscuring portions of β have been removed to reveal the inside of the main channel. The outline of β is shown as a cyan line. The active-site Mg²⁺ is shown as a magenta sphere. The disordered segment of σ is denoted by orange dots, connecting σ residues 336 to 346, which are labeled. The NH₂-terminus of the σ fragment (corresponding to *Taq* σ^A residue 93), which points into the RNAP channel toward the active-site Mg²⁺, is indicated (N). The figure was made with the program GRASP (41).

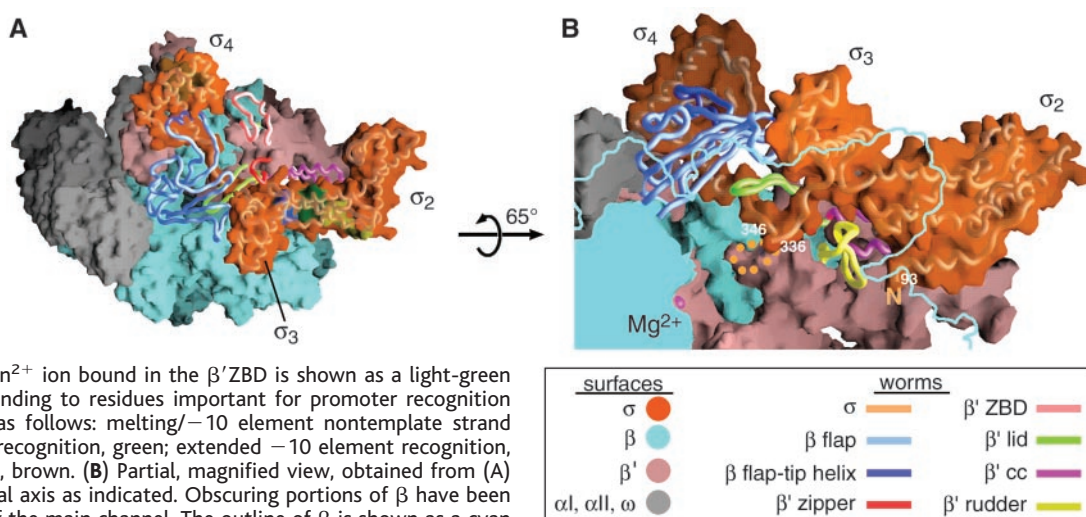


Table 1. Crystallographic analysis and structural model.

Diffraction data								
Data set*	Wavelength (Å)	Resolution (Å)	Redundancy (total/last shell)	Completeness (%) (total/last shell)	<i>I</i> / σ (<i>I</i>) (total/last shell)	<i>R</i> _{sym} † (%) (total/last shell)	No. of sites	Phasing power‡ (50 to 5.5 Å) (acentric/centric)
Native	1.03320	50–4.0	1.9/1.7§	79.9/70.1	5.4/1.8	15.8/35.7	–	–
Ta ₆ Br ₁₄	1.03320	50–5.5	2.3	96.6/95.4	6.6/2.42	10.1/31.5	8	0.76/1.09

Figure of merits||: Ta₆Br₁₄ (50 to 5.5 Å), 0.19; molecular replacement (50 to 4.0 Å), 0.58; combined (50 to 4.0 Å), 0.59.

Structural model

Protein subunit¶	Number of residues		Regions modeled
	Sequence	Model	
β'	1525	1171	3–145, 453–1240, 1257–1409, 1413–1499
β	1119	1085	2–56, 63–212, 220–245, 254–291, 298–649, 653–1116
$\Delta 1.1\sigma^A$	347	322	93–155, 171–336, 346–438
αI	313	224	6–229
αII	313	220	6–225
ω	99	92	2–93
Total	3716	3114	

*Data sets were collected at APS-SBC 19ID, Advanced Photon Source, Structural Biology Center, Argonne, IL. †*R*_{sym} = $\sum |I - \langle I \rangle| / \sum I$, where *I* is observed intensity and $\langle I \rangle$ is average intensity obtained from multiple observations of symmetry-related reflections. ‡Phasing power was calculated by MLPHARE (43). §Final shell: 4.14 to 4.0 Å. ||Figure of merit (combined phase between model and isomorphous phases) was calculated by CNS (44). ¶Also included in the model were one Mg²⁺ and two Zn²⁺ ions.

active site, and upstream single-stranded RNA extruded through the RNA exit channel formed by the β flap. The structure of elongating yeast RNAP II confirmed the modeled downstream DNA and RNA-DNA hybrid (15). The exiting upstream RNA was not resolved.

The σ_3 - σ_4 linker occupies the same space as the exiting RNA transcript of the elongation complex (Fig. 5A). At most, a 5- to 6-nt transcript can be accommodated without steric clash. If the nine-residue disordered segment in the σ_3 - σ_4 linker (residues 337 to 345) were modeled as a loop extending toward the RNAP active-site Mg^{2+} (Fig. 5A), it could interfere with transcripts of only 2 or 3 nt. Upstream, the clash persists until the transcript fills the exit channel at a length of 13 or 14 nt (Fig. 5A). Thus, there is a marked correspondence between the range of transcript lengths with which the σ_3 - σ_4 linker would clash and the typical lengths of transcripts that would be subject to abortive release. We suggest that, in the initiating com-

plex, the σ_3 - σ_4 linker must be displaced from the RNA exit channel as the transcript is extended. Competition for a binding site in the RNAP would hinder the initiation process and destabilize the transcripts, leading to abortive initiation.

Consistent with this hypothesis, substitution of two highly conserved residues in region 3.1 near the beginning of the σ_3 - σ_4 linker (Pro³²⁹ and Ser³³¹ in *Taq* σ^A numbering) led to a profound reduction in the ratio of abortive to full-length transcripts (36). The substitutions may disrupt important contacts between the linker and RNAP, making the linker easier to displace by the advancing transcript. The 33-residue linker contains nine conserved acidic residues, giving an overall charge of -8 to -9 among group 1 σ subunits. The negatively charged, extended polypeptide chain may serve as a nucleic acid mimic.

To test the hypothesis about abortive initiation, we compared the distribution of RNAs produced by *E. coli* holoenzyme with that of wild-type σ^{70} or of a σ^{70} COOH-terminal truncation lacking region 3.2 [σ^{70} (residues 1 to 503) or $\sigma^{70}\Delta C$] initiating from an extended -10 promoter (12). Wild-type σ^{70} holoenzyme produced 71-nt run-off transcripts, as well as an

11-fold molar excess of abortive transcripts (Fig. 5B). On the other hand, production of abortive products by the $\sigma^{70}\Delta C$ -holoenzyme was not significant (Fig. 5B), which supports the hypothesis.

This model posits the sequential displacement of the σ_3 - σ_4 linker from the RNA exit channel by the advancing RNA transcript. The σ_3 - σ_4 linker would be completely displaced by a 12- to 14-nt transcript, and this in turn may weaken the interaction between σ_4 and the β flap, causing release of σ_4 and ultimately the rest of σ . This is consistent with the two-step model of Shimamoto *et al.* (37) for σ release, in which a rapid triggering (displacement of the σ_2 - σ_3 linker by the RNA transcript) is followed by slow dissociation.

Relation to eukaryotic initiation. The catalytic core of eukaryotic RNAPs is homologous in structure and function with the bacterial enzyme (13). Although details of initiation in bacteria and eukaryotes are quite different, the tasks required of the general initiation apparatus are the same: to recognize the promoter and recruit RNAP, to specify the start site, and to melt the DNA to create the transcription bubble. Very little data exist

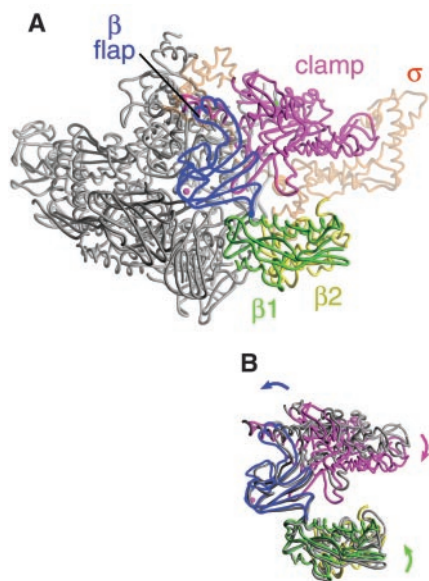


Fig. 3. Mobile modules and conformational changes. The viewing angle is the same as in Fig. 1A. (A) View of the RNAP holoenzyme α -carbon backbone, shown as a worm. The σ subunit is colored orange and rendered faint by making it transparent. The five mobile modules of core RNAP [table S1 (12)] are color coded as follows: core module, gray; $\beta 1$, green; $\beta 2$, yellow; β flap, blue; clamp, magenta. The active-site Mg^{2+} is shown as a magenta sphere; the β' ZBD Zn^{2+} ion is shown as a green sphere. (B) The holoenzyme mobile modules that move with respect to the core domain ($\beta 1$, $\beta 2$, β flap, clamp) are shown [color coded as in (A)] superimposed on the same segments from the core RNAP crystal structure (3, 10), colored gray. The two structures were aligned according to the core module, which is not shown. The active-site Mg^{2+} is shown as a magenta sphere. The movements of the mobile modules from the core RNAP structure to their positions in the holoenzyme are indicated by arrows. The figure was made with the program RIBBONS (42).

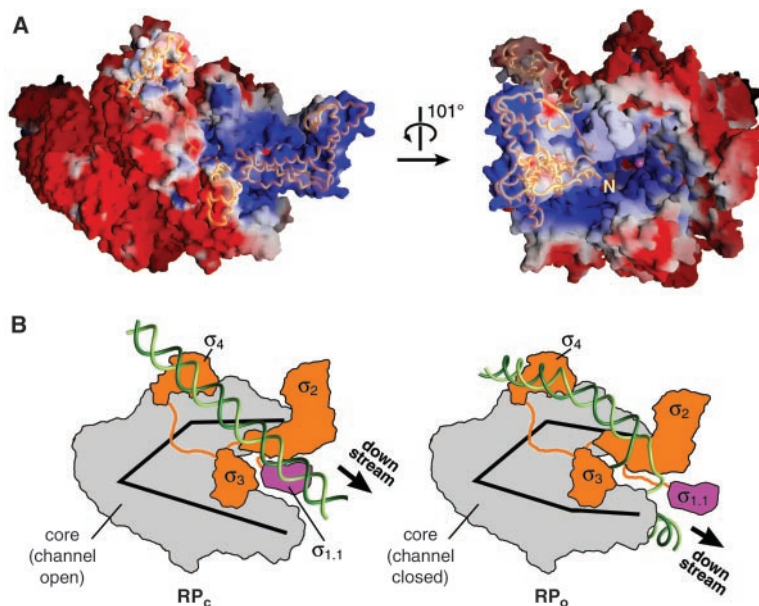


Fig. 4. Electrostatic distribution of holoenzyme and σ region 1.1. (A) Two surface views of the RNAP holoenzyme, color coded according to electrostatic surface potential (negative, red; neutral, white; positive, blue). The transparent α -carbon backbone worm of σ (orange) is superimposed. (Left) Same view as in Fig. 1A. (Right) View obtained from the left view by a rotation about the vertical axis as indicated; the active-site Mg^{2+} in the back of the main RNAP channel is visible as a magenta sphere. The NH_2 -terminus of the σ fragment (corresponding to *Taq* σ^A residue 93), which points into the RNAP channel toward the active-site Mg^{2+} , is indicated (N). (B) Schematic diagram illustrating the proposed mechanism of the negatively charged σ region 1.1 in promoting open complex formation. The viewing angle is similar to that in Fig. 1A. Two states of the RNAP holoenzyme–promoter DNA complex are illustrated. The positioning of the DNA is according to (35). The core RNAP is colored gray, and σ is colored orange, except region 1.1, which is colored magenta. (Left) The initial closed promoter complex (RP_c), where we propose that σ region 1.1 is positioned inside the positively charged RNAP channel [protecting it from hydroxyl-radical cleavage (34)], holding the channel open (indicated by thick black lines) to allow entry of double-stranded DNA. (Right) Final open promoter complex (RP_o), where DNA has entered the RNAP main channel and the channel has closed, ejecting σ region 1.1, where it is exposed in solution to proteases (35) and hydroxyl-radical cleavage (34). (A) was made with the program GRASP (41).

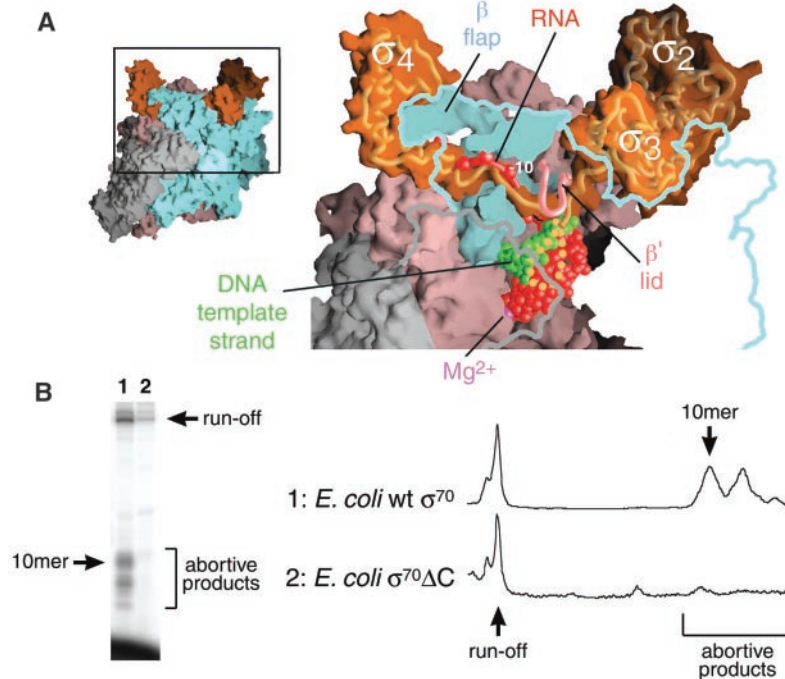


Fig. 5. Structural basis of abortive initiation. **(A)** (Left) View of the RNAP holoenzyme structure, represented as a molecular surface and color coded as follows: α , α II, ω : gray; β : cyan; β' : pink; σ : orange. (Right) The boxed area on the left is magnified. The σ surface is rendered transparent, revealing the orange α -carbon backbone worm inside; the β' lid is shown as a pink worm without the corresponding surface. Also shown in the structure is the RNA-DNA hybrid (from +1 to -7) and the upstream single-stranded RNA (from -8 to -14) from the elongation complex model (RNA, red; DNA template strand, green) (16). The base-paired RNA-DNA hybrid is shown as atoms. The phosphate backbone of the upstream single-stranded RNA is shown as a worm, with phosphate atoms shown as spheres. In the magnified view on the right, obscuring portions of α I and β have been removed, revealing the active-site Mg^{2+} (shown as a magenta sphere), the inside of the main RNAP channel filled by the RNA-DNA hybrid, and the RNA exit channel. The outlines of the α I and β subunits are shown as gray and cyan lines, respectively. Selected positions of the RNA (with respect to the transcription start site at +1) are labeled. **(B)** (Left) Autoradiograph showing the radioactive RNA produced by *E. coli* wild-type σ^{70} -holoenzyme (lane 1) and $\sigma^{70}\Delta C$ -holoenzyme (lane 2) transcribing from an extended -10 promoter and analyzed by denaturing gel electrophoresis (12). (Right) Profiles of the two lanes from the gel on the left, quantitated by phosphorimager. The two profiles have been normalized according to the run-off peak. (A) was made with the program GRASP (47).

to identify epitopes for eukaryotic general initiation factors on their RNAPs, but we expect analogies with σ -core RNAP interactions. For example, a eukaryotic initiation factor may place a segment near the RNAP active site and in the RNA exit channel, analogous to σ region 3.2, to assist binding the initiating nucleotide and to effect abortive initiation, which is a prominent feature of eukaryotic initiation (38, 39).

Conclusions. The bacterial RNAP holoenzyme structure provides a view of an intact σ subunit and delineates the role of core RNAP elements in positioning σ within the initiation complex. Features of the σ -core RNAP interaction indicate profound conformational changes in establishing the complex.

The structure also provides insight into the role of σ in binding the initiating nucleotide substrate and in the processes of abortive initiation and open complex formation. A holoenzyme with a COOH-terminal truncation of σ^A lacking region 3.2 was active on extended -10

promoters but had a defect in initiating nucleotide binding (11). We propose that the disordered segment of σ^A at the 3.1/3.2 junction directly participates in binding the initiating nucleotide in the RNAP i-site. Once the first phosphodiester bond forms and RNAP undergoes the initial translocation, nucleotide binding in the i-site is no longer required as it is occupied by the 3' end of the transcript, explaining the logic of fulfilling this one-time role by an initiation-specific subunit. This segment of σ near the active site then becomes an obstacle to RNA extension and induces abortive initiation.

References and Notes

- P. Cramer, D. A. Bushnell, R. D. Kornberg, *Science* **292**, 1863 (2001).
- S. A. Darst, *Curr. Opin. Struct. Biol.* **11**, 155 (2001).
- G. Zhang *et al.*, *Cell* **98**, 811 (1999).
- R. D. Kornberg, *Trends Cell Biol.* **9**, M46 (1999).
- Reviewed in C. A. Gross *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **63**, 141 (1998).
- P. L. deHaseth, M. L. Zupancic, M. T. J. Record, *J. Bacteriol.* **180**, 3019 (1998).
- A. J. Carpousis, J. D. Gralla, *Biochemistry* **19**, 3245 (1980).

- W. R. McClure, C. L. Cech, D. E. Johnston, *J. Biol. Chem.* **253**, 8941 (1978).
- A. Mustaeve, N. Korzheva, *Curr. Opin. Microbiol.* **4**, 119 (2001).
- E. A. Campbell *et al.*, *Cell* **104**, 901 (2001).
- E. A. Campbell *et al.*, *Mol. Cell*, **9**, 527 (2002).
- Supplementary material is available on Science Online at www.sciencemag.org/cgi/content/full/296/5571/1280/DC1.
- T. M. Gruber, D. A. Bryant, *J. Bacteriol.* **179**, 1734 (1997).
- M. Lonetto, M. Gribskov, C. A. Gross, *J. Bacteriol.* **174**, 3843 (1992).
- A. L. Gnatt, P. Cramer, J. Fu, D. A. Bushnell, R. D. Kornberg, *Science* **292**, 1876 (2001).
- N. Korzheva *et al.*, *Science* **289**, 619 (2000).
- K. Severinov *et al.*, *J. Biol. Chem.* **269**, 20826 (1994).
- Reviewed in S. A. Darst *et al.*, in *Nucleic Acids and Molecular Biology*, F. Ekstein, D. M. J. Lilley, Eds. (Springer, London, 1997), vol. 11, pp. 27-40.
- K. A. Barne, J. A. Bown, S. J. W. Busby, S. D. Minchin, *EMBO J.* **16**, 4034 (1997).
- T. Gardella, T. Moyle, M. M. Susskind, *J. Mol. Biol.* **206**, 579 (1989).
- D. A. Siegle, J. C. Hu, W. A. Walter, C. A. Gross, *J. Mol. Biol.* **206**, 591 (1989).
- P. Cramer *et al.*, *Science* **288**, 640 (2000).
- J. Fu *et al.*, *Cell* **98**, 799 (1999).
- S. A. Darst *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 4296 (2002).
- L. L. Conte, C. Chothia, J. Janin, *J. Mol. Biol.* **285**, 2177 (1999).
- J. Janin, S. Miller, C. Chothia, *J. Mol. Biol.* **204**, 155 (1988).
- M. M. Sharp *et al.*, *Genes Dev.* **13**, 3015 (1999).
- T. M. Arthur, L. C. Anthony, R. R. Burgess, *J. Biol. Chem.* **275**, 23113 (2000).
- B. A. Young *et al.*, *Cell* **105**, 935 (2001).
- D. M. Joo, N. Ng, R. Calendar, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4907 (1997).
- J. A. Camarero *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
- A. J. Dombroski, W. A. Walter, M. T. Record, D. A. Siegle, C. A. Gross, *Cell* **70**, 501 (1992).
- S. Vuthoori, C. W. Bowers, A. McCracken, A. J. Dombroski, D. M. Hinton, *J. Mol. Biol.* **309**, 561 (2001).
- H. Nagai, N. Shimamoto, *Genes Cells* **2**, 725 (1997).
- K. S. Murakami, S. Masuda, E. A. Campbell, O. Muzzin, S. A. Darst, *Science* **296**, 1285 (2002).
- R. Sen, H. Nagai, V. J. Hernandez, N. Shimamoto, *J. Biol. Chem.* **273**, 9872 (1998).
- N. Shimamoto, T. Kamigochi, H. Utiyama, *J. Biol. Chem.* **261**, 11859 (1986).
- P. Bhargava, G. A. Kassavetis, *J. Biol. Chem.* **274**, 26550 (1999).
- D. S. Luse, G. A. Jacob, *J. Biol. Chem.* **262**, 14990 (1987).
- J. Archambault, J. D. Friesen, *Microbiol. Rev.* **57**, 703 (1993).
- A. Nicholls, K. A. Sharp, B. Honig, *Prot. Struct. Funct. Genet.* **11**, 281 (1991).
- M. Carson, *J. Appl. Crystallogr.* **24**, 958 (1991).
- Z. Otwinowski, in *Proc. CCP4 Study Weekend*, W. Wolf, P. R. Evans, A. G. W. Leslie, Eds. (Science and Engineering Research Council Daresbury Laboratory, Warrington, UK, 1991), pp. 80-86.
- P. D. Adams, N. S. Pannu, R. J. Read, A. T. Brunger, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5018 (1997).
- We thank G. Schneider for providing Ta₈Br₄ T. Gaal and R. Gourse for the -10 con promoter fragment, and R. Landick for invaluable discussions. We are indebted to A. Joachimiak, S. L. Ginell, and N. Duke (Advanced Photon Source Structural Biology Center) and to M. Becker and L. Berman (NSLS X25) for support during data collection. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, under contract W-31-109-ENG-38. K.S.M. was supported by a Norman and Rosita Winston Postdoctoral Fellowship and a Human Frontier Science Program Postdoctoral Fellowship. Supported in part by NIH grants GM53759 and GM61898 (S.A.D.).

7 January 2002; accepted 21 March 2002