

# Accepted Manuscript

Translocation by multi-subunit RNA polymerases

Maria Kireeva, Mikhail Kashlev, Zachary F. Burton

PII: S1874-9399(10)00018-0  
DOI: doi: [10.1016/j.bbagr.2010.01.007](https://doi.org/10.1016/j.bbagr.2010.01.007)  
Reference: BBAGRM 234

To appear in: *BBA - Gene Regulatory Mechanisms*

Received date: 28 October 2009  
Revised date: 14 January 2010  
Accepted date: 14 January 2010



Please cite this article as: Maria Kireeva, Mikhail Kashlev, Zachary F. Burton, Translocation by multi-subunit RNA polymerases, *BBA - Gene Regulatory Mechanisms* (2010), doi: [10.1016/j.bbagr.2010.01.007](https://doi.org/10.1016/j.bbagr.2010.01.007)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

TRANSLOCATION BY MULTI-SUBUNIT RNA POLYMERASES

Maria Kireeva<sup>a</sup>, Mikhail Kashlev<sup>a</sup> and Zachary F. Burton<sup>b,\*</sup>

<sup>a</sup>Gene Regulation and Chromosome Biology Laboratory, National Cancer Institute—Frederick, Bldg. 539, Room 222, Frederick, MD 21702-1201

<sup>b</sup>Department of Biochemistry and Molecular Biology, Michigan State University, E. Lansing, MI 48824-1319

\*corresponding author: [burton@cns.msu.edu](mailto:burton@cns.msu.edu), 517-353-0859, fax: 517-353-9334

DNA template and RNA/DNA hybrid movement through RNA polymerase (RNAP) is referred to as “translocation”. Because nucleic acid movement is coupled to NTP loading, pyrophosphate release, and conformational changes, the precise ordering of events during bond addition is consequential. Moreover, based on several lines of experimental evidence, translocation, pyrophosphate release or an associated conformational change may determine the transcription elongation rate. In this review we discuss various models of translocation, the data supporting the hypothesis that translocation rate determines transcription elongation rate and also data that may be inconsistent with this point of view. A model of the nucleotide addition cycle accommodating available experimental data is proposed. On the basis of this model, the molecular mechanisms regulating translocation and potential routes for NTP entry are discussed.

### Structural features of the multi-subunit RNAPs

Multi-subunit RNA polymerases (RNAPs) are large, dynamic molecular machines ubiquitous to the three kingdoms of life, eubacteria, archaea and eukarya, that share a significant degree of homology [1-3]. A multitude of biochemical and structural evidence suggests that the basic mechanism of transcription elongation is common between RNAPs from different organisms. Therefore, most of the structural and functional properties determined for one multi-subunit RNAP are generally applicable to other multi-subunit RNAPs.

RNAPs extend a RNA chain by processively reading a DNA template strand [4]. In doing so, RNAPs function as powerful molecular motors [5] translocating along DNA and only releasing the template and its RNA transcript in response to specific termination signals [4,6,7]. In Figure 1, images of the *Saccharomyces cerevisiae* RNAP II ternary elongation complex (TEC) [8] are shown. In contrast to DNA polymerases, which can release and re-bind DNA during replication, even temporary dissociation of the RNAP TEC leads to the collapse of the transcription bubble and abortion of transcription. To prevent this from happening, RNAPs encircle DNA with a protein clamp (Figure 1A) limiting accessibility of the active center to incoming substrates and hindering release of inorganic pyrophosphate – a bi-product of phosphodiester bond formation. In this review, we discuss the mechanisms and structural elements in RNAP (protein channels, helices and flexible loops) that are involved in substrate delivery, catalysis, pyrophosphate release and translocation along DNA.

The long “bridge helix” (BH) [9,10], which borders the active site, is a defining feature of RNAP. Because the BH abuts the RNA/DNA hybrid, a possible molecular mechanism for translocation (Figures 1B and 2) might involve thermal fluctuations and/or local unwinding or bending of the BH [9,11-14] (Figure 3). Indeed, identification of hyperactive BH mutant RNAPs by saturation mutagenesis supports

the potential importance of BH dynamics in translocation [14]. The two “trigger helices” (TH), connected by a short flexible “trigger loop” (TL), form a three-helix bundle with the BH. The TH-TL part of this bundle undergoes a major conformational change (TL→TH closing) upon binding of the NTP substrate to the active center of RNAP [15] (Figure 1B). Notably, because of close proximity between TH and BH, opening and closing of the TL-TH is expected to modify BH dynamics and thus regulate translocation [11,13,14,16]. The existence of the specific conformation of the TL, wedged against the central part of the BH and stabilized by the transcription inhibitor  $\alpha$ -amanitin [13], supports the possibility of a coordinated conformational change of the TL and BH during translocation (Figure 3, structure 2). The concept of thermally driven translocation [11,13,17-22] has been a dominant theme in the field, with the basic idea that RNAP spontaneously steps between pre- and post-translocation states (Figure 2).

Because the RNAP active site is buried deep within the RNAP structure, the question arises of the route(s) of NTP entry into the TEC. One possible route is the secondary pore, a solvent accessible channel [8,16,23-25] that appears to be gated by opening and closing of the TL [8], folding of the TL to the TH [16] and by the translocation state of the TEC [8,9] (Figures 1 and 2). When loaded through the secondary pore, it appears that NTPs can bind only to the post-translocated TEC because, in the pre-translocated TEC, the DNA template base projects into the main enzyme channel rather than the secondary pore [9] (Figure 2, upper panel). In addition, the close proximity of the 3' end of the nascent transcript against the BH in the pre-translocated TEC (Figures 2 and 3) leaves little space for the incoming NTP in the active center [9].

### **The role of translocation in the nucleotide addition cycle**

Transcription elongation occurs by a repetitive nucleotide addition cycle. The TEC, ready for catalysis, has the 3' end of the nascent RNA base paired to the template and positioned in the *i* site

(Figure 2, lower panel); the substrate NTP enters the active center and is paired to the next DNA base in the  $i+1$  site. The NTP is positioned for catalysis by a conformational change (isomerization) of the RNAP. In single-subunit RNAPs, isomerization comprises a major closing of RNAP domains [26]; in multi-subunit RNAPs, isomerization includes movement of the TL, a mobile element of the catalytic subunit [8,16], TL folding into the TH [16], and formation of multiple contacts of the TH with the NTP in the RNAP active site (Figure 3) [8,16]. The closed, catalytic TH structure has a primarily helical conformation (Figure 3; structure 3), while the relaxed TL has substantially more loop (structures 1 and 2) [16]. Characterization of TL mutants [27,28] and kinetic analyses of transcription elongation [28-32] established that isomerization may represent one of the slower steps of the nucleotide addition cycle. It remains to be determined the extent to which isomerization and reverse isomerization are regulated by external factors, including incoming NTP substrates. When isomerization is completed, phosphodiester bond formation is likely to occur rapidly.

Notably, upon completion of phosphodiester bond formation, the TEC does not immediately become available to accept the next substrate NTP. Prior to the positioning of the next NTP in the RNAP active site, pyrophosphate, the by-product of elongation, is expelled, and RNAP undergoes translocation, moving one base pair downstream along the DNA template. The 3' end of the RNA is thus transferred from the  $i+1$  (substrate) site to the  $i$  site within the RNAP active center (see Figure 2). The possibility that pyrophosphate could remain TEC-associated after translocation and that pyrophosphate release could then be stimulated by binding of the next NTP to the  $i+1$  site was also proposed [33], although this model appears to require stable retention of pyrophosphate in a post-translocated TEC, and, assuming secondary pore NTP loading, with an open TL. Although translocation of RNAPs has been extensively studied, the detailed molecular mechanism of this process and its exact kinetic

characteristics remain to be established. For instance, the rate of the transition from the pre-translocated to the post-translocated state (see Figure 2) has not been experimentally determined. While translocation of bacteriophage T7 DNA polymerase [34], poliovirus RNA polymerase [35], and T7 RNAP [36] is rapid and does not affect the apparent polymerization rate, the shifts of the translocation equilibrium of HIV reverse transcriptase do affect RT catalytic activity [37]. Translocation by multi-subunit RNAPs has been proposed to be a target for regulation of transcription elongation, and several models of translocation currently exist in the field.

#### **Power stroke and thermal ratchet translocation**

Three models of translocation: a power stroke model (not shown), a thermal ratchet model (Figures 2 and 4, left panel), and two variants of allosteric models (see Figure 4, right two panels) have been considered in the field. The power stroke model, originally proposed on the basis of structural analyses of the bacteriophage T7 RNAP TEC [26] postulated that translocation requires the energy of NTP hydrolysis, and is tightly coupled to NMP incorporation into the growing RNA chain and to pyrophosphate release. Currently this model is disfavored, because it does not appear consistent with most biochemical, single molecule, and structural data. Specifically, the power stroke mechanism does not explain the ability of multi-subunit RNAPs to undergo catalytic inactivation by backtracking, during which the 3' end of the nascent RNA is misplaced from the RNAP active site [38,39]; to spontaneously return to the active state in the absence of RNA synthesis [40] and the ability of RNAP to translocate backwards by a large distance during processive pyrophosphorolysis and endonucleolytic shortening of the 3' end of the transcript stimulated by transcript cleavage factors [41]. Successful crystallization of several post-translocated TECs [8,15,16,25,42-44] indicates that RNAP preferentially dwells in a post-translocated state in the absence of NTP substrates and NTP hydrolysis. At least one crystal of the RNAP

II TEC captured the pre-translocated state [9], and one structure showed equilibrium between the pre-translocated and post-translocated states in the crystal matrix [13]. The most convincing functional data ruling out the power stroke mechanism of elongation were obtained from single molecule analyses of transcription [17,45,46].

Earlier investigations of transcription elongation in a single-molecule experimental setup [47] demonstrated that transcription by bacterial RNAP at saturating NTP concentrations stalls when an applied resisting force reaches 14 piconewtons. Subsequent research suggested that high opposing force induces backtracking of bacterial RNAP [46] and yeast RNAP II [45]. Importantly, Abbondanzieri et al. [17] established that at limiting NTP concentrations RNAP advances in single base pair increments, indicating that either NTP loading or a NTP-dependent conformational change becomes rate-limiting for elongation. In that work, the best fits to force-velocity curves were obtained from an elongation model in which NTP substrates interact with both the pre- and post-translocated TEC.

The question of whether on-pathway elongation rates are strictly force-dependent (which would indicate that the equilibrium between the pre- and post-translocation states is reached at each nucleotide addition cycle and determines the transcription elongation rate) remained controversial. On the one hand, the Wang [48,49] and the Bustamante [45,50,51] groups reported that transcription rates by bacterial RNAP and yeast RNAP II are strictly force-dependent; on the other hand, Block, Gelles and co-authors reported that, at moderate opposing force loads or with application of assisting force, transcription velocity may be largely independent from the applied force [17,52-54]. The current consensus appears to be that differences in conclusions may largely be attributed to varying approaches to normalization of the experimental data [55]; currently, transcription velocity at non-saturating NTP

concentrations is believed to be dependent on applied force, consistent with a simple thermal ratchet translocation mechanism.

### **Ratchet translocation and transcriptional pausing**

A thermal ratchet mechanism of translocation has been used to explain numerous aspects of transcription pausing. Von Hippel and co-authors suggested that some transcription pauses are thermodynamically induced [22]. The idea of the thermodynamic nature of pausing has been further developed based on catalytic inactivation of RNAP induced through backtracking [40]. Upstream movement of RNAP, referred to as backtracking, occurs without a change of the RNA-DNA register, and maintains a consistent size of the transcription bubble. As a result, the RNA 3' end extrudes from the active site into the secondary pore [38,56]. The backtracked state is incompatible with substrate NTP loading and RNA extension. Backtracking extent and efficiency is believed to be largely dependent on the lower free energy of the TEC at an upstream position, compared to the active conformation in which the 3' end of the RNA is located in the active site [40,57,58]. However, the free energy of the transcription bubble is not the only factor that determines the propensity of RNAP to backtrack. Other factors, including, but not limited to, the nature of the 3' NMP and downstream DNA sequence, can affect backtracking [59].

It has been shown that bacterial RNAP and eukaryotic RNAP II undergo backtracking at some sequence specific pause sites [57,58]. Furthermore, a theoretical model was proposed by Wang and co-workers suggesting that sequence-specific transcriptional pausing is determined by the propensity of RNAP to remain in the pre-translocated state or to undergo reverse translocation (backtracking) to a more thermodynamically favorable position on the template [18,19]. This model considers translocation a strictly thermal-driven process, the efficiency of which depends on the difference in the free energy of

a TEC in the pre- and post-translocated state. In order to achieve a fairly high accuracy in the prediction of localization and sequence-specific pauses on various DNA templates, this model postulates that there is a significant energy barrier for the transition of the pre-translocated TEC to the backtracked state. Therefore, according to this model, most short-lived sequence-specific pauses are due to a delay in thermal ratchet translocation [18].

A distinct theoretical model based on the thermodynamics of the TEC and considering translocation as the main step of the nucleotide addition cycle that determines the elongation dynamics has been proposed by Ruckenstein and co-workers [21]. Their model predicts that most transcriptional pauses are caused by limited backtracking of RNAP. It argues that there is no structural basis for the existence of a significant thermodynamic barrier between the pre-translocated and backtracked state; instead, backtracking (and, therefore, pausing) extent is thought to be limited by formation of nascent RNA hairpins, which limit backward translocation of RNAP. Otherwise, the models suggested by Bai et al. [18,19] and by Tadigotla et al. [21] are fundamentally similar. Both models predicted pausing patterns on several previously characterized DNA sequences with comparable accuracy, arguing that RNAP translocation could directly regulate elongation and transcriptional pausing.

The link between translocation and pausing is suggested, not only by the thermodynamic analyses and theoretical modeling of transcriptional pausing, but also by biochemical analyses of hairpin-dependent pausing intermediates. Landick and co-workers established that sequence-specific pausing in the leader region of the *his* operon is determined by a multi-partite signal that includes the RNA hairpin, the downstream DNA sequence, and the nature of the 3'NMP base of the nascent RNA in the paused TEC [60,61]. The TEC at the *his* pause site preferentially dwells in the pre-translocated state [62], and it has been suggested that translocation is blocked because the 3' end of the RNA in the

paused TEC acquires a specific “frayed” conformation [62-64]. It is not clear, however, whether the fraying of the 3' NMP occurs in a hairpin-dependent manner [62,64,65]. Importantly, hairpin-dependent and backtracked pauses might share a common precursor [57], which was presumed to be in the pre-translocated state [57,66]. Furthermore, based on single-molecule data, the biochemically characterized hairpin-dependent *his* leader pause and several sequence-specific “ubiquitous” pauses, also appear to share a common precursor [53]. Taken together, the unified transcription pause hypothesis (reviewed in [66]) suggested that all types of transcriptional pauses share a common pre-translocated precursor and thus emphasized the importance of translocation as a regulatory step in transcription elongation.

#### **Thermal (Brownian) ratchet translocation in transcription elongation**

The thermal ratchet model suggests that translocation occurs by thermal fluctuations of RNAP along the template with one base pair increments [22,67]. Published TEC structures provide no immediate clue to identify which domains in RNAP constitute the ratchet and how translocation is constrained to a single base pair. However, despite the extensive number of protein-nucleic acid contacts that are broken during the process, translocation is practically isoenergetic, because reannealing of the upstream DNA compensates for the melting of the downstream DNA and contacts that are broken are re-formed [18,19,22]. Specifically, RNAP displaces about 12 nucleotides of the non-template DNA, and nascent RNA forms an 8-9 base pair RNA/DNA hybrid [7,68,69]. Downstream DNA is tightly held by the RNAP clamp [10,23,69] (see Figure 1). During translocation, the RNAP (which maintains extensive contacts with DNA and RNA) moves forward, rewinding the DNA at the upstream edge of the transcription bubble, melting the upstream base pair in the RNA-DNA hybrid, and unwinding the DNA at the downstream edge of the transcription bubble [7].

The ratchet translocation model postulates that during the nucleotide addition cycle, phosphodiester bond formation is followed by a rapid pyrophosphate release and by a spontaneous, thermally driven diffusion of the TEC from the pre-translocated to the post-translocated state (Figure 2). In the thermal ratchet model, the equilibrium between the pre-translocated and post-translocated states is established before the next NTP diffuses into the active site [18-20,22,67] (Figure 4, left panel, 2→3). A ratchet mechanism of translocation may be conserved between multi-subunit RNAPs, single-subunit RNAPs, and reverse transcriptases. In agreement with the ratchet mechanism predictions, the post-translocation conformation is stabilized by a complementary incoming NTP in the multi-subunit RNAPs [11,28,70,71], in the T7 RNAP [20], and in the HIV-1 reverse transcriptase [72]. Furthermore, pyrophosphate shifts the translocation equilibrium toward the pre-translocated state in T7 RNAP [20], and pyrophosphate analogues shift the translocation equilibrium toward the pre-translocated state in HIV reverse transcriptase [37]; the effect of pyrophosphate or a pyrophosphate analogue on the translocation equilibrium of multi-subunit RNAPs has not yet been addressed.

Additional insight into the control of the translocation equilibrium has been obtained from the study of two mutations located in the vicinity of the G loop (an alternate name for the TL) of the *E. coli* RNAP  $\beta'$  subunit [11]. One mutation caused slow transcription elongation, significantly enhancing sequence-specific pausing. The second mutation increased the overall transcription rate, suppressing sequence-specific pausing. TECs formed by wild type and mutant variants of RNAP were analyzed using Exo III footprinting, which allowed for estimation of the translocation equilibrium in stalled TECs. The translocation equilibrium can be visualized as one base pair heterogeneity of the Exo III footprints of RNAP in a TEC halted at a unique template position by substrate deprivation [11,28,71,73]. Nudler and co-workers [11] observed that in a TEC formed by a slow mutant, the translocation equilibrium was

shifted toward the backtracked and pre-translocated states. By contrast, the mutation increasing the overall transcription rate and decreasing transcriptional pausing shifted the translocation equilibrium toward the post-translocated state. Based on these observations, it was concluded that the rate of RNAP oscillations between the pre- and post-translocation states, and the translocation equilibrium are essential determinants of the transcription elongation rate. Furthermore, TL and TH, along with the incoming NTP, were proposed to play a crucial role in regulation of the translocation equilibrium. The two-pawl ratchet mechanism became a broadly accepted model for RNAP translocation (for example, see [74]), but some details of the mechanism have been challenged [65,71]. Specifically, the ratchet translocation model could not explain a number of observations from pre-steady-state analyses of transcription elongation, which indicated that the dynamics of individual NTP additions could be influenced by NTPs complementary to the downstream template base [29,31,32,75-78]. Therefore, alternative models suggesting that translocation presents a target for allosteric regulation [77,78] are considered in the field.

#### **Incoming NTP as an allosteric regulator of translocation**

The first report on pre-steady-state analyses of NTP incorporation by *E. coli* RNAP established that the time course for RNA extension is biphasic [75], and this observation was interpreted as an indication of the existence of fast and slow (activated and unactivated) pathways of nucleotide addition (Figure 4, right panel). Furthermore, a non-hydrolyzable analogue of a substrate NTP promotes incorporation of the substrate NTP by *E. coli* RNAP. It has been suggested that the non-hydrolyzable analogue binds to an allosteric site (the streptolydigin-binding loop in *E. coli* RNAP; fork loop 2 (FL2) in yeast RNAP II, see Figure 6A), and induces a conformational change of the TEC from the unactivated to the activated state [75] (Figure 4, 1→7). According to this model, the activated TEC incorporates the

regular NTP more efficiently than the unactivated TEC. It is necessary to emphasize that the ratchet model, which only allows NTP binding to the  $i+1$  site or to the site overlapping with the  $i+1$  site (see Figure 2), predicts that a non-hydrolyzable analogue can only decrease the transcription elongation rate, by acting as a competitive inhibitor of NTP binding to the post-translocated TEC.

Notably, an experiment was done with the non-hydrolyzable ATP analogue AMPcPP on a template encoding two consecutive AMP residues. In addition to the AMPcPP analogue, limiting ATP was present to support slow AMP incorporation. Addition of AMPcPP strongly stimulated incorporation of the second AMP compared to the first AMP [75]. These observations, taken together with the dependence of the NTP incorporation rate on the nature of the downstream template base, established in subsequent work from the same group suggested that the “activation” of RNAP through an “allosteric” site promotes translocation [77]. Based on this interpretation, the fast fraction of the TECs observed in the biphasic time course was assumed to be a post-translocated TEC and the slow fraction a pre-translocated TEC. The model of the nucleotide addition cycle, in which translocation is stimulated by the incoming NTP, differs from the original thermal ratchet model in an essential aspect: the substrate NTP must bind to the pre-translocated TEC (Figure 4, right panel), at an allosteric site distinct from the catalytic center (Figure 5C).

### **NTP-driven translocation**

Burton and co-workers advanced the pre-steady state analyses of transcription elongation, applying a variety of functional probes in attempts to dissect the transcription elongation cycle. NTP incorporation by human RNAP II appeared biphasic [78], similar to NTP incorporation kinetics by *E. coli* RNAP [75]. The fast fraction was interpreted as a post-translocated TEC, “poised” for NTP binding and bond formation. The amplitude of the fast fraction was dependent on the presence of elongation factor

TFIIF and hepatitis delta antigen, suggesting that these factors stimulate translocation [78]. Notably, in the presence of TFIIF, the NTP complementary to the  $i+2$  base of the template DNA increased the fast burst fraction in the incorporation of the preceding NMP [29]. This observation was similar to the effects reported for bacterial RNAP [77].

Additionally, a potent transcription elongation inhibitor, the mushroom-derived toxin  $\alpha$ -amanitin, was used as a probe to analyze rate-limiting steps of NMP addition by human RNAP II and to reveal the effects of downstream NTP substrates. Fast NMP addition of a single base was observed even in the presence of  $\alpha$ -amanitin, suggesting that the toxin inhibits translocation (Figure 4,  $2 \rightarrow 3$ ), but does not affect NTP binding and phosphodiester bond formation [76]. Notably, the amplitude of the fast amanitin-resistant fraction depended not only on the presence of TFIIF, but also on the presence of the NTP complementary to the  $i+2$  template base, and this effect was further enhanced by NTPs complementary to the  $i+3$  and  $i+4$  bases [29,76]. Furthermore, in the presence of  $\alpha$ -amanitin, the presence of NTPs complementary to downstream template bases induced a reversal of phosphodiester bond formation [29] (Figure 4, right panel,  $6 \rightarrow 5$ ) requiring the presence of the substrate NTP complementary to the  $i+2$  template base [29,31] (Figure 4,  $10 \rightarrow 9$ ). In summary, pre-steady-state analyses of elongation by human RNAP II provided several lines of experimental evidence that substrate NTPs interact in a template-dependent manner with the pre-translocated TEC and affect the dynamics of the previous bond formation (see Figure 5C).

Another crucial technical breakthrough achieved by the Burton group is characterization of the synthesis of two consecutive phosphodiester bonds by RNAP II in a pre-steady-state setup [32,78]. Decrease in the rate of formation of the second bond compared to the rate of the first bond was proposed as a test of whether translocation and pyrophosphate release could be rate-limiting [22].

Notably, synthesis of the second phosphodiester bond by human RNAP II was significantly slower than the synthesis of the first phosphodiester bond [32,78]. The difference between the rates of the escape from the stall and processive synthesis has been confirmed in the analyses of phosphodiester bond formation by yeast RNAP II and *E. coli* RNAP. In later experiments, the rate of a given phosphodiester bond formation has been measured as the escape from the stall (first bond) and as processive synthesis (second bond), ruling out effects of local sequence context on the rate of phosphodiester bond formation. In this setup, the rate of processive synthesis was at least twice as slow as the rate of the escape from the stall [79]. Therefore, at least one of the steps manifesting completion of the nucleotide addition cycle, which includes reverse isomerization, pyrophosphate release and translocation, is rate-limiting during transcription elongation.

Based on these observations, a NTP-driven translocation model was proposed [29,32,76,78,80-82]. According to this model, pre-hybridization of downstream NTPs promotes translocation [32,78]; during processive RNA synthesis, NTP-driven translocation is coupled to pyrophosphate release [29,31,32,78] (Figure 4, 7→8). The potential rate-limiting nature of pyrophosphate release (1→2) is specifically emphasized in the NTP-driven translocation model based on the analyses of formation of two phosphodiester bonds [32,78], on rates of completion of phosphodiester bonds and on the observation of bond reversal with  $\alpha$ -amanitin [29,31].

Strong support for the view that pyrophosphate release can be rate-limiting (Figure 4, right panel, 1→2) and could be modulated by the next incoming substrate NTP (7→8) was obtained from pre-steady-state analyses of pyrophosphate release by *E. coli* RNAP [33]. Pyrophosphate, generated as the bi-product of phosphodiester bond formation, lingered within the TEC for seconds, and such a slow rate of pyrophosphate release is incompatible with the bulk elongation rate (10-20 nt/s). Notably, in the

presence of the next substrate NTP encoded by the template, the observed rate of pyrophosphate release increased about 200-fold. This observation indicates that the next NTP binds to a pre-translocated TEC with bound pyrophosphate in the active site, and the incoming NTP dramatically stimulates the completion of the previous nucleotide addition cycle. However, the alternative scenario must be considered that pyrophosphate may survive translocation to be displaced from the post-translocated TEC by the incoming NTP. It is important to emphasize that, according to a broadly accepted interpretation of structural data, the substrate NTP enters the active site of the post-translocated TEC through the secondary pore, but it appears that a NTP cannot be loaded into the pre-translocated TEC via the pore (see Figure 2). A pure thermal ratchet translocation mechanism demands that NTPs load directly to the active site, presumably through the secondary pore, but such a model is not consistent with the effects of an incoming NTP substrate on formation of the preceding bond and on the release of pyrophosphate as observed in the pre-steady-state kinetic experiments discussed above. To resolve the apparent contradiction, we re-examine the available structural information and identify alternative NTP loading routes, by which the incoming NTP might interact with the pre-translocated TEC.

### **Potential routes of NTP loading**

The problem of NTP substrate entry appears unique to multi-subunit RNAPs (see Figure 1). In most DNA polymerases, single-subunit RNAPs and reverse transcriptases, the active site appears readily accessible for substrate (d)NTPs. By contrast, the active site of multi-subunit RNAPs is deeply buried within the protein structure and is not readily accessible from outside [8,15,16,23-25,43]. Sequestration and enclosure of nucleic acids ensures high processivity of transcription but presents a barrier to NTP access. As noted above, the active site of RNAP is made up of the  $i$  and  $i+1$  sites. In the post-translocated TEC, the  $i$  site is the position of the 3'-end of the RNA, and the incoming NTP substrate should ultimately

be loaded into the  $i+1$  site (also defined as the A site, see Figure 5A). In the pre-translocated TEC, the 3'-end of the RNA occupies the  $i+1$  site. Therefore, to explain the incoming NTP effect on pre-translocated TECs, described above, it is necessary to identify another NTP binding site(s), distinct and downstream from the  $i+1$  site (see Figure 5). We consider two potential modes of NTP loading to the RNAP active site: 1) NTPs enter through the secondary pore directly to the active site ( $i+1$ ) (Figure 5A, B, D); and 2) NTPs enter through the main enzyme channel, past the DNA non-template strand, to a downstream template site(s) and eventually to the active site (i.e.  $i+2 \rightarrow i+1$ ) (Figure 5C).

### Secondary pore

The secondary pore is a solvent accessible channel to the active site, which has been suggested to be the sole or major route of NTP entry based on multiple sets of crystallographic data obtained for bacterial and yeast RNAP TECs [15,16,23,24,43]. In order to load NTPs through the secondary pore, however, the TEC must be post-translocated and the TL must be in an open conformation (see Figures 1 and 2). Therefore, significant constraints on timing, routes, and consequences of NTP loading through the pore are evident. Indeed, theoretical simulation of diffusion of a negatively charged NTP-Mg<sup>2+</sup> (charge of -2) into the pore, which itself has negative electrostatics, suggests that even at physiological (0.2 to >3 mM) NTP concentrations, the complementary NTPs can enter into the active site through the pore and bind there at a rate of about 20 s<sup>-1</sup> [83]. It is necessary to emphasize that this estimated rate is about 50 times slower than the experimentally observed rate of the template-specific NTP sequestration for human [32,78] and yeast [28,79] RNAP II and *E. coli* RNAP [75,77,79]. Additionally, pre-binding of the NTP substrate in the entry (E) site (Figure 5E), which is believed to occur significantly faster than the NTP entry into the catalytic center (Figure 5A, B) [83], cannot explain the rapid template-specific

sequestration of substrate NTP[28,32,79], because a NTP in the E site cannot pair to the DNA template base [15] (Figure 5E).

Another argument against NTP entry through the secondary pore is based on the pre-steady-state analyses of transcription elongation in the presence of elongation factor TFIIIS, which binds in the secondary pore and occludes it [43,56] without showing significant effects on NTP affinity and on NTP loading [32,81]. Finally, NTP loading through the pore does not explain how downstream NTPs affect completion of the catalytic cycle, changing the pattern of bond reversal and promoting pyrophosphate release. Once again, observed incoming NTP effects are strictly template-dependent; therefore, they cannot readily be explained by NTP binding to the non-templated E site (Figure 5E). In summary, the hypothesis of NTP loading through the RNAP pore appears consistent with structural observations but not with kinetic data.

#### **Main channel**

A potential alternative to NTP loading through the secondary pore is NTP diffusion through the downstream DNA binding channel, and this route is consistent with template-dependent effects of downstream NTPs. Indeed, Holmes and Erie (2003) [77] proposed that the allosteric template-specific NTP binding site is formed by the streptolydigin-binding domain of the second largest subunit of *E. coli* RNAP ( $\beta$ -D Loop I or fork loop 2 (FL2)), a flexible structure located in the main enzyme channel that is proximal to the active site (Figure 6). Later, Burton and co-workers hypothesized that FL2 might be involved in the direct transfer of the NTP hybridized to the template DNA base from the downstream DNA binding channel to the active site [80]. The surprising flexibility of the DNA base, which in different crystal structures (containing similar conformations of the DNA backbone) is found in the downstream

DNA binding channel [9], over the BH [84], or in the active site [13,15,44,85] in principle, supports the possibility of such a transfer (Figure 6).

Loading NTPs to the TEC through the main enzyme channel or “cleft” also raises unresolved issues (see Figure 1). The main argument against downstream NTP loading is that NTPs have not been visualized in x-ray crystal structures of TECs. It may be that blocking chemistry at the  $i+1$  active site, which is necessary to maintain a NTP substrate at  $i+1$  for crystallography, distorts TECs to prevent downstream NTP loading. Furthermore, specific conditions required for crystallization might be unfavorable for downstream NTP binding. Currently, no electrostatic or diffusion modeling is available to indicate how NTPs might load through the main channel. As with the secondary pore NTP loading model, there does not appear to be much space to load NTPs through the clamp. Also, some partially folded conformations of FL2 might be expected to sterically interfere with the transfer of a template base-NTP pair over the BH to the active site. Even if FL2 is unfolded, the space for the NTP transfer from the main channel to the active site is constrained (Figure 6B). FL2 is either not resolved in the majority of crystals of *S. cerevisiae* RNAP II TECs, or it partially occludes the primary channel.

In the crystal of yeast RNAP II with the inhibitory RNA aptamer FC (a short synthetic RNA selected *in vitro* for its high affinity to RNAP II [86]), however, FL2 is found in a fully ordered conformation [84]. FC aptamer is comprised of two short RNA hairpins. One hairpin occupies the catalytic cleft with a similar placement to the RNA-DNA hybrid in the regular TEC. The other hairpin binds to the primary DNA channel in a slightly tilted configuration compared to the regular DNA duplex. In this crystal, FL2 occupies the most open position (Figure 6), which provides sufficient space for NTP diffusion from the  $i+2$  site in the main channel to the  $i+1$  site (Figure 6B).

The primary channel hypothesis considers that transfer of the NTP from the  $i+2$  to the  $i+1$  site occurs exclusively from the pre-translocated TEC and that NTP transfer is coupled to translocation (Figure 5C). Indeed, the  $i+2$  NTP in the post-translocated TEC hybridizes to the template in the primary channel generating a gap between the 3' end of the RNA in the  $i$  site and the NTP in the  $i+2$  site. In this complex, the  $i+2$  NTP occludes the channel, and it has to dissociate before the  $i+1$  site becomes accessible for NTP entry. The recently published crystal structures of a yeast RNAP II TEC in the presence of  $\alpha$ -amanitin [13], however, revealed a new “wedged” post-translocated intermediate lacking these spatial constraints. In the new TEC structure the RNA and DNA strands adopt the post-translocated configuration, whereas the  $i+1$  base in the template DNA is flipped out into the primary channel (Figure 5D). In this TEC structure, therefore, the NTP pre-loading site is moved forward close to the  $i+1$  position. Apparently, this TEC can accept the NTP in the primary channel followed by its loading to the active center as a base pair with the  $i+1$  base in DNA. A similar “pre-insertion” transition state was previously observed in a bacteriophage T7 RNAP TEC, in the absence of transcription inhibitors [85], arguing that such an intermediate may belong to the main elongation pathway rather than resulting from interaction of RNAP II with  $\alpha$ -amanitin.

It should be noted that a remarkable structural plasticity of the DNA-protein interactions in the main channel of RNAP may provide an additional space for NTP diffusion from the  $i+2$  to the  $i+1$  site. A big part of the channel is made from a mobile clamp in the Rpb1 subunit occupying highly variable positions in different crystals [23,44,87]. RNAP has been reported to transcribe to within a very short distance of obstacles in DNA such as bulky DNA lesions [88] and Gal repressor [89], Lac repressor [90] or EcoRI enzyme [91] bound to their cognate sites on DNA. Taking into account that the primary channel can accommodate ~18 base pairs of the downstream DNA, the ability of RNAP to transcribe to within 1-

2 base pairs of the lesion and to within 8-10 base pairs of the *EcoRI*-DNA or *LacI*-DNA complex argues that DNA can be lifted up from the channel or the channel can be substantially widened during elongation by opening of the protein clamp. The reported ability of RNAP II molecules transcribing DNA in tandem to approach each other closely [92] further supports this idea.

Another argument put forward against NTP loading through the downstream channel is that the non-template DNA strand might interfere with template-specific NTP loading. Clearly, downstream template opening is necessary for pre-loading of NTPs to template DNA; however, based on biochemical and structural studies, the extent of downstream bubble opening remains controversial. On the one hand, melting of 2-3 base pairs downstream from the RNA 3' end has been detected by chemical footprinting of the TECs formed by *E. coli* RNAP [38,93], early TECs of human RNAP II [94] and mature TECs of *S. cerevisiae* RNAP II [68,92]. On the other hand, in the *T. thermophilus* RNAP structure, downstream DNA is double-stranded up to and including the *i*+2 base pair [25]. It is necessary to mention, however, that the nucleic acid scaffold in the *T. thermophilus* TEC does not represent a complete transcription bubble, because in crystal construction the non-template DNA strand was truncated precisely at the *i*+2 position. Therefore, it is plausible that the non-template DNA is not, in fact, normally base paired to the template so close to the active site [25]. An alternative explanation of the observed downstream bubble closing in *T. thermophilus* RNAP is that the non-conserved  $\beta$  R422 residue plays a key role in maintaining the DNA duplex immediately downstream from the active site. By contrast, *S. cerevisiae* RNAP II has a glycine substituted at the homologous position (Rpb2 G506). Glycine could not perform a similar function to arginine in closing the downstream DNA, because the positively charged  $\beta$  R422 head group contacts the *i*+1 phosphate of the DNA template strand. Recent single-molecule FRET (fluorescence resonance energy transfer) studies used to map the trajectory of the non-

template DNA strand in a stalled *S. cerevisiae* RNAP II TEC indicate that  $i+2$  template bases are unpaired but  $i+3$  bases are paired [95].

The extent of downstream DNA melting in the vicinity of the active site will be resolved when structures of TECs formed on fully complementary dsDNA become available. Unfortunately, all the TECs of yeast RNAP II that were used for structural analyses so far, contain gaps or non-complementary DNA immediately downstream from the active site, and, therefore, the structure of the DNA in these TECs does not reflect the melting of the downstream DNA on a fully double stranded template. In summary, the results of the functional analyses of downstream DNA melting support the possibility of NTP hybridization to the  $i+2$  template base, while the results of structural analyses are inconclusive for this particular issue.

The hypothesis of FL2-mediated NTP loading has been challenged by site directed mutagenesis (Figure 6). The NTP-driven translocation model [78,80] and allosteric models [77] predicted that FL2 residues would likely be critical for elongation, and a few mutation phenotypes are consistent with this view. Specifically, *P. furiosus*  $\beta$  R445A (homologous position to *Sc Rpb2* R504) substitution has a significant effect on transcription [96]. Furthermore, *Sc Rpb2* R512C/A substitutions are very slow in elongation[97,98], which could be explained by changes in the conformation or the mobility of FL2. However, arginine 512 is adjacent both to FL2 and to the BH, and might directly affect the conformation of the active center; therefore, its substitution to cysteine or alanine might have a direct effect on catalysis. The latter possibility appears to be likely, because substitutions of *S. cerevisiae* RNAP II FL2 residues, *Rpb2* R504A, D505A, K510A, and Q513A do not have very strong effects on transcription [97] (Figure 6A). In summary, judging from current mutagenic studies, it does not appear that FL2 can have a very direct or specific role in forming a  $i+2$  NTP binding site in the main enzyme channel.

While the results of site-directed mutagenesis do not appear consistent with FL2 providing the specific contacts for substrate loading, these results do not rule out the general possibility of specific NTP hybridization to a template DNA base(s) in the downstream DNA channel. This possibility is also supported by the downstream DNA melting established by permanganate footprinting of TECs formed by prokaryotic and eukaryotic RNAPs (see references above). Therefore, despite the lack of direct experimental proof of NTP loading through the main enzyme channel, we incorporate the template-specific NTP hybridization to the template DNA base in the downstream DNA channel into our revised version of the model for the regulation of the nucleotide addition cycle (Figures 7 and 8) to explain the allosteric effects of the incoming NTP on transcription elongation.

#### **Nucleotide addition cycle during processive transcription: an updated model**

The spectrum of results obtained through comprehensive structural and functional analyses of transcription elongation does not appear entirely consistent with any of the previously proposed transcription elongation models. Here we attempt to find consensus between the Brownian ratchet, allosteric, and NTP-driven translocation models, which would accommodate current knowledge of TEC properties, kinetics of transcription, and effects of RNAP mutations (see Figures 4-8). Figures 2 and 7 show schematics that describe many of the proposed TEC structures, and Figures 4 and 8 present detailed pathway diagrams indicating each intermediate and elemental step.

In Figure 7, intermediate 1 is the product TEC after phosphodiester bond formation. Intermediate 1 has a closed TH and bound pyrophosphate in the active site ( $i+1$ ; green shading). Considering the intermediate 1 structure, pyrophosphate could be retained after chemistry until TL opening. Step **a** and intermediate 2 illustrate one of the two assumptions of our proposed model: the incoming NTP substrate binds the TEC with a closed TH and bound pyrophosphate (step **a**) at the  $i+2$

downstream site. In step **b**, the TEC undergoes reverse isomerization, which is indicated by TL opening, enabling pyrophosphate release. The second assumption of our model is that the complementary NTP at  $i+2$  might induce the TH $\rightarrow$ TL transition, thus opening the active site for pyrophosphate release and freeing the thermal ratchet for translocation. The assumptions about the  $i+2$  NTP binding in the downstream DNA channel and its allosteric effect on active site opening allow us to explain all the experimental effects of the  $i+2$  NTP on transcription elongation and pyrophosphate release. The results demonstrating slow pyrophosphate release, which is stimulated by the next complementary NTP [33], strongly suggest that pyrophosphate release is rate-limiting in elongation. As a further test of this idea, determination of the pyrophosphate release rate should be done in the presence and absence of an allosteric NTP analogue, within the same sequence context in which an allosteric effect of the incoming NTP on transcription elongation was reported [75]. Such an experiment would establish a direct link between the stimulation of pyrophosphate release and stimulation of transcription elongation by the incoming complementary NTP and confirm the rate-limiting nature of reverse isomerization and pyrophosphate release during processive transcription. Ultimately, however, experimental proof of this part of the model will require direct detection of the  $i+2$  NTP in the TEC by crystallographic methods or by fluorescence detection techniques.

It is not clear whether the NTP originally bound to the  $i+2$  site could remain hybridized to the template during transfer to the  $i+1$  site, as proposed in the original NTP-driven translocation model [78], or whether it diffuses to the active site independently from the template base, as proposed in the allosteric model [77]. Our model does not specify the detailed mechanism of NTP entry in the active site, but we propose that this step is not rate-limiting (Figures 7 and 8, step **c**). Note that the estimated rates of the NTP diffusion through different routes were discussed above. Intermediate 4 shows the NTP in

the active site in a pre-insertion conformation (also depicted in Figure 5B) prior to TL→TH isomerization (step **d**). Intermediate 5 shows the catalytic intermediate with a NTP loaded in the insertion site and a closed TH conformation (see also Figure 5A). The isomerized TEC (intermediate 5) is ready for formation of the phosphodiester bond (step **e**), which completes the nucleotide addition cycle of RNAP. Figure 8 shows the complete kinetic pathway describing the relationship between the NTP addition and translocation. It also depicts steps in the cycle that can be rate-limiting for elongation.

The key feature of the proposed model, similar to the NTP-driven translocation and allosteric models, is the functional coupling of template-dependent binding of the substrate NTP with the completion of the previous nucleotide addition cycle (Figure 4, the middle and right-side panels). Such coupling is required to explain the effects of downstream substrate NTPs on transcription kinetics [29,31,32,75,77,78], and to accommodate the dramatic stimulation of pyrophosphate release by the incoming substrate NTP [33]. We hypothesize that the structural basis for cross-talk between a NTP interacting with the template DNA at  $i+2$  and the active site ( $i+1$ ) could be found in the recently proposed TL-TH hypothesis [16] and structural data revealing translocation intermediates [13] (Figure 3).

According to the trigger loop-trigger helices (TL-TH) hypothesis, each nucleotide addition cycle is governed by conformational changes of the TL (Figures 1 and 3). When the substrate NTP is positioned in the active site for catalysis, the TL becomes structurally re-organized into a three-helix bundle with the bridge helix (BH), supporting optimal conformation of the substrate NTP relative to the RNA 3'NMP for catalysis, and this part of the nucleotide addition cycle is verified by structural data [16]. Upon completion of a phosphodiester bond, pyrophosphate is released and template translocation must then occur (Figure 7, step c). The closed, catalytic conformation of the TH, however, does not appear to allow for completion of the nucleotide addition cycle because the closed TH would be expected to suppress

pyrophosphate release and constrain translocation (Figure 7); therefore, after phosphodiester bond formation, it appears that the TH must open to the relaxed TL conformation to support pyrophosphate release and translocation (Figure 7; step b).

Detection of a putative translocation intermediate with a partially open or “wedged” conformation of the TL has shed some light on the mechanism of nucleotide addition cycle completion (Figure 3). Specifically, the structure of TECs observed in the presence of the transcription elongation inhibitor  $\alpha$ -amanitin [13] was compared to the structure of the free TEC [56]. Apparently,  $\alpha$ -amanitin restricts complete unfolding of the TL, and traps it wedged into the center of the BH, distorting the active site conformation (Figure 3, intermediate 2). It was proposed, therefore, that wedging may drive forward translocation of RNAP along the RNA and DNA. Consistent with this view, the nucleic acid scaffold in the TEC containing  $\alpha$ -amanitin was observed in the post-translocated conformation. The template DNA base, however, has been trapped in an intermediate position between the main DNA channel and the active site, hovering right above the BH, with base-pairing groups directed into the main channel. Based on these observations, Brueckner and Cramer proposed that  $\alpha$ -amanitin may inhibit completion of the translocation process [13].

It is necessary to emphasize that putative translocation intermediates were characterized in TECs deprived of substrate NTPs. Therefore, it remains unknown how the product of phosphodiester bond formation (NMP·PPi) affects opening of the TL, translocation and pyrophosphate release. We hypothesize that the reaction product might stabilize the closed conformation of the TH, impeding completion of the nucleotide addition cycle. Notably, mutation of Glu1103 to glycine, which promotes NTP sequestration by the stalled TEC [28], also slows processive RNA synthesis in some sequence contexts (unpublished), arguing that destabilization of the open conformation of the TL might have a

negative impact on processive RNA synthesis. Next, we propose that hybridization of the next substrate NTP to the template base might promote TL opening, and, indeed, the base of the TL is located in the vicinity of the potential downstream NTP binding site. Binding of a NTP at  $i+2$ , therefore, might act as an allosteric effector to stimulate the opening of the TH (Figure 7; steps a-c), thus enabling translocation and pyrophosphate release. Currently, however, it cannot be excluded that NTP binding might stabilize the wedged configuration of the TL and BH, promoting translocation per se.

Furthermore, the fate of the NTP hybridized to the  $i+2$  template DNA base is also not clear. It might be transferred to the active site simultaneously with the template base during translocation, similar to what has been proposed in the latest version of the NTP-driven translocation model. Alternatively, base pairing with the template might be broken during translocation (as was proposed in [75]), in which case the NTP would enter the active site by passive diffusion over the BH, and restore base pairing with the template in the active site. The rapid and efficient sequestration of the substrate NTP, observed in a variety of TECs, suggests that a significant fraction of NTPs, pre-hybridized to the template in the downstream DNA binding channel, is committed to future synthesis. Specifically, rapid quench flow studies [28] indicate that these NTPs do not exchange with the free NTP pool nor do they become accessible to EDTA inactivation.

The present model does not specify which of several steps in the nucleotide addition cycle is affected by NTP hybridization to the template in the downstream DNA binding channel. In principle, it could be reverse isomerization of the closed TH to the open TL, release of pyrophosphate, or forward translocation. Below we discuss the evidence suggesting that reverse isomerization, rather than ratchet translocation of the substrate-free and pyrophosphate-free TEC, is rate-limiting during transcription elongation and is a likely target of the allosteric regulation by the incoming NTP substrate.

**Potential targets for allosteric regulation in the nucleotide addition cycle: reverse isomerization, pyrophosphate release, or translocation?**

Contrary to interpretations of many biochemical and single-molecule experiments and molecular modeling, structural analyses of free RNAPs and TECs did not require that translocation (sliding of RNAP along the RNA and DNA) be rate-limiting for elongation. Indeed, most of the structurally characterized TECs reside in the post-translocated state. Based on constrained dimensions and negative pore electrostatics, Kornberg and co-workers proposed that NTP diffusion and binding, rather than translocation, should be rate-limiting [83]. Furthermore, equilibrium between the pre- and post-translocated state has been observed in at least one TEC crystal by detection of the position of a bromine atom incorporated into the template DNA relative to the protein mass, suggesting that RNAP II freely (and perhaps rapidly) fluctuates between the pre- and post-translocated states even when the TEC is crystallized [13]. To resolve the apparent contradictions between structural and functional data interpretations, we separate the issue of the rate-limiting nature of post-incorporation events during processive transcription and the effect of the ratchet translocation equilibrium on the activity of the TEC deprived of substrate NTPs and pyrophosphate. Below we discuss the experimental evidence arguing against the rate-limiting nature of thermal ratchet equilibrium of the free TEC. It is necessary to emphasize that the data discussed below do not challenge the existence of a post-chemistry rate-limiting step during processive transcription.

**Some mutations that increase transcription elongation rate in RNAP II shift the ratchet translocation equilibrium toward the pre-translocated state**

Characterization of the catalytic properties and translocation equilibrium of two mutant variants of *S. cerevisiae* RNAP II, one carrying a point *rpb1-E1103G* mutation [28], and another lacking the entire

Rpb9 subunit [99], provided additional arguments against any strict correlation between the ability of the stalled TEC to acquire a post-translocated state and to possess a high catalytic activity. Both variants of RNAP II showed an elevated elongation rate both in bulk elongation tests and in pre-steady-state analyses of single phosphodiester bond formation [28,99]. At the same time, the translocation equilibrium of the stalled TECs formed by these “fast” RNAP II variants was distinctly and reproducibly shifted toward the pre-translocated state [28,99]. Taken together, these observations suggest that the correlation between the preferential positioning of the stalled TEC in the pre-translocated state and the decreased catalytic activity of RNAP, reported by Bar-Nahum et al. [11], and used as a foundation of the translocation-centered model, is not universal. In the case of *rpb1-E1103G* and  $\Delta$ *rpb9* RNAP II, the apparent suppression of forward translocation of the stalled TEC does not interfere with an efficient sequestration of the NTP in the RNAP II active site and with rapid catalysis [28,99].

The results of pre-steady-state characterization of *rpb1-E1103G* RNAP II, taken together with the analyses of the translocation equilibrium of the same TEC by Exo III footprinting, suggested some changes in the interpretation of the previously reported data obtained from pre-steady-state analyses of transcription by *E. coli* RNAP [77] and human RNAP II [32,78]. Most of the previous results that were interpreted in terms of the induction of the forward translocation of the stalled TEC by elongation factors or by downstream NTPs were obtained using EDTA as the reaction quench agent [29,32,76-78]. Using EDTA quenching, however, it is the sequestration of the substrate NTP from exchange with the free NTP pool, rather than phosphodiester bond formation that is observed [28,32,79,99,100]. It was initially assumed that stable sequestration of the substrate NTP could only occur in the post-translocated TEC; therefore, the amplitude of the fast (“burst”) fraction observed in EDTA quench was considered a direct measure of the post-translocated fraction in the stalled TEC [32,78]. This

interpretation, however, is not consistent with an increased NTP sequestration by the preferentially pre-translocated *rpb1-E1103G* and *rpb9* RNAP II [28,99]. It appears plausible that moderate activation of the escape from the stall by the downstream NTP or by elongation factors does not involve stimulation of translocation, but rather promotes NTP positioning in the active site after (rapid) translocation that might increase the rate of phosphodiester bond formation. Alternatively, the downstream NTP might stimulate pyrophosphate release, thus suppressing bond reversal which could potentially contribute to the decreased apparent rate of the preceding NTP incorporation.

**Transcriptional pausing is not necessarily associated with stabilization of the pre-translocated state**

The first evidence challenging the view that transcriptional pausing is caused exclusively by the inability of RNAP to forward translocate appeared in a report addressing the TL-TH role in hairpin-dependent transcriptional pausing [65]. RNAP lacking a TL is very slow in elongation. It was observed, moreover, that deletion of the TL minimized the distinction between the paused and non-paused elongation rates; most notably, ratchet translocation of the TEC was not significantly affected by full deletion of the TL. Therefore, Landick and co-workers [65] proposed that the paused state may depend on a distinct conformation of the TL-TH. Furthermore, they suggested the possibility that the paused conformation of the TEC persists after translocation, during NTP binding and catalysis. The kinetic model based on these results better describes the NTP dependence of the escape from the pause, than the classic model, according to which the paused TEC dwells in a pre-translocated state (Supplemental Material in [65] and [101]).

The hypothesis that TECs pause only in a pre-translocated or backtracked state was further challenged by the recent characterization of sequence-specific transcription pauses. Because pauses of

*E. coli* RNAP on the T7 D111 template were predicted to be caused by a translocation block [18,19] or by backtracking [21], Kireeva and Kashlev [70] characterized the translocation states and catalytic activities of *E. coli* RNAP TECs stalled at two prominent pause sites on the T7 D111 template. In contrast to the ratchet model predictions, the translocation state and even the ability to bind the incoming NTP were indistinguishable in paused and active TECs. Evidently, the paused conformation of the TEC determined the catalytic activity at the step of bond formation, rather than at the prior translocation step. In other words, the ability of the TEC to translocate and to bind the incoming NTP does not guarantee a high rate of phosphodiester bond formation because *E. coli* TECs can pause in the post-translocation state in a conformation that is responsive to NTP addition. These observations appear to rule out an obligatory direct correlation between efficient translocation and high catalytic activity. Importantly, even if the TEC is stalled at the pause site for a prolonged time, pausing persists, indicating that pausing is unlikely to occur because of an inefficient release of pyrophosphate. It is likely that isomerization and/or chemistry steps are slow in the paused *E. coli* TEC.

### Summary

Translocation of the TEC has been considered a key step determining the rate of nucleotide addition based on several lines of experimental evidence. We review these data and conclude that during processive RNA synthesis, bond completion, which involves reverse isomerization of the TEC from the closed catalytic to the open state, release of pyrophosphate, and translocation, can be rate-limiting and strongly dependent upon the incoming NTP. Consistent with the previously proposed NTP-driven translocation and allosteric models, we propose a model in which bond completion is stimulated by binding of a substrate NTP hybridized to template in the downstream DNA binding channel (at  $i+2$ ) (Figure 7). At some or many DNA template positions, it appears that a NTP-stimulated conformational

change (reverse isomerization) is rate-limiting for phosphodiester bond completion. We conclude that the role of passive thermal ratchet translocation has probably been overestimated (Figures 2 and 4) in the control of transcription elongation rate. A definitive resolution of the question of whether ratchet translocation can be rate-limiting for elongation awaits direct measurements of translocation rates, by FRET or other biophysical approaches. In the meantime, we propose focusing research efforts on better understanding regulation of biologically relevant translocation during processive RNA synthesis (Figures 7 and 8), occurring in the presence of substrate NTPs and the reaction bi-product pyrophosphate.

#### Acknowledgements

We thank Adam Parks for helpful discussions, and the anonymous reviewers for their insightful comments. The contents of this publication do not necessarily reveal the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial product, or organizations imply endorsement by the U.S. Government. ZFB has support from Michigan State University, The College of Osteopathic Medicine, Michigan State University, and the Initiative in Gene Expression in Development and Disease, Michigan State University.

#### Abbreviations

RNA polymerase (RNAP); ternary elongation complex (TEC); transcription factor RNA polymerase II (TFII); bridge helix (BH); trigger helices (TH); trigger loop (TL); fork loop 2 (FL2); pre-insertion site (P); insertion site (A).

#### Figure Legends

Figure 1. Features of the RNAP TEC. A) *S. cerevisiae* RNAP II TEC. Protein, gray; “clamp”, cyan; nucleic acids, yellow; NTP, red. B) Detailed images of the active site in open (above) and closed (below) TL-TH conformations. DNA template, gray; RNA, pink; NTP, chemistry (C, cyan; N, blue; O, red; P, tan); bridge helix, yellow; TL or TH, cyan. At the right are images of the secondary pore. Protein, cyan; NTP, chemistry;  $Mg^{2+}$ , green.

Figure 2. Thermal ratchet translocation. DNA template, gray; RNA, pink and red;  $i+1$  site, green shading; BH, yellow; TL in the open conformation, cyan. In the pre-translocated state, the 3' end of the RNA (red) resides in the  $i+1$  site. In the post-translocated state, the 3' end of the RNA resides in the  $i$  site. The picture represents a view along the long axis of the BH. The right-side cartoon shows the space-filled representation of the structures of the  $i$  and  $i+1$  sites in the pre-translocated and post-translocated TEC.

Figure 3. Three conformations of the TL. Colors are the same as in Figure 1. Open, “wedged”, and closed TL conformations detected in crystal structures are shown. The conformation of the TL affects the conformation of the BH. A superposition of the three conformations of the BH depicting its increased bending upon folding of the TH is shown on the right.

Figure 4. Detailed kinetic pathways describing alternative models for the NTP addition cycle of RNAPs.

Figure 5. The binding sites for NTPs in the RNAP TEC. All colors are the same as in Figure 1 except the RNA is gray. A) A site: NTP forms a chemistry-ready base pair in the active center ( $i+1$  site). B) P site: NTP forms base pairs in the active center in a tilted conformation with the pyrophosphate moiety misaligned with the 3' OH group of the RNA. This binding precedes transfer of the NTP to the A site. C)  $i+2$  site: NTP base pairs with the  $i+2$  base in the template DNA strand in the primary channel of RNAP. The base pairing occurs in the pre-translocated TEC. Translocation drags the NTP to the active center

together with the  $i+2$  DNA base. D)  $i+2$  site made by  $i+1$  base of DNA: the same as in C except the  $i+1$  base in DNA occupies the flipped out position facing the primary channel in the post-translocated TEC. The incoming NTP hybridizes with the  $i+1$  base in the primary channel and flips into the empty active center. E)  $E$  site: NTP binds in a flipped-out configuration with the pyrophosphate moiety partially occluding the active center. The binding occurs in the post-translocated TEC. There is no base pairing between the NTP and  $i+1$  base in DNA. This configuration was observed in co-crystals of TECs with non-complementary NTPs [15].

Figure 6. A) The “Fork” region of the Rpb2 subunit (465-546 aa, green) makes a substantial part of the proposed NTP entry channel connecting the  $i+2$  and  $i+1$  sites. The FL2 region of the fork (cyan) contains amino acid residues (pink) harboring mutations that affect catalytic activity of RNAP II. B) Surface view of the NTP entry path connecting the primary channel with the active center. The Rpb1 and Rpb2 parts of the path are shown in gray. The path containing the opened FL2 can support unconstrained NTP entry. Other colors are as in Figure 1. The model was generated by replacement of the structure of the FC RNA aptamer in the binary RNAP II/FC RNA complex [84] with the structure of the regular RNA-DNA scaffold from the structure of the regular TEC [8].

Figure 7. A model for processive RNA elongation catalyzed by multi-subunit RNAPs. Colors as in Figure 2. Step a: bond formation results in retention of pyrophosphate in the active center caused by occlusion of the secondary pore by the closed TH. Ratchet translocation is blocked. Step b:  $i+2$  NTP binding stimulates a conformational change leading to TL opening that frees the thermal ratchet for translocation, stimulating pyrophosphate release. Step c: thermal ratchet translocation of the NTP to the  $P$  site. Steps d and e: closing of the TH transfers the NTP to the  $A$  site followed by bond formation.

Figure 8. Detailed kinetic pathway of processive RNA elongation. Steps and intermediates are labeled as in Figure 7.

#### References

- [1] L.A. Allison, M. Moyle, M. Shales, C.J. Ingles, Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases, *Cell* 42 (1985) 599-610.
- [2] J. Huet, R. Schnabel, A. Sentenac, W. Zillig, Archaeobacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type, *EMBO J.* 2 (1983) 1291-1294.
- [3] D. Langer, J. Hain, P. Thuriaux, W. Zillig, Transcription in archaea: similarity to that in eucarya, *Proc. Natl. Acad. Sci. U. S. A* 92 (1995) 5768-5772.
- [4] R.L. Dedrick, C.M. Kane, M.J. Chamberlin, Purified RNA polymerase II recognizes specific termination sites during transcription in vitro, *J. Biol. Chem.* 262 (1987) 9098-9108.
- [5] J. Gelles, R. Landick, RNA polymerase as a molecular motor, *Cell* 93 (1998) 13-16.
- [6] P.H. von Hippel, T.D. Yager, Transcript elongation and termination are competitive kinetic processes, *Proc. Natl. Acad. Sci. U. S. A* 88 (1991) 2307-2311.
- [7] T.D. Yager, P.H. von Hippel, A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*, *Biochemistry* 30 (1991) 1097-1118.
- [8] D. Wang, D.A. Bushnell, K.D. Westover, C.D. Kaplan, R.D. Kornberg, Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis, *Cell* 127 (2006) 941-954.
- [9] A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution, *Science* 292 (2001) 1876-1882.
- [10] G. Zhang, E.A. Campbell, L. Minakhin, C. Richter, K. Severinov, S.A. Darst, Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution, *Cell* 98 (1999) 811-824.
- [11] G. Bar-Nahum, V. Epshtein, A.E. Ruckenstein, R. Rafikov, A. Mustaev, E. Nudler, A ratchet mechanism of transcription elongation and its control, *Cell* 120 (2005) 183-193.
- [12] V. Epshtein, A. Mustaev, V. Markovtsov, O. Bereshchenko, V. Nikiforov, A. Goldfarb, Swing-gate model of nucleotide entry into the RNA polymerase active center, *Mol. Cell* 10 (2002) 623-634.

- [13] F. Brueckner, P. Cramer, Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation, *Nat. Struct. Mol. Biol.* 15 (2008) 811-818.
- [14] L. Tan, S. Wiesler, D. Trzaska, H.C. Carney, R.O. Weinzierl, Bridge helix and trigger loop perturbations generate superactive RNA polymerases, *J. Biol.* 7 (2008) 40.
- [15] K.D. Westover, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center, *Cell* 119 (2004) 481-489.
- [16] D.G. Vassylyev, M.N. Vassylyeva, J. Zhang, M. Palangat, I. Artsimovitch, R. Landick, Structural basis for substrate loading in bacterial RNA polymerase, *Nature* 448 (2007) 163-168.
- [17] E.A. Abbondanzieri, W.J. Greenleaf, J.W. Shaevitz, R. Landick, S.M. Block, Direct observation of base-pair stepping by RNA polymerase, *Nature* 438 (2005) 460-465.
- [18] L. Bai, A. Shundrovsky, M.D. Wang, Sequence-dependent kinetic model for transcription elongation by RNA polymerase, *J. Mol. Biol.* 344 (2004) 335-349.
- [19] L. Bai, R.M. Fulbright, M.D. Wang, Mechanochemical kinetics of transcription elongation, *Phys. Rev. Lett.* 98 (2007) 068103.
- [20] Q. Guo, R. Sousa, Translocation by T7 RNA polymerase: a sensitively poised Brownian ratchet, *J. Mol. Biol.* 358 (2006) 241-254.
- [21] V.R. Tadigotla, O. Maoileidigh, A.M. Sengupta, V. Epshtein, R.H. Ebright, E. Nudler, A.E. Ruckenstein, Thermodynamic and kinetic modeling of transcriptional pausing, *Proc. Natl. Acad. Sci. U. S. A* 103 (2006) 4439-4444.
- [22] D.A. Erie, T.D. Yager, P.H. von Hippel, The single-nucleotide addition cycle in transcription: a biophysical and biochemical perspective, *Annu. Rev. Biophys. Biomol. Struct.* 21 (1992) 379-415.
- [23] P. Cramer, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution, *Science* 292 (2001) 1863-1876.
- [24] D.G. Vassylyev, S. Sekine, O. Laptenko, J. Lee, M.N. Vassylyeva, S. Borukhov, S. Yokoyama, Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, *Nature* 417 (2002) 712-719.
- [25] D.G. Vassylyev, M.N. Vassylyeva, A. Perederina, T.H. Tahirov, I. Artsimovitch, Structural basis for transcription elongation by bacterial RNA polymerase, *Nature* 448 (2007) 157-162.
- [26] Y.W. Yin, T.A. Steitz, The structural mechanism of translocation and helicase activity in T7 RNA polymerase, *Cell* 116 (2004) 393-404.

- [27] C.D. Kaplan, K.M. Larsson, R.D. Kornberg, The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin, *Mol. Cell* 30 (2008) 547-556.
- [28] M.L. Kireeva, Y.A. Nediaklov, G.H. Cremona, Y.A. Purtov, L. Lubkowska, F. Malagon, Z.F. Burton, J.N. Strathern, M. Kashlev, Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation, *Mol. Cell* 30 (2008) 557-566.
- [29] X.Q. Gong, C. Zhang, M. Feig, Z.F. Burton, Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II, *Mol. Cell* 18 (2005) 461-470.
- [30] M. Kireeva, Y.A. Nediaklov, X.Q. Gong, C. Zhang, Y. Xiong, W. Moon, Z.F. Burton, M. Kashlev, Millisecond phase kinetic analysis of elongation catalyzed by human, yeast, and *Escherichia coli* RNA polymerase, *Methods* (2009).
- [31] Y. Xiong, Z.F. Burton, A tunable ratchet driving human RNA polymerase II translocation adjusted by accurately templated nucleoside triphosphates loaded at downstream sites and by elongation factors, *J. Biol. Chem.* 282 (2007) 36582-36592.
- [32] C. Zhang, Z.F. Burton, Transcription factors IIF and IIS and nucleoside triphosphate substrates as dynamic probes of the human RNA polymerase II mechanism, *J. Mol. Biol.* 342 (2004) 1085-1099.
- [33] R.S. Johnson, M. Strausbauch, R. Cooper, J.K. Register, Rapid kinetic analysis of transcription elongation by *Escherichia coli* RNA polymerase, *J. Mol. Biol.* 381 (2008) 1106-1113.
- [34] S.S. Patel, I. Wong, K.A. Johnson, Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant, *Biochemistry* 30 (1991) 511-525.
- [35] J.J. Arnold, C.E. Cameron, Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mg<sup>2+</sup>, *Biochemistry* 43 (2004) 5126-5137.
- [36] V.S. Anand, S.S. Patel, Transient state kinetics of transcription elongation by T7 RNA polymerase, *J. Biol. Chem.* 281 (2006) 35677-35685.
- [37] B. Marchand, E.P. Tchesnokov, M. Gotte, The pyrophosphate analogue foscarnet traps the pre-translocational state of HIV-1 reverse transcriptase in a Brownian ratchet model of polymerase translocation, *J. Biol. Chem.* 282 (2007) 3337-3346.
- [38] N. Komissarova, M. Kashlev, Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded, *Proc. Natl. Acad. Sci. U. S. A* 94 (1997) 1755-1760.
- [39] T.C. Reeder, D.K. Hawley, Promoter proximal sequences modulate RNA polymerase II elongation by a novel mechanism, *Cell* 87 (1996) 767-777.

- [40] N. Komissarova, M. Kashlev, RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA, *J. Biol. Chem.* 272 (1997) 15329-15338.
- [41] G.H. Feng, D.N. Lee, D. Wang, C.L. Chan, R. Landick, GreA-induced transcript cleavage in transcription complexes containing *Escherichia coli* RNA polymerase is controlled by multiple factors, including nascent transcript location and structure, *J. Biol. Chem.* 269 (1994) 22282-22294.
- [42] F. Brueckner, U. Hennecke, T. Carell, P. Cramer, CPD damage recognition by transcribing RNA polymerase II, *Science* 315 (2007) 859-862.
- [43] H. Kettenberger, K.J. Armache, P. Cramer, Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS, *Mol. Cell* 16 (2004) 955-965.
- [44] K.D. Westover, D.A. Bushnell, R.D. Kornberg, Structural basis of transcription: separation of RNA from DNA by RNA polymerase II, *Science* 303 (2004) 1014-1016.
- [45] R.J. Davenport, G.J. Wuite, R. Landick, C. Bustamante, Single-molecule study of transcriptional pausing and arrest by *E. coli* RNA polymerase, *Science* 287 (2000) 2497-2500.
- [46] M.D. Wang, M.J. Schnitzer, H. Yin, R. Landick, J. Gelles, S.M. Block, Force and velocity measured for single molecules of RNA polymerase, *Science* 282 (1998) 902-907.
- [47] H. Yin, M.D. Wang, K. Svoboda, R. Landick, S.M. Block, J. Gelles, Transcription against an applied force, *Science* 270 (1995) 1653-1657.
- [48] L. Bai, T.J. Santangelo, M.D. Wang, Single-molecule analysis of RNA polymerase transcription, *Annu. Rev. Biophys. Biomol. Struct.* 35 (2006) 343-360.
- [49] A. Shundrovsky, T.J. Santangelo, J.W. Roberts, M.D. Wang, A single-molecule technique to study sequence-dependent transcription pausing, *Biophys. J.* 87 (2004) 3945-3953.
- [50] E.A. Galburt, S.W. Grill, A. Wiedmann, L. Lubkowska, J. Choy, E. Nogales, M. Kashlev, C. Bustamante, Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner, *Nature* 446 (2007) 820-823.
- [51] Y.X. Mejia, H. Mao, N.R. Forde, C. Bustamante, Thermal probing of *E. coli* RNA polymerase off-pathway mechanisms, *J. Mol. Biol.* 382 (2008) 628-637.
- [52] J. Gelles, H. Yin, L. Finzi, O.K. Wong, R. Landick, Single-molecule kinetic studies on DNA transcription and transcriptional regulation, *Biophys. J.* 68 (1995) 73S.
- [53] K.M. Herbert, P.A. La, B.J. Wong, R.A. Mooney, K.C. Neuman, R. Landick, S.M. Block, Sequence-resolved detection of pausing by single RNA polymerase molecules, *Cell* 125 (2006) 1083-1094.

- [54] K.C. Neuman, E.A. Abbondanzieri, R. Landick, J. Gelles, S.M. Block, Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking, *Cell* 115 (2003) 437-447.
- [55] K.M. Herbert, W.J. Greenleaf, S.M. Block, Single-molecule studies of RNA polymerase: motoring along, *Annu. Rev. Biochem.* 77 (2008) 149-176.
- [56] D. Wang, D.A. Bushnell, X. Huang, K.D. Westover, M. Levitt, R.D. Kornberg, Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution, *Science* 324 (2009) 1203-1206.
- [57] I. Artsimovitch, R. Landick, Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals, *Proc. Natl. Acad. Sci. U. S. A* 97 (2000) 7090-7095.
- [58] M. Palangat, R. Landick, Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II, *J. Mol. Biol.* 311 (2001) 265-282.
- [59] P.J. Hawryluk, A. Ujvari, D.S. Luse, Characterization of a novel RNA polymerase II arrest site which lacks a weak 3' RNA-DNA hybrid, *Nucleic Acids Res.* 32 (2004) 1904-1916.
- [60] C.L. Chan, R. Landick, Dissection of the his leader pause site by base substitution reveals a multipartite signal that includes a pause RNA hairpin, *J. Mol. Biol.* 233 (1993) 25-42.
- [61] C.L. Chan, R. Landick, The Salmonella typhimurium his operon leader region contains an RNA hairpin-dependent transcription pause site. Mechanistic implications of the effect on pausing of altered RNA hairpins, *J. Biol. Chem.* 264 (1989) 20796-20804.
- [62] I. Touloukhonov, R. Landick, The flap domain is required for pause RNA hairpin inhibition of catalysis by RNA polymerase and can modulate intrinsic termination, *Mol. Cell* 12 (2003) 1125-1136.
- [63] S. Kyzer, K.S. Ha, R. Landick, M. Palangat, Direct versus limited-step reconstitution reveals key features of an RNA hairpin-stabilized paused transcription complex, *J. Biol. Chem.* 282 (2007) 19020-19028.
- [64] I. Touloukhonov, I. Artsimovitch, R. Landick, Allosteric control of RNA polymerase by a site that contacts nascent RNA hairpins, *Science* 292 (2001) 730-733.
- [65] I. Touloukhonov, J. Zhang, M. Palangat, R. Landick, A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing, *Mol. Cell* 27 (2007) 406-419.
- [66] R. Landick, The regulatory roles and mechanism of transcriptional pausing, *Biochem. Soc. Trans.* 34 (2006) 1062-1066.
- [67] R. Guajardo, R. Sousa, A model for the mechanism of polymerase translocation, *J. Mol. Biol.* 265 (1997) 8-19.

- [68] M.L. Kireeva, N. Komissarova, D.S. Waugh, M. Kashlev, The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex, *J. Biol. Chem.* 275 (2000) 6530-6536.
- [69] I. Sidorenkov, N. Komissarova, M. Kashlev, Crucial role of the RNA:DNA hybrid in the processivity of transcription, *Mol. Cell* 2 (1998) 55-64.
- [70] M.L. Kireeva, M. Kashlev, Mechanism of sequence-specific pausing of bacterial RNA polymerase, *Proc. Natl. Acad. Sci. U. S. A* (2009).
- [71] S. Tuske, S.G. Sarafianos, X. Wang, B. Hudson, E. Sineva, J. Mukhopadhyay, J.J. Birktoft, O. Leroy, S. Ismail, A.D. Clark, Jr., C. Dharia, A. Napoli, O. Laptenko, J. Lee, S. Borukhov, R.H. Ebright, E. Arnold, Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation, *Cell* 122 (2005) 541-552.
- [72] B. Marchand, M. Gotte, Site-specific footprinting reveals differences in the translocation status of HIV-1 reverse transcriptase. Implications for polymerase translocation and drug resistance, *J. Biol. Chem.* 278 (2003) 35362-35372.
- [73] E. Kashkina, M. Anikin, T.H. Tahirov, S.N. Kochetkov, D.G. Vassilyev, D. Temiakov, Elongation complexes of *Thermus thermophilus* RNA polymerase that possess distinct translocation conformations, *Nucleic Acids Res.* 34 (2006) 4036-4045.
- [74] F. Brueckner, J. Ortiz, P. Cramer, A movie of the RNA polymerase nucleotide addition cycle, *Curr. Opin. Struct. Biol.* 19 (2009) 294-299.
- [75] J.E. Foster, S.F. Holmes, D.A. Erie, Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation, *Cell* 106 (2001) 243-252.
- [76] X.Q. Gong, Y.A. Nediaklov, Z.F. Burton, Alpha-amanitin blocks translocation by human RNA polymerase II, *J. Biol. Chem.* 279 (2004) 27422-27427.
- [77] S.F. Holmes, D.A. Erie, Downstream DNA sequence effects on transcription elongation. Allosteric binding of nucleoside triphosphates facilitates translocation via a ratchet motion, *J. Biol. Chem.* 278 (2003) 35597-35608.
- [78] Y.A. Nediaklov, X.Q. Gong, S.L. Hovde, Y. Yamaguchi, H. Handa, J.H. Geiger, H. Yan, Z.F. Burton, NTP-driven translocation by human RNA polymerase II, *J. Biol. Chem.* 278 (2003) 18303-18312.
- [79] M. Kireeva, Y.A. Nediaklov, X.Q. Gong, C. Zhang, Y. Xiong, W. Moon, Z.F. Burton, M. Kashlev, Millisecond phase kinetic analysis of elongation catalyzed by human, yeast, and *Escherichia coli* RNA polymerase, *Methods* 48 (2009) 333-345.

- [80] Z.F. Burton, M. Feig, X.Q. Gong, C. Zhang, Y.A. Nedialkov, Y. Xiong, NTP-driven translocation and regulation of downstream template opening by multi-subunit RNA polymerases, *Biochem. Cell Biol.* 83 (2005) 486-496.
- [81] C. Zhang, H. Yan, Z.F. Burton, Combinatorial control of human RNA polymerase II (RNAP II) pausing and transcript cleavage by transcription factor IIF, hepatitis delta antigen, and stimulatory factor II, *J. Biol. Chem.* 278 (2003) 50101-50111.
- [82] C. Zhang, K.L. Zobeck, Z.F. Burton, Human RNA polymerase II elongation in slow motion: role of the TFIIF RAP74 alpha1 helix in nucleoside triphosphate-driven translocation, *Mol. Cell Biol.* 25 (2005) 3583-3595.
- [83] N.N. Batada, K.D. Westover, D.A. Bushnell, M. Levitt, R.D. Kornberg, Diffusion of nucleoside triphosphates and role of the entry site to the RNA polymerase II active center, *Proc. Natl. Acad. Sci. U. S. A* 101 (2004) 17361-17364.
- [84] H. Kettenberger, A. Eisenfuhr, F. Brueckner, M. Theis, M. Famulok, P. Cramer, Structure of an RNA polymerase II-RNA inhibitor complex elucidates transcription regulation by noncoding RNAs, *Nat. Struct. Mol. Biol.* 13 (2006) 44-48.
- [85] D. Temiakov, V. Patlan, M. Anikin, W.T. McAllister, S. Yokoyama, D.G. Vassylyev, Structural basis for substrate selection by t7 RNA polymerase, *Cell* 116 (2004) 381-391.
- [86] M. Thomas, S. Chedin, C. Carles, M. Riva, M. Famulok, A. Sentenac, Selective targeting and inhibition of yeast RNA polymerase II by RNA aptamers, *J. Biol. Chem.* 272 (1997) 27980-27986.
- [87] J. Fu, A.L. Gnatt, D.A. Bushnell, G.J. Jensen, N.E. Thompson, R.R. Burgess, P.R. David, R.D. Kornberg, Yeast RNA polymerase II at 5 Å resolution, *Cell* 98 (1999) 799-810.
- [88] S. Tornaletti, P.C. Hanawalt, Effect of DNA lesions on transcription elongation, *Biochimie* 81 (1999) 139-146.
- [89] D.E. Lewis, N. Komissarova, P. Le, M. Kashlev, S. Adhya, DNA sequences in gal operon override transcription elongation blocks, *J. Mol. Biol.* 382 (2008) 843-858.
- [90] D.E. Syroid, J.P. Capone, RNA chain elongation and termination by mammalian RNA polymerase III. Analysis of tRNA gene transcription by imposing a reversible factor-mediated block to elongation using a sequence-specific DNA binding protein, *J. Mol. Biol.* 244 (1994) 482-493.
- [91] P.A. Pavco, D.A. Steege, Elongation by *Escherichia coli* RNA polymerase is blocked in vitro by a site-specific DNA binding protein, *J. Biol. Chem.* 265 (1990) 9960-9969.
- [92] H. Saeki, J.Q. Svejstrup, Stability, flexibility, and dynamic interactions of colliding RNA polymerase II elongation complexes, *Mol. Cell* 35 (2009) 191-205.

- [93] E. Zaychikov, L. Denissova, H. Heumann, Translocation of the Escherichia coli transcription complex observed in the registers 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble", Proc. Natl. Acad. Sci. U. S. A 92 (1995) 1739-1743.
- [94] F.C. Holstege, U. Fiedler, H.T. Timmers, Three transitions in the RNA polymerase II transcription complex during initiation, EMBO J. 16 (1997) 7468-7480.
- [95] E. Kashkina, M. Anikin, F. Brueckner, E. Lehmann, S.N. Kochetkov, W.T. McAllister, P. Cramer, D. Temiakov, Multisubunit RNA polymerases melt only a single DNA base pair downstream of the active site, J. Biol. Chem. 282 (2007) 21578-21582.
- [96] S. Naji, M.G. Bertero, P. Spitalny, P. Cramer, M. Thomm, Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement, Nucleic Acids Res. 36 (2008) 676-687.
- [97] C. Domecq, M. Kireeva, J. Archambault, M. Kashlev, B. Coulombe, Z.F. Burton, Site-directed mutagenesis, purification and assay of Saccharomyces cerevisiae RNA polymerase II, Protein Expr. Purif. 69 (2010) 83-90.
- [98] D.L. Pappas, Jr., M. Hampsey, Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in Saccharomyces cerevisiae, Mol. Cell Biol. 20 (2000) 8343-8351.
- [99] C. Walmacq, M.L. Kireeva, J. Irvin, Y. Nedialkov, L. Lubkowska, F. Malagon, J.N. Strathern, M. Kashlev, Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II, J. Biol. Chem. 284 (2009) 19601-19612.
- [100] J.J. Arnold, D.W. Gohara, C.E. Cameron, Poliovirus RNA-dependent RNA polymerase (3Dpol): pre-steady-state kinetic analysis of ribonucleotide incorporation in the presence of Mn<sup>2+</sup>, Biochemistry 43 (2004) 5138-5148.
- [101] R. Landick, Transcriptional pausing without backtracking, Proc. Natl. Acad. Sci. U. S. A 106 (2009) 8797-8798.

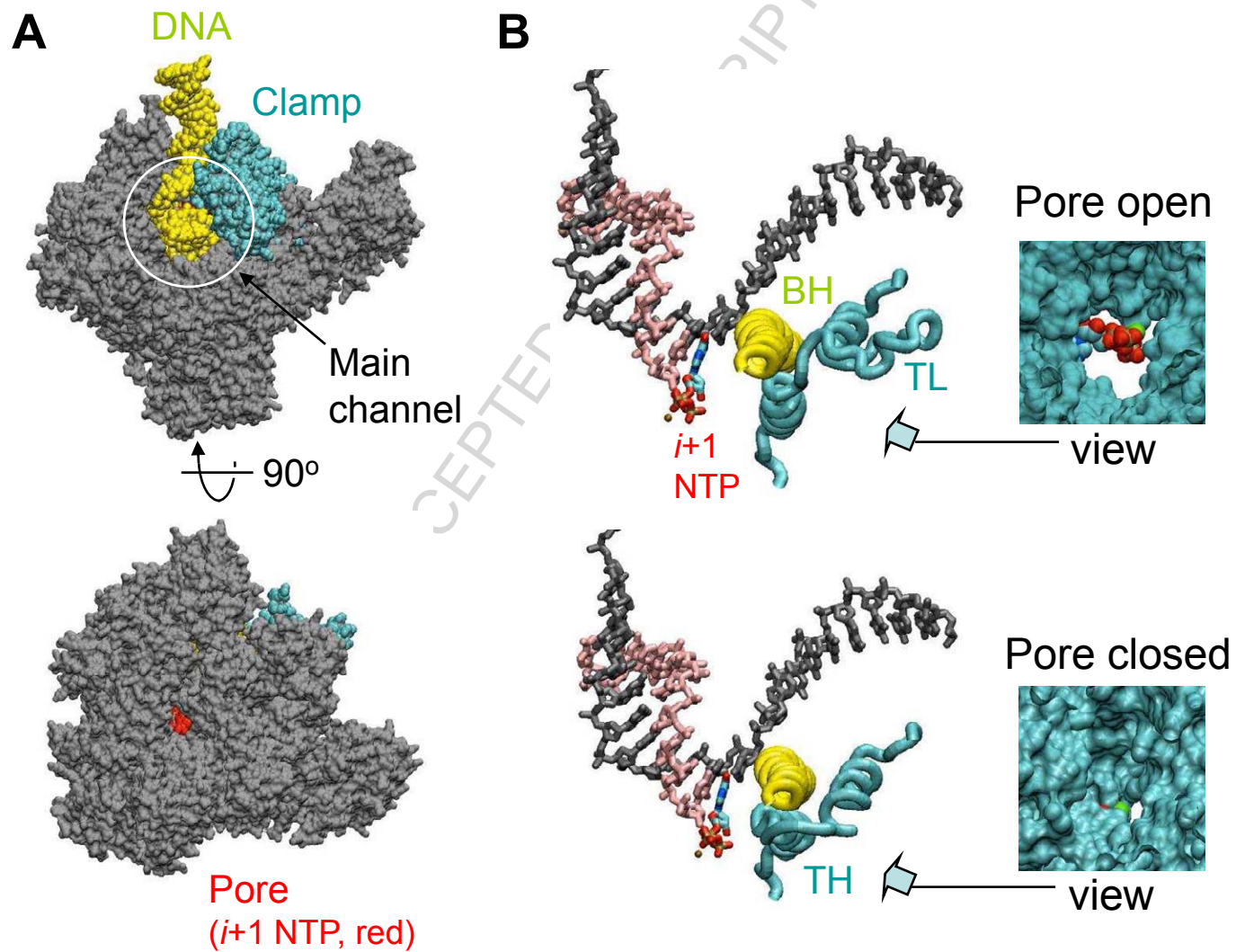


Fig. 1

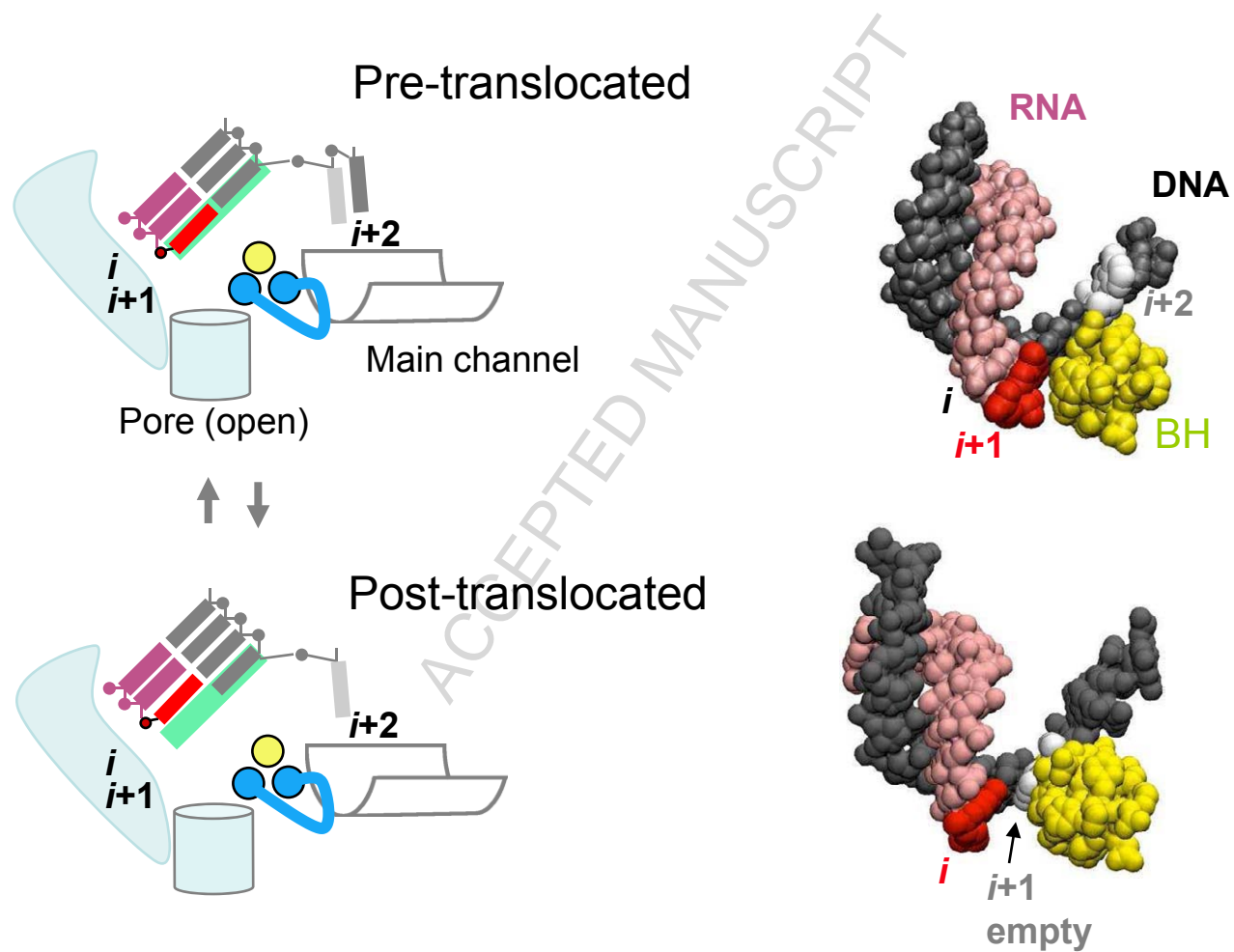


Fig. 2

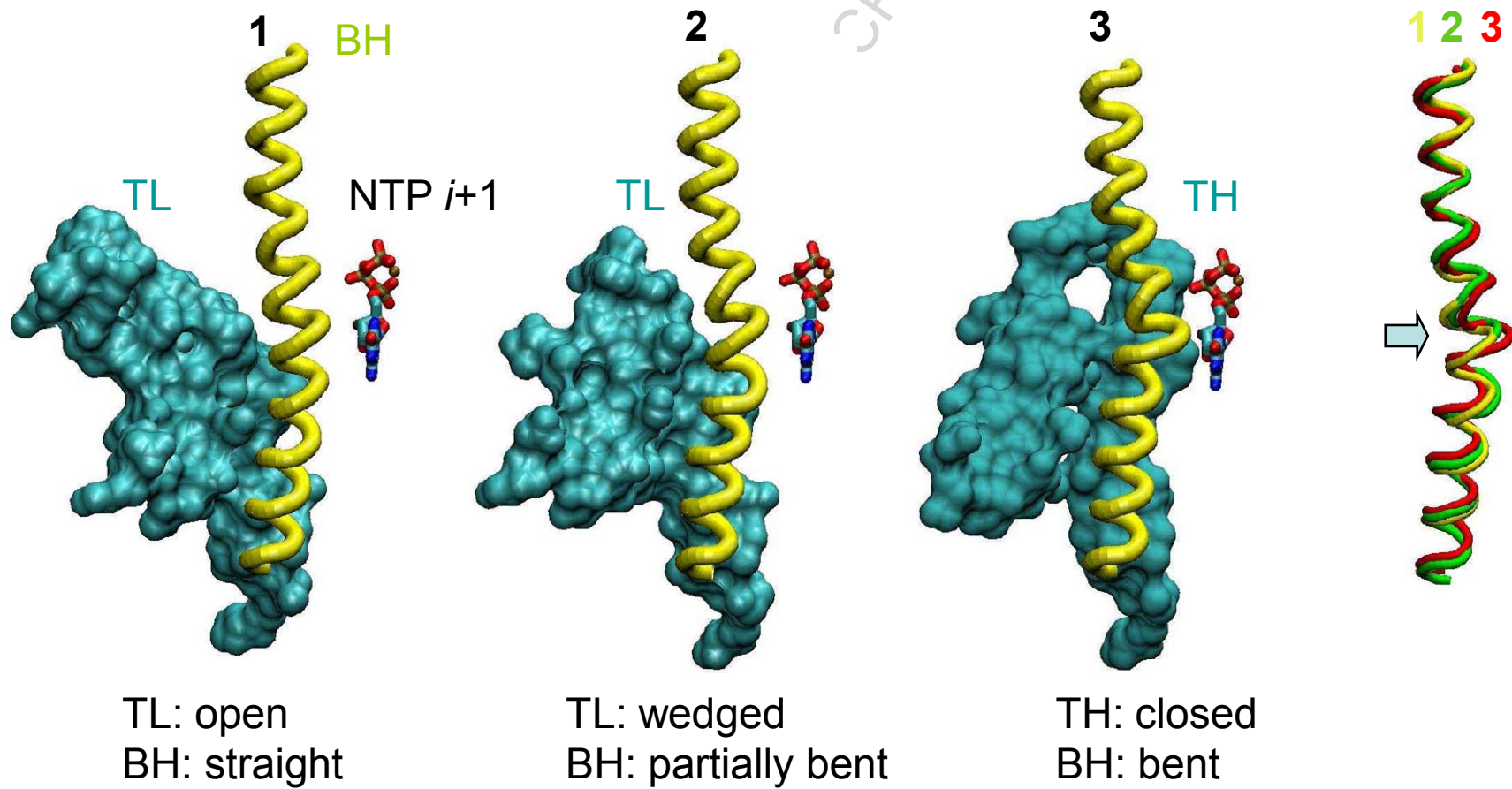
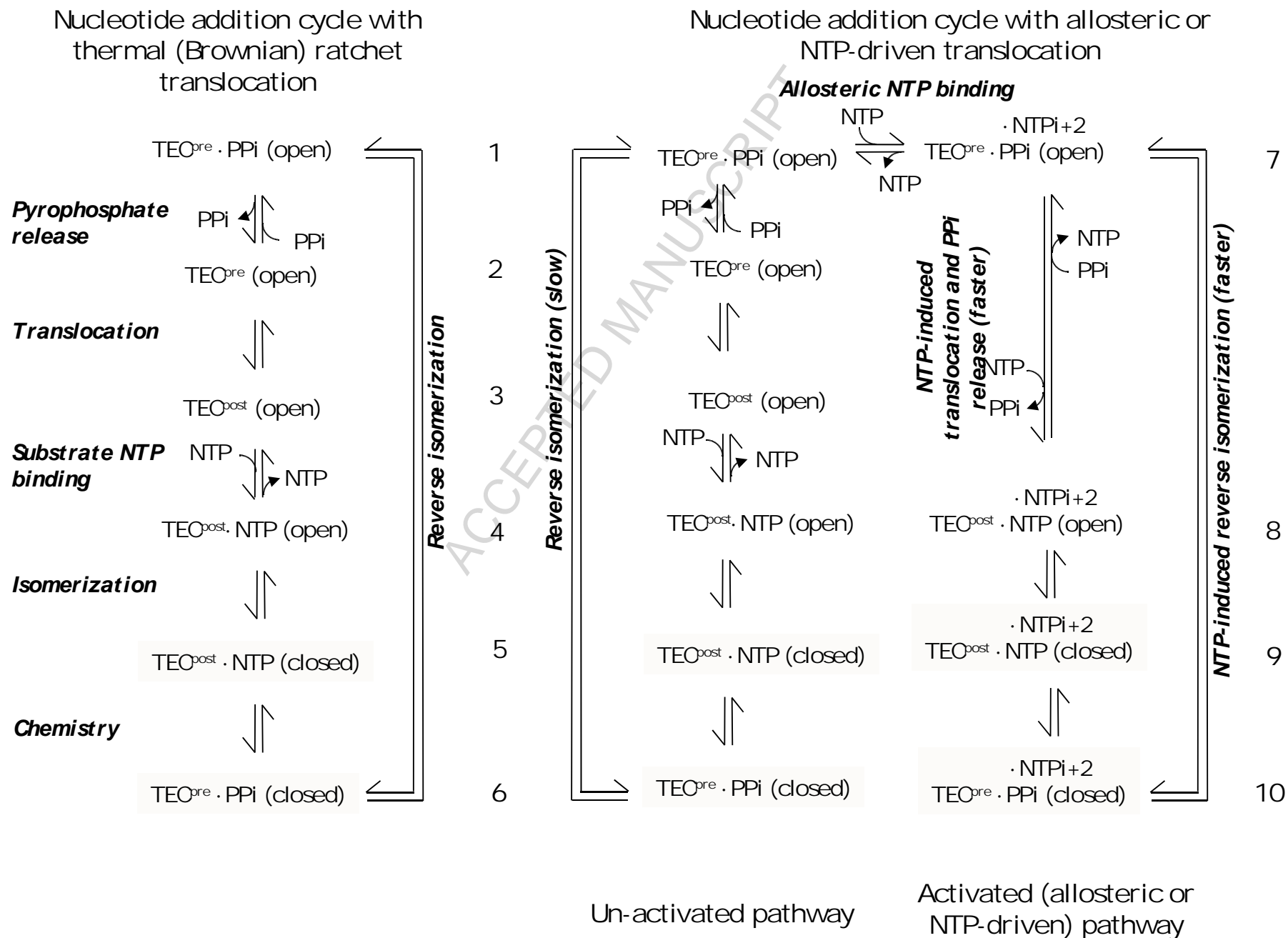


Fig. 3



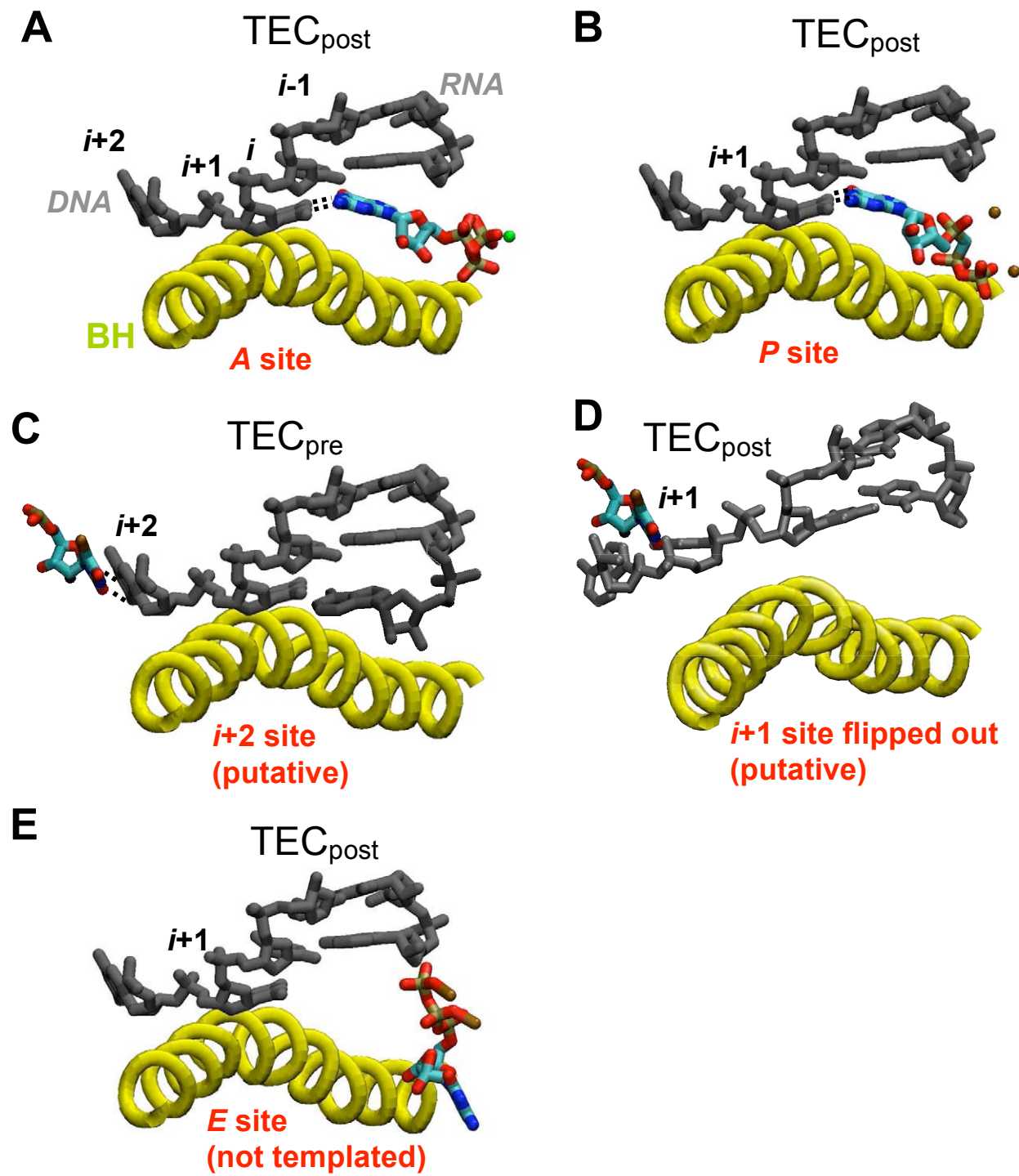


Fig. 5

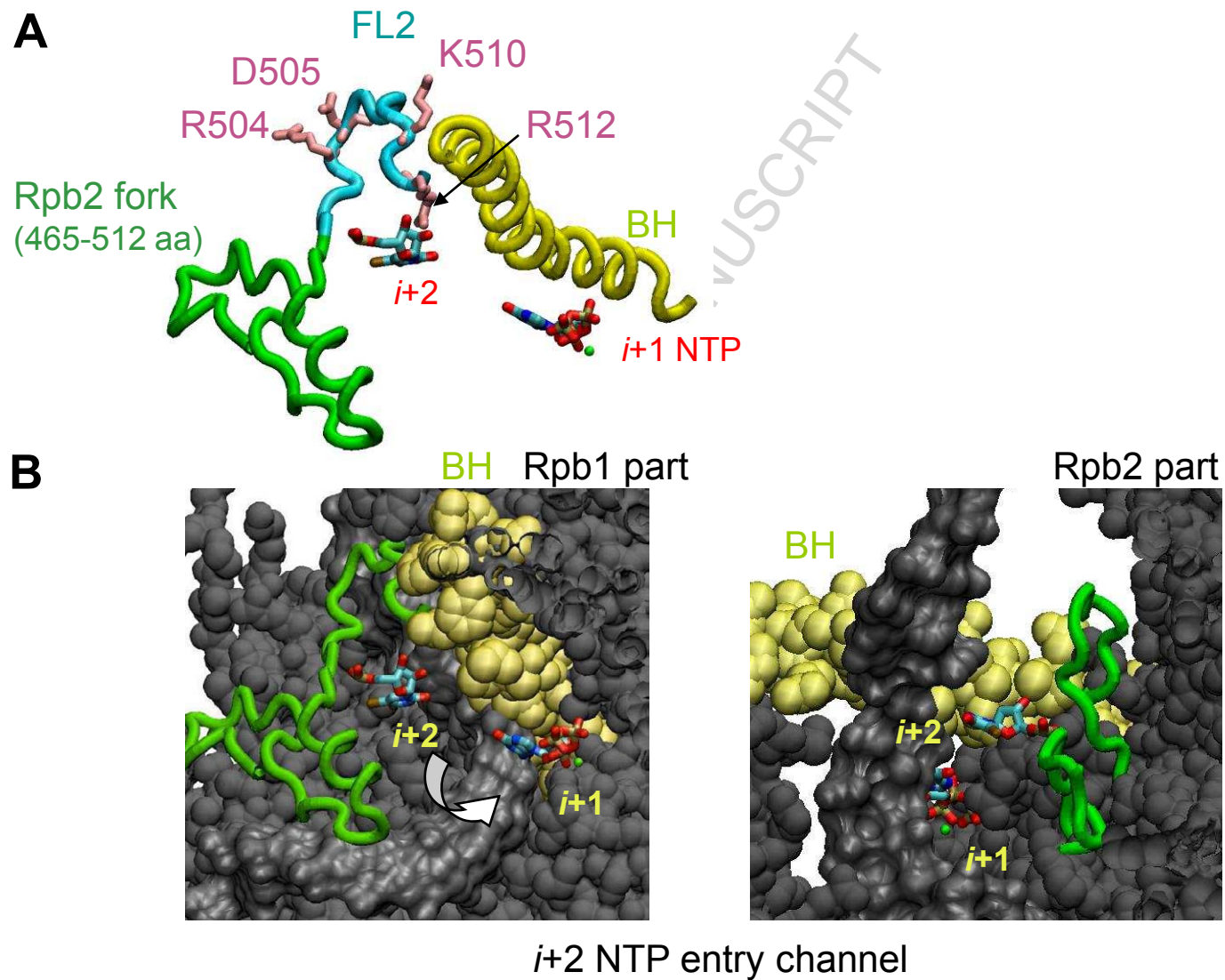


Fig. 6

Figure\_7

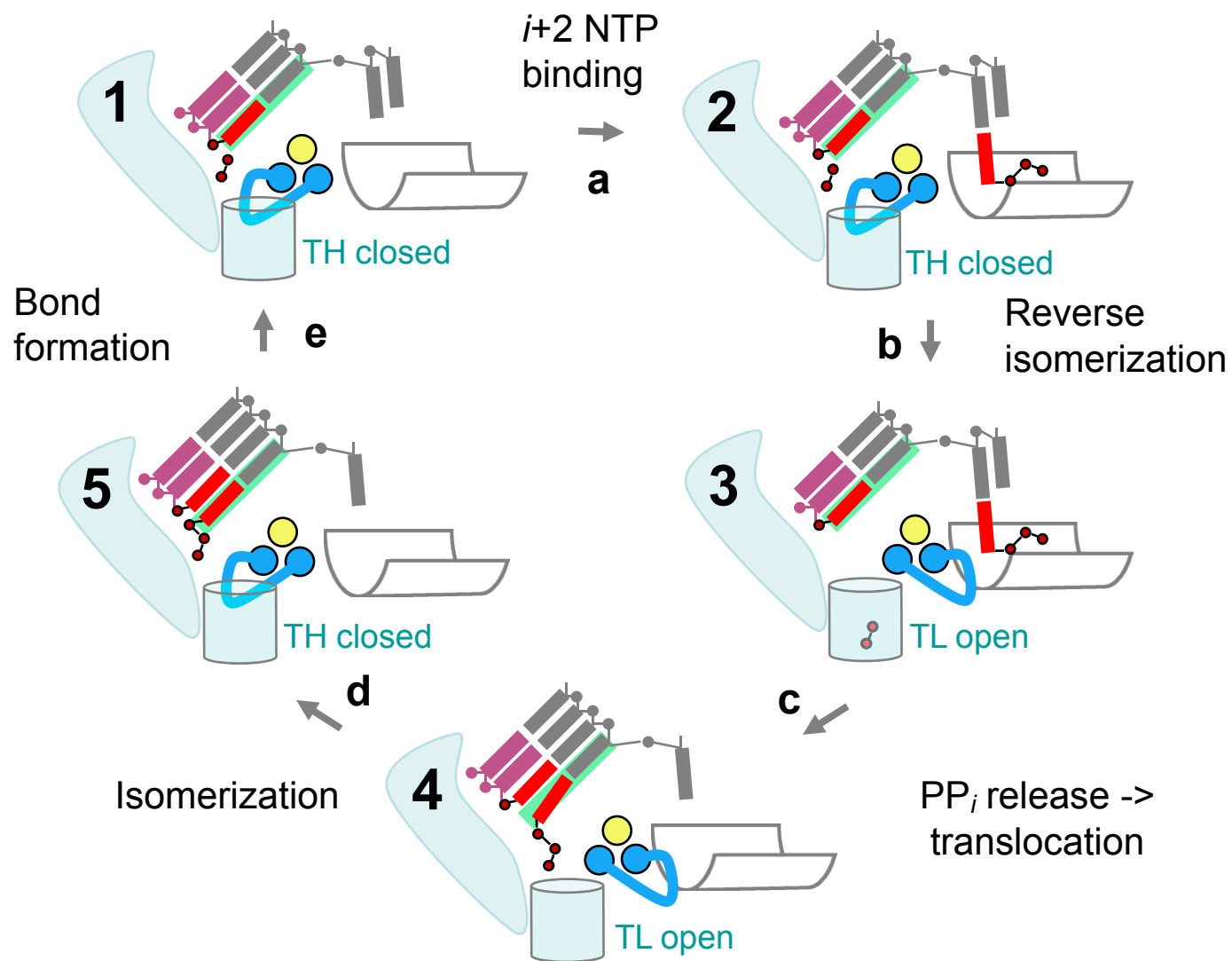


Fig. 7

