Cell, Vol. 118, 281–284, August 6, 2004, Copyright ©2004 by Cell Press

Regulation of RNA Polymerase through the Secondary Channel

Minireview

Bryce E. Nickels and Ann Hochschild*
Department of Microbiology
and Molecular Genetics
Harvard Medical School
200 Longwood Avenue, Blg. D-1
Boston, Massachusetts 02115

Summary

High-resolution crystal structures have highlighted functionally important regions in multisubunit RNA polymerases, including the secondary channel, or pore, which is postulated to allow the diffusion of small molecules both into and out of the active center of the enzyme. Recent work from several groups has illustrated how regulatory factors and small molecules can exploit the secondary channel to gain access to the active site and modify the transcription properties of RNA polymerase.

With the recent determination of the first high-resolution crystal structures of multisubunit RNA polymerases (Zhang et al., 1999; Cramer et al., 2000), the study of transcription has entered a new age. These structures, which reveal a remarkable degree of conservation between the prokaryotic and eukaryotic enzymes, have illuminated the catalytic activities of RNA polymerase (RNAP) and provided new insight into the mechanisms employed by transcription regulators. In bacteria, the transcription process can be divided into a number of distinct steps. First, the RNAP holoenzyme (consisting of a catalytically competent core enzyme in complex with a sigma factor) binds to duplex promoter DNA to form the closed RNAP-promoter complex. Next, a series of conformational changes leads to the formation of the initiation-competent open complex in which the DNA is locally melted to expose the transcription start site. RNAP can then initiate transcription, typically directing the synthesis of short abortive RNA products that are released and resynthesized before RNAP breaks its interactions with promoter DNA and escapes into productive elongation. During its translocation along the DNA, the transcription elongation complex encounters pause sites and potential arrest sites at which the nascent RNA remains stably bound to the enzyme. Finally, upon reaching a termination site, RNAP releases the RNA transcript and dissociates from the DNA.

Figure 1A shows the three-dimensional crystal structure of a bacterial core enzyme. Overall, the enzyme has a crab claw shape, the claws comprising the large β and β′ subunits (corresponding to Rpb1 and Rpb2 of the yeast RNAP II). The two claws define the enzyme's main channel, which accommodates the DNA template and the RNA–DNA hybrid that forms during transcription. The active center, which is marked by a stably bound Mg^{2+} ion, lies at the back of the main channel (see Figures 1A and 1B). A second, smaller channel (see Figure 1B) also links the external milieu with the active site of the enzyme. Because the DNA largely blocks access to the active center via the main channel during transcription, this secondary channel is the presumed entryway for substrate nucleotides (NTPs) to reach the site of catalysis (Zhang et al., 1999; Cramer et al., 2000; Korzheva et al., 2000).

Here we focus on an exciting series of recent studies that reveals how regulatory factors can exploit the secondary channel to gain direct access to the catalytic center of the enzyme. When bound within the secondary channel, these factors affect transcription either by modifying the catalytic properties of RNAP or by potentiating the action of small molecule effectors that themselves bind near the RNAP active center. In addition, we highlight studies revealing that inhibitors of RNAP can target the secondary channel.

Transcript Cleavage Factors: GreA/B and TFIIS

Highly homologous in sequence and structure, GreA and GreB are bacterial transcription elongation factors that can prevent the formation of (or reactivate) arrested transcription complexes, which can arise at transcription pause sites through a process of reverse translocation, known as backtracking (reviewed in Fish and Kane [2002]). During backtracking, the RNAP catalytic center slides back relative to the 3′ end of the nascent transcript and the 3′ tail of the RNA is displaced into the secondary channel. For elongation to resume, an endonucleolytic cleavage event must occur to create a new RNA 3′-OH correctly positioned at the active center of the enzyme. This cleavage reaction is carried out by the RNAP active site itself; although the reaction is ordinarily disfavored, it can be stimulated by high pH or the addition of a protein cofactor (reviewed in Fish and Kane [2002]).

The structure of GreA/B consists of an N-terminal antiparallel coiled coil and a C-terminal globular domain (Figure 1C) (see Fish and Kane [2002] and references therein). The modeling of high-resolution component structures onto a low-resolution structure of an *E. coli* RNAP-GreB complex revealed that the C-terminal domain of GreA/B is bound to the rim of the secondary channel, while the coiled coil extends into the secondary channel such that its tip approaches the active center of the enzyme (Opalka et al., 2003). Two other structural models, constructed based on biochemical data, also place the coiled-coil tip at or near the catalytic center (Laptenko et al., 2003; Sosunova et al., 2003). A pair of invariant acidic residues that protrude from the coiled-coil tip plays a critical role in the transcript cleavage reaction. The specific function of these acidic residues was inferred in the context of the proposal that all the catalytic activities of RNAP proceed by a two-metal-ion mechanism (Steitz, 1998; Sosunov et al., 2003). One of the two metal ions (MgI) is tightly bound within the active center of the enzyme, being chelated by three invariant aspartate residues, whereas the second (MgII) is bound only weakly by the enzyme and requires additional stabilization. Modeling suggests that during RNA synthesis the phosphates of the incoming NTP substrate are suitably positioned to provide the required stabilization (Sosunov et al., 2003). The chemistry of the endonucleolytic cleavage reaction that can be stimulated by the Gre factors similarly requires the participation of both MgI and MgII; in this case, the acidic residues at the tip of

*Correspondence: ahochschild@hms.harvard.edu

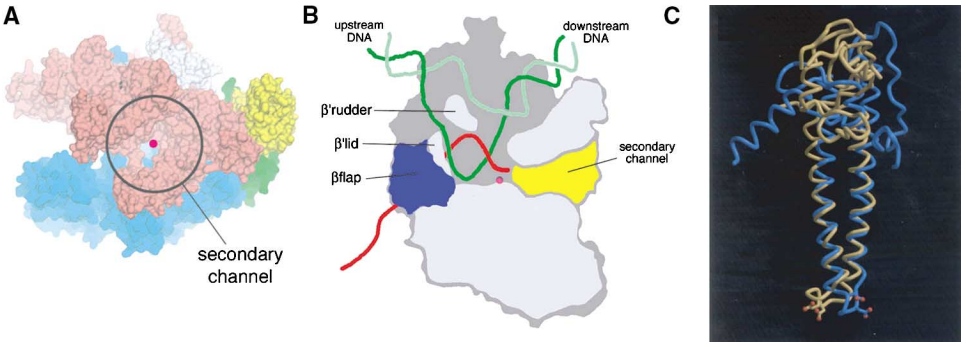

Figure 1. The RNAP Secondary Channel and the Structures of GreA and DksA

(A) Space-filling representation of the three-dimensional structure of core RNAP from *Thermus aquaticus* (courtesy of S.A. Darst). The β′ and β subunits are colored pink and cyan, respectively, the α subunits are green and yellow, and the ω subunit is gray. Shown through the secondary channel is MgI (magenta sphere).

(B) Cross-sectional view of a transcription elongation complex (adapted from Opalka et al. [2003]). To generate this view, the image in (A) was rotated ~90° clockwise about the z axis and then ~90° about the *y* axis so that the β pincer lies on top of the β′ pincer. Then the model was sliced in half, parallel to the plane of the page, allowing the β′ pincer to be viewed from inside the main channel. The template and nontemplate DNA strands are colored dark and light green, respectively, and the nascent RNA is red. The active site Mg^{2+} (MgI) is shown as a magenta sphere. Areas of the structure that were sliced in the plane of the page are light colored (except for the β flap, which is dark blue) and areas of the structure that lie below the plane of the page are darker colored, except for the secondary channel, which is highlighted in yellow.

(C) GreA and DksA superimposed (from Perederina et al. [2004]). The α-carbon backbones of GreA (tan) and DksA (blue) were superimposed by their coiled-coil domains. The conserved acidic residues at the coiled-coil tips are shown in ball-and-stick representation (bottom).

the Gre coiled coil appear well positioned to stabilize the binding of MgII (Laptenko et al., 2003; Opalka et al., 2003; Sosunova et al., 2003). This proposal is supported by mutagenesis experiments that confirmed the functional importance of the conserved acidic residues in stimulating endonucleolytic cleavage (Laptenko et al., 2003; Opalka et al., 2003; Sosunova et al., 2003).

The mechanism of Gre action implies that GreA or GreB and an extruded 3′ RNA segment can simultaneously occupy the secondary channel. In fact, the coiled-coil domains of both GreA and GreB contain a cluster of positively charged residues that define a surface-exposed basic patch well positioned to contact the extruded RNA in backtracked transcription complexes (Kulish et al., 2000; Sosunova et al., 2003). The interaction between this basic patch and the extruded RNA likely stabilizes the binding of Gre to the elongation complex, raising the possibility that the Gre factors are recruited specifically to backtracked elongation complexes in vivo.

Remarkably, the eukaryotic transcription elongation factor TFIIS, which is functionally analogous to the Gre factors but structurally unrelated, also protrudes into the RNAP secondary channel and positions a pair of absolutely conserved acidic residues near the active center of the enzyme (Kettenberger et al., 2003). Structural analysis suggests that these two acidic residues stimulate transcript cleavage by stabilizing the binding of MgII (Kettenberger et al., 2003). Thus, both the Gre factors and TFIIS gain access to the RNAP active center via the secondary channel and donate specific residues that function in conjunction with conserved residues in the active center to modify the catalytic properties of RNAP.

DksA and the Small Molecule Effectors ppGpp and Initiating Nucleotide

Recent advances in the understanding of the regulation of ribosomal RNA (rRNA) transcription have led to the discovery of another transcription factor, DksA, that binds within the secondary channel to modulate RNAP function. DksA bears a striking structural resemblance to the Gre factors, possessing a coiled-coil domain (with two conserved acidic amino acid residues at its tip) that is nearly superimposable on the coiled-coil domains of the Gre factors (Figure 1C). However, DksA is not a transcript cleavage factor and instead sensitizes RNAP to the opposing effects of two small molecules, ppGpp and the nucleotide that will be incorporated at the 5′ end of the transcript (the initiating nucleotide, iNTP).

The transcription of the rRNA genes is the rate-limiting step in ribosome synthesis, and the rRNA promoters have unusual kinetic properties that render them uniquely sensitive to changes in the intracellular concentrations of the small molecule effectors ppGpp (a negative regulator) and iNTP (a positive regulator), which serve as signals of the cell's nutritional status (reviewed in Paul et al. [2004b]). In particular, the effects of ppGpp and iNTP on rRNA transcription depend on the unusual instability of rRNA open complexes. Thus, although ppGpp destabilizes all promoter open complexes in vitro, the specific inhibitory effect of ppGpp on rRNA transcription in vivo can be explained by postulating that, in the cell, open complex lifetime is rate limiting for transcription initiation at the rRNA promoters. Similarly, the specific stimulatory effect of elevated concentrations of iNTPs on rRNA transcription has been attributed to the tendency of rRNA open complexes to collapse before transcription can initiate. Despite this explanation for the specific effects of changing concentrations of ppGpp and iNTPs on rRNA transcription in vivo, two puzzling discrepancies remained to be addressed (see Paul et al., 2004a [this issue of *Cell*]). First, the inhibitory effects of ppGpp on promoter activity in vivo (10- to 20-fold) were considerably larger than those measured in vitro under comparable conditions (2- to 3-fold). Second, the rRNA promoters apparently respond to a different range of iNTP concentrations in vivo than in vitro. Specifically, the iNTP concentrations measured in vivo are higher than the iNTP concentrations required for maximal promoter activity under typical in vitro conditions. The identification of DksA as a missing cofactor in the regulation

of rRNA transcription now provides a resolution of both discrepancies.

Previously associated with a variety of in vivo phenotypes, DksA was not before recognized as a transcription factor. In their article in this issue of *Cell*, Paul et al. (2004) show that *dksA* is strictly required for the regulation of rRNA transcription by ppGpp and iNTPs in vivo and that purified DksA binds to RNAP and exerts direct effects on rRNA transcription. In fact, like ppGpp, DksA on its own destabilizes promoter open complexes in general, and rRNA open complexes in particular. Moreover, DksA was found to enhance significantly the inhibitory effect of ppGpp on rRNA transcription in vitro and to increase the iNTP concentrations required to achieve half-maximal promoter activity. These observations suggest that, in vivo, DksA is required to further destabilize rRNA promoter open complexes so that the system is appropriately poised to respond to physiologically relevant concentrations of both ppGpp and iNTPs. In addition to amplifying the magnitude of ppGpp-dependent transcription inhibition in vitro, DksA was found to reduce the concentration of ppGpp required for half-maximal inhibition. The latter observation suggests that DksA may directly stabilize the binding of ppGpp to RNAP, a conjecture supported by the results of structural studies, also reported in this issue of *Cell* (Perederina et al., 2004).

The unexpected structural correspondence between DksA and the Gre factors (see Figure 1C) immediately suggested that DksA inserts its coiled-coil domain into the RNAP secondary channel, an inference supported by biochemical data placing the DksA coiled-coil tip near the active center of the enzyme (Perederina et al., 2004). Furthermore, the presence of two conserved acidic residues at the tip of the DksA coiled-coil domain suggested, by analogy with the Gre factors, that DksA may function, at least in part, to stabilize the binding of a functionally significant Mg^{2+} ion. With these considerations in mind, Perederina et al. (2004) sought to construct a model of the DksA-RNAP-ppGpp ternary complex based on the recently reported crystal structure of an RNAP-ppGpp complex (Artsimovitch et al., 2004). This crystal structure revealed two independent RNAP molecules, each of which contained one ppGpp molecule bound at the bottom of the secondary channel, close to the active center of the enzyme (Artsimovitch et al., 2004). Although the two ppGpp molecules were bound to the same site within the enzyme, they were oriented differently, with either the 5′ or the 3′ diphosphate making interactions with conserved residues at the active center. However, regardless of the orientation of the bound ppGpp molecule, each diphosphate apparently binds one Mg^{2+} ion. While the Mg^{2+} ion that is closest to the active center of the enzyme is well coordinated in the structure and seems relatively inaccessible, the distal Mg^{2+} ion is not seen to make any direct interactions with RNAP amino acid residues and appears accessible from the secondary channel. Accordingly, Perederina et al. (2004) constructed a ternary complex model in which the distal tip of the DksA coiled coil is positioned to permit interaction between the conserved acidic residues of DksA and the Mg^{2+} ion bound to the distal disphosphate of the bound ppGpp molecule. In this model, the globular domain of DksA binds the same structural motif at the rim of the secondary channel that serves as the primary binding site for the globular domain of the Gre factors, reinforcing the structural parallel between DksA and the Gre factors. Furthermore, consistent with the proposed role of DksA's conserved acidic residues, elimination of the negatively charged side chains abolished the ability of DksA to potentiate the effect of ppGpp on transcription from a test promoter in vitro (Perederina et al., 2004).

The DksA-RNAP-ppGpp structural model nicely accounts for the apparent stabilization of ppGpp binding in the presence of DksA. Nevertheless, neither the RNAP-ppGpp structure nor the DksA-RNAP-ppGpp structural model accounts fully for the effects of ppGpp and DksA on transcription. Thus, although the available structural information provides a critical framework for developing and testing specific hypotheses, many questions remain to be answered. First, further evidence is needed to support the biological relevance of the ppGpp binding site as defined by the RNAP-ppGpp crystal structures. Because most of the well-defined interactions between the ppGpp moiety and RNAP involve highly conserved amino acid residues in the vicinity of the catalytic center, the challenge is to design mutant enzymes that retain at least some catalytic activity but are specifically defective in binding ppGpp. One such mutant was designed and found to exhibit reduced sensitivity to ppGpp in vitro (Artsimovitch et al., 2004); it will be interesting to learn whether this mutant retains its insensitivity to ppGpp in the presence of bound DksA.

Second, it will be important to determine the biological relevance of the two orientations for the binding of ppGpp near the RNAP active center. In this regard, it is interesting to note that ppGpp can serve as both a negative and a positive regulator of transcription. Thus, under conditions of amino acid starvation, ppGpp stimulates transcription initiation from certain amino acid biosynthetic promoters, an effect that is also potentiated by DksA (Paul et al., 2004a). Accordingly, Artsimovitch et al. (2004) have proposed that the two observed orientations of RNAP bound ppGpp correspond to an inhibitory and a stimulatory orientation (each of which could apparently be stabilized by DksA). To test this proposal, identifying specific amino acid substitutions in RNAP that might be predicted to exert differential effects on the two proposed binding modes for ppGpp would be very important. Similarly, a screen for DksA mutants that potentiate negative control but not positive control (or vice versa) would be potentially informative, though uncoupling the negative and positive effects of ppGpp may not be possible.

Third, the detailed structural basis for ppGpp-mediated destabilization of promoter open complexes remains to be established. Noting that rRNA promoters in *E. coli* are characterized by a G+C-rich discriminator region just upstream of the transcription start site, Artsimovitch et al. (2004) have proposed that ppGpp may base pair with cytosine residues exposed on the nontemplate strand in the promoter open complex. In particular, they suggest that when bound in one of the two observed orientations (with the 3′ diphosphate closest to the active center), ppGpp can base pair with a cytosine at position −1 on the nontemplate strand and so perturb the open complex. Nevertheless, this model requires further testing, in part because ppGpp has been shown to affect open complex stability at some promoters lacking the requisite C residues (Barker et al., 2001).

Fourth, the structural basis for the ability of DksA to destabilize promoter open complexes in the absence of ppGpp also remains to be understood. In this regard, it

will be interesting to learn whether the side chains of the conserved acidic residues at the coiled-coil tip play any role in the effect of DksA on open complex lifetime (independent of their role in stabilizing the binding of ppGpp to RNAP). For example, the conserved acidic residues may participate in specific interactions near the catalytic center of RNAP even in the absence of bound ppGpp. Any such interactions might affect open complex stability and/or alter the catalytic properties of the enzyme.

Gene-Specific Regulation by Factors that Bind RNAP

Unlike classical regulators of transcription initiation that bind to specific DNA sites associated with their target promoters, the regulators described here employ more subtle mechanisms to achieve specificity. DksA, which together with its cofactor ppGpp destabilizes promoter open complexes, provides an illustrative example: the unusual kinetic properties of its target promoters render them sensitive to its effects in vivo. The case of the Gre factors is less clear. Whether or not they are recruited specifically to backtracked elongation complexes, which may tend to form on particular transcription units under particular conditions, remains to be determined.

Inhibition via the Secondary Channel: Microcin J25

As well as providing a binding site for regulators that affect RNAP function, the secondary channel also presents a target for RNAP inhibitors. Two recent studies (Mukhopadhyay et al., 2004; Adelman et al., 2004) provide strong evidence that the peptide antibiotic Microcin J25 (MccJ25) binds within and essentially plugs the secondary channel. Produced from strains of *E. coli* containing a plasmid that carries genes encoding a synthesis, maturation, and export system, MccJ25 is a 21 amino acid bactericidal peptide that adopts an unusual configuration referred to as a "lassoed tail" or "lariat-protoknot" (see Mukhopadhyay et al. [2004] for references). Previous data indicated that MccJ25 toxicity results from its ability to inhibit transcription (Delgado et al., 2001; Yuzenkova et al., 2002), but the mechanistic basis for this inhibition was unknown.

Multiple lines of evidence support the conclusion that MccJ25 binds within the RNAP secondary channel and blocks nucleotide entry into the active site. First, numerous amino acid substitutions in RNAP that confer resistance to MccJ25 have been isolated, and all of these substitutions localize to the secondary channel (Delgado et al., 2001; Yuzenkova et al., 2002; Mukhopadhyay et al., 2004). Second, increasing concentrations of NTPs counteract the inhibitory effect of MccJ25 (Mukhopadhyay et al., 2004). Third, MccJ25 inhibits backtracking, presumably by binding within the secondary channel and blocking the extrusion of the 3′ end of the nascent transcript into the channel (Adelman et al., 2004); in addition, MccJ25 can be crosslinked to the 3′ end of nascent RNA (Adelman et al., 2004). Fourth, MccJ25 interferes with GreA/B-dependent transcript cleavage (Adelman et al., 2004). Fifth, the binding of MccJ25 to RNAP was demonstrated directly (Mukhopadhyay et al., 2004). Finally, biophysical analysis and structural modeling places MccJ25 in the secondary channel (Mukhopadhyay et al., 2004; Adelman et al., 2004).

These studies raise the possibility that new compounds designed to obstruct the RNAP secondary channel may be effective antimicrobials. In addition, the analysis of MccJ25 function provides the only experimental support to date for the postulate that nucleotides pass through the secondary channel to reach the active site.

Conclusion

The studies discussed here establish the RNAP secondary channel as an important target that can be exploited by regulators that modulate the activities of RNAP and inhibitors that prevent transcription elongation. Among the regulators, the prokaryotic Gre factors and the eukaryotic TFIIS perform analogous functions despite the fact that they resemble one another in neither amino acid sequence nor structure. Another prokaryotic regulator, DksA, bears striking structural resemblance to the Gre factors, but no discernable similarity at the level of amino acid sequence. These correspondences in function and/or structure suggest that additional proteins will be uncovered that affect RNAP function by binding within the secondary channel. Dissecting the strategies they employ to modify the transcription apparatus will likely reveal new variations on the theme of regulation through the secondary channel.

Selected Reading

Adelman, K., Yuzenkova, J., La Porta, A., Zenkin, N., Wang, M.D., and Severinov, K. (2004). Mol. Cell *14*, 753–762.

Artsimovitch, I., Patlan, V., Sekine, S., Vassylyeva, M.N., Hosaka, K., Yokoyama, S., and Vassylyev, D.G. (2004). Cell *117*, 299–310.

Barker, M.M., Gaal, T., Josaitis, C.A., and Gourse, R.L. (2001). J. Mol. Biol. *305*, 673–688.

Cramer, P., Bushnell, D.A., Fu, J., Gnatt, A.L., Maier-Davis, B., Thompson, N.E., Burgess, R.R., Edwards, A.M., David, P.R., and Kornberg, R.S. (2000). Science *288*, 640–649.

Delgado, M.A., Rintoul, M.R., Farias, R.N., and Salomon, R.A. (2001). J. Bacteriol. *183*, 4543–4550.

Fish, R.N., and Kane, C.M. (2002). Biochim. Biophys. Acta *1577*, 287–307.

Kettenberger, H., Armache, K.-J., and Cramer, P. (2003). Cell *114*, 347–357.

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

Kulish, D., Lee, J., Lomakin, I., Nowicka, B., Das, A., Darst, S.A., Normet, K., and Borukhov, S. (2000). J. Biol. Chem. *275*, 12789–12798.

Laptenko, O., Lee, J., Lomakin, I., and Borukhov, S. (2003). EMBO J. *22*, 6322–6334.

Mukhopadhyay, J., Sineva, E., Knight, J., Levy, R.M., and Ebright, R.H. (2004). Mol. Cell *14*, 739–751.

Opalka, N., Chlenov, M., Chacon, P., Rice, W.J., Wriggers, W., and Darst, S.A. (2003). Cell *114*, 335–345.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W., and Gourse, R.L. (2004a). Cell *118*, this issue, 311–322.

Paul, B.J., Ross, W., Gaal, T., and Gourse, R.L. (2004b). Annu. Rev. Genet., in press.

Perederina, A., Svetlov, V., Vassylyeva, M., Yokoyama, S., Artsimovitch, I., and Vassylyev, D.G. (2004). Cell *118*, this issue, 297–309.

Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). EMBO J. *22*, 2234–2244.

Sosunova, E., Sosunov, V., Kozlov, M., Nikiforov, V., Goldfarb, A., and Mustaev, A. (2003). Proc. Natl. Acad. Sci. USA *100*, 15469–15474.

Steitz, T.A. (1998). Nature *391*, 231–232.

Yuzenkova, J., Delgado, M., Nechaev, S., Savalia, D., Epshtein, V., Artsimovitch, I., Mooney, R.A., Landick, R., Farias, R.N., Salomon, R., and Severinov, K. (2002). J. Biol. Chem. *277*, 50867–50875.

Zhang, G., Campbell, E.A., Minakhin, L., Richter, C., Severinov, K., and Darst, S.A. (1999). Cell *98*, 811–824.