

# Structure and Function of the Transcription Elongation Factor GreB Bound to Bacterial RNA Polymerase

Natacha Opalka,<sup>1</sup> Mark Chlenov,<sup>1</sup> Pablo Chacon,<sup>2,4</sup> William J. Rice,<sup>3</sup> Willy Wriggers,<sup>2,5</sup> and Seth A. Darst<sup>1,\*</sup>

<sup>1</sup>The Rockefeller University  
1230 York Avenue  
New York, New York 10021

<sup>2</sup>Department of Molecular Biology  
The Scripps Research Institute  
La Jolla, California 92037

<sup>3</sup>Skirball Institute of Biomolecular Medicine  
and Department of Cell Biology  
New York University Medical Center  
New York, New York 10016

## Summary

Bacterial GreA and GreB promote transcription elongation by stimulating an endogenous, endonucleolytic transcript cleavage activity of the RNA polymerase. The structure of *Escherichia coli* core RNA polymerase bound to GreB was determined by cryo-electron microscopy and image processing of helical crystals to a nominal resolution of 15 Å, allowing fitting of high-resolution RNA polymerase and GreB structures. In the resulting model, the GreB N-terminal coiled-coil domain extends 45 Å through a channel directly to the RNA polymerase active site. The model leads to detailed insights into the mechanism of Gre factor activity that explains a wide range of experimental observations and points to a key role for conserved acidic residues at the tip of the Gre factor coiled coil in modifying the RNA polymerase active site to catalyze the cleavage reaction. Mutational studies confirm that these positions are critical for Gre factor function.

## Introduction

During each phase of the transcription cycle, RNA polymerase (RNAP) activity is modulated by interactions with a wide variety of regulatory factors. Among these are transcription elongation factors, which stimulate the overall elongation rate (Uptain et al., 1997).

Despite the absolute processivity of cellular RNAPs, transcript elongation is far from a smooth, continuous process. Due to specific regulatory signals or nonspecific obstacles to the forward progress of the enzyme, elongating RNAP frequently pauses at sites on the DNA template. In addition to transient halting, some RNAP molecules may undergo reverse translocation (a process termed backtracking), creating arrested complexes, a

transcriptionally inactive state where the 3'-OH of the RNA transcript is disengaged from the enzyme active site (Komissarova and Kashlev, 1997a, 1997b; Nudler et al., 1997; Reeder and Hawley, 1996). A special class of elongation factors, which includes prokaryotic GreA and GreB, and eukaryotic TFIIS (SII), increases the overall elongation rate by greatly mitigating pausing and reactivating arrested complexes (Borukhov and Goldfarb, 1996; Borukhov et al., 1992, 1993; Fish and Kane, 2002; Reines et al., 1989).

Backtracked, arrested complexes can be rescued by internal hydrolytic cleavage and release of the transcript 3' fragment, generating a new 3'-OH in register with the RNAP catalytic center, allowing for renewed RNA synthesis (Borukhov et al., 1993; Izban and Luse, 1992; Surratt et al., 1991). This surprising activity is intrinsic to the RNAP enzyme (Awrey et al., 1997; Orlova et al., 1995), is catalyzed by the RNAP active site itself (Rudd et al., 1994; Wang and Hawley, 1993), and is greatly stimulated by GreA/B acting on bacterial RNAP (Borukhov et al., 1992, 1993) or TFIIS acting on eukaryotic RNAP II (Reines, 1992; Sluder et al., 1989). Although the requirement of the transcript cleavage factors for the natural progression of RNAP in vivo has been established (Marr and Roberts, 2000; Toulme et al., 2000), their biological role remains unclear. In addition to rescuing arrested complexes and increasing the overall elongation rate, the cleavage factors may play a role in modulating RNAP behavior at pause signals (Marr and Roberts, 2000), increasing transcriptional fidelity (Erie et al., 1993), and in stimulating promoter clearance (Hsu et al., 1995).

*Escherichia coli* GreA and GreB are sequence homologs (Borukhov et al., 1993) and have homologs in every known bacterial genome. The 2.2 Å resolution crystal structure of *E. coli* GreA comprises an N-terminal antiparallel  $\alpha$ -helical coiled coil, and a C-terminal globular domain (Stebbins et al., 1995). While the C-terminal domain (CTD) binds RNAP (Koulich et al., 1997, 1998; Polyakov et al., 1998), the N-terminal coiled coil interacts with the transcript 3' end and is responsible for stimulating the transcript cleavage reaction (Koulich et al., 1997, 1998; Stebbins et al., 1995). GreA and GreB stimulate transcript cleavage in different ways; GreA induces cleavage of 3'-RNA fragments 2–3 nt in length and can only prevent the formation of arrested complexes, whereas GreB induces cleavage of fragments up to 18 nt in length and can rescue preexisting arrested complexes (Borukhov et al., 1993). These functional differences correlate with a distinctive structural feature, the distribution of positively charged residues on one face of the N-terminal coiled coil (Kulish et al., 2000). Remarkably, despite close functional similarity, the prokaryotic Gre factors have no sequence or structural similarity with eukaryotic TFIIS (Morin et al., 1996; Olmsted et al., 1998).

Previously, we reported a 15 Å resolution structure of *E. coli* core RNAP, determined by cryo-electron microscopy (cryo-EM) and image processing from helical crystals (Darst et al., 2002). The structure showed excellent correspondence with the X-ray structure of *Thermus*

\*Correspondence: darst@rockefeller.edu

<sup>4</sup>Present address: Centro de Investigaciones Biológicas, Calle Velázquez 144, Madrid 28006, Spain.

<sup>5</sup>Present address: School of Health Information Sciences and Institute of Molecular Medicine, University of Texas Health Science Center, 7000 Fannin, Suite 600, Houston, Texas 77030.

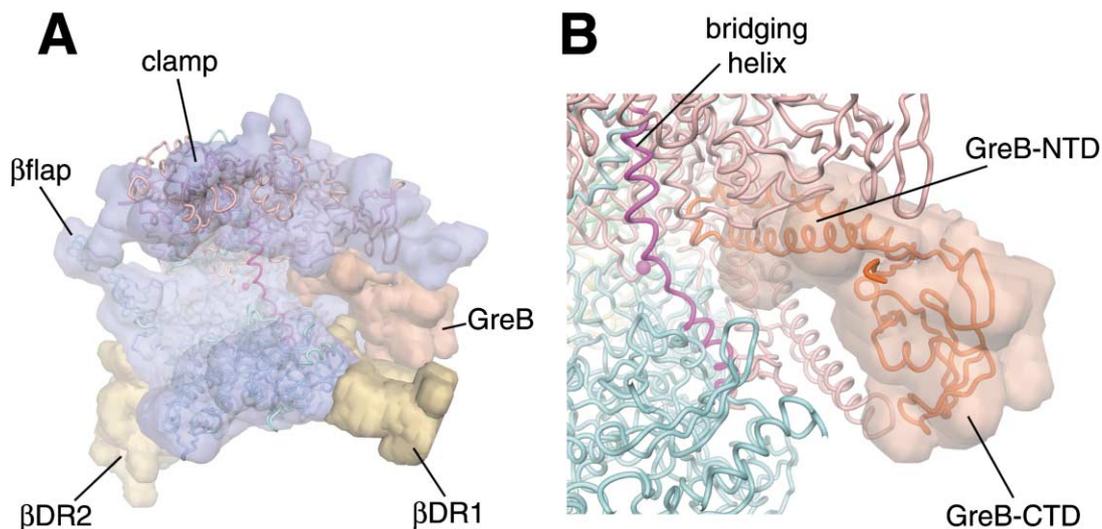


Figure 1. Cryo-EM Density Maps

(A) Cryo-EM reconstruction of the *E. coli* core RNAP/GreB complex and fit of the *Taq* core RNAP X-ray structure. Density corresponding to a single complex was extracted from the helical reconstruction and contoured at  $1.1\sigma$  (transparent surface). The *Taq* core RNAP X-ray structure, flexibly fit to the cryo-EM density (Darst et al., 2002; Wrighers and Chacon, 2001), is shown as an  $\alpha$  carbon backbone worm ( $\beta$  subunit, cyan;  $\beta'$ , pink, except the  $\beta'$ -bridging helix is magenta; active site  $Mg^{2+}$ , magenta sphere). Cryo-EM density corresponding to the crystal structure is blue, while density not accounted for by the RNAP structure is colored and labeled as follows: Yellow densities  $\beta$ DR1 and  $\beta$ DR2 correspond to large insertions present in *E. coli* but missing in *Taq* RNAP (Borukhov et al., 1991b; Severinov et al., 1994) that were observed previously (Darst et al., 2002; Opalka et al., 1999). The red density (GreB) was not observed previously.

(B) The cryo-EM density attributed to GreB is shown, along with the flexibly fit *Taq* core RNAP structure (color-coded as in A) and the  $\alpha$  carbon backbone of the GreB homology model (Koulich et al., 1997) fit into the density (orange worm).

*aquaticus* (*Taq*) core RNAP, allowing interpretation of the cryo-EM map in terms of the high-resolution structure. Here, we present a 15 Å resolution helical reconstruction of the *E. coli* core RNAP/GreB complex that allows fitting of high-resolution RNAP and GreB structures. The model of the complex reveals a remarkable binding mode for GreB; the globular CTD binds RNAP at the edge of the active site channel (as observed previously from an analysis in negative stain; Polyakov et al., 1998), while the N-terminal coiled-coil domain extends 45 Å into a channel directly to the RNAP active site. Analysis of the model leads us to propose a detailed mechanism for Gre factor activity that explains a wide range of experimental observations. The results point to a key role for conserved acidic residues at the tip of the Gre factor coiled coil in modifying the RNAP active site to catalyze the transcript cleavage reaction, and mutational studies confirm that these positions are critical for Gre factor function.

## Results and Discussion

### Structure of the *E. coli* Core RNAP/GreB Complex

Cryo-EM and image processing were used to determine the structure of the *E. coli* core RNAP/GreB complex from helical crystals (Darst et al., 2002; Opalka et al., 1999; Polyakov et al., 1995, 1998) to a nominal resolution of 15 Å (see Experimental Procedures and Supplemental Data available at <http://www.cell.com/cgi/content/full/114/3/335/DC1>). The density corresponding to a single RNAP/GreB complex was delineated by comparison with the *Taq* core RNAP crystal structure (Campbell et al., 2001; Zhang et al., 1999; PDB ID 1I6V; Figure 1A).

Previously, comparison of the RNAP crystal structure and the 15 Å reconstruction of *E. coli* core RNAP alone (Darst et al., 2002) required an alignment procedure based on difference mapping combined with rigid body and flexible fitting (allowing conformational changes between domains of the X-ray structure). A very large conformational change of the *Taq* RNAP structure, corresponding mainly to opening of the clamp domain (Fu et al., 1999) to widen the main active site channel by nearly 25 Å, was required to fit the *E. coli* cryo-EM map. A similar conformation of the RNAP was observed here, but the flexible fitting procedure was performed anew because of differences in the disposition of RNAP domains, particularly the clamp (see Supplemental Data available on website). The resulting fit between the cryo-EM map and the “flexed” *Taq* core RNAP structure showed excellent correspondence of structural features, except that three significant densities were not accounted for by the X-ray structure (labeled  $\beta$ DR1,  $\beta$ DR2, and GreB; Figure 1A).

The densities labeled  $\beta$ DR1 (Severinov et al., 1994) and  $\beta$ DR2 (Borukhov et al., 1991b) were observed and identified in earlier reconstructions of *E. coli* core RNAP (Darst et al., 2002; Opalka et al., 1999). These correspond to large insertions present in the *E. coli* RNAP  $\beta$  subunit but absent in the *Taq* enzyme. The density-labeled GreB was not observed in the structure of *E. coli* core RNAP alone. We attribute this density to GreB for two reasons. First, the size and shape of the density corresponds to the distinctive shape of the homology-modeled GreB structure (Koulich et al., 1997), with its globular CTD and extended coiled-coil N-terminal domain (NTD), leading to an unambiguous fit that did not require any conforma-

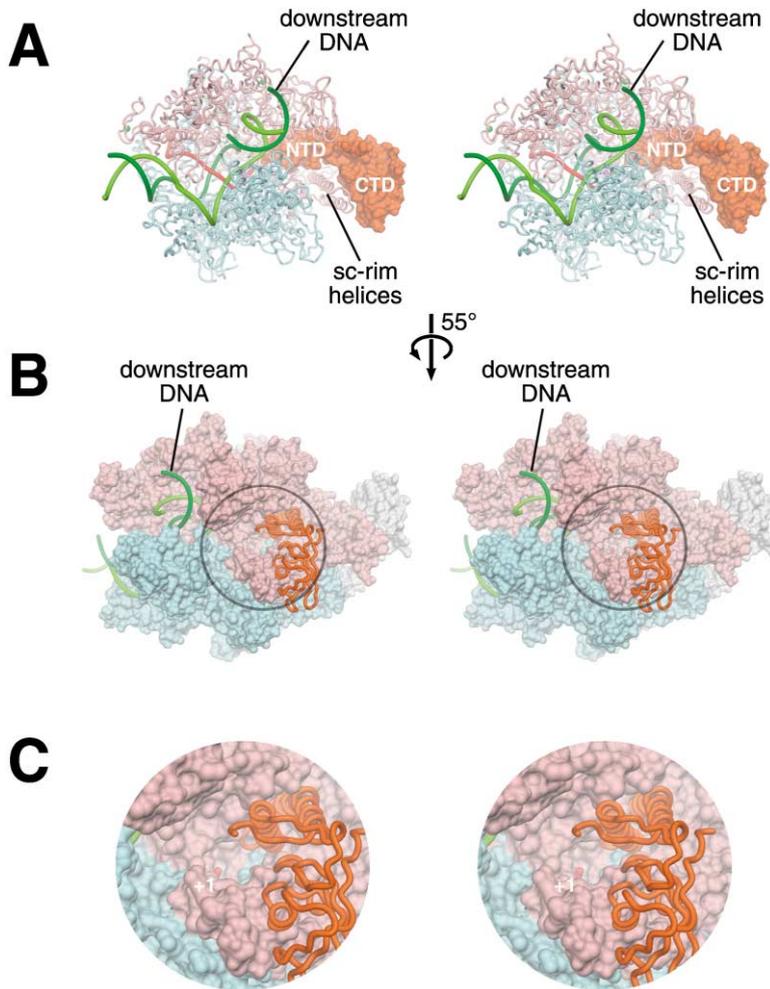


Figure 2. Model of a TEC with GreB

Stereo views of the TEC model with GreB, color-coded as follows: RNAP,  $\alpha$  and  $\omega$  subunits, gray;  $\beta$ , cyan;  $\beta'$ , pink; active site  $Mg^{2+}$ , magenta sphere; DNA, non-template strand, light green; template strand, dark green/RNA transcript (shown from +1 to -14), red; and GreB, orange. The nucleic acids (Gnatt et al., 2001; Korzheva et al., 2000) are shown as phosphate-backbone worms.

(A) RNAP is shown as an  $\alpha$  carbon backbone worm (with the  $\beta'$  bridging helix colored magenta), GreB (with labeled NTD and CTD) as a molecular surface. The GreB-CTD interacts primarily with the secondary channel-rim helices (sc-rim helices) of the  $\beta'$  subunit.

(B) RNAP is shown as a molecular surface and GreB as a backbone worm. The view, oriented with respect to the view in (A) as shown, is directly into the secondary channel. The entrance to the secondary channel is circled. The circled area is also magnified in part (C).

(C) Magnified view into the secondary channel (circled area from B). The view extends all the way to the active site, where the 3' end of the RNA (in the +1 position as labeled) is visible.

tional changes of the GreB molecule (Figure 1B). Second, the location of the CTD, bound at the outside edge of the RNAP active site channel near the entrance of the secondary channel, corresponds to the location of the CTD binding site localized in an earlier study of negatively stained samples (Polyakov et al., 1998). The N-terminal coiled coil of GreB extends from the entrance of the secondary channel (Darst, 2001; Korzheva et al., 2000; Zhang et al., 1999), along the roof of the secondary channel (comprising exclusively the RNAP  $\beta'$  subunit), behind the  $\beta'$  bridging helix nearly to the RNAP active site  $Mg^{2+}$  deep within the RNAP active site channel (Figure 1).

#### Model of a Ternary Elongation Complex with GreB

The fit of the RNAP and GreB structures into the cryo-EM map provides an atomic model for the complex. However, the conformation of the RNAP, with its open-clamp domain and nearly 45 Å wide active site channel, is determined by crystal packing (Darst et al., 2002) and is far from the closed-clamp conformation (Gnatt et al., 2001) of the ternary elongation complex (TEC) with which GreA/B functionally interact. The bacterial RNAP structure that corresponds most closely to the conformation of yeast RNAP II observed in an elongating complex (Gnatt et al., 2001) is the core portion ( $\alpha_2\beta\beta'\omega$ ) of the *Taq*

RNAP holoenzyme observed in complex with a promoter DNA fragment (Murakami et al., 2002). In this structure, the clamp domain is slightly closed compared with the yeast RNAP II elongating complex, but still easily accommodates the 9 bp RNA/DNA hybrid, placed according to (Gnatt et al., 2001), and the remaining nucleic acids of the TEC, placed according to (Korzheva et al., 2000).

GreB was placed in the context of the TEC by maintaining the interactions with the roof of the secondary channel observed in the cryo-EM structure. This resulted in an excellent complementary fit of GreB to the TEC with no steric clashes (Figures 2, 3, and 4). The GreB-CTD binds outside the RNAP active site channel, interacting primarily with two  $\alpha$  helices that form part of a rim at the entrance to the secondary channel (sc-rim helices, Figure 2). The GreB N-terminal coiled coil extends 45 Å into the secondary channel nearly to the RNAP active site, fitting snugly in a groove on one side of the secondary channel (Figure 2). This model of the functional interaction between the TEC and GreB served as the structural basis for further analysis of the complex.

#### TEC Backtracking and GreB

The backtracked TEC is an obligate intermediate in Gre-induced transcript cleavage (Komissarova and Kashlev,

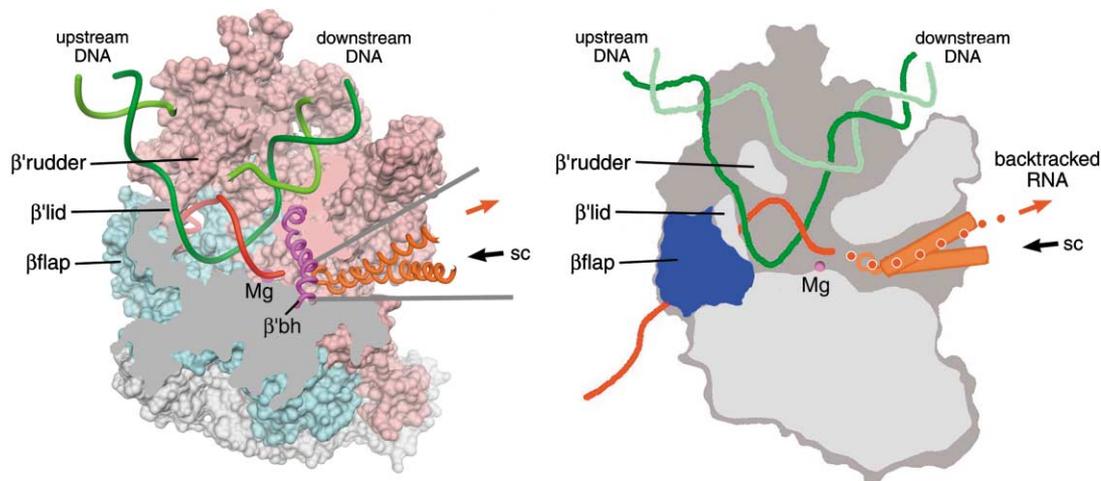


Figure 3. RNA Backtracking and GreB

Cross-sectional view of the TEC/GreB model, generated by slicing the structure in half, and viewing the  $\beta'$ -pincer from inside the active site channel. On the left, RNAP is shown as a molecular surface. The  $\alpha$  and  $\omega$  subunits are colored gray;  $\beta$ , cyan;  $\beta'$ , pink, except the  $\beta'$  bridging-helix ( $\beta'$ bh), which extends out of the plane toward the viewer, is shown as a magenta backbone worm. The active site  $Mg^{2+}$  is shown as a magenta sphere. The GreB coiled coil is shown as an orange backbone worm. The nucleic acids are shown as phosphate-backbone worms (nontemplate DNA, light green; template DNA, dark green; RNA, red). The secondary channel (sc) is roughly demarcated by the gray lines. The paths of incoming nucleotide substrates and backtracking RNA transcript are shown (black and red arrows, respectively). On the right is a schematic cartoon illustrating the same features, with the path of the backtracked RNA 3' end denoted by red dots.

1997b). In normal forward progression during transcription elongation, the RNAP enzyme translocates in the downstream direction along the DNA template, one bp per step. Each forward step is associated with the melting of one bp at the downstream edge of the transcription bubble and reannealing of one bp at the upstream edge. A 9 bp RNA/DNA hybrid is maintained at the growing end of the transcript by the melting of one RNA/DNA bp at the upstream end of the hybrid, and the addition of a bp at the downstream end by the extension of the RNA chain. During backtracking, the RNAP enzyme reverse translocates in the upstream direction on the DNA template, with the melting of bps at the upstream edge of the transcription bubble and reannealing at the downstream edge. The RNA transcript simultaneously reverse translocates through the enzyme, maintaining the 9 bp RNA/DNA hybrid and the base-pairing register with the template strand of the DNA, resulting in the formation of a single-stranded, 3'-RNA fragment. Unlike forward translocation, which proceeds one bp at a time, backtracking RNAP can slide along the DNA template over larger distances.

Within the TEC, the disposition of the RNA/DNA hybrid is such that the single-stranded 3'-RNA fragment, generated during backtracking, has no place to go except behind the  $\beta'$  bridging helix and out the secondary channel, the reverse of the presumed route of incoming nucleotide substrates (Cramer et al., 2000, 2001; Korzheva et al., 2000; Zhang et al., 1999; Figure 3). This path for the 3'-RNA fragment during backtracking is supported by crosslinking studies (Borukhov et al., 1991a; Epshtein et al., 2002; Markovtsov et al., 1996).

In the TEC/GreB complex, although the GreB N-terminal coiled coil extends from outside the RNAP active site channel into the active site via the secondary channel, it does not block the secondary channel (Figure 2).

Even in the presence of GreB, the narrowest diameter of the secondary channel in the model is roughly 10 Å, leaving room for the backtracked RNA, or for incoming nucleotide substrates to enter the active site during elongation. This is possible because one edge of the GreB coiled coil fits snugly into a narrow groove in the roof of the secondary channel (Figures 2 and 4A). The opposite face of the coiled coil is exposed to the surface of the secondary channel, in an optimal position to interact with the backtracked RNA (Figures 3 and 4).

#### RNAP/Gre Factor Interactions

It is informative to divide GreB into three structural features, the N-terminal coiled coil (residues 1–36 and 49–73), the coiled-coil tip (residues 37–48), and the CTD (residues 74–159). The interactions of these GreB features with elements of the RNAP structure within the TEC/GreB model are tabulated in Table 1 and illustrated in Figures 4A and 4B.

Outside the RNAP active site channel, the GreB-CTD interacts exclusively with two  $\alpha$  helices from  $\beta'$  that form a part of the entrance rim of the secondary channel (Figures 2 and 4A). These  $\alpha$  helices lie within a region of nonconserved sequence found between  $\beta'$  conserved regions D and E. Hydroxyl radical protein-protein footprinting identified a region of GreB-CTD (residues 116–127) as being an important determinant for binding core RNAP, and Ala substitutions of two residues within this region (D121A and P123A) caused binding defects (Loizos and Darst, 1999). The orientation of the GreB-CTD is such that this region (and these two residues in particular) is within interaction distance of  $\beta'$  (Table 1; Figure 4A).

Moving toward the RNAP active site, one face of the GreB coiled coil interacts with nonconserved residues of  $\beta'$  lining the secondary channel. Finally, the coiled-

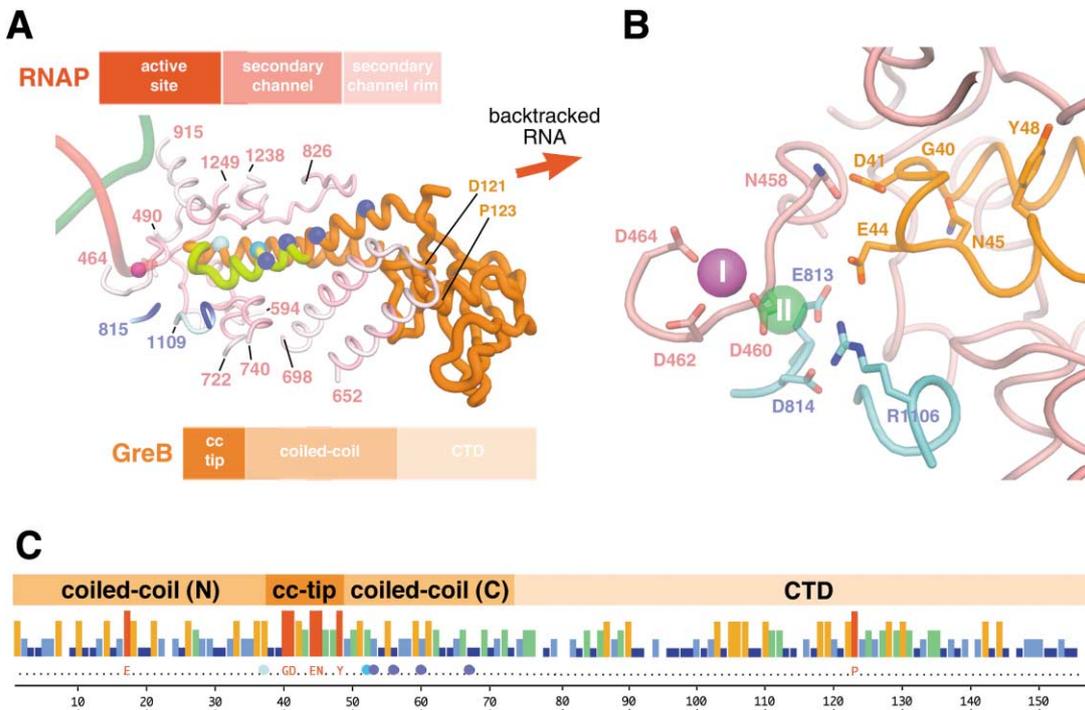


Figure 4. RNAP/GreB Interactions and Gre Factor Sequence Conservation

(A) RNAP/GreB interactions. Segments of RNAP within 10 Å of the modeled GreB (also see Table 1) are shown as backbone worms colored cyan ( $\beta$ ) or light pink ( $\beta'$ ), but segments within 4 Å of GreB are colored blue ( $\beta$ ) or dark pink ( $\beta'$ ). The active site  $Mg^{2+}$  is shown as a magenta sphere. GreB is shown as a thick orange backbone worm, but a segment that crosslinks to backtracked RNA with a 3' crosslinkable nucleotide analog is yellow (Koulich et al., 1997; Stebbins et al., 1995). The blue sphere denotes  $\alpha$  carbons of conserved basic patch residues. From left to right, these are: GreA R37 (cyan); GreA R52/GreB K52 (medium blue); GreB R53, R56, R60, and R67 (dark blue). Two positions in the GreB-CTD that are important for RNAP binding (D121 and P123) are also shown (Loizos and Darst, 1999). The RNA (red) and template strand DNA (dark green) are shown as transparent worms. The path of backtracking RNA from left to right, over the GreB coiled coil and out the secondary channel rim; red, top), while GreB is divided into the coiled-coil tip, the coiled coil, and CTD (orange, bottom).

(B) RNAP/GreB interactions near the RNAP active site. RNAP and GreB backbone worms are color-coded as in (A). Absolutely conserved side chains of the Gre-coiled-coil tip, and selected absolutely conserved side chains of the RNAP active site, are shown, with carbon atoms colored according to the backbones: N, blue and O, red. The active site  $Mg^{2+}$  consistently observed in crystal structures (Cramer et al., 2001; Vassilyev et al., 2002; Zhang et al., 1999) is shown as a magenta sphere (I). The modeled position of MgII (Sosunov et al., 2003) is shown as a transparent green sphere.

(C) Sequence conservation among Gre factor sequences. The histogram represents sequence identity (red bar, 100%; no bar, <20%) from an alignment of 56 GreA and 10 GreB sequences (the full alignment is available as Supplemental Data). Above the histogram, the structural features of GreA/B (coiled-coil N-terminal helix, coiled-coil tip, coiled-coil C-terminal helix, CTD) are denoted. Below the histogram, the identity of residues conserved among all sequences is shown in red. Conserved basic patch residues are shown in light blue (GreA) or dark blue (GreB). The sequence numbering (using *E. coli* GreB) is shown at the bottom.

coil tip makes interactions with some of the most highly conserved regions of the  $\beta$  and  $\beta'$  subunits surrounding the RNAP active site, including the universally conserved "NADFDGD" motif ( $\beta'$  458–464; Cramer et al., 2001; Mustaev et al., 1997; Zhang et al., 1999) that chelates a  $Mg^{2+}$ -ion essential for catalytic activity (Table 1; Figure 4).

In a sequence alignment of all available GreA/B sequences (56 GreA and 10 GreB sequences, Figure 4C; the full alignment is available as Supplemental Data), the sequence conservation is relatively low. Within the CTD, only Pro123 is absolutely conserved, and a Pro at this position is important for RNAP binding (Loizos and Darst, 1999). The low sequence conservation, combined with the fact that most regions of GreB interact with nonconserved regions of  $\beta'$  (Table 1) explains why, in cases tested, Gre factor-stimulated cleavage is genus-specific (Hogan et al., 2002).

The most highly conserved region of Gre factors is clearly the coiled-coil tip. Over all residues, only 7 are absolutely conserved, and 5 of those (Gly40, Asp41, Glu44, Asn45, and Tyr48) are found in the 12 residue tip (Figure 4C). The disposition of these absolutely conserved residues is detailed in relation to selected, absolutely conserved residues in the RNAP active site in Figure 4B.

#### Gre Factor Basic Patch and Transcript Cleavage

The distal region of the Gre factor coiled coil interacts with the transcript 3' end (Koulich et al., 1998; Stebbins et al., 1995; Figure 4A), and biochemical and genetic experiments show that the coiled coil is responsible for inducing the transcript cleavage reaction (Koulich et al., 1997, 1998). Analysis of the surface charge distribution revealed the presence of a "basic patch" on one face of the coiled coil (Koulich et al., 1997; Stebbins et al.,

Table 1. RNAP/Gre Factor Interactions

GreB Structural Feature	Interacting RNAP Segment <sup>a</sup>	RNAP Conserved Region	RNAP Structural Feature
coiled coil <sup>b</sup>	β'597–603	-	secondary channel
coiled coil <sup>b</sup>	β'827–838	-	secondary channel
coiled coil <sup>b</sup>	β'1241–1249	-	secondary channel
coiled-coil tip <sup>c</sup>	β812–814 <sup>f</sup>	βF	active site
coiled-coil tip <sup>c</sup>	β1105–1107 <sup>e</sup>	βH	active site
coiled-coil tip <sup>c</sup>	β'454–459 <sup>g</sup>	β'D	active site
coiled-coil tip <sup>c</sup>	β'496–505	β'D	active site
coiled-coil tip <sup>c</sup>	β'725–739 <sup>e</sup>	β'E	active site
coiled-coil tip <sup>c</sup>	β'927–929 <sup>e</sup>	β'G	active site
CTD <sup>d</sup>	β'655–658	-	secondary channel rim
CTD (near Pro123) <sup>d</sup>	β'662–664	-	secondary channel rim
CTD (near Asp 121) <sup>d</sup>	β'680–688	-	secondary channel rim

<sup>a</sup>RNAP residues within 4 Å of GreB, expressed in *E. coli* RNAP numbering.

<sup>b</sup>GreB residues 1–36 and 49–73.

<sup>c</sup>GreB residues 37–48.

<sup>d</sup>GreB residues 74–158.

<sup>e</sup>RNAP segment cleaved by hydroxylradicals generated from Fe<sup>2+</sup> in the active site (Mustaev et al., 1997).

<sup>f</sup>RNAP β subunit segment containing absolutely conserved βE813–D814 motif.

<sup>g</sup>Beginning of absolutely conserved active site motif, NADFDGD (β'458–464).

1995). GreA has a small basic patch localized near the distal end of the coiled coil, while the GreB basic patch comprises the entire length of the coiled coil. The basic residues responsible for these different features are conserved among GreA or GreB sequences, but not between them (Koulich et al., 1997; Figure 4C). Structure-based mutagenesis showed that the extent of the basic patch determined the mode of transcript cleavage (GreA-like, with 2–3 nt RNA cleavage products, or GreB-like, with 2–18 nt cleavage products; Kulish et al., 2000). This led to the proposal that the role of the basic patch was to anchor the single-stranded 3' fragment of the backtracked RNA, and that the basic patch served as a “molecular ruler” to detect the length of the 3'-RNA fragment (Kulish et al., 2000). The TEC/Gre structural model corroborates these proposals.

The disposition of the coiled coil along the secondary channel is such that the conserved basic residues involved in forming the GreA or GreB basic patches (Figure 4C) are exposed in a line along the direction of the expected path for the backtracked RNA 3' fragment (Figure 4A). Near the coiled-coil tip, the GreA basic patch residues (R37 and R52) are located a distance of roughly 20 Å along the secondary channel from the RNAP active site Mg<sup>2+</sup>. On the other hand, the GreB basic patch residues (K52, K53, R56, R60, and K67) define a line along the exposed edge of the coiled-coil α helix, starting roughly 20 Å from the active site Mg<sup>2+</sup> and ending nearly 45 Å away, just at the secondary channel entrance (Figure 4A). This correlates with the ability of GreB to induce cleavage of backtracked RNA segments anywhere from 2–18 nt in length, while GreA induces cleavage of only 2–3 nt segments.

#### The RNAP Active Site and Gre Factor Mechanism

The distribution of basic patch residues along the exposed face of the coiled-coil addresses how GreA and GreB affect different modes of transcript cleavage, but does not address the stimulation of the cleavage reaction itself. For this, we turn to a description of the RNAP active site and its enzymatic activities detailed in Sosu-

nov et al. (2003). The principal activity of the RNAP is, of course, polymerization, the formation of a phosphodiester bond through the transfer of a nucleotidyl moiety from the NTP substrate to the 3'-OH of the RNA terminus. In addition, the RNAP can catalyze RNA degradation reactions, *exo-* or *endonucleolytic* phosphodiester bond cleavage, by either pyrophosphorolysis or hydrolysis. The multisubunit cellular RNAPs likely catalyze all of these reactions through the same two-metal-ion mechanism proposed for single-subunit DNA and RNAPs (Steitz, 1998). In the polymerization reaction, one metal (MgI) coordinates the attacking 3'-OH. MgII ligates the leaving group, facilitating bond breakage between the α- and β-phosphates of the NTP substrate. Both metals stabilize the pentacovalent transition state. In the reverse reactions, the roles of the metal ions switch.

In the available high-resolution crystal structures of multisubunit cellular RNAPs (Cramer et al., 2001; Gnatt et al., 2001; Vassylyev et al., 2002; Zhang et al., 1999), only one active site metal ion (presumed to be MgI) is observed in a reliable position (Figure 4B), consistent with the observation that only one metal ion (corresponding to MgI) is chelated tightly by the enzyme, with a dissociation constant ( $K_d$ ) on the order of 10 μM, through three absolutely conserved Aspartates of the β'-N<sup>458</sup>AD<sup>460</sup>FD<sup>462</sup>GD<sup>464</sup> motif (Mustaev et al., 1997). In the absence of stabilizing factors, MgII is not chelated effectively by the enzyme ( $K_d > 100$  mM). Based on the strict requirements of the RNAP SN2 reaction mechanism (Yee et al., 1979), Sosunov et al. (2003) modeled the position of MgII (Figure 4B). Consistent with the weak binding of MgII, the metal is likely coordinated by two of the three conserved aspartates (D460 and D462), pointing to the role of stabilizing factors to assist the enzyme in coordinating MgII. In the case of the polymerization reaction, the phosphates of the NTP substrate stabilize MgII binding. In the case of pyrophosphorolysis, pyrophosphate stabilizes MgII. In the case of the hydrolytic reactions, Sosunov et al. (2003) propose that the phosphates of a nonbase-paired NTP bound in a nonproductive site adjacent to the active site (toward

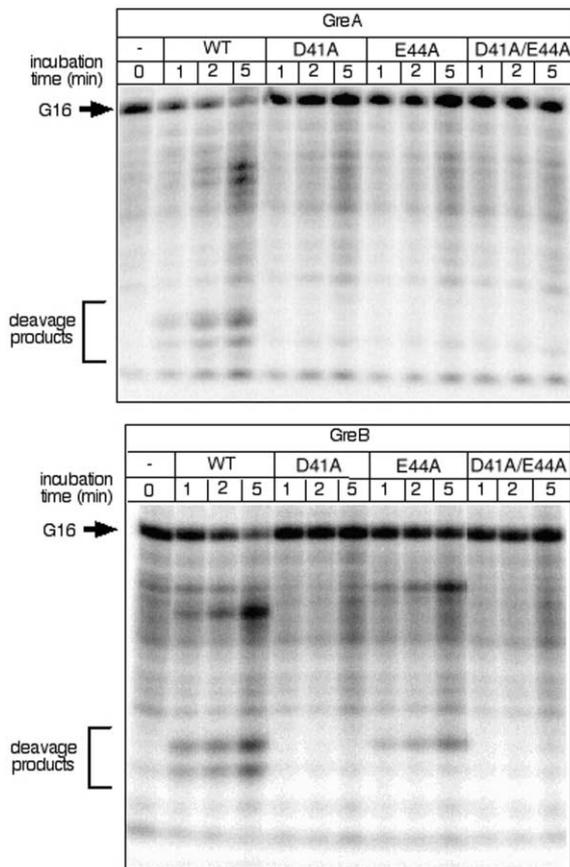


Figure 5. Effect of Alanine Substitutions on Gre Factor-Induced Cleavage Activity

G16 complexes (Feng et al., 1994) were generated and incubated without (-) or with wild-type (wt) or mutant GreA (top image) or GreB (bottom image), and then labeled RNA was separated by electrophoresis and visualized by phosphorimager. Cleavage activity results in the disappearance of the starting G16 complexes, and the appearance of shorter RNAs (cleavage products) with time.

the secondary channel) serve to stabilize MgII binding (evidence for such a binding site has been observed in NTP-soaked crystals of *Taq* RNAP; S.A.D., unpublished data). However, the nonspecific NTP stimulates the exonucleolytic reaction, generating NMP as a degradation product, but not the endonucleolytic reaction stimulated by GreA/B (Sosunov et al., 2003). Steric considerations indicate that the 3'-single-stranded RNA fragment of the backtracked transcript within the secondary channel would prevent the binding of the MgII-stabilizing nonspecific NTP. Thus, we propose that the primary role for the Gre factor coiled-coil tip in interacting with the RNAP active site is to stabilize the binding of MgII, allowing catalysis of the endonucleolytic cleavage of the 3'-RNA fragment.

Examination of the active site details in the TEC/GreB model (Figure 4B) suggests a critical role for two absolutely conserved acidic residues in the Gre factor coiled-coil tip (Asp41 and Glu44) positioned very near the RNAP active site. Ala substitutions at either of these positions caused severe defects in the ability of GreA or GreB to stimulate transcript cleavage in an in vitro assay (Figure 5). The D41A substitution in GreA or GreB, and the E44A

substitution in GreA, appeared to completely abolish Gre factor activity, whereas the E44A substitution in GreB retained a low level of activity. Native gel shift assays indicated that the mutants bound RNAP with a similar affinity as the wild-type factors (data not shown). Thus, the mutations caused defects in the ability of the Gre factors to stimulate the transcript cleavage reaction, not in their ability to bind RNAP, consistent with findings that the Gre-NTD contributes little to overall Gre factor binding (Koulich et al., 1998).

In the TEC/GreB model (Figure 4B), conserved Asp41 and Glu44 are situated to effect the coordination of MgII by either a direct or an indirect mechanism. Small shifts in the positioning of the Gre-coiled coil (taking into account the low-resolution of the cryo-EM map that formed the basis for the model) would allow direct coordination to MgII.

Alternatively, an indirect mechanism for Gre factor stabilization of MgII binding is possible. In addition to the three aspartates of the  $\beta'$ -N<sup>458</sup>AD<sup>460</sup>FD<sup>462</sup>GD<sup>464</sup> motif, two additional, universally conserved acidic residues,  $\beta$ E813 and  $\beta$ D814, are located within angstroms of MgI (Figure 4B).  $\beta$ D814 in particular might be capable of chelating MgII except that  $\beta$ D814 is captured in a salt bridge with another absolutely conserved residue,  $\beta$ R1106 (Figure 4B). High pH (>9–10) or Ala substitution of  $\beta$ R1106 stimulated the RNAP exonuclease activity and removed the dependence on the nonspecific NTP (Sosunov et al., 2003). These results, in combination with other mutational studies, were attributed to the loss of the  $\beta$ D814- $\beta$ R1106 salt bridge, which would free  $\beta$ D814 to bind MgII. This mechanism also explains how high pH stimulates the intrinsic RNAP transcript cleavage activity in the absence of Gre factors (Awrey et al., 1997; Orlova et al., 1995). This suggests an alternative, indirect mechanism for the effect of the Gre factor conserved acidic residues by swapping the salt bridge with  $\beta$ R1106, thereby freeing  $\beta$ D814 to bind MgII.

The TEC/GreB model suggests a role for another conserved Gre factor residue of the coiled-coil tip, Tyr48. It is located, along with the GreA basic patch, on the exposed face of the Gre-coiled coil, in the secondary channel, roughly 20 Å from MgI (Figure 4B). This suggests Tyr48 may directly interact with backtracked RNA to anchor its position, in collaboration with the electrostatic interactions from the basic patch residues.

#### Functional Relationship of Prokaryotic Gre Factors with Eukaryotic TFIIIS

The bacterial Gre factors and eukaryotic TFIIIS show striking functional similarity, yet no sequence or structural relationship (Olmsted et al., 1998). TFIIIS consists of three domains (Morin et al., 1996). Domains II and III, along with a 26 residue unstructured linker, are necessary and sufficient for transcript cleavage activity (Agarwal et al., 1991; Nakanishi et al., 1995; Olmsted et al., 1998). Structures for domains II (a 3 helix bundle) and III (a zinc ribbon) have been determined by solution NMR (Morin et al., 1996; Qian et al., 1993). Genetic and functional studies have shown that the role of domain II is to bind RNAP II (Agarwal et al., 1991; Awrey et al., 1998), probably at a site outside the entrance to the RNAP II pore (analogous to the bacterial RNAP secondary chan-

nel; Archambault et al., 1992; Wu et al., 1996). Domain III plays a critical role in stimulating the transcript cleavage reaction. Most interestingly, domain III contains two absolutely conserved acidic residues (Asp290 and Glu291, using yeast TFIIIS numbering), and mutation of either of these residues to any other residue abolishes transcript cleavage activity (Awrey et al., 1998; Cipres-Palacin and Kane, 1995; Jeon et al., 1994). Residues within the unstructured linker, as well as the linker length, are also critical for activity, suggesting it may interact with RNAP II and serve as a spacer to correctly position domains II and III (Awrey et al., 1998).

These facts suggest a mechanistically analogous model for TFIIIS function, where TFIIIS-domain II takes the role of the Gre-CTD, the domain II-III linker takes the role of the Gre coiled coil, and domain III takes the role of the Gre coiled-coil tip. In this model, domain II anchors TFIIIS to the RNAP II elongation complex by binding near the entrance to the pore. The linker extends into the pore, probably binding to one surface of the pore, allowing domain III to position itself so that the absolutely conserved DE motif can interact with the RNAP II active site to effect the transcript cleavage reaction, probably through stabilization of MgII binding. New structural results now strongly support this model (Kettenberger et al., 2003 [this issue of *Ce/I*]).

## Conclusions

Insights derived from the TEC/GreB model explain a wide range of experimental observations. The model confirms suspected roles for some Gre factor structural features, such as the role of the basic patch in directly interacting with, and detecting the length of the backtracked RNA to determine the mode of cleavage. In addition, the model provides unexpected insights; such as the role of conserved acidic residues at the Gre coiled-coil tip in remodeling the RNAP active site to catalyze the cleavage reaction. Further insights into the structural and chemical mechanism for Gre factor-mediated transcript cleavage will come from additional mutational and biochemical studies, aided by the model, as well as high-resolution structural information on the complex.

Sequence analysis indicates that the eukaryotic and prokaryotic RNAPs are clearly linked through evolution (Archambault and Friesen, 1993; Minakhin et al., 2001; Zhang et al., 1999; Zhang and Darst, 1998). Moreover, structures of the eukaryotic and prokaryotic enzymes are even more closely related than expected from their sequences (Cramer, 2002; Cramer et al., 2001; Darst, 2001; Ebright, 2000; Zhang et al., 1999). Both prokaryotic GreA/B and eukaryotic TFIIIS are important regulatory factors that interact extensively with their respective RNAPs, affecting highly conserved structural and functional features of the enzymes. Thus, it is surprising to find that GreA/B and TFIIIS apparently evolved independently, having no sequence or structural similarity. Nevertheless, the factors were known to be close functional homologs, and we now suggest that the structurally divergent factors stimulate transcript cleavage by their respective RNAPs using essentially the same molecular mechanism.

Whereas the core catalytic component of the RNAP is highly conserved from bacteria to man, many regulatory factors that act directly on the enzyme are not. For instance, many of the eukaryotic basal initiation factors show little or no sequence similarity with their prokaryotic functional counterpart, the  $\sigma$  subunit (Gross et al., 1998; Kornberg, 1999), and available structures of eukaryotic basal initiation factors do not suggest structural relationships with  $\sigma$  (Campbell et al., 2002; Geiger et al., 1996; Juo et al., 2003; Kamada et al., 2001; Malhotra et al., 1996; Nikolov et al., 1992, 1995; Tan et al., 1996). As we learn more about the structural and functional mechanism of transcription regulation in prokaryotes and eukaryotes, it may not be surprising to find additional examples of unrelated regulatory factors that manipulate the conserved core transcription machinery using identical strategies.

## Experimental Procedures

### Protein Preparation

*E. coli* core RNAP, assembled in vivo, was prepared from a coexpression system (pGEMABC, K.S. Murakami, personal communication), expressing *rpoA*, *B*, and *C* genes simultaneously (coding for the RNAP  $\alpha$ ,  $\beta$ , and  $\beta'$  subunits, respectively). After growing transformed cells in LB medium with ampicillin (100  $\mu$ g/ml), expression was induced by adding IPTG to a final concentration of 1 mM, and cells were grown for an additional 3 hr at 37°. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl, [pH 8.0], 1 mM EDTA, 5 mM DTT, and 0.5 mM PMSF), lysed by passage through a continuous flow French press (Avestin), then clarified by centrifugation. Polymin P was added to the supernatant to a final concentration of 0.6 % v/v. The precipitate was pelleted and washed by two successive cycles of suspension and centrifugation using buffer A (10 mM Tris-HCl, [pH 8], 0.1 mM EDTA, and 5% glycerol) + 0.5 M NaCl, and a third wash with buffer A + 1 M NaCl to elute RNAP. The supernatant was precipitated by adding  $(\text{NH}_4)_2\text{SO}_4$ . The precipitate was pelleted by centrifugation and resuspended in buffer A + 0.1 M NaCl. The sample was loaded onto a BioRex 70 column (Amersham Pharmacia Biotech) and eluted with a linear NaCl gradient. The eluate was then loaded onto a Superdex 200 gel filtration column (Amersham Pharmacia Biotech) equilibrated with buffer A + 0.5 M NaCl. Fractions containing pure protein were pooled and stored at -80°C.

Cloning, expression, and purification of wild-type and mutant GreA/B were performed as described (Loizos and Darst, 1999). GreA/B mutants were constructed using the QuikChange site-directed mutagenesis system (Stratagene). The mutations were confirmed by DNA sequencing (The Rockefeller University DNA Sequencing Resource Center) and in some cases by electrospray mass spectrometry.

### Crystallization

The complex of core RNAP with GreB was studied since GreB forms a more stable complex with RNAP (Koulich et al., 1997). Wild-type GreB was added to *E. coli* core RNAP (10:1 molar ratio) and tubular crystals were grown on positively charged liposomes as described (Darst et al., 2002; Opalka et al., 1999; Polyakov et al., 1995, 1998). The presence of GreB in the crystals was confirmed by a binding assay (Polyakov et al., 1998).

### Data Collection and Structure Determination

Specimen preparation, cryo-EM, and analysis were performed as described (Darst et al., 2002; Opalka et al., 1999). For each individual tube, a density map was calculated. Tubes showing similar unit cells were averaged in Fourier space to generate an averaged map. Two averaged maps were calculated combining 6 and 4 tubes, respectively (Supplemental Data available on website) using little-g averaging (DeRosier et al., 1998). These two maps were then aligned and averaged using a real-space procedure (Rice et al., 2001). Layerlines extending to 13 Å resolution were included in the final recon-

struction. Flexible fitting of the *Taq* core RNAP crystal structure and the GreB homology model were performed as described previously (Darst et al., 2002; Wriggers and Chacon, 2001), with some modifications (see Supplemental Data).

#### Gre Factor-Induced Cleavage Reactions

Gre factor-induced transcript cleavage was assayed by monitoring the cleavage of stalled complexes containing a 16-mer RNA transcript (G16) prepared from a modified T7A1 transcription unit (Feng et al., 1994). Reactions were performed in 30  $\mu$ l of standard transcription buffer (STB; 40 mM Tris-HCl, [pH 8.0], 100 mM KCl, and 10 mM MgCl<sub>2</sub>). *E. coli* core RNAP (His<sub>6</sub>-tagged at the  $\beta$ '-C terminus, 0.3  $\mu$ M) was incubated with 1  $\mu$ M  $\sigma^{70}$  for 10 min at 37°C, followed by the addition of 1 mM ApU dinucleotide (Sigma-Aldrich) and 0.1  $\mu$ M of pCL185T7A1 promoter DNA (Feng et al., 1994) for an additional 10 min at 37°C. Reactions were initiated by the addition of 15  $\mu$ M ATP, CTP, and 0.4  $\mu$ M [ $\alpha$ -<sup>32</sup>P]GTP (3000 Ci/mmol; New England Nuclear) to obtain the stalled G16 complex. After 5 min at 37°C, complexes were immobilized on 10  $\mu$ l of Ni-NTA agarose (Qiagen), washed twice with 1 ml ice-cold STB + 1 M NaCl, followed by a 1 ml wash with STB. After a 30 min incubation at 4°C, complexes were eluted from the beads with 50  $\mu$ l STB + 100 mM imidazole and used in cleavage assays. Cleavage reactions were performed in 5  $\mu$ l STB + 100 mM imidazole at 37°C with 15 molar excess of wild-type or mutant GreA/B over the core RNAP. Reactions were terminated by the addition of 5  $\mu$ l loading buffer containing 7 M urea in TBE buffer, boiled for 1 min, separated by 20% polyacrylamide gel electrophoresis under denaturing conditions, and visualized and quantitated using PhosphorImager analysis (Molecular Dynamics).

#### Acknowledgments

We are grateful to D. Stokes for use of the CM200 field emission gun (Skirball Institute, New York) and T. Muir for use of the electrospray mass spectrometer. We thank R. Landick for insightful discussion. Figures were generated using DINO (<http://www.dino3d.org>). N.O. was supported by funds from the Philippe Foundation. P.C. was supported by funds from the La Jolla Interfaces in Science Program/Burroughs Wellcome Fund. This work was supported by National Institutes of Health Grant GM62968 (W.W.), and GM58020 and GM61898 (S.A.D.).

Received: May 21, 2003  
Revised: June 20, 2003  
Accepted: July 3, 2003  
Published: August 7, 2003

#### References

Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Yoon, H. (1991). Stimulation of transcript elongation requires both the zinc finger and RNA polymerase II binding domains of human TFIIIS. *Biochemistry* 30, 7842–7851.

Archambault, J., and Friesen, J.D. (1993). Genetics of RNA polymerases I, II, and III. *Microbiol. Rev.* 57, 703–724.

Archambault, J., Lacroute, F., Ruet, A., and Friesen, J.D. (1992). Genetic interaction between transcription elongation factor TFIIIS and RNA polymerase II. *Mol. Cell. Biol.* 12, 4142–4152.

Awrey, D.E., Weilbacher, R.G., Hemming, S.A., Orlicky, S.M., Kane, C.M., and Edwards, A.M. (1997). Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIIS. *J. Biol. Chem.* 272, 14747–14754.

Awrey, D.E., Shimasaki, N., Koth, C., Weilbacher, R.G., Olmsted, V.K., Kazanis, S., Shan, X., Arellano, J., Arrowsmith, C.H., Kane, C.M., and Edwards, A.M. (1998). Yeast transcript elongation factor (TFIIIS), structure and function. II. RNA polymerase binding, transcript cleavage, and read-through. *J. Biol. Chem.* 273, 22595–22605.

Borukhov, S., and Goldfarb, A. (1996). Purification and assay of *Escherichia coli* transcript cleavage factors GreA and GreB. *Methods Enzymol.* 274, 315–326.

Borukhov, S., Lee, J., and Goldfarb, A. (1991a). Mapping of a contact

for the RNA 3' terminus in the largest subunit of RNA polymerase. *J. Biol. Chem.* 266, 23932–23935.

Borukhov, S., Severinov, K., Kashlev, M., Lebedev, A., Bass, I., Rowland, G.C., Lim, P.-P., Glass, R.E., Nikiforov, V., and Goldfarb, A. (1991b). Mapping of trypsin cleavage and antibody-binding sites and delineation of a dispensable domain in the  $\beta$  subunit of *Escherichia coli* RNA polymerase. *J. Biol. Chem.* 266, 23921–23926.

Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992). GreA protein: a transcription elongation factor from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 89, 8899–8902.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. *Cell* 72, 459–466.

Campbell, E.A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S.A. (2001). Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* 104, 901–912.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma factor. *Mol. Cell* 9, 527–539.

Cipres-Palacin, G., and Kane, C.M. (1995). Alanine-scanning mutagenesis of human transcript elongation factor TFIIIS. *Biochemistry* 34, 15375–15380.

Cramer, P. (2002). Multisubunit RNA polymerases. *Curr. Opin. Struct. Biol.* 12, 89–97.

Cramer, P., Bushnell, D.A., Fu, J., Gnat, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640–649.

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

Darst, S.A. (2001). Bacterial RNA polymerase. *Curr. Opin. Struct. Biol.* 11, 155–162.

Darst, S.A., Opalka, N., Chacon, P., Polyakov, A., Richter, C., Zhang, G., and Wriggers, W. (2002). Conformational flexibility of bacterial RNA polymerase. *Proc. Natl. Acad. Sci. USA* 99, 4296–4301.

DeRosier, D., Stokes, D.L., and Darst, S.A. (1998). Averaging data derived from images of helical structures with different symmetries. *J. Mol. Biol.* 289, 159–165.

Ebright, R.H. (2000). RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *J. Mol. Biol.* 293, 199–213.

Epshtein, V., Mustaev, A., Markovtsov, V., Bereshchenko, O., Nikiforov, V., and Goldfarb, A. (2002). Swing-gate model of nucleotide entry into the RNA polymerase active center. *Mol. Cell* 10, 623–634.

Erie, D.A., Hajiseyedjavadi, O., Young, M.C., and von Hippel, P.H. (1993). Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* 262, 867–873.

Feng, G., Lee, D.N., Wang, D., Chan, C.L., and Landick, R. (1994). GreA-induced transcript cleavage in transcription complexes containing *E. coli* RNA polymerase is controlled by nascent transcript location and structure. *J. Biol. Chem.* 269, 22282–22294.

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

Fu, J., Gnat, A.L., Bushnell, D.A., Jensen, G.J., Thompson, N.E., Burgess, R.R., David, P.R., and Kornberg, R.D. (1999). Yeast RNA polymerase II at 5 Å resolution. *Cell* 98, 799–810.

Geiger, J.H., Hahn, S., Lee, S., and Sigler, P.B. (1996). Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* 272, 830–836.

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

Gross, C.A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., and Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harb. Symp. Quant. Biol.* 63, 141–155.

Hogan, B.P., Hartsch, T., and Erie, D.A. (2002). Transcript cleavage

- by *Thermus thermophilus* RNA polymerase. Effects of GreA and anti-GreA factors. *J. Biol. Chem.* 277, 967–975.
- Hsu, L.H., Vo, N.V., and Chamberlin, M.J. (1995). *Escherichia coli* transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression *in vivo* and *in vitro*. *Proc. Natl. Acad. Sci. USA* 92, 11588–11592.
- Izban, M.G., and Luse, D.S. (1992). The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'→5' direction in the presence of elongation factor SII. *Genes Dev.* 6, 1342–1356.
- Jeon, C.J., Yoon, H., and Agarwal, K. (1994). The transcription factor TFIIIS zinc ribbon dipeptide Asp-Glu is critical for stimulation of elongation and RNA cleavage by RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 91, 9106–9110.
- Juo, Z.S., Kassavetis, G.A., Wang, J., Geiduschek, E.P., and Sigler, P.B. (2003). Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* 422, 534–539.
- Kamada, K., De Angelis, J., Roeder, R.G., and Burley, S.K. (2001). Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc. Natl. Acad. Sci. USA* 98, 3115–3120.
- Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Architecture of the RNA polymerase II-TFIIIS complex and implications for mRNA cleavage and proofreading. *Cell* 114, this issue, 347–357.
- Komissarova, N., and Kashlev, M. (1997a). RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* 272, 15329–15338.
- Komissarova, N., and Kashlev, M. (1997b). Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc. Natl. Acad. Sci. USA* 94, 1755–1760.
- Kornberg, R.D. (1999). Eukaryotic transcriptional control. *Trends Cell Biol.* 9, M46–M49.
- Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). A structural model of transcription elongation. *Science* 289, 619–625.
- Koulich, D., Orlova, M., Malhotra, A., Sali, A., Darst, S.A., Goldfarb, A., and Borukhov, S. (1997). Domain organization of transcript cleavage factors GreA and GreB. *J. Biol. Chem.* 272, 7201–7210.
- Koulich, D., Nikiforov, V., and Borukhov, S. (1998). Distinct functions of N- and C-terminal domains of GreA, an *Escherichia coli* transcript cleavage factor. *J. Mol. Biol.* 276, 379–389.
- Kulish, D., Lee, J., Lomakin, I., Nowicka, B., Das, A., Darst, S.A., Normet, K., and Borukhov, S. (2000). The functional role of basic patch, a structural element of *Escherichia coli* transcript cleavage factors GreA and GreB. *J. Biol. Chem.* 275, 12789–12798.
- Loizos, N., and Darst, S.A. (1999). Mapping interactions of *Escherichia coli* GreB with RNA polymerase and ternary elongation complexes. *J. Biol. Chem.* 274, 23378–23386.
- Malhotra, A., Severinova, E., and Darst, S.A. (1996). Crystal structure of a  $\sigma^{70}$  subunit fragment from *Escherichia coli* RNA polymerase. *Cell* 87, 127–136.
- Markovtsov, V., Mustaev, A., and Goldfarb, A. (1996). Protein-RNA interactions in the active center of transcription elongation complex. *Proc. Natl. Acad. Sci. USA* 93, 3221–3226.
- Marr, M.T., and Roberts, J.W. (2000). Function of transcription cleavage factors GreA and GreB at a regulatory pause site. *Mol. Cell* 6, 1275–1285.
- Minakhin, L., Bhagat, S., Brunning, A., Campbell, E.A., Darst, S.A., Ebricht, R.H., and Severinova, K. (2001). 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* 98, 892–897.
- Morin, P.E., Awrey, D.E., Edwards, A.M., and Arrowsmith, C.H. (1996). Elongation factor TFIIIS contains three structural domains: solution structure of domain II. *Proc. Natl. Acad. Sci. USA* 93, 10604–10608.
- Murakami, K., Masuda, S., and Darst, S.A. (2002). Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296, 1280–1284.
- Mustaev, A., Kozlov, M., Markovtsov, V., Zaychikov, E., Denissova, L., and Goldfarb, A. (1997). Modular organization of the catalytic center of RNA polymerase. *Proc. Natl. Acad. Sci. USA* 94, 6641–6645.
- Nakanishi, T., Shimoaraiso, M., Kubo, T., and Natori, S. (1995). Structure-function relationship of yeast S-II in terms of stimulation of RNA polymerase II, arrest relief, and suppression of 6-azauracil sensitivity. *J. Biol. Chem.* 270, 8991–8995.
- Nikolov, D.B., Hu, S.H., Lin, J., Gasch, A., Hoffmann, A., Horikoshi, M., Chua, N.H., Roeder, R.G., and Burley, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. *Nature* 360, 40–46.
- Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119–128.
- Nudler, E., Mustaev, A., Lukhtanov, E., and Goldfarb, A. (1997). The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89, 33–41.
- Olmsted, V.K., Awrey, D.E., Koth, C., Shan, X., Morin, P.E., Kazanin, S., Edwards, A.M., and Arrowsmith, C.H. (1998). Yeast transcript elongation factor (TFIIS), structure and function. I: NMR structural analysis of the minimal transcriptionally active region. *J. Biol. Chem.* 273, 22589–22594.
- Opalka, N., Mooney, R.A., Richter, C., Severinova, K., Landick, R., and Darst, S.A. (1999). Direct localization of a  $\beta$  subunit domain on the three-dimensional structure of *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. USA* 97, 617–622.
- Orlova, M., Newlands, J., Das, A., Goldfarb, A., and Borukhov, S. (1995). Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl. Acad. Sci. USA* 92, 4596–4600.
- Polyakov, A., Severinova, E., and Darst, S.A. (1995). Three-dimensional structure of *Escherichia coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* 83, 365–373.
- Polyakov, A., Richter, C., Malhotra, A., Koulich, D., Borukhov, S., and Darst, S.A. (1998). Visualization of the binding site for the transcript cleavage factor GreB on *Escherichia coli* RNA polymerase. *J. Mol. Biol.* 281, 262–266.
- Qian, X., Jeon, C., Yoon, H., Agarwal, K., and Weiss, M.A. (1993). Structure of a new nucleic-acid-binding motif in eukaryotic transcriptional elongation factor TFIIIS. *Nature* 365, 277–279.
- Reeder, T.C., and Hawley, D.K. (1996). Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism. *Cell* 87, 767–777.
- Reines, D. (1992). The RNA polymerase II elongation complex. Factor-dependent transcription elongation involves nascent RNA cleavage. *J. Biol. Chem.* 267, 3795–3800.
- Reines, D., Chamberlin, M.J., and Kane, C.M. (1989). Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene *in vitro*. *J. Biol. Chem.* 264, 10799–10809.
- Rice, W.J., Young, H.S., Martin, D.W., Sachs, J.R., and Stokes, D.L. (2001). Structure of the Na<sup>+</sup>, K<sup>+</sup>-ATPase at 11-Å resolution: Comparison with Ca<sup>2+</sup>-ATPase in E<sub>1</sub> and E<sub>2</sub> states. *Biophys. J.* 80, 2187–2197.
- Rudd, M.D., Izban, M.G., and Luse, D.S. (1994). The active site of RNA polymerase II participates in transcript cleavage within arrested ternary complexes. *Proc. Natl. Acad. Sci. USA* 91, 8057–8061.
- Severinova, K., Kashlev, M., Severinova, E., Bass, I., McWilliams, K., Kutter, E., Nikiforov, V., Snyder, L., and Goldfarb, A. (1994). A non-essential domain of *E. coli* RNA polymerase required for the action of the termination factor Alc. *J. Biol. Chem.* 269, 14254–14259.
- Sluder, A.E., Greenleaf, A.L., and Price, D.H. (1989). Properties of a *Drosophila* RNA polymerase II elongation factor. *J. Biol. Chem.* 264, 8963–8969.
- Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* 22, 2234–2244.
- Stebbins, C.E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb,

- A., and Darst, S.A. (1995). Crystal structure of the GreA transcript cleavage factor from *Escherichia coli*. *Nature* 373, 636–640.
- Steitz, T.A. (1998). A mechanism for all polymerases. *Nature* 391, 231–232.
- Surratt, C.K., Milan, S.C., and Chamberlin, M.J. (1991). Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. *Proc. Natl. Acad. Sci. USA* 88, 7983–7987.
- Tan, S., Hunziker, Y., Sargent, D.F., and Richmond, T.J. (1996). Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature* 381, 127–151.
- Toulme, F., Mosrin-Huaman, C., Sparkowski, J., Das, A., Leng, M., and Rahmouni, A.R. (2000). GreA and GreB proteins revive backtracked RNA polymerase in vivo by promoting transcript trimming. *EMBO J.* 19, 6853–6859.
- Uptain, S.M., Kane, C.M., and Chamberlin, M.J. (1997). Basic mechanisms of transcript elongation and its regulation. *Annu. Rev. Biochem.* 66, 117–172.
- Vassilyev, D.G., Sekine, S., Laptenko, O., Lee, J., Vassilyeva, M.N., Borukhov, S., and Yokoyama, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- Wang, D., and Hawley, D.K. (1993). Identification of a 3′–5′ exonuclease activity associated with human RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 90, 843–847.
- Wriggers, W., and Chacon, P. (2001). Modeling tricks and fitting techniques for multi-resolution structures. *Structure* 9, 779–788.
- Wu, J., Awrey, D.E., Edwards, A.M., Archambeault, J., and Friesen, J. (1996). In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS. *Proc. Natl. Acad. Sci. USA* 93, 11552–11557.
- Yee, D., Armstrong, V.W., and Eckstein, F. (1979). Mechanistic studies on deoxyribonucleic acid dependent ribonucleic acid polymerase from *Escherichia coli* using phosphorothioate analogues. 1. Initiation and pyrophosphate exchange reactions. *Biochemistry* 18, 4116–4120.
- Zhang, G., and Darst, S.A. (1998). Structure of the *Escherichia coli* RNA polymerase  $\alpha$  subunit amino-terminal domain. *Science* 281, 262–266.
- Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811–824.