ELSEVIER

Elongation by RNA polymerase: a race through roadblocks

Dmitry G Vassylyev

Transcription is the first and most regulated step of gene expression. RNA polymerase (RNAP) is the heart of the transcription machinery and a major target for numerous regulatory pathways in living cells. The crystal structures of transcription complexes formed by bacterial RNAP in various configurations have provided a number of breakthroughs in understanding basic, universal mechanisms of transcription and have revealed regulatory 'hot spots' in RNAP that serve as targets and anchors for auxiliary transcription factors. In combination with biochemical analyses, these structures allow feasible modeling of the regulatory complexes for which experimental structural data are still missing. The available structural information suggests a number of general mechanistic predictions that provide a reference point and direction for future studies of transcription regulation.

Address
Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Schools of Medicine and Dentistry, 402B KAUL Genetics Building, 720 20th Street South, Birmingham, AL 35294, United States

Corresponding author: Vassylyev, Dmitry G (dmitry@uab.edu)

Introduction

In all organisms, transcription is carried out by DNA-dependent RNA polymerases (RNAP) and can be divided into three mechanistically and structurally distinct stages: initiation, elongation and termination. In bacteria and eukaryotes, each phase of transcription is a target for numerous regulatory factors. The interplay among the unstable initiation phase, transition to a stable elongation complex (EC) followed by processive synthesis, transient halting at numerous 'roadblocks', and RNA release depends on the intricate network of interactions between RNAP, nucleic acid (NA) signals, and/or auxiliary transcription factors. The past few years have been characterized by an 'explosion' of structural studies that resulted in detailed structural characterization of several key transcription intermediates formed by both bacterial and eukaryotic multi-subunit RNAPs [1,2,3••,4••,5,6••,7,8,9••,10,11,12]. These structures shed significant light on such general mechanisms as downstream (dw) DNA and RNA/DNA hybrid strand separation [3••,6••,7], substrate selection and loading [4••,6••,9••], DNA translocation [4••,13•,14•], formation and rescue of the paused and backtracked/arrested complexes [6••,11,12], inhibition by small molecules [4••,13•,14•,15,16,17,18,19•,20,21], and so on.

This review is focused on the mechanistic insights gained from recent crystallographic analysis of the bacterial system that revealed regulatory 'hot spots' in RNAP and implied common mechanisms utilized by structurally and/or functionally divergent transcription factors.

Structural organization of the bacterial EC

The high resolution structure of the *T. thermophilus* (tt) EC provided the first detailed view of the bacterial EC and identified the determinants of its stability, processivity, and response to nucleic acid signals, effectors, and auxiliary proteins [3••]. In the ttEC, the downstream (dw) DNA binding cavity accommodates 13 bp of the dwDNA duplex. The 9 bp RNA/DNA hybrid resides in the RNAP main channel, and the nascent RNA transcript, which is displaced from the template (T) DNA, is threaded through the RNA exit channel (Figure 1a, b). Access to the RNAP active site through the main channel is blocked suggesting that the widely open secondary channel (SC) is the major substrate entry pore. In the crystals, the complex is in the post-translocated state with the acceptor DNA template (register +1) available for base pairing with the incoming NTP.

The dwDNA duplex is melted immediately upstream of the active site (register +2) implying that only one substrate at a time may bind to the EC [3••,4••,22]. The +2 dwDNA base pair stacks on the fork loop-2 (fork-2), which most probably plays a crucial role in dwDNA strand separation and proper positioning of the open acceptor T-base in the active site [23].

At the upstream edge of the transcription bubble, the last (9th) bp of the RNA/DNA hybrid stacks on the β′-subunit 'lid' loop that sterically blocks the nascent RNA/DNA duplex, reminiscent of the downstream fork-2 [24,25]. The first displaced RNA base is trapped in the hydrophobic pocket formed by the β-subunit switch-3 segment implying the DNA-dependent mechanism of RNA displacement. The switch-3 pocket may possess certain sequence specificity and emerge as a crucial checkpoint mediating initiation, translocation,

Figure 1
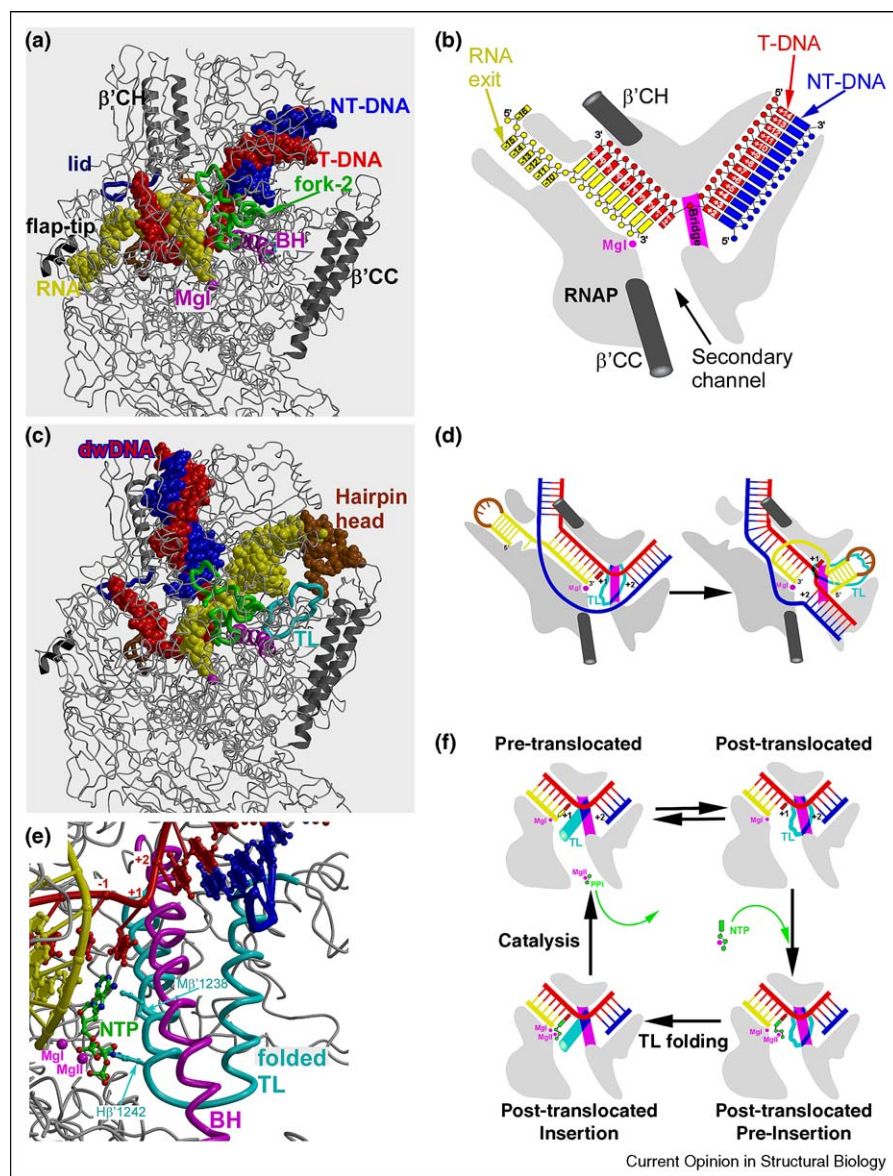
Structure of the bacterial elongation complex. **(a, b)** The overall three-dimensional X-ray structure (a) and schematic drawing (b). The same color scheme is used for all Figures in this article. **(c, d)** The 'DNA displacement model' of transcription termination; the three-dimensional model (c) and schematic drawing (d). **(e)** Folding of the TL closes the RNAP active site in the substrate bound, 'insertion' ttEC. **(f)** The model of the nucleotide addition cycle.

pausing, and/or termination [24–26]. It is also possible that a similar, yet unknown, pocket accommodates the first displaced non-template (NT) DNA base (+1) additionally stabilizing the bubble. The mechanisms of the dwDNA and RNA/DNA hybrid strand separation are most probably universal for all multi-subunit RNAPs [6••,7].

Insights into termination

Modeling based on the ttEC structure shows that the intact RNA exit channel may accommodate and stabilize the short (5 bp) RNA hairpins that are characteristics of paused transcription complexes. On the contrary, there does not appear to be enough space for the bulky (8–12 bp) termination hairpins suggesting substantial widening of the RNA exit channel. The recently proposed allosteric model of termination suggests that these alterations open the passage to the main channel and allow hairpins to invade the main channel, partially (4–5 bp) melt the upstream RNA/DNA hybrid and travel by ∼70 Å towards the RNAP active site where it interacts with the catalytic trigger loop (TL) [26]. Altogether, migration of the RNA hairpin compromises the hybrid and induces conformational changes in RNAP that trigger

NA dissociation; DNA is static and plays no active role. This RNA-dependent mechanism is thus at odds with the earlier 'translocation' model in which multiple steps of unproductive (no RNA synthesis) forward DNA translocation result in transcript release and complex dissociation [27].

A combined, 'DNA displacement' model may reconcile these conflicting mechanisms (Figure 1c, d). We suggest that positioning of the hairpin both near the RNAP active site and the dwDNA-binding cavity as specified in the allosteric model would result in competition with and displacement of the dwDNA duplex. Displacement presumably results from high affinity interactions of the hairpin head with the TL; these interactions are not evident if the dwDNA is not displaced because it is located between the hairpin and TL in the crystal structure [26]. Displacement of dwDNA most probably dislodges the acceptor T-base from the active center to produce a catalytically inactive complex in which the remaining downstream hybrid (4–5 bp) can be unzipped as specified by the translocation model [27]. Notably, the translocation model alone does not explain of how forward translocation can proceed without RNA synthesis.

Figure 2
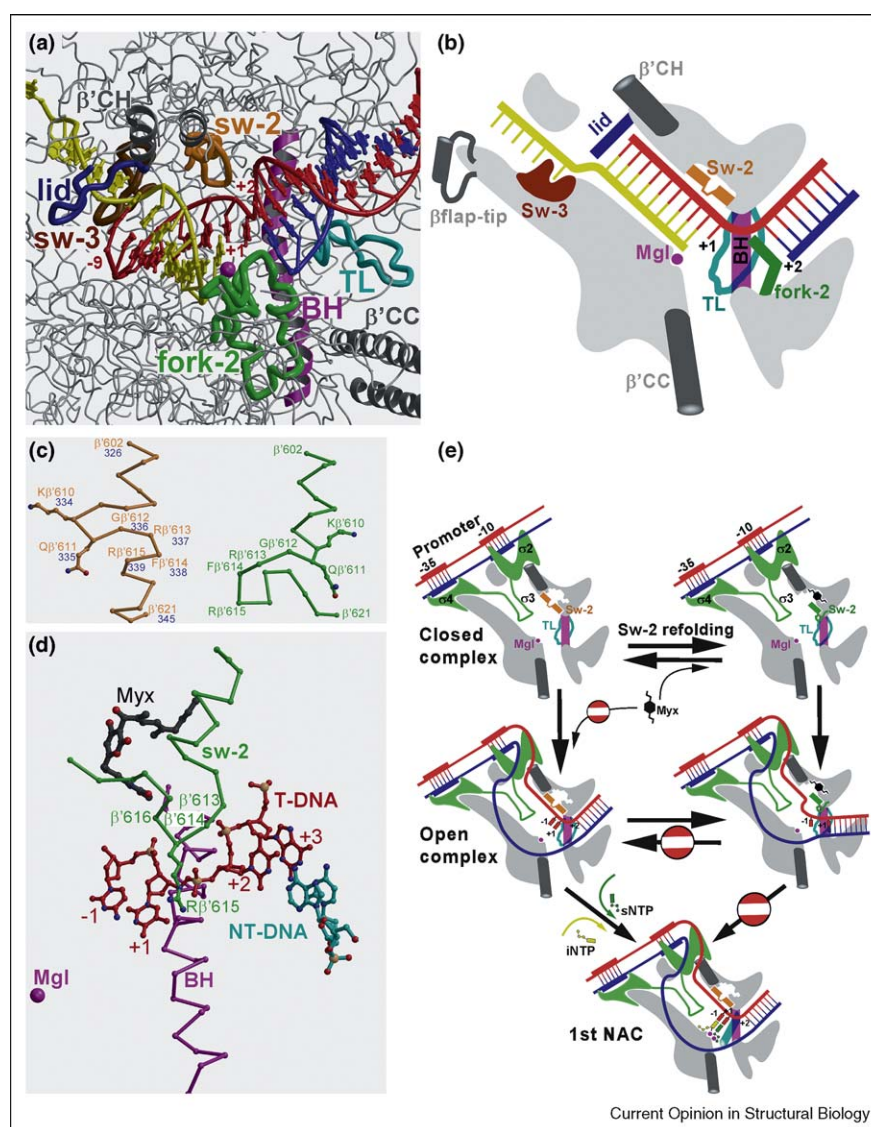
Regulatory targets in RNAP. **(a, b)** Conformational 'switches' in RNAP; the crystal structure of the ttEC (**a**) and schematic drawing (**b**). Sw-2 and sw-3 are the switches 2 and 3, respectively. **(c)** Refolding of the β'-subunit switch-2 segment stabilized by myxopyronin (Myx). The original and refolded configurations are in orange and green, respectively. The *T. thermophilus* and *E. coli* sequence numbers are shown in orange/green and blue, respectively. **(d)** The refolded sw-2 appears to clash with the DNA T-strand near the active site (register +1) in the transcription bubble. **(e)** The effect of the sw-2 refolding on the open complex formation and mechanism of the myxopyronin action.

Nucleotide addition cycle

Our two structures of substrate-bound ttECs revealed that NTP-induced refolding of the TL mediates formation of the closed, catalytically active intermediate [4$^{\bullet\bullet}$] (Figure 1e). This finding allowed us to propose the nucleotide addition cycle (NAC) model that is most probably relevant for all multi-subunit RNAPs [6$^{\bullet\bullet}$,9$^{\bullet\bullet}$,13$^{\bullet}$] (Figure 1f). Before NTP loading, the EC exists in equilibrium between pre-translocated and post-translocated states. Substrate loading occurs in two steps. First, the NTP binds to the open (unfolded TL) post-translocated EC in a template-dependent manner forming an inactive, pre-insertion intermediate. Secondly, NTP-induced displacement of the bridge helix (BH) and fork-2 facilitates folding of the TL in the α-helical hairpin. Upon TL folding, the complex isomerizes to the catalytically competent, closed 'insertion' state. The NAC culminates with the catalytic reaction that results in transcript extension and pyrophosphate release. The antibiotic streptolydigin (Stl) binds in a pocket formed by the BH and fork-2 and blocks their NTP-dependent displacement, thereby preventing TL folding and freezing the substrate in the inactive pre-insertion state [4$^{\bullet\bullet}$,17].

Regulatory targets in RNAP

Structural studies have revealed that alterations (displacement/refolding) of several crucial RNAP domains modulate activity and/or stability of the transcription complexes. The domains involved in these conformational switches are likely targets of auxiliary transcription factors that enhance or inhibit these switches (Figure 2).

The TL, which undergoes dramatic substrate-induced refolding emerges as a central regulatory element, as well as a key determinant for the fidelity and processivity of transcription in multi-subunit RNAPs [3$^{\bullet\bullet}$,4$^{\bullet\bullet}$,9$^{\bullet\bullet}$,13$^{\bullet}$, 14$^{\bullet}$,28] (Figures 1e, f and 2a, b). Given the high sensitivity of its structure to even subtle alterations of adjacent structural domains and a possibility that it may adopt multiple conformations, the modulation of TL refolding by various transcription factors provides numerous degrees of freedom in transcription regulation. In addition to Stl, the TL is a plausible target for tagetitoxin (Tgt), DksA/ppGpp system, Gre-factors and Gfh1.

The fork-2 loop appears to play multiple functional roles mediating dwDNA melting and substrate loading, and maintaining bubble stability through interactions with the RNA/DNA hybrid (Figure 2a and b). This loop, which bridges the upstream and downstream RNAP domains and forms a part of the catalytic center, may also serve as a circuit for transmitting allosteric signals generated by remotely bound transcription factors (RfaH, NusG, NusA, etc.) to the RNAP active site. Interestingly, the fork-2 forms part of binding pocket for rifamycins (Rifs) suggesting that Rifs most probably block the func-

tionally significant NTP-induced displacement of this loop in an Stl-like manner, thereby allosterically affecting active site configuration and/or substrate loading in agreement with the proposed allosteric model [16].

Recently, specific conformations of the BH and TL induced by a toxin, α-amanitin, were shown to stabilize a pre-templated intermediate in the eukaryotic enzyme in which the overall complex is in the post-translocated state, while the open acceptor T-base occupies an inactive site over the BH between pre-translocated and post-translocated registers [13$^{\bullet}$]. The two conformations of the BH that have been observed in bacterial RNAP (straight and 'flipped') were also proposed to mediate DNA translocation [1,16]. In principle, the flipped BH may indeed stabilize the pre-templated state. Both, Stl and α-amanintin affect the BH and TL conformations and thus in the absence of the substrate Stl may theoretically trap the pre-templated bacterial EC [4$^{\bullet\bullet}$,13$^{\bullet}$,17]. However, an intact post-translocated complex is evident in the substrate-free ttEC/Stl structure (DGV, unpublished data). While the exact role of the bacterial BH in translocation remains to be elucidated, this helix stacks on the 'active' acceptor T-base, mediates the TL folding and is accessible from both, the secondary and main channels suggesting that its configuration may be modulated by auxiliary factors to regulate the NAC.

The lid loop is located at the junction between the main channel and the RNA exit channels, and the upstream DNA pore. In addition to its commonly accepted role in RNA/DNA hybrid strand separation, the lid loop may also function as a 'valve' that opens a passageway between these functional 'chambers' during the major structural transitions of the transcription cycle. Indeed, according to the allosteric model of termination, displacement of the lid seems indispensable for intrusion of the termination hairpin in the main channel [26]. On the contrary, in the holoenzyme, the lid locks the extended inter-domain linker of the σ-subunit (region 3.1) inside the core enzyme structure suggesting that opening of the lid accompanies/triggers the σ-factor release during transition from the initiation to elongation phase [24]. Interestingly, the structure of the lid resembles that of the specificity loop in the single-subunit T7 RNAP that undergoes drastic structural rearrangement (repositioning and refolding) upon transition from initiation and elongation complexes and plays essential, but distinct roles in both complexes [29–31]. Therefore, it is possible that the lid also possesses 'chameleon' properties and that these reconfigurations are regulated by external transcription factors.

Most recently, we demonstrated that the antibiotic myxopyronin stabilizes refolding of the β′-subunit switch-2 segment and, therefore, may sterically block downstream propagation of the nascent transcription bubble during

open complex formation [19•] (Figure 2c, d, e). Mutations in switch-2 mimic the antibiotic effects on promoter complexes suggesting that this region may serve as a natural molecular checkpoint for DNA loading in response to regulatory factors/signals. Consistently with this hypothesis, transcription factor DksA appears to potentate this structural switch [32]. While switch-2 refolding seems unlikely to affect a stable EC, it may play a regulatory role in unstable termination complexes and be a target for some termination/anti-termination factors.

A second, low affinity Mg^{2+} ion bound to the RNAP active site is known to be a 'catalytic' metal required for all reactions catalyzed by RNAP. The 'catalytic' transcription factors (Gre-factors, non-template NTPs, pyrophosphate) stimulate intrinsic catalytic activities of RNAP presumably through direct coordination of this catalytic ion [33]. On the contrary, structural and biochemical data indicate that the antibiotic tagetitoxin and ppGpp appear to convert the second catalytic ion into an inhibitory one in essentially the same binding site [15,18]. These results suggest that other 'inhibitory' transcription factors may modulate transcription by targeting this 'regulatory' metal. The competitive interplay of these catalytic and inhibitory factors may provide exquisite regulatory control of the transcription process.

Regulatory anchors in RNAP

In the absence of direct structural data on RNAP complexes with protein transcription factors, localization of the major binding sites on RNAP for these factors and elucidation of the mechanisms of their recruitment to the cognate transcription complexes are of central importance for understanding general and specific mechanisms of transcription regulation. To this end, two solid 'anchors' for transcription factors were identified. The protruding coiled-coils (CCs) in the β′-subunit of RNAP, the 'upstream' clamp helices (CH) and the C-terminal β′CC at the rim of the SC appear to serve as the major binding sites for distinct ('upstream' and SC) subsets of competing transcription factors. Notably, while these factors possess different regulatory mechanisms, most of them are recruited to the two RNAP CCs in a very similar fashion—the hydrophobic tips of the CCs are inserted in the open hydrophobic cavities of the proteins.
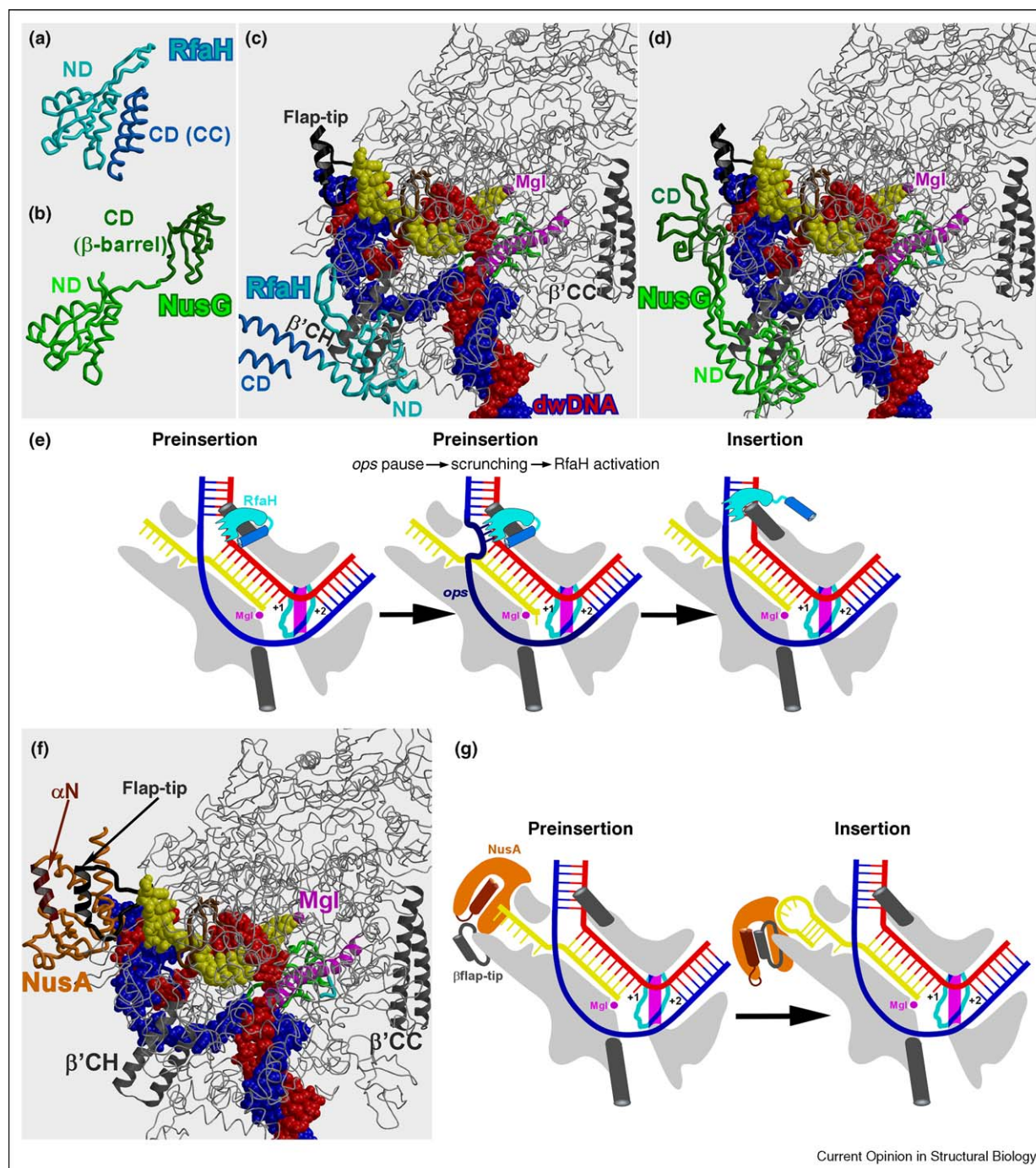
The 'upstream' transcription factors

So far, the three major 'upstream' transcription factors were shown to bind to the CH in a competitive manner: the initiation factor σ, and the two elongation factors, RfaH and NusG [1,33,34•,35]. RfaH and NusG are paralogs that regulate transcriptional pausing and termination and possess sequence and structural similarities [34•,36]. The key difference between NusG and RfaH is that NusG acts as a sequence-independent general elongation

factor, while RfaH is operon specific and its action depends on the ops site in DNA; this sequence alone induces transcription pausing. The structures of RfaH and NusG revealed two (N-terminal and C-terminal) domains. The N-domains displayed high similarity, while the C-domains, retaining sequence homology, appeared strikingly different; the β-barrel in NusG, and an α-helical CC in RfaH [34•] (Figure 3a and b). Both N-domains possess a vast hydrophobic cavity that is closed by the C-domain in RfaH but is exposed in NusG. This cavity most probably constitutes the RNAP-binding site in both proteins. This cavity in RfaH becomes unmasked only upon sequence-specific binding to the ops NT–DNA that triggers domain dissociation. Identification of CH as a common anchor for RfaH and NusG allowed structural modeling of the RfaH-bound and NusG-bound ECs (Figure 3c and d). An interesting implication from the RfaH/EC model is that RfaH may interact with only a few (2–3) NT–DNA bases, while the ops element consists of 9 essential nts. One possible scenario is that the ops-induced pause, which presumably requires the entire ops sequence, is accompanied by DNA scrunching and that scrunching exposes the 2–3 NT bases recognized by RfaH (Figure 3e). Another important prediction is that, in addition to binding to the major, functionally active 'insertion' binding site on RNAP, RfaH also binds to the 'pre-insertion' site in an 'inactive' configuration and is converted to an 'active' configuration when it meets and recognizes the specific DNA target.

The global transcription factor NusA is recruited to transcription complexes upon formation of RNA hairpins. Biochemical data suggest that the N-terminal domain of NusA is essential for binding to RNAP and that binding occurs near the RNA exit channel, presumably in the β-subunit flap domain [33,37]. Following a general prediction concerning the hydrophobic mode of binding of regulatory factors to RNAP, we have identified in the NusA structure the prominent hydrophobic cavity in the N-terminal domain [38] that is blocked in the structure by the N-terminal amphiphilic α-helix. Thus, RfaH-like activation may be required to displace this N-terminal helix and to open the cavity for the RNAP target (Figure 3f and g). We speculate that upon activation this groove will accommodate some RNAP α-helix resembling the 'self-inhibitory' NusA α-helix. The protruding amphiphilic flap–tip helix from the β-subunit flap domain is a promising candidate, this helix is located at the rim of the RNA exit channel where the RNA hairpins fold [37]. Similar to RfaH, 'pre-insertion' recognition of the hairpin may trigger NusA activation. In the holoenzyme, the flap–tip helix is trapped in the hydrophobic groove formed by region 4 of the σ-subunit [1]; consistently the σ-factor and NusA compete for binding to RNAP [33]. The flap–tip helix is thus another potential regulatory anchor for the 'upstream' group of the transcription regulators.

Figure 3

The 'upstream' transcription factors. **(a, b)** The crystals structures of RfaH (**a**) and NusG (**b**) reveal the similar N-terminal domains (NDs) and strikingly distinct folds of the C-terminal domains (CDs). **(c, d)** Structural modeling suggests that RfaH (**c**) and NusG (**d**) bind to the same site (β'CH) on the EC. **(e)** DNA-dependent activation of RfaH may occur through DNA scrunching induced by pausing on the *ops* site. **(f, g)** Structural model (f) and schematic diagram (g) of NusA recruitment to the EC through the displacement of the NusA N-terminal α-helix (αN). Only the N-terminal domain of NusA was modeled.

The secondary channel transcription factors

All structurally characterized SC protein transcription factors with known function (Gre-factors, DksA, Gfh1) [33,39,40] possess a two-domain (globular and CC) archi- tecture (Figure 4a) and are thought to directly modulate the RNAP catalytic site through their 'functional' CC domains. The CC domains presumably penetrate the SC, while the primary role of the 'structural' globular domains

Figure 4
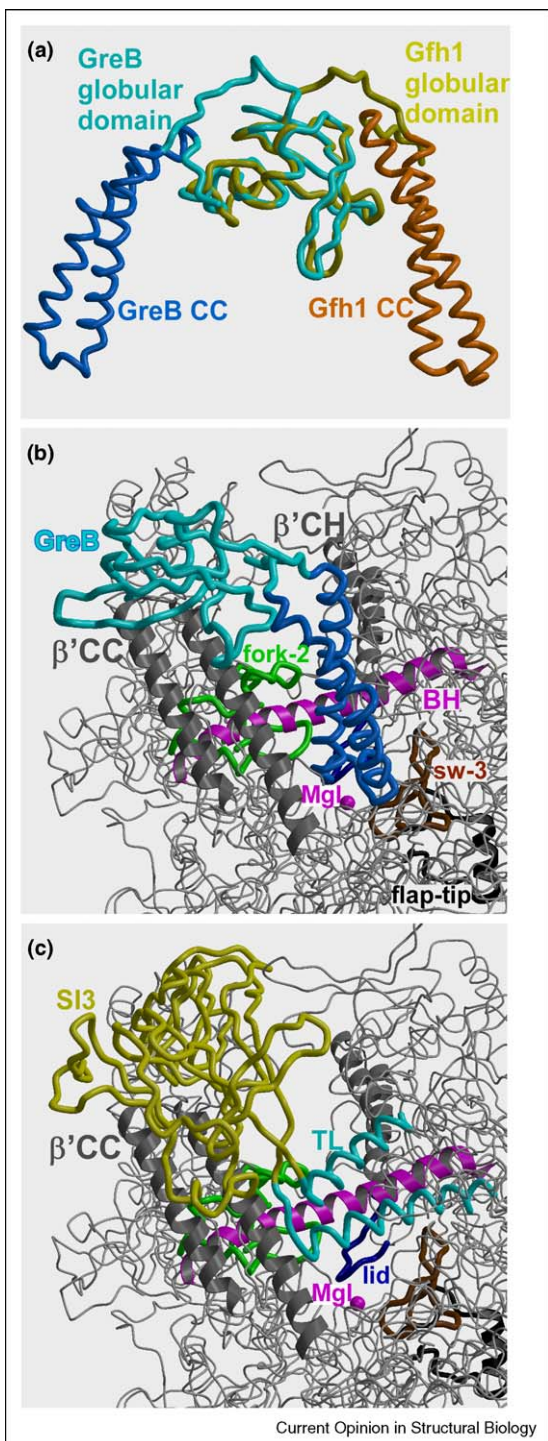
The secondary channel transcription factors. (a) The structures of GreB and Gfh1 are superimposed by the globular domains. (b, c) Structural models of the RNAP/GreB (b) and RNAP/SI3 (c) complexes. Hydrophobic tip of the RNAP β′CC is inserted in the open hydrophobic cavities of GreB and the SI3 domain.

is thought to form a stable and specific complex with RNAP.

Bacterial Gre-factors (GreA/B) stimulate the intrinsic endonucleolytic activity of RNAP and assist RNA polymerase (RNAP) in rescuing backtracked and/or arrested ECs [41]. A commonly accepted mechanism suggests that the two invariant acidic side chains at the tip of the CC domain coordinate the second catalytic Mg^{2+} ion [33]. However, several conflicting and controversial structural models have been proposed for the Gre/RNAP complex [42–44]. The X-ray structure of the GreB protein and mutational analysis identified a vast open hydrophobic cavity in the GreB globular domain and a complementary hydrophobic patch at the tip of the β′CC as the major binding partners; these results allowed plausible modeling of the RNAP/GreB complex [45•] (Figure 4b). This model is now being confirmed by the crystal structure of GreB complexed with isolated β′CC (DGV, in preparation).

The *T. thermophilus* Gfh1 protein belongs to the Gre-family of transcription factors, yet possesses inhibitory, rather than catalytic activity. The Gfh1 structure reveals two Gre-like domains with strikingly distinct interdomain orientation [46–48] (Figure 4a, c). Similar to the Gre-factors, Gfh1 most probably binds to the β′CC and accesses the catalytic center of RNAP. Its recruitment to RNAP, however, appears to require activation through domain rearrangement to adopt an active, Gre-like conformation. By analogy with RfaH, this implies the pre-insertion binding mode and a specific target(s) that triggers activation.

DksA binds to RNAP and greatly stimulates the activity of 'magic spot', ppGpp during stringent control. DksA possesses the two (globular and CC) domain architecture reminiscent of the Gre-factors. At the same time, ppGpp was shown to bind near the active site of ttRNAP [15]. Together, the structural and biochemical data [39,40] suggest a synergetic model in which the accessory DksA protein folds around the β′CC (but does not bind to the β′CC tip) and stabilizes ppGpp binding to RNAP [39] through its CC-domain to enhance the effect of magic spot. However, recently, two DksA variants with the single substitutions in the globular and CC-domain were reported to possess high ppGpp-independent activity [49]. There are the two major alternative interpretations of these results: (i) DksA is a major player in the DksA/ppGpp synergetic tandem, that is, ppGpp is not a 'magic spot' but rather is an accessory molecule (DksA 'activator'); (ii) both regulators possess specific, but somewhat distinct effects on transcription suggesting that their mutual activation is triggered by direct interactions. Both mechanisms assume activation of the system that most probably occurs through ppGpp-induced/stabilized conformational alterations of DksA and/or DksA-induced/

stabilized reconfiguration of the ppGpp binding site. One possible scenario is that DksA binds to the β′CC through the hydrophobic cavity of its globular domain in a Gre-like fashion. However, in DksA this cavity is masked by a CC domain similar to that of the inactive, apo-RfaH protein [34●]. If this prediction is correct, a ppGpp-dependent domain opening would mediate DksA recruitment to RNAP that may occur via the two step, pre-insertion-to-insertion mode.

The two other SC transcription factors have recently been characterized. The Rnk protein conserves the Gre-like globular domain, but has a short, rudimentary CC suggesting thereby a similar binding mechanism, but distinct functional targets in RNAP [50]. While the TraR protein shares significant sequence and functional similarities with DksA, there are two striking differences: (i) TraR effects on transcription are ppGpp-independent and (ii) the TraR sequence is truncated from the N-terminus resulting in deletion of a portion of the globular domain and the first α-helix of the CC without which the protein can hardly adopt a stable DksA-like conformation [51]. Thus, TraR most probably functions as a dimer R [52].

Finally, in *E. coli* a bulky (188 residues, SI3) β′-subunit domain, which is dispensable, is inserted in the catalytic TL. Although the structure of the isolated SI3 domain was recently determined [53], the location of this domain in the RNAP structure and its functional role remain elusive. One intriguing possibility is that this domain may reversibly bind to the β′CC in a Gre-like fashion (Figure 4c) and the on-bound and off-bound states may confer the regulatory effects through competition/interactions with the SC factors and/or via modulation of TL conformation.

Conclusions

Analysis of the available structural information allows us to make several provocative mechanistic predictions concerning general principles of transcription regulation. These predictions may provide a foundation for future studies of the bacterial transcription machinery. First, there are a limited number of regulatory anchors in RNAP; multiple functionally distinct transcription factors compete for binding to a single anchor. One important role of such competition for the 'upstream' anchors may be to demarcate the different phases of transcription. Second, the known anchors contain protruding structural elements with hydrophobic patches complementary to the hydrophobic cavities of the cognate transcription factors. *Third*, structural (inactive-to-active) isomerization of the transcription factors, presumably triggered by recognition of the target complexes, in many cases appears to mediate recruitment to RNAP. Therefore, recruitment may occur in two, pre-insertion-to-insertion steps. Finally, a number of flexible structural 'switches' in

RNAP may serve as the regulatory targets for various auxiliary transcription factors. Regulatory events are thus often accompanied by unpredictable structural alterations/refolding of RNAP and/or transcription factors suggesting that structural studies are the major, and in some cases the only, tool to elucidate basic mechanisms of transcription and regulation of gene expression.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
●● of outstanding interest

1. Vassylyev DG, Sekine S, Laptenko O, Lee J, Vassylyeva MN, Borukhov S, Yokoyama S: **Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution**. *Nature* 2002, **417**:712-719.

2. Bushnell DA, Westover KD, Davis RE, Kornberg RD: **Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms**. *Science* 2004, **303**:983-988.

3. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH,
●● Artsimovitch I: **Structural basis for transcription elongation by bacterial RNA polymerase**. *Nature* 2007, **448**:157-162.
This paper reports the 2.5 Å resolution structure of the *T. thermophilus* elongation complex that provided a number of crucial mechanistic insights in the basic principles of transcription and revealed the key determinants in RNAP essential for fidelity, stability and processivity of transcription elongation.

4. Vassylyev DG, Vassylyeva MN, Zhang J, Palangat M,
●● Artsimovitch I, Landick R: **Structural basis for substrate loading in bacterial RNA polymerase**. *Nature* 2007, **448**:163-168.
The 3.0A resolution structures of the two bacterial ECs with the substrate analog revealed the inactive (pre-insertion) and active (insertion) substrate intermediates and provided an evidence for the two-step mechanism of substrate loading in the bacterial enzyme. The structures showed that isomerization from the inactive, open (pre-insertion) complex to the catalytically active closed (insertion) state occurs via substrate-induced folding of the TL. The antibiotic streptolydigin prevents proper folding of the TL thereby freezing the substrate complex in the inactive, pre-insertion configuration.

5. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD: **Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution**. *Science* 2001, **292**:1876-1882.

6. Kettenberger H, Armache KJ, Cramer P: **Complete RNA
●● polymerase II elongation complex structure and its interactions with NTP and TFIIS**. *Mol Cell* 2004, **16**:955-965.
The authors present a medium resolution structure of the EC formed by yeast RNA polymerase II. The structure of the EC with bound NTP analog provided the first experimental evidence for the pre-insertion substrate intermediate in the eukaryotic system.

7. Westover KD, Bushnell DA, Kornberg RD: **Structural basis of transcription: separation of RNA from DNA by RNA polymerase II**. *Science* 2004, **303**:1014-1016.

8. Westover KD, Bushnell DA, Kornberg RD: **Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center**. *Cell* 2004, **119**:481-489.

9. Wang D, Bushnell DA, Westover KD, Kaplan CD, Kornberg RD:
●● **Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis**. *Cell* 2006, **127**:941-954.
In this work, the authors reported a number of the eukaryotic ECs with different (cognate and non-cognate NTPs) and demonstrated that structural reconfiguration of the TL mediates closure of the active site in the presence of the cognate NTP in the yeast enzyme in a similar fashion of that of the bacterial system.

10. Brueckner F, Hennecke U, Carell T, Cramer P: **CPD damage recognition by transcribing RNA polymerase II**. *Science* 2007, **315**:859-862.

11. Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD: **Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution**. *Science* 2009, **324**:1203-1206.

12. Sydow JF, Brueckner F, Cheung ACM, Damsma GE, Dengl S, Lehmann E, Vassylyev DG, Cramer P: **Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA**. *Mol. Cell.* 2009, **34**:710-721.

13. Brueckner F, Cramer P: **Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation**. *Nat Struct Mol Biol* 2008, **15**:811-818.
The structure of the yeast EC with bound toxin, α-amanitin revealed the important, pre-templated translocation intermediate and suggested the essential roles of the bridge helix and TL in translocation.

14. Kaplan CD, Larsson KM, Kornberg RD: **The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin**. *Mol Cell* 2008, **30**:547-556.
The structure of eukaryotic RNAP complexed with α-amanitin and subsequent biochemical data demonstrated that the TL is an important determinant of transcription fidelity.

15. Artsimovitch I, Patlan V, Sekine S, Vassylyeva MN, Hosaka T, Ochi K, Yokoyama S, Vassylyev DG: **Structural basis for transcription regulation by alarmone ppGpp**. *Cell* 2004, **117**:299-310.

16. Artsimovitch I, Vassylyeva MN, Svetlov D, Svetlov V, Perederina A, Igarashi N, Matsugaki N, Wakatsuki S, Tahirov TH, Vassylyev DG: **Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins**. *Cell* 2005, **122**:351-363.

17. Temiakov D, Zenkin N, Vassylyeva MN, Perederina A, Tahirov TH, Kashkina E, Savkina M, Zorov S, Nikiforov V, Igarashi N et al.: **Structural basis of transcription inhibition by antibiotic streptolydigin**. *Mol Cell* 2005, **19**:655-666.

18. Vassylyev DG, Svetlov V, Vassylyeva MN, Perederina A, Igarashi N, Matsugaki N, Wakatsuki S, Artsimovitch I: **Structural basis for transcription inhibition by tagetitoxin**. *Nat Struct Mol Biol* 2005, **12**:1086-1093.

19. Belogurov GA, Vassylyeva MN, Sevostyanova A, Appleman JR, Xiang AX, Lira R, Webber SE, Klyuyev S, Nudler E, Artsimovitch I et al.: **Transcription inactivation through local refolding of the RNA polymerase structure**. *Nature* 2009, **457**:332-335.
The structure of the bacterial RNAP holoenzyme in complex with antibiotic myxopyronin revealed that refolding of the switch-2 segment mediates downstream propagation of the transcription bubble during open complex formation. The data suggest that the conformational transitions of the switch-2 may play an essential regulatory role.

20. Campbell EA, Pavlova O, Zenkin N, Leon F, Irschik H, Jansen R, Severinov K, Darst SA: **Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase**. *EMBO J* 2005, **24**:674-682.

21. Campbell EA, Korzheva N, Mustaev A, Murakami K, Nair S, Goldfarb A, Darst SA: **Structural mechanism for rifampicin inhibition of bacterial RNA polymerase**. *Cell* 2001, **104**:901-912.

22. Kashkina E, Anikin M, Brueckner F, Lehmann E, Kochetkov SN, McAllister WT, Cramer P, Temiakov D: **Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site**. *J Biol Chem* 2007, **282**:21578-21582.

23. Naji S, Bertero MG, Spitalny P, Cramer P, Thomm M: **Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement**. *Nucleic Acids Res* 2008, **36**:676-687.

24. Toulokhonov I, Landick R: **The role of the lid element in transcription by *E. coli* RNA polymerase**. *J Mol Biol* 2006, **361**:644-658.

25. Kent T, Kashkina E, Anikin M, Temiakov D: **Maintenance of RNA-DNA hybrid length in bacterial RNA polymerases**. *J Biol Chem* 2009, **284**:13497-13504.

26. Epshtein V, Cardinale CJ, Ruckenstein AE, Borukhov S, Nudler E: **An allosteric path to transcription termination**. *Mol Cell* 2007, **28**:991-1001.

27. Santangelo TJ, Roberts JW: **Forward translocation is the natural pathway of RNA release at an intrinsic terminator**. *Mol Cell* 2004, **14**:117-126.

28. Toulokhonov I, Zhang J, Palangat M, Landick R: **A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing**. *Mol Cell* 2007, **27**:406-419.

29. Yin YW, Steitz TA: **Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase**. *Science* 2002, **298**:1387-1395.

30. Tahirov TH, Temiakov D, Anikin M, Patlan V, McAllister WT, Vassylyev DG, Yokoyama S: **Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution**. *Nature* 2002, **420**:43-50.

31. Cheetham GM, Jeruzalmi D, Steitz TA: **Structural basis for initiation of transcription from an RNA polymerase–promoter complex**. *Nature* 1999, **399**:80-83.

32. Rutherford ST, Villers CL, Lee JH, Ross W, Gourse RL: **Allosteric control of Escherichia coli rRNA promoter complexes by DksA**. *Genes Dev* 2009, **23**:236-248.

33. Roberts JW, Shankar S, Filter JJ: **RNA polymerase elongation factors**. *Annu Rev Microbiol* 2008, **62**:211-233.

34. Belogurov GA, Vassylyeva MN, Svetlov V, Klyuyev S, Grishin NV, Vassylyev DG, Artsimovitch I: **Structural basis for converting a general transcription factor into an operon-specific virulence regulator**. *Mol Cell* 2007, **26**:117-129.
The structure of RfaH allowed to elucidate mechanism of action of this transcription factor in which the protein undergoes DNA-dependent activation via the domain dissociation that is required to open the binding cavity and to form a stable complex with RNAP. The structural analysis also allowed to identify the β′-subunit clamp helices as the RfaH major binding site on RNAP to which the protein is recruited through the hydrophobic interactions.

35. Sevostyanova A, Svetlov V, Vassylyev DG, Artsimovitch I: **The elongation factor RfaH and the initiation factor sigma bind to the same site on the transcription elongation complex**. *Proc Natl Acad Sci U S A* 2008, **105**:865-870.

36. Steiner T, Kaiser JT, Marinkovic S, Huber R, Wahl MC: **Crystal structures of transcription factor NusG in light of its nucleic acid- and protein-binding activities**. *EMBO J* 2002, **21**:4641-4653.

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

38. Gopal B, Haire LF, Gamblin SJ, Dodson EJ, Lane AN, Papavinasasundaram KG, Colston MJ, Dodson G: **Crystal structure of the transcription elongation/anti-termination factor NusA from Mycobacterium tuberculosis at 1.7 A resolution**. *J Mol Biol* 2001, **314**:1087-1095.

39. Perederina A, Svetlov V, Vassylyeva MN, Tahirov TH, Yokoyama S, Artsimovitch I, Vassylyev DG: **Regulation through the secondary channel–structural framework for ppGpp-DksA synergism during transcription**. *Cell* 2004, **118**:297-309.

40. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Foster JW, Gourse RL: **DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP**. *Cell* 2004, **118**:311-322.

41. Borukhov S, Laptenko O, Lee J: ***Escherichia coli* transcript cleavage factors GreA and GreB: functions and mechanisms of action**. *Methods Enzymol* 2001, **342**:64-76.

42. Sosunova E, Sosunov V, Kozlov M, Nikiforov V, Goldfarb A, Mustaev A: **Donation of catalytic residues to RNA polymerase active center by transcription factor Gre**. *Proc Natl Acad Sci U S A* 2003, **100**:15469-15474.

43. Laptenko O, Lee J, Lomakin I, Borukhov S: **Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase**. *EMBO J* 2003, **22**:6322-6334.

44. Opalka N, Chlenov M, Chacon P, Rice WJ, Wriggers W, Darst SA: **Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase**. *Cell* 2003, **114**:335-345.

45. Vassylyeva MN, Svetlov V, Dearborn AD, Klyuyev S, Artsimovitch I,
• Vassylyev DG: **The carboxy-terminal coiled-coil of the RNA polymerase beta'-subunit is the main binding site for Gre factors**. *EMBO Rep* 2007, **8**:1038-1043.
The structure of the *E. coli* GreB protein allowed to identify the β'-subunit C-terminal coiled-coil (β′CC) of RNAP as the major anchor for the protein to which GreB binds through predominantly the hydrophobic interactions.

46. Symersky J, Perederina A, Vassylyeva MN, Svetlov V, Artsimovitch I, Vassylyev DG: **Regulation through the RNA polymerase secondary channel. Structural and functional variability of the coiled-coil transcription factors**. *J Biol Chem* 2006, **281**:1309-1312.

47. Lamour V, Hogan BP, Erie DA, Darst SA: **Crystal structure of Thermus aquaticus Gfh1, a Gre-factor paralog that inhibits rather than stimulates transcript cleavage**. *J Mol Biol* 2006, **356**:179-188.

48. Laptenko O, Kim SS, Lee J, Starodubtseva M, Cava F, Berenguer J, Kong XP, Borukhov S: **pH-dependent conformational switch activates the inhibitor of transcription elongation**. *EMBO J* 2006, **25**:2131-2141.

49. Blankschien MD, Lee JH, Grace ED, Lennon CW, Halliday JA, Ross W, Gourse RL, Herman C: **Super DksAs: substitutions in DksA enhancing its effects on transcription initiation**. *EMBO J* 2009, **28**:1720-1731.

50. Lamour V, Rutherford ST, Kuznedelov K, Ramagopal UA, Gourse RL, Severinov K, Darst SA: **Crystal structure of Escherichia coli Rnk, a new RNA polymerase-interacting protein**. *J Mol Biol* 2008, **383**:367-379.

51. Blankschien MD, Potrykus K, Grace E, Choudhary A, Vinella D, Cashel M, Herman C: **TraR, a homolog of a RNAP secondary channel interactor, modulates transcription**. *PLoS Genet* 2009, **5**:e1000345.

52. Pinto UM, Winans SC: **Dimerization of the quorum-sensing transcription factor TraR enhances resistance to cytoplasmic proteolysis**. *Mol Microbiol* 2009, **73**:32-42.

53. Chlenov M, Masuda S, Murakami KS, Nikiforov V, Darst SA, Mustaev A: **Structure and function of lineage-specific sequence insertions in the bacterial RNA polymerase beta' subunit**. *J Mol Biol* 2005, **353**:138-154.