Published in final edited form as:

Biochim Biophys Acta. 2013 January ; 1829(1): 39–54. doi:10.1016/j.bbagr.2012.09.007.

Basic Mechanisms of RNA Polymerase II Activity and Alteration of Gene Expression in *Saccharomyces cerevisiae*

Craig D. Kaplan^{a,§}

^aDepartment of Biochemistry and Biophysics, Texas A&M University, TAMU 2128, College Station, TX, 77843-2128, Telephone (979) 845-0429, Fax (979) 862-4718

Abstract

Transcription by RNA Polymerase II (Pol II), and all RNA polymerases for that matter, may be understood as comprising two cycles. The first cycle relates to the basic mechanism of the transcription process wherein Pol II must select the appropriate nucleoside triphosphate (NTP) substrate complementary to the DNA template, catalyze phosphodiester bond formation, and translocate to the next position on the DNA template. Performing this cycle in an iterative fashion allows the synthesis of RNA chains that can be over one million nucleotides in length in some larger eukaryotes. Overlaid upon this enzymatic cycle, transcription may be divided into another cycle of three phases: initiation, elongation, and termination. Each of these phases has a large number of associated transcription factors that function to promote or regulate the gene expression process. Complicating matters, each phase of the latter transcription cycle are coincident with cotranscriptional RNA processing events. Additionally, transcription takes place within a highly dynamic and regulated chromatin environment. This chromatin environment is radically impacted by active transcription and associated chromatin modifications and remodeling, while also functioning as a major platform for Pol II regulation. This review will focus on our basic knowledge of the Pol II transcription mechanism, and how altered Pol II activity impacts gene expression *in vivo* in the model eukaryote *Saccharomyces cerevisiae*.

Keywords

RNA Polymerase II; transcription elongation; yeast; gene expression; genetics

1. Pol II is a highly conserved molecular machine

Pol II is a member of a family of multisubunit RNA polymerases (msRNAPs) that are highly conserved in structure (Table 1). This family includes Pol I and Pol III in eukaryotes, additionally Pol IV and Pol V in plants (Pol II homologs with roles dedicated to gene

© 2012 Elsevier B.V. All rights reserved.

[§]To whom correspondence should be addressed: cdkaplan@tamu.edu, 322A Biochemistry, TAMU 2128, College Station, TX, 77843-2128.

Publisher's Disclaimer: 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 citable 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.

silencing [1-3]), and bacterial and archaeal RNA polymerases (RNAPs) [4]. We have structural information on this enzyme family from all domains of life, *Saccharomyces cerevisiae* (extensive structures, with revolutionary structures from the Kornberg group [5, 6], and a large number subsequently from Kornberg and Cramer groups, please see recent structural movie from Cramer's group [7]) and *Schizosaccharomyces pombe* [8] Pol II enzymes from eukaryotes, *Sulfolobus solfataricus* RNAP [9] from archaea, and *Thermus aquaticus* [10-12] and *Thermus thermophilus* [13-15] RNAPs from bacteria. An increasing complement of msRNAP structures with different nucleic acids, substrates and transcription factors are adding to our understanding of how these enzymes work and how they might be regulated (Table I). Core structural scaffolds of msRNAPs from all domains of life are highly similar, regardless of total subunit numbers, with archaeal and eukaryotic enzymes being most similar as would be expected from their evolutionary relationship. Additional subunits found in eukaryotic family members appear to be added over evolutionary time to a highly conserved structure that has remained essentially unchanged between extant eukaryotes and archaea [9]. A major difference between most Pol II enzymes from all other msRNAPs is the C-terminal domain (CTD) of the largest subunit, comprising a collection of heptapeptide repeats with multiple residues capable of being post-translationally modified. The CTD functions as a highly dynamic regulatory scaffold for the recruitment of a large number of factors, from RNA processing to chromatin modifying, at different stages of the transcription process. The CTD will not be discussed further here, but see recent papers and reviews on CTD modification and function for more [16-20].

2. Substrate selection and catalysis by Pol II

2.1 Basic organization of the Pol II elongation complex

All data presented for msRNAPs are consistent with use of a two-metal mechanism for synthesis of RNA chains, similar to other cellular nucleic acid polymerases [21, 22]. Focusing on the Pol II structure, the conserved architecture of the enzyme comprises a “crab claw” shape with pincers formed by the two largest protein subunits, Rpb1 and Rpb2 (Figure 1). These two subunits also comprise the vast majority of nucleic acid interaction surfaces of Pol II along with the entirety of the active site. A productive elongation complex contains ~9 nucleotides (nt) of synthesized RNA hybridized to the DNA template (the RNA-DNA hybrid), downstream double stranded DNA, a region of non-template strand DNA, and has the ability to separate RNA transcript from the DNA template allowing template and non-template strands to anneal. The region of DNA melting with or without RNA is referred to as the transcription bubble, and encompasses the region of the nascent transcript that upon synthesis remains bound to the template.

Crystal structures of directly assembled elongation complexes (ECs) for *S. cerevisiae* Pol II [23-25] and *T. thermophilus* RNAP [13, 14] inform our understanding of critical substrate interactions for substrate discrimination and catalysis. It is important to note that all EC-substrate crystal structures are compromised in one fashion or another to preclude phosphodiester bond formation. The two methods for reaction inhibition are use of a chain-terminating 3'-dNMP at the 3' end of the RNA substrate or use of a non-hydrolysable substrate analog, such as AMPCPP. There are multiple, non-mutually exclusive models for

substrate association with Pol II prior to catalysis. First, Pol II has the ability to non-specifically interact with NTP substrates in an orientation nonproductive for catalysis through binding in a crystallographically supported “entry” (E) site [24], though of course this site might also represent a step in the evacuation of mismatched substrates. Second, the location of a templated, non-hydrolysable NTP analog present in a conformation not thought to efficiently support catalysis has been interpreted as a pre-insertion (PI) site [25]. Third biochemical data from kinetic analyses on RNAP and Pol II examining next-nucleotide-dependent substrate effects on transcription have been interpreted to support pre-loading of substrates base-pairing with the DNA template in positions downstream of the active site [26-28]. Binding of substrates in downstream templated positions may allosterically affect msRNAP activity without necessarily transferring those exact substrates to the active site. Though the mode of substrate entry in the Pol II active site poses a very difficult question to answer definitively, it seems likely that the biochemical environment of the Pol II active site may support association of NTP(s) in a number of conformations. Critical contacts to the NTP substrate with the triphosphate moiety of the nucleotide to be added are made through Mg^{2+} ions that are essential for catalysis. Mg^{2+} ions are coordinated through an essential triad of glutamic acid residues, with likely contribution from the 3'-OH of the nascent RNA. The deprotonation of the 3'-OH is another critical role for Mg^{2+} in the phosphodiester bond formation reaction.

2.2 Pol II-substrate interactions

Pol II and other msRNAPs function as Brownian ratchets (discussed below, this section) wherein thermal motions allow the enzymes to move forwards and backwards relative to the sequence of the DNA template [29, 30]. During elongation, the translocation state of the Pol II determines whether the active site is available for incorporation of the next NMP into the nascent RNA (Figure 2). Upon phosphodiester bond formation, the NMP added to the nascent chain occupies the position of catalysis as the product of the reaction (Figure 2A). Upon translocation, the movement of the enzyme one base along the DNA template allows incoming NTPs to survey the DNA template for complementary base-pairing interactions (Figure 2B-C). Figure 2B illustrates non-specific association of substrate in E-site prior to entry into templated conformation, but is not necessarily a required event, and substrates may arrive from elsewhere (as discussed in 2.1). Several steps ensure faithful addition of correct substrates. First, there are interactions between active site residues such as Rpb1 N479 [23] (analogous residues tested biochemically in *E. coli* RNAP [31], *T. thermophilus* RNAP [32]) and Rpb1 R446 (an essential residue in Pol II, but identified through structural analyses [23, 24]) and 2' and 3' hydroxyl (OH) groups on the NTP ribose moiety. Pol II must select against naturally occurring 2'-dNTPs and does so efficiently (likely to be 1000-10000 fold selection *in vivo*, taking into account catalytic inefficiency with 2'-dNTPs coupled with competition from NTPs and lower levels of 2'-dNTPs *in vivo*). However, to the best of my knowledge, the *in vivo* incorporation rate of 2'-dNTPs into RNA has not been ascertained.

Consistent with recognition of the 2'-OH being important for efficient catalysis of NTPs over 2'-dNTPs, an *E. coli* RNAP EC with a 2'-deoxy 3'-NMP is less stably pre-translocated than the same EC with a 2'-OH 3'-NMP, illustrating functional importance of active site

contacts between the 2'-OH for both NTP substrates and RNA (product of the reaction) [33]. Second, basepairing between substrate and template DNA contributes to stabilization of NTP in the Pol II active site, and it is likely that basepairing cooperates with Pol II-substrate 2'- and 3'-OH interactions to promote active site occupancy. The contribution of substrate basepairing to fidelity has recently been examined elegantly by the Wang and Kool groups, through the use of NTP substrate analogs [34]. They show that basepairing hydrogen bonds, and not merely shape of basepairs, are critical for efficient catalysis. Finally, movement of a mobile Pol II active site loop, the "trigger loop" (TL), brings a suite of substrate-interacting amino acid sidechains into contact with basepaired NTP and these TL-NTP interactions are required for efficient catalysis and substrate selection (Pol II studies [23, 35-37], *E. coli* RNAP studies [13, 30, 38, 39], *T. aquaticus* RNAP studies [32, 40, 41], *M. jannaschii* RNAP studies [42]). TL movement comprises a method for kinetic selection for matched substrates, analogous to movement of "fingers" domains of DNA polymerases, rather than greatly contributing to Pol II active site-substrate affinity (Figure 2D-F). Kinetic selection of substrates allows multiple substrates to associate with Pol II (matched or mismatched) with similar affinity yet have catalysis vastly favor appropriately templated substrates. Discrimination solely based on affinity instead would require distinct active site conformations for any particular substrate and more binding energy than that provided by complementary basepairing interactions. The coupling of TL-NTP contacts that support catalysis to conformation states much more likely to be populated by matched NTP substrates allows the TL to be a major component of Pol II fidelity.

2.3 Multiple conformation states and movement of the Pol II trigger loop

How does the TL fold or move into substrate-interacting conformations? While molecular dynamics simulations are able to simulate amino acid conformations and limited movement, description of the full TL movement from structurally observed endpoints is not yet possible on simulation timescales. Time-resolved single molecule FRET studies might represent a fruitful approach to describing TL dynamics and their timescales, but would require modification of the TL and would be technically challenging. Examination of TL conformations from a number of crystal structures coupled with genetic analyses of TL substitutions does suggest functionally separable steps in TL-substrate interactions (Figure 3A). In the absence of substrate NTP, the TL cannot be fully described in crystal structures of Pol II ECs. Residues within the N-terminal TL helix to approximately Q1078 are observed in similar positions in the presence or absence of matched NTP substrate, while the C-terminal TL helix is bent at a hinge within Rpb1 residues 1096-1097. Q1078 is positioned to interact with the 2'- and 3'-OH groups of the substrate ribose as well as another residue implicated in substrate ribose OH-interactions (N479) (Figure 3B). In some ECs with matched NTP substrates, the TL is not observed in a completely folded orientation [24, 43]; however, additional residues within the NTP-interacting region, L1081 [24, 43] and N1082 [43] are moved into substrate interacting positions (Figure 3C). L1081 stacks against the substrate base, putatively recognizing the NTP base when planar with template base, while N1082 is positioned to interact with the substrate 3'-OH and possibly phosphate groups. In these structures, the C-terminal TL helix remains bent, consistent with incomplete TL movement. Finally, in the Pol II EC structure with a complete folded TL, H1085 and the rest of the TL loop have moved into position to interact with a matched substrate NTP (Figure

3D, same conformation as magenta TL in Figure 3A)[23]. H1085 interacts with the substrate β -phosphate in this conformation and the C-terminal TL helix is unbent. Genetic analyses examining mutant phenotypes of a large number of TL residue substitution combinations, as well as substitution of a residue outside of the TL that interact with substrate (N479S), support the crystal structure putative TL folding intermediates as being functional [36]. Combinations among Q1078, N1082 and N479 substitutions tend to show epistasis suggesting they might function at the same step, while combinations between substitutions in these residues and substitutions of H1085 show strong double mutant phenotypes. These results suggest that H1085 acts at a distinct step from other substrate-adjacent residues, consistent with a multi-step model for TL folding and substrate interaction [36].

Movement of the TL has a number of consequences for Pol II function. First, interaction between TL and NTP closes the Pol II active site, which is detectable as transient resistance of Pol II to quenching of transcription through chelation of active site Mg $2+$ ions by EDTA [37]. Second, molecular dynamics simulations suggest that TL interaction with matched NTP in basepairing conformation stabilizes said interaction and limiting substrate dynamics may promote catalysis [44]. Additionally, TL-active site closure has been predicted to alter the chemical environment of the Pol II active through dehydration in other molecular dynamics simulations, which may have a catalytic function [45]. Third, molecular dynamics simulations of a product-bound Pol II structure, which is pre-translocated but retains pyrophosphate, indicate that the TL has increased dynamic properties in the presence of pyrophosphate relative to NTP, and that TL movement from closed to open conformations may aid in evacuation of pyrophosphate [46]. Fourth, this active site closure is also likely to represent a conformation of Pol II (and in other msRNAPs) resistant to factors that function through the Pol II secondary channel [25, 47, 48]. The secondary channel (also referred to as the “funnel”) is a region adjacent to msRNAP active sites where regulatory factors, including factors that promote cleavage of the nascent RNA 3' end if it is backtracked, bind and access the active site. The closed Pol II active site is incompatible with backtracking or with insertion of cleavage factor (TFIIS) because the TL occupies the region where the RNA must be located in the backtracked state and where TFIIS would insert into the active site [25, 49, 50].

3. Fidelity of Pol II transcription

Since the TL promotes fast catalysis of template-matched NTP substrates, it is a major contributor to transcription fidelity. Mutations in the TL can alter fidelity in a few ways. First, amino acid substitutions of substrate-interacting residues generally decreases msRNAP activity with template-specified (matched) NTPs more than mismatched NTPs, thus reducing net substrate selectivity [13, 32, 35, 36, 38, 39, 41]. It should be kept in mind that the 12 different misincorporation events for the canonical A, G, U and C triphosphates have disparate rates (and disparate rates of mismatch extension subsequent to misincorporation)[32, 51], so conclusions from assays measuring any particular misincorporation event might not necessarily be broadly applicable. Indeed data from *T. aquaticus* RNAP indicate that different misincorporation events have different levels of requirement for the TL, yet for all substrates, the measured TL contribution to misincorporation is less than its contribution to rates of correct NTP incorporation [32].

Second, some mutations in Pol II TL residues distal to NTP interacting residues but adjacent to TL hinge regions increase activity for all substrates, yet do so more for mismatched or inappropriate substrates than for matched NTPs, thus also reducing net selectivity [35, 37]. Because of increased incorporation rates for mismatched substrates, it seems likely that the increased activity class of Pol II TL mutants more meaningfully contribute to *in vivo* misincorporation than reduced activity fidelity mutants. The reason for this is that there are additional mechanisms for insuring transcriptional fidelity post-incorporation, including decreased rate of mismatch extension because misalignment of the mismatched RNA 3'-end. Misalignment of the end of the RNA and delay in its extension in turn allows backtracking of msRNAPs to be favored [52]. Such a delay in extension can be observed in single molecule optical trapping assays where conditions that favor misincorporation also increase frequency of long pauses, which are interpreted as delays in extension due to misincorporation [53]. Upon backtracking of enzyme along the template, a mismatched end will be further displaced and then can be cleaved from the transcript due to intrinsic hydrolysis activity of Pol II [54, 55] or TFIIS-mediated cleavage [55, 56], though intrinsic cleavage by Pol II is extremely inefficient (discussed in section 5.2) [50]. Cleavage by TFIIS would remove any mismatches present in the RNA end being cleaved, thus functioning in post-incorporation editing. Misincorporation rates by slower, less selective Pol II enzymes would be relatively reduced compared to rates of backtracking and RNA cleavage and therefore less likely to be an *in vivo* problem for slower enzymes versus inaccurate but fast Pol II mutants, which might favor mismatch extension at the expense of backtracking and cleavage. Finally, the TL has also been implicated in reducing msRNAP active site affinity for mismatched substrates, possibly through incompatibility of some TL conformations and non-basepaired or non-Watson-Crick basepaired substrate-template interactions (this is from a detailed analysis of the *Thermus aquaticus* RNAP TL [32]).

Substitutions of residues throughout the Pol II active site or the absence of the Rpb9 subunit are observed to alter enzyme fidelity *in vitro* [37, 57-59]. The role of Rpb9 in fidelity has been attributed to possible function through the TL [57], as part of Rpb9 is proximal to the TL for some TL conformations [8, 25]. Rpb9 is also required for efficient TFIIS function [60]; therefore Rpb9 function in fidelity might be multifold, affecting mismatch and mismatch extension through alteration of TL function, and additionally transcript cleavage. Alteration of a number of TL residues [35-37] or residues in the closely TL-apposed Rpb1 bridge helix (BH) [59], or in Rpb1 residues suggested to function in fidelity through *in vivo* assays [58, 61], show wide ranges of effects on catalysis or fidelity, illustrating that alteration of catalytic rates and fidelity (relative rates of matched NTP and mismatched NTP incorporation) can be uncoupled. It is not known what other residues outside the TL also function through the TL as has been proposed for Rpb9. It seems likely that there are multiple paths for connecting Pol II residues, loops and domains to active site function, and the diversity of molecular phenotypes exhibited by different substitutions suggests possibilities for communication from active site distal elements. Such elements include Fork Loop 2 [62, 63], the "F" loop [64, 65], and Link domain [65], among others (Figure 4).

4. Pol II Dynamics and Translocation

4.1 Pol II as Brownian ratchet and control of translocation

Pol II and other msRNAPs function as Brownian ratchets wherein thermal motion allows the enzymes to diffuse forwards and backwards along the DNA template. The transcription cycle is diagrammed in Figure 2. Pol II forward translocation results from substrate binding and catalysis of phosphodiester bond formation [29, 30, 53]. Evidence from single molecule optical trapping studies on *S. cerevisiae* Pol II, where effects of hindering or assisting force on enzyme velocity over a range of substrate concentrations were determined, could be fit to a model for the transcription mechanism that places the translocation step subsequent to catalysis [53]. NTP dependence of force-velocity curves supports NTP binding prior (either in E-site or other location) or subsequent to translocation (E-site, A-site or other site) but not generally being required for translocation, limiting possible allosteric roles for NTP binding on Pol II activity. It has been proposed that NTP binding directly promotes translocation [66], possibly through allosteric effects, however force-velocity curves of single Pol II molecules do not show the level of NTP dependence predicted by such a model.

Translocation may not be equally favorable for all Pol II conformations, consistent with a proposed “two ratchet” model for msRNAPs [30]. In the two ratchet model, one ratchet is the binding of NTP substrate in the Pol II active site where catalysis of phosphodiester bond formation gives directionality to translocation, and the second ratchet relates to conformational states of the enzyme that permit or disfavor translocation (a possibility discussed in molecular dynamics studies on Pol II from Feig and Burton and coworkers [67, 68]). Occupancy of the Pol II active site by a basepaired NTP indeed stabilizes Pol II in the post-translocated state when it cannot be hydrolyzed [37]. Conversely to promoting a post-translocated state by TL closure over substrate, prolonged active site closure subsequent to catalysis should inhibit pyrophosphate release and stabilize the pre-translocated state through interactions with product 3'-NMP. Recent work from the Belogurov group elegantly demonstrates the ability of TL-3'-NMP interactions to stabilize the pre-translocated state when it is induced by the antibiotic Tagetitoxin [33]. Their work supports a model where in the absence of TL folding and TL-3'-NMP contacts, which would also be destabilized by the release of pyrophosphate upon partial TL unfolding, the post-translocated state has a higher affinity for the 3'-NMP than the pre-translocated state.

A TL mutant, Rpb1 E1103G, which increases Pol II elongation rate relative to wild type [35, 37, 69] and was previously shown to promote active site closure in EDTA quench assays [37](discussed above), appears to have the property of altering Pol II translocation properties by stabilizing the pretranslocated state. Behavior of this mutant in single molecule optical trapping assays is consistent with increased catalytic efficiency, but decreased propensity for translocation [53]. This decreased propensity for translocation is evidenced by an increase in sensitivity of E1103G Pol II to assisting force (indicating an altered average ratio of pre-translocated/translocated enzyme over at a high fraction of positions on the template). These results support the model that the closed conformation of the Pol II active site may not be as permissive for translocation as the open conformation. This result is consistent with a second ratchet controlling translocation of Pol II, functioning to restrict translocation to

particular enzyme conformations, and limiting translocation to single base steps. Limiting translocation to catalytically deficient conformations (open active site) could favor an ordering of biochemical events in the transcription cycle, if each translocation step required conformational dynamics that passed through conformations favoring catalysis. If catalysis does not happen to occur, or substrate does not bind and stabilize the post-translocated state, then the nucleotide-dependent ratchet would not promote forward translocation.

Conformational changes occurring that would be permissive for translocation would then likely favor reverse translocation or no translocation at all, because additional forward translocation (“hypertranslocation”) is not energetically favorable. This is because RNA-DNA hybrid contacts would be reduced and an RNA-DNA basepair at the upstream edge of the RNA-DNA hybrid would be melted by the apparatus that separates the nascent RNA from template DNA, without being replaced by a newly synthesized RNA-DNA basepair at the 3′ end of the nascent RNA. Hypertranslocation has been proposed as a pathway for intrinsic termination for *E. coli* RNAP [70], but is not required at all intrinsic terminators [71], and precise mechanisms of termination for msRNAPs remain less well understood.

4.2 Nature of conformational changes underlying Pol II translocation

What conformational changes might control translocation in Pol II? As discussed above, folding of the TL to close the active site would physically restrict backtracking, as the TL occupies the region where a backtracked RNA would be extruded from the active site. The TL proximal BH has long been a candidate for contribution to translocation because of distinct conformational states (straight vs. bent or kinked) observed in crystal structures of msRNAPs from different species (*Thermus aquaticus* RNAP [10] and *S. cerevisiae* Pol II [5]). Since the DNA template strand must pass over the BH and the base present in the *i*+1 position (occupying the active site when in the post-translocated state) must flip in orientation from its *i*+2 position to insert into the active site, it is easy to imagine that alteration in BH conformation might alter Pol II contacts with DNA. Consistent with BH flexibility and functional conformational changes, ultra-high scale mutagenesis studies on an archaeal msRNAP from *Methanocaldococcus jannaschii* (*Mja*) reveal two regions of the BH where mutations predicted to stabilize or promote kinking increase enzyme activity, consistent with promotion of translocation or catalysis [42, 65]. Molecular dynamics simulations support the propensity of these alpha-helical regions to adopt non-helical conformations [65, 72]. Combination of super-activating BH-kinking substitutions (for the C-terminal BH kink) and super-activating TL substitutions in *Mja* RNAP had no further effect on activity, suggesting these BH and TL substitutions might alter the same step in RNA synthesis [42]. A mutation in the Pol II BH has also recently been shown to increase Pol II activity [59].

A number of molecular dynamics studies have been performed on Pol II and other msRNAPs to gain insight into possible conformational changes underlying translocation [45, 59, 67, 68]. Reference [59], cited above for studies on a the BH, coupled molecular dynamics on *Thermus thermophilus* RNAP with extensive mutagenesis in the *S. cerevisiae* Pol II system and described a number of interactions that develop stably in molecular dynamics simulations and are mutually exclusively to either open or closed TL structures. These alternate conformations, described as “switches” (not to be confused with flexible

“switch” regions 1 through 5 described structurally as promoting closure of the Pol II clamp upon binding DNA [6]) are proposed to relate to the suite of conformational changes that occur during the translocation cycle, and are underpinned by flexible “hinges” at the base of the TL helices and within the BH. In fact, DNA is observed to partially translocate in simulations of the open *Th* RNAP structure, consistent with hypotheses that the open TL enzyme conformation is permissive for translocation. Molecular dynamics simulations are not yet able to cover timescales where large domain movements such as full TL opening or closing or downstream DNA or RNA/DNA hybrid translocation can be observed, but smaller scale conformational changes are likely to contribute to the thermal fluctuations underlying distinct Pol II conformations.

4.3 Role of DNA sequence in Pol II translocation

Melting of DNA ahead of the Pol II and re-annealing of DNA in the wake of the enzyme render this aspect of translocation somewhat energetically neutral but subject to some sequence considerations. The energetics of the hybrid region versus upstream and downstream DNA sequences will impact the propensity of msRNAPs to backtrack. This is because enzymes would be expected to occupy local energy minima, such as over a region where the RNA-DNA hybrid would be A-T(U) rich and the sequence upstream G-C rich. Nucleic acid contacts between Pol II and the transcription scaffold have been thought to be minimally sequence specific, thus template positions would not necessarily be easily distinguishable, allowing Pol II access to the many template positions with caveats for sequence energetics. However, recent intriguing work from the Landick [73], Belogurov ([33], introduced in 4.1) and Zenkin [74] labs has revealed important sequence preferences for msRNAPs, and some of these observations have been experimentally extended to Pol II. All of these observations are likely applicable to Pol II, should conservation of mechanism hold. The Landick group has shown sequence-specific effects on pyrophosphorolysis (discussed in section 5.1 below), which they use as a probe for RNAP translocation state (pyrophosphorolysis requires that RNAP be in the pre-translocated state) [73]. Belogurov and colleagues used parallel kinetic analyses of chain elongation, translocation state and pyrophosphate release to show sequence preferences of *E. coli* RNAP, with some RNA 3'-NMPs favoring positioning in the post-translocated state [33]. They additionally showed that pyrophosphate release precedes or coincides with translocation, consistent with the model that TL opening (allowing pyrophosphate release) permits translocation [33]. Zenkin and colleagues showed that *E. coli* RNAP distinguishes between RNA-DNA hybrid sequences, showing sequence preference whereby a particular NMP is preferred in one register, thus can stabilize pre-translocated, post-translocated or “backstepped” states depending on the translocation register of the enzyme [74]. For one pause sequence, they observed pausing of eukaryotic Pol I, II, and III enzymes to be similar to *E. coli* and *S. pneumonia* RNAPs, suggesting possible universality of sequence recognition within the RNA-DNA hybrid. Recognition of sequence within the RNA-DNA hybrid region might also contribute to polymerase fidelity because loss of such interactions for a mismatched RNA 3' end might allow backtracking to be favored.

5. Other enzymatic activities of the Pol II active site

5.1 Pyrophosphorolysis

As discussed above, Pol II can participate in factor-mediated RNA cleavage through action of TFIIIS. Naturally, the reversal of phosphodiester bond formation through the reconstitution of an NTP via pyrophosphate-mediated cleavage of the nascent RNA 3'-NMP (pyrophosphorolysis) is supported by the Pol II active site. Pyrophosphorolysis is much less efficient relative to the forward reaction, nucleotide addition, and requires both high levels of pyrophosphate and that Pol II be pre-translocated. Pyrophosphorolysis is defective in bacterial RNAPs with substitutions in TL NTP-interacting residues, consistent with roles for the TL in catalysis (that is both forward and reverse reactions are defective) [13, 39]. Interestingly, differential sensitivity of *E. coli* or *T. aquaticus* RNAP ECs with nascent RNAs varying in the *i* (terminal) and *i*-1 (penultimate) bases to pyrophosphorolysis indicates some sequence specificity within the RNA-DNA hybrid region on the RNA side, as mentioned above in 4.3 [73]. Hein et al [73] interpreted resistance to pyrophosphorolysis as a particular RNAP complex being prone to the post-translocated state, and thus unable to undergo pyrophosphorolysis. When the *i* NMP was GMP or AMP, the post-translocated state was favored more than if it was CMP or UMP, and in RNAP ECs where the RNA was already short (slowing pyrophosphorolysis), complexes were more resistant to pyrophosphorolysis. Conversely, when the *i* NMP was CMP or UMP and the *i*-1 was GMP or AMP, the pre-translocated state was favored, and measured pyrophosphorolysis rates were faster. Their experiments also showed this sequence preference for Pol II. Intriguingly, the preference for purines in the *i* position (when post-translocated), leaving the *i*+1 site open for an NTP substrate may explain partially the preference of Pol II for initiation using ATP or GTP [75]. During initial phosphodiester bond formation, two NTPs must be bound, including one in the post-translocated *i* position (sometimes called the P site for some polymerases [76], but not to be confused with Proofreading "P" site [50]), which normally during elongation would contained a base of the RNA-DNA hybrid, and the other in the *i*+1 position (Addition or A site according to [24], or sometimes Nucleotide (N) site [76]). Sequence specificity might function additionally to stabilize the critical initiation conformation. Start site preference in *E. coli* and *S. cerevisiae* is for purines [75, 77], with a pyrimidine at position -1 (also meaning a purine opposite this position on the non-template strand), consistent with results from a T7-like RNAP initial transcribing complex, which has a +1 GTP (in *i* position) showing stacking interactions with a non-template strand purine at -1 [76].

5.2 Intrinsic hydrolysis

MsRNAPs themselves are able to hydrolyze the nascent RNA transcript in a 3' to 5' direction, independent of other factors, and with low efficiency [55, 78]. Although this endonucleolytic cleavage is supported by the same Pol II active site as pyrophosphorolysis, intrinsic hydrolysis is a chemically separate activity, resulting in oligonucleotides with 5'-NMP, whereas the product of pyrophosphorolysis is an NTP [79]. First detected for *E. coli* RNAP [78], msRNAPs from other organisms possess the ability to intrinsically hydrolyze a phosphodiester bond but with different rates. *E. coli* RNAP is faster than *S. cerevisiae* Pol II, with the latter also being about two orders magnitude slower than *T. aquaticus* RNAP [50,

80]. Almost exclusively, the 2nd phosphodiester bond is cleaved at the 3' end of RNA transcript within 1-bp backtracked and pre-translocated elongation complexes [80]. The role of the TL in RNA hydrolysis awaits full elucidation, but recent work of Yuzenkova and Zenkin [81] has revealed the essential role for *T. aquaticus* β' H1242 (analogous to Rpb1 H1085) as a general base, directly participating in the reaction. Alternatively, this histidine was suggested to orient the 3'-NMP that facilitates cleavage [81]. It is commonly accepted that transcript cleavage serves to release Pol II from transcriptional arrest. The Svejstrup group found that an *S. cerevisiae* TFIIS double amino-acid loop mutant (D290A, E291A), which is inactive for TFIIS cleavage stimulatory activity, also was inhibitory for intrinsic cleavage when added to Pol II *in vitro* [82]. This result would be consistent with a TL function in Pol II intrinsic cleavage, as TFIIS might restrict TL to a non-active conformation, blocking intrinsic cleavage. Furthermore, it was suggested that extrication from backtracking through transcript hydrolysis in the absence of TFIIS cleavage was required for transcript elongation and cell viability, as overexpression of the mutant TFIIS was lethal *in vivo* [82]. In this study, the mutant TFIIS did not block transcription in an *in vitro* system consistent with its effects being restricted to trapping backtracked Pol II complexes by blocking both pathways for transcript cleavage, factor-dependent and intrinsic hydrolysis. It also remains possible that additional, cryptic dominant negative effects of this TFIIS mutant might contribute to its lethality *in vivo*.

6. *In vivo* phenotypes controlled by Pol II activity

6.1 Identification of Pol II mutants in genetic screens

S. cerevisiae has long been an essential model system for understanding eukaryotic transcription mechanisms for all eukaryotic msRNAPs through genetic, biochemical, structural and genomic means. Subsequent portions of this review will focus on how genetic analyses of Pol II subunits has and continues to inform our understanding of the transcription process *in vivo*. Mutations in Pol II subunits have been identified in a number of genetic screens for altered transcription, in many cases linking Pol II subunits to gene-specific transcription phenotypes. With increased understanding of the Pol II mechanism and its regulation, revisiting these alleles and phenotypes will be productive for the understanding of how alterations in Pol II subunits control gene expression *in vivo*. An extremely valuable review from Coulombe and colleagues collates most if not all Pol I, II, III, and *E. coli* RNAP mutants identified through 2006, discussing them from a structural perspective [83]. This review itself was an update of an earlier highly useful review from Archambault and Friesen similarly covering molecular genetics of Pol I, II, and III through 1993 [84].

In the late 1980s, *sit1* and *sit2* mutants (suppressors of initiation of transcription) were identified as alleles of *RPO21* (official name of *RPB1*) and *RPB2*, respectively, through screens for mutations that allowed bypass of mutants in the transcription activators of *HIS4* [85]. Cloning of the *RPO21/RPB1 sit* mutants identified a number of mutants encoding substitutions within conserved regions of *RPO21*, especially conserved region F (described in [4]), which comprises the F Loop and Bridge Helix, suggesting possible effects on Pol II activity [86]. Indeed, *sit1-7* encodes Rpb1 G730D [86], independently identified by Koyama

et al as a suppressor of *rpb9* phenotypes [61]. Biochemical characterization of Rpb1 G730D by the Kashlev lab indicates that this mutation reduces Pol II substrate selectivity but does so by reducing Pol II activity with matched substrates [58]. Another *RPO21/RPB1* *sit* allele, *sit1-8* contained two substitutions, one encoding Rpb1 N479Y within an NTP-interacting residue and the other encoding Rpb1 A1076V (*sit1-8G*) in the Rpb1 TL [86], then described as Rpb1 homology block G. Work from my lab suggests that A1076V has gain-of-function properties in its ability to suppress growth defects of other, catalytically defective TL substitutions [36]. Friesen and coworkers identified N479Y as lethal in certain strain backgrounds in the absence of A1076V [86], and taken together with the fact that other substitutions at N479 cause elongation defects [23, 36], allow the suggestion that N479Y is likely a reduction of function allele, as G730D is, and that N479Y lethality is suppressed because A1076V counteracts some of the reduced catalytic activity. These connections allow us now to interpret the *sit* phenotype as relating to a reduction in function of Pol II, even though the net result is increased Pol II transcription at *HIS4*. The *sit1* class of Pol II alleles may be distinct from other classes of Pol II allele whose phenotypes derive from reduced function because patterns of altered start site selection in different contexts are distinct [86].

Similarly to *sit* alleles, *RPO21/RPB1* alleles were identified as spontaneous upregulators of a gene, in this case *URA2* under conditions when it is not normally expressed [87]. Surprisingly, these alleles all clustered in a small region of *RPO21/RPB1* adjacent to the C-terminal end of the BH and template DNA (in Pol II ECs) known as the “Switch 1” region (Figure 4). The mode of action of these alleles at *URA2* is complicated, because *URA2* transcription is complex, with a small, unstable RNA normally transcribed from upstream start sites under conditions when no full length *URA2* RNA is expressed [87, 88]. Upon induction, functional *URA2* is expressed from downstream transcription start sites [87, 88]. *URA2* therefore represents a kind of complex locus with multiple transcription units interacting. *URA2* regulation is related to an increasing number of genes in *S. cerevisiae* whose regulation integrates the production of different transcripts into regulatory decisions (see Figure 5). This class of genes appears to be especially sensitive to Pol II activity changes. The sequence of the *URA2* upstream unstable RNA is intriguing, as it is T-rich on the non-template strand, encoding a U-rich RNA, which further suggests that compromise of transcription in this region upon uracil depletion might be part of the normal activation of *URA2*. Indeed, uracil limitation activates *URA2* transcription from downstream start sites, allowing the sequence of the upstream unstable RNA to be bypassed. This is critical for *URA2* expression because transcription from upstream starts allows sequences within the upstream unstable RNA to recruit Nab3/Nrd1 to promote premature termination [87, 88]. Two models, not mutually exclusive, might explain usage of the downstream *URA2* start sites in Pol II mutants and under uracil limitation. First, altered start site selection at *URA2* might favor downstream start sites; indeed a start site selection shift to downstream start sites at *ADH1* has been shown for the *rpo21/rpb1* Switch 1 substitution strain *rpb1-S1401P*, which constitutively expresses *URA2* [88]. Second, mutants in Pol II with reduced catalytic activity might mimic the effects of uracil limitation on wild type Pol II and cause a transcription defect in the upstream small RNA over the T-rich sequence, favoring or allowing access to downstream start sites. This second model is distinct from the first

because it posits that downstream starts are inaccessible due to elongation from upstream positions, in contrast to upstream starts being biochemically favored due to sequence or location. Consistent with both of these models, the elongation defective *rpb1-H1085Y* TL substitution mutant, which confers downstream start site shifts at *ADH1*, constitutively upregulates *URA2* [36].

In addition to genome-wide screens, targeted mutagenesis has been performed on Pol II subunits to identify mutations that confer a number of phenotypes, such as temperature sensitivity [89, 90], sensitivity to nucleotide-depleting drugs [36, 69, 91], and suppression of transposon-insertion dependent transcriptional defects (the Spt⁻ phenotype) [35, 36, 92], among others. The Spt⁻ phenotype (Suppressor of Ty) derives from suppression of one or more transcription defects induced by insertion of retrotransposons into genes or their promoters [93, 94]. A number of these retroelement insertions have undergone recombination between their long terminal repeats (LTRs) leaving just a solo LTR at the site of insertion (the yeast Ty retrotransposon LTR is called a ψ element). The ψ contains the Ty-element promoter, therefore insertion of a ψ element into the promoter or 5' end of a gene creates a complex transcription unit with both the native and inserted promoters juxtaposed, allowing for possible interaction or interference. In addition to Pol II alleles with the Spt⁻ phenotype, a large number of chromatin regulators, and chromatin components themselves confer the Spt⁻ phenotype (reviewed in [95]). The identification of transcription and chromatin components that together conferred altered transcription phenotypes was one of the earliest indications of chromatin functioning as a malleable platform for gene regulation. Among the transcription regulators identified as *SPT* genes were what we now recognize as general Pol II elongation factors with some having roles in cotranscriptional nucleosome assembly and disassembly. The initially described *rpo21/rpb1* Spt⁻ alleles were also located in the TL [92], but have not been further characterized. Additional identified Pol II Spt⁻ alleles encoding substitutions in the Rpb1 TL have increased biochemical activity or are predicted by genetic interactions to have increased function [36]. It seems likely that the Spt⁻ phenotype for Pol II TL alleles is predictive of increased or altered activity that is distinct from defective catalysis.

6.2 Start site selection or utilization

In *S. cerevisiae*, a number of mutants have been identified in Pol II subunits or in subunits of Pol II general transcription factors that alter position or distribution of transcription start sites downstream of core promoter elements (Reviewed by Corden here [96]). These include mutations in genes encoding the Rpb1 [36, 88, 97, 98], Rpb2 [99] and Rpb9 [100-102] subunits of Pol II, and in genes encoding GTFs TFIIF [103-105], TFIIB [106-110] and TFIID [111]. Some alleles of *RPO21/RPB1* were identified in a global screen for mutations that suppressed expression defects in the *CYC1* gene due to *cyc1* 5' UTR mutations introducing an upstream ATG [112]. The short out of frame upstream ORF renders *cyc1* transcription sensitive to nonsense-mediated decay. Mutations in a number of genes, including *SUA7* (TFIIB) [106, 107] and *SUA8* (*RPO21/RPB1*) [97], suppress this defect by allowing transcription start site selection downstream of the mutant ATG, precluding nonsense-mediated decay. An allele of *RPB2* was discovered as a mutant that itself suppressed the start site defect of certain *sua7* alleles [99]. As mentioned above, *SUA7*

encodes for TFIIB, and certain alleles of *SUA7* cause growth defects and downstream shifts in start site selection. In addition to suppressing growth and start site defects of *sua7* alleles, *rpb2-101* on its own confers upstream shifts in transcription start site selection, consistent with flexibility in start site selection in both directions and with contributions of multiple Pol II subunits [99].

S. cerevisiae is somewhat unique among model eukaryotic systems in that transcriptional start sites are at a more variable and greater distance downstream of core promoter elements such as the TATA box [113]. As in other organisms, not all *S. cerevisiae* promoters contain consensus TATA elements and promoters containing or lacking consensus TATA elements have differential requirements for or contributions of regulatory factors (for example see [114, 115]). Very recent work from the Pugh lab indicates that TATA-like elements, containing 1-2 mismatches from consensus TATA element, are recognizable within a large number of *S. cerevisiae* promoters through analysis of high-resolution chromatin IP studies pinpointing sites of Pol II GTF-DNA crosslinks [116]. With such high resolution data for identifying putative locations of pre-initiation complex (PIC) formation, it will now become possible to discern rules and relationships of productively used transcriptional start sites and sites of PIC formation or other promoter features such as promoter sequence elements, size of promoter nucleosome-free region and stability or composition of promoter chromatin structure.

Originally, mutants in Pol II subunits conferring changes in start site distribution were interpreted generally as altering initiation through GTF functions or inputs in promoting initiation. This is reasonable for Pol II mutants adjacent to predicted GTF binding surfaces, for which we await-high resolution structural information. Recent, extensive crosslinking studies indicate that TFIIF makes a number of contacts with the *RPB2* side of the Pol II cleft [105, 117, 118]. A number of mutations in TFIIF cause upstream shifts in start site selection [104, 105], as do *rpb9* [100-102] and *rpb2-101* [99]. The predicted binding surface for TFIIF is near the Rpb2 residue substituted in *rpb2-101* (Rpb2 G369), while Rpb9 binds under the Rpb2 side of the Pol II cleft. *rpb2-101* encodes Rpb2 G369S [99] and targets the same residue as previously identified *Spt⁻* alleles of *rpb2*, *rpb2-504/505* (alleles have the identical change) [92]. The *Spt⁻* phenotype of *rpb2-504/505* may relate to an initiation function through TFIIF as at least one allele of TFIIF confers a weak *Spt⁻* phenotype (discussed in 6.1), similarly to *rpb2-504/505* (H. Jin and C.D. Kaplan, unpublished observations). Taken together, this surface of Pol II may relate to the integration and function of TFIIF in initiation and start site selection.

Work on additional Pol II mutants indicates that alteration of transcription start site selection may have a fundamental relationship to efficiency of RNA chain elongation activity, therefore functioning in parallel with GTF activity [36]. Substitutions in the Pol II TL that confer increased elongation rates also cause upstream transcription start site shifts at a number of genes, while substitutions that confer decreased elongation rates, cause downstream shifts at *ADHI*, a widely used and sensitive model gene for Pol II start site in *S. cerevisiae* [36]. To explain how Pol II activity alters transcription start site selection, it has been proposed that reduction of activity increases abortive initiation, which might be translated into downstream translocation [36, 98, 119]. Such proposals are consistent with

flux models for Pol II scanning in *S. cerevisiae*, where start site usage at downstream positions is controlled by efficiency of upstream start utilization. Thus, polar effects of start site selection mutants, which shift start sites upstream or downstream to the relative disadvantage of distal or proximal starts, respectively, are observed [120]. These models also explain why continuous transcription between upstream promoter positions and downstream positions does not seem to be required for Pol II promoter scanning. This conclusion is due to the observation that chain-terminating nucleotides do not block Pol II promoter scanning *in vitro* [98, 121]. Conversely, work in human systems (reviewed in this issue) on Pol II indicates that abortive initiation is not common for promoters examined.

An alternate possibility for scanning (if it does not involve abortive initiation), is that transcription bubble translocation passes sequences through the active site and the probability that any particular start site will be used is determined by the following: the duration of time a template base is present in the initiation conformation, the probability that that sequence can support initiation, and the catalytic rate of Pol II. An appropriate analogy here would be to a shooting gallery, where the rate at which a target passes and the rate of firing together contribute to the probability a target is hit. If Pol II is able to sufficiently catalyze enough phosphodiester bonds such that translocation due to synthesis coordinates with or disengages transcription from translocation due to scanning, transcription may be productive. If Pol II has a lower probability of initiating synthesis within such a window due to reduced activity, scanning translocation will end up further downstream on average before initiation occurs. In such a model, with no abortive initiation, the synthesis of the first phosphodiester bond would have to be enough for productive initiation otherwise abortive initiation would be observed. In such a model, TFIIF would likely provide the energy for scanning, as has been proposed previously (reviewed in Corden [96]) and would be consistent with recent results characterizing alleles of the critical TFIIF component, Ssl2, which is the subunit required for template opening [111]. In this work, an *ssl2* allele was shown to suppress a downstream shifting start site allele of *sua7* (TFIIB), and a number of *ssl2* alleles were shown to shift starts site distribution upstream at *ADH1*. An interpretation of these alleles would be that they could be defective or slow for TFIIF-driven scanning, favoring upstream starts over downstream starts.

6.3 Sensitivity of Pol II mutants to nucleotide synthesis inhibiting drugs and altered expression of *IMD2* in *S. cerevisiae*

Just as Pol II mutants with disparate effects on elongation rate alter start site selection in different ways, they also show disparate phenotypes *in vivo* when exposed to drugs that deplete pools of certain nucleotides. These drugs, 6-azauracil (6-AU) and mycophenolic acid (MPA), have been widely used to probe functions of transcription related mutants in *S. cerevisiae* (discussed extensively in [36]). The connection between use of these drugs and transcription came about when it was determined that the 6-AU sensitive *ppr2-3* mutant was an allele of the gene encoding the *S. cerevisiae* TFIIS gene, *DST1* [122]. *ppr2-3* was selected in a genetic screen for mutants defective in the uracil metabolic pathway using sensitivity to 6-AU, a known inhibitor of orotidylic acid decarboxylase (encoded by *URA3*). Subsequently, the 6-AU sensitivity of *dst1* (TFIIS) was determined to be partially conferred by 6-AU inhibition of a GTP metabolic pathway activity, inosine monophosphate

dehydrogenase (IMPDH), in addition to *URA3* inhibition [123]. When it was found that Pol II mutants could also confer 6-AU sensitivity, and that some of these Pol II alleles could be suppressed by overexpression of TFIIS, an attractive “elongation inhibitor” model for 6-AU/MPA function was proposed [91].

The model, widely embraced and extended in the literature, proposes that IMPDH inhibitors 6-AU and MPA reduce GTP levels and in doing so promote transcription elongation defects such as pausing or backtracking by limiting a transcription substrate. Therefore, mutants with putatively compromised elongation, such as *dst1*, or Pol II mutants defective for TFIIS interaction, or in a number of other transcription factors genes, are sensitized to reduction of GTP through 6-AU or MPA treatment, because the combination of transcription defects becomes hypothetically synergistic or extends beyond a threshold for normal growth. There is no doubt that these drugs *do* confer transcriptional phenotypes to wild type cells (some examples here [124-127]), and in many cases mutant drug sensitivities can be suppressed by exogenous guanine, indicating observed phenotypes are GTP-related metabolite dependent and not additional drug target effects (of which there might be a number based on broad sensitivities to these drugs, see [128, 129]). Extensive and detailed work from the Reines lab has gone to great lengths to determine the mechanisms by which wild type cells are resistant to 6-AU and MPA, and why transcription mutants might be sensitive [130-136]. They have found that wild type *S. cerevisiae* tolerate IMPDH inhibition by 6-AU and MPA because they upregulate a resistant form of the enzyme encoded by *IMD2*. Furthermore, the Reines lab discovered that *IMD2* upregulation in presence of 6-AU and MPA is generally defective in transcription-related drug sensitive mutants, including *dst1* [136]. The inability to induce *IMD2* in the presence of MPA or 6-AU would then be sufficient to explain drug sensitivity of transcription mutants, because these mutants would be severely starved for GTP relative to wild type cells similarly treated with drug. Any transcription phenotypes of drug-sensitive mutants ascribed to synergistic effects of loss of function of such factors and reduction in GTP levels should be immediately suspect because wild type controls may not be similarly GTP starved, and global transcription defects might solely be attributed to GTP starvation differences between wild type and mutant strains and not loss of elongation-promoting activity of the drug-sensitive mutant in question. It is possible that some drug-dependent phenotypes of sensitive mutants relate to *IMD2*-independent transcriptional defects, as overexpression of *IMD2* does not completely rescue some drug-dependent transcription phenotypes of certain mutants [135], though in many cases it does [129, 135]. Interestingly, *dst1* 6-AU and MPA sensitivity has recently been shown to be suppressed by *sfp1*, a mutation that does not restore *IMD2* expression to *dst1* strains, indicating that *IMD2* expression defects might not be completely explanatory for the *dst1* phenotype [137].

All the results above raise the question: what makes *IMD2* expression so sensitive to transcription defects? This conundrum has been unraveled to a great extent by the Brow and Reines labs, with an intriguing first hint from the Ares group. The Ares group identified an RNA stabilized in the *rrp6* exosome mutant just upstream of *IMD2* [138]. This upstream *IMD2* transcript illustrates analogy between *IMD2* and *URA2* regulation, which though distinct in many ways, still share the requirement for an upstream transcript over the

promoter region to integrate regulatory input derived from nucleotide limitation. Steinmetz and Brow discovered that the *IMD2* promoter was highly transcribed by performing genome-wide chromatin IP and identifying a prominent peak of Pol II occupancy over the *IMD2* promoter under conditions where *IMD2* was not induced [139]. They then mapped the start sites of the *IMD2* promoter transcript by primer extension, showing they mapped to Gs, while the downstream functional *IMD2* start mapped to an A position. Furthermore, they showed that the *IMD2* promoter transcript required Sen1/Nrd1/Nab3-termination function because in a *sen1* mutant, full-length transcripts originating at the upstream, G start sites were observed [139].

In subsequent work from the Brow and Reines groups, it was shown that usage of upstream G starts was specifically sensitive to MPA treatment, causing a preference for downstream start sites instead when GTP was limiting [130, 140]. Because there is a start site switch involved in *IMD2* regulation, start site selection mutants were examined for their effects on *IMD2* transcription. Strathern and colleagues [69] had previously identified 6-AU sensitive alleles of *RPB1*, *rpb1-D260N* and *rpb1-N488D*, as constitutive expressers of *IMD2*, while Kuehner and Brow showed that *IMD2* expression in these alleles could be partially suppressed by guanine. This result indicated that guanine sensing (proposed to be through GTP) by Pol II might be shifted in these mutants, such that in normal media, their defects in transcription mimic GTP starvation and activate *IMD2*. Interestingly, this class of Pol II allele represents the rare transcription mutants that are sensitive to MPA or 6-AU without having an *IMD2* activation defect, indicating their sensitivity may be downstream of *IMD2* expression and possibly through a general transcription defect. Further linking *IMD2* expression to Pol II initiation activity, downstream start site shifting alleles of *SUA7* (TFIIB) also constitutively express *IMD2*, suggesting that GTP levels might be communicate to the *IMD2* promoter directly through Pol II during initiation and start site selection [140]. Intriguingly, other 6-AU sensitive alleles within Pol II identified by Malagon *et al* [69], and all MPA-sensitive Pol II alleles thus far identified within the Rpb1 TL [35, 36](C. Kaplan, unpublished observation) do not constitutively upregulate *IMD2*. Instead, where tested, these alleles are defective for induction of *IMD2*. These alleles are also known to confer increased activity to Pol II *in vitro* while also shifting start sites upstream at *ADH1* *in vivo*, thus they might compromise start site switching at *IMD2* because they are less efficient at communicating acute GTP starvation after MPA treatment [36].

6.4 Nature of Pol II transcription defects *in vivo* in *S. cerevisiae*

The majority of Pol II mutants discussed so far have *in vivo* phenotypes that may be broadly due to initiation defects and not transcriptional elongation defects. Why might this be? First, many of transcriptional phenotypes for which Pol II mutants have been screened are now known to be highly sensitive to initiation defects, in many cases to start site selection. Start site selection appears to be directly responsive to Pol II catalytic activity, explaining how Pol II mutants can directly alter initiation (not surprisingly, because the first measurable committed step in start site selection is phosphodiester bond formation of two NTPs) [36]. The fact that many *in vivo* transcriptional phenotypes relate to phenotypes exquisitely sensitive to small changes in biosynthetic enzyme expression (*sit* mutants, *spt* mutants), or genes whose expression sensitivity may be amplified by compound transcription units (Spt-

phenotype reporters, *URA2*, *IMD2*, *SER3*, etc.) (Figure 5), suggests that genetic assays have been finely balanced to detect transcription defects in unique gene-specific configurations. Compound transcription units have the ability to communicate defects in initiation, elongation, or termination of one transcript on the initiation of another transcript, which gives a measurable change in output, detected as a phenotype.

The *SRG1-SER3* system (Figure 5E) has been extensively studied as a locus where an upstream non-coding RNA, *SRG1*, regulates expression of the downstream *SER3* gene in a number of ways. First, expression of *SRG1* is required for inhibition of *SER3* [141, 142]. Second, nucleosome assembly/positioning that functions during *SRG1* transcription is also required for *SER3* repression [143-145]. Because this system has demonstrated ability to be responsive to numerous transcription defects, it appears analogous to *IMD2*, *URA2*, and Spt⁻-phenotype reporter alleles. Indeed, unpublished observations from my lab indicate that *SER3* repression is exquisitely sensitive to changes in Pol II function, though showing a different pattern of effects from *lys2-128*, *IMD2*, or *URA2* because both fast and slow TL mutants derepress *SER3*.

6.5 Pausing, arrest and Pol II processivity *in vivo* in *S. cerevisiae*

S. cerevisiae genes are relatively short, due to the paucity of introns. That being said, it is generally assumed that Pol II should efficiently transcribe the full extent of templates once the promoter has been cleared. To measure Pol II elongation rate *in vivo* and indirectly measure apparent Pol II processivity, chromatin IP has been used at very long genes [126, 146] to determine occupancy at steady state and the kinetics of loss of chromatin IP signal upon gene shut off [126]. Differential kinetics of gene clearance from downstream sequences relative to sequences nearer the promoter, which will be cleared first as the last wave of Pol II moves downstream, can be used to infer apparent elongation rate. Similarly, normalized ChIP signal for Pol II between mutant and wild type strains in steady state can be used to infer apparent processivity defects, when looking at the ratio of ChIP signal between 5' and 3' ends of a gene. Mutants in a number of factors show defects that are inferred as processivity defects [126]. Comparison of run-on data with ChIP data suggests that ChIP occupancy is not well correlated with actual transcription rate, and that some genes might be transcribed at different rates [147]. Intriguingly, 5'-to-3' Pol II ChIP ratios change at a number of genes depending on growth conditions of the strains, indicating there might be global mechanisms for altering Pol II transcription during the elongation phase [148]. Since Pol II ECs are highly stable, it seems active processes might be removing Pol II from templates. Such processes might include ubiquitination and degradation of Rpb1 of stalled Pol II as happens upon Pol II arrest due to DNA damage (for review see [149] and references therein), or premature termination.

At least two termination pathways operate in *S. cerevisiae*, Poly(A)-dependent and Sen1/Nab3/Nrd1-dependent (reviewed in [150]). The former functions subsequent to transcript cleavage at 3'-ends of genes, while the latter functions in 3'-end formation of some non-coding RNA genes as well as at short RNAs such as CUTs, including *IMD2* and *URA2* promoter RNAs. Much of Pol II termination is regulated through phosphorylation of the Pol II Rpb1 C-terminal domain, containing a heptapeptide repeat sequence with up to five

phosphoaccepting residues per repeat (discussed [17, 150]) including the recently described CTD repeat position 1 Tyrosine [151]. Tyrosine-1 has long been known as phosphorylatable in higher eukaryotes, but now also appears so in *S. cerevisiae*, with the intriguing fact that there had been thought to be minimal tyrosine phosphorylation in the yeast proteome. Recent work examining effects of termination defective alleles of Pcf11 or Nrd1, etc., and genome-wide localization of these factors indicates that generally recruitment of Pcf11 and Nrd1 have different patterns, but sometimes can be recruited to function at the same transcripts, while showing specificity for others [17, 151-153]. Altogether, these results suggest complex relationships between termination pathways.

It has long been known that nucleosomes can represent robust barriers to Pol II elongation *in vitro* [154], yet conversely chromatin templates are apparently efficiently transcribed *in vivo*. A large collection of factors functions to remodel and modify chromatin during the transcription process (recently reviewed extensively [95], with focus on histone chaperones Spt6 and Spt16 here [155]), and transcription can lead to reduced nucleosome occupancy over highly transcribed regions, consistent with some disassembly of nucleosomes during transcription [146, 156]. High resolution mapping of Pol II-associated RNA 3'-ends (as markers of Pol II EC position in the genome), using a technique developed by Churchman and Weissman called NET-seq, has indicated that Pol II-associated RNA ends periodically map downstream or promoters to locations coincident with the average positions of promoter-proximal nucleosomes [157]. This accumulation was observed in strains that had TFIIS deleted (*dst1*), to prevent TFIIS rescue of backtracked ECs. These periodic putative positions of Pol II elongation complexes were not apparent at the single gene level, but on the level of metagenome analysis, peaking just prior to the nucleosome dyad, which is the position of high-probability Pol II arrest on a nucleosome *in vitro* [158]. This result is consistent with some fraction of Pol II ECs accumulating at with a higher probability within nucleosomal DNA, likely through pausing followed by backtracking or arrest. Also consistent with backtracking being common *in vivo*, gene-specific peaks of Pol II-associated RNA 3' ends shift between WT and *dst1* strains, indicative of widespread TFIIS-mediated RNA cleavage [157]. Basepair-resolution mapping of nucleosome positions in comparison with the NET-seq data from has indicated that Pol II may pause throughout the nucleosome even in the distal half (instead of elongation engendering complete unwrapping of DNA once into the downstream half of the nucleosome) [159]. This apparent nucleosome-induced pausing or arrest shows ~10 nucleotide periodicity that shifts ~5 nt between WT and *dst1* NET-seq data. These results indicated that Pol II has a greater chance of pausing at DNA positions with specific DNA helical orientation relative to the histones, but backtracks one half helical turn to DNA positions with the opposite orientation relative to the histones [159].

7. Conclusions

Extensive progress has been made thus far using molecular genetic, structural and biophysical tools to advance our understanding of Pol II mechanisms. The model organism *S. cerevisiae* has been incredibly useful for understanding the *in vivo* consequences of altered Pol II activity and how gene expression may be altered, which represents an important goal for a complete knowledge of the transcription process. Increasingly

sophisticated biophysical and genomics approaches are likely to come to the fore in Pol II studies.

Acknowledgments

Funding for work in the Kaplan laboratory at Texas A&M on eukaryotic transcription mechanisms *in vitro* and *in vivo* from NIH grant R01 GM097260 and Welch Foundation grant A-1763 to C.D.K is gratefully acknowledged. Pavel Abart is kindly acknowledged for discussion of topics in review and comments on the manuscript.

References

1. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000; 408:796–815. [PubMed: 11130711]
2. Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell*. 2005; 120:613–622. [PubMed: 15766525]
3. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. *Science*. 2005; 308:118–120. [PubMed: 15692015]
4. Allison LA, Moyle M, Shales M, Ingles CJ. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell*. 1985; 42:599–610. [PubMed: 3896517]
5. Gnatt AL, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*. 2001; 292:1876–1882. [PubMed: 11313499]
6. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science*. 2001; 292:1863–1876. [PubMed: 11313498]
7. Cheung AC, Cramer P. A Movie of RNA Polymerase II Transcription. *Cell*. 2012; 149:1431–1437. [PubMed: 22726432]
8. Spahr H, Calero G, Bushnell DA, Kornberg RD. *Schizosaccharomyces pombe* RNA polymerase II at 3.6-Å resolution. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:9185–9190. [PubMed: 19458260]
9. Hirata A, Klein BJ, Murakami KS. The X-ray crystal structure of RNA polymerase from Archaea. *Nature*. 2008; 451:851–854. [PubMed: 18235446]
10. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell*. 1999; 98:811–824. [PubMed: 10499798]
11. Murakami KS, Masuda S, Darst SA. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science*. 2002; 296:1280–1284. [PubMed: 12016306]
12. Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA. Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*. 2002; 296:1285–1290. [PubMed: 12016307]
13. 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. [PubMed: 17581591]
14. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature*. 2007; 448:157–162. [PubMed: 17581590]
15. 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 Å resolution. *Nature*. 2002; 417:712–719. [PubMed: 12000971]
16. Bataille AR, Jeronimo C, Jacques PE, Laramée L, Fortin ME, Forest A, Bergeron M, Hanes SD, Robert F. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular cell*. 2012; 45:158–170. [PubMed: 22284676]

17. Kim H, Erickson B, Luo W, Seward D, Graber JH, Pollock DD, Megee PC, Bentley DL. Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nature structural & molecular biology*. 2010; 17:1279–1286.
18. Egloff S, Dienstbier M, Murphy S. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends in genetics : TIG*. 2012; 28:333–341. [PubMed: 22622228]
19. Drogat J, Hermand D. Gene-specific requirement of RNA polymerase II CTD phosphorylation. *Molecular microbiology*. 2012; 84:995–1004. [PubMed: 22553990]
20. Buratowski S. Progression through the RNA polymerase II CTD cycle. *Molecular cell*. 2009; 36:541–546. [PubMed: 19941815]
21. Steitz TA, Smerdon SJ, Jager J, Joyce CM. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science*. 1994; 266:2022–2025. [PubMed: 7528445]
22. Yang W, Lee JY, Nowotny M. Making and breaking nucleic acids: two-Mg²⁺-ion catalysis and substrate specificity. *Molecular cell*. 2006; 22:5–13. [PubMed: 16600865]
23. 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. [PubMed: 17129781]
24. 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. [PubMed: 15537538]
25. 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. [PubMed: 15610738]
26. Kennedy SR, Erie DA. Templated nucleoside triphosphate binding to a noncatalytic site on RNA polymerase regulates transcription. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:6079–6084. [PubMed: 21447716]
27. Foster JE, Holmes SF, Erie DA. Allosteric binding of nucleoside triphosphates to RNA polymerase regulates transcription elongation. *Cell*. 2001; 106:243–252. [PubMed: 11511351]
28. Gong XQ, Zhang C, Feig M, Burton ZF. Dynamic error correction and regulation of downstream bubble opening by human RNA polymerase II. *Mol Cell*. 2005; 18:461–470. [PubMed: 15893729]
29. Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, Block SM. Direct observation of base-pair stepping by RNA polymerase. *Nature*. 2005; 438:460–465. [PubMed: 16284617]
30. Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E. A ratchet mechanism of transcription elongation and its control. *Cell*. 2005; 120:183–193. [PubMed: 15680325]
31. Svetlov V, Vassilyev DG, Artsimovitch I. Discrimination against deoxyribonucleotide substrates by bacterial RNA polymerase. *J Biol Chem*. 2004; 279:38087–38090. [PubMed: 15262972]
32. Yuzenkova Y, Bochkareva A, Tadigotla VR, Roghanian M, Zorov S, Severinov K, Zenkin N. Stepwise mechanism for transcription fidelity. *BMC Biol*. 2010; 8:54. [PubMed: 20459653]
33. Malinen AM, Turtola M, Parthiban M, Vainonen L, Johnson MS, Belogurov GA. Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic acids research*. 2012
34. Kellinger MW, Ulrich S, Chong J, Kool ET, Wang D. Dissecting Chemical Interactions Governing RNA Polymerase II Transcriptional Fidelity. *Journal of the American Chemical Society*. 2012; 134:8231–8240. [PubMed: 22509745]
35. Kaplan CD, Larsson KM, Kornberg RD. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Molecular cell*. 2008; 30:547–556. [PubMed: 18538653]
36. Kaplan CD, Jin H, Zhang IL, Belyanin A. Dissection of Pol II Trigger Loop Function and Pol II Activity-Dependent Control of Start Site Selection In Vivo. *PLoS genetics*. 2012; 8:e1002627. [PubMed: 22511879]
37. Kireeva ML, Nedialkov YA, Cremona GH, Purtov YA, Lubkowska L, Malagon F, Burton ZF, Strathern JN, Kashlev M. Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. *Mol Cell*. 2008; 30:557–566. [PubMed: 18538654]

38. Touloukhonov 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. [PubMed: 17679091]
39. Zhang J, Palangat M, Landick R. Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol*. 2010; 17:99–104. [PubMed: 19966797]
40. Temiakov D, Zenkin N, Vassilyeva MN, Perederina A, Tahirov TH, Kashkina E, Savkina M, Zorov S, Nikiforov V, Igarashi N, Matsugaki N, Wakatsuki S, Severinov K, Vassilyev DG. Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol Cell*. 2005; 19:655–666. [PubMed: 16167380]
41. Kaplan CD. The architecture of RNA polymerase fidelity. *BMC Biol*. 2010; 8:85. [PubMed: 20598112]
42. Tan L, Wiesler S, Trzaska D, Carney HC, Weinzierl RO. Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J Biol*. 2008; 7:40. [PubMed: 19055851]
43. Cheung AC, Sainsbury S, Cramer P. Structural basis of initial RNA polymerase II transcription. *The EMBO journal*. 2011; 30:4755–4763. [PubMed: 22056778]
44. Huang X, Wang D, Weiss DR, Bushnell DA, Kornberg RD, Levitt M. RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. *Proc Natl Acad Sci U S A*. 2010; 107:15745–15750. [PubMed: 20798057]
45. Seibold SA, Singh BN, Zhang C, Kireeva M, Domecq C, Bouchard A, Nazione AM, Feig M, Cukier RI, Coulombe B, Kashlev M, Hampsey M, Burton ZF. Conformational coupling, bridge helix dynamics and active site dehydration in catalysis by RNA polymerase. *Biochim Biophys Acta*. 2010; 1799:575–587. [PubMed: 20478425]
46. Da LT, Wang D, Huang X. Dynamics of pyrophosphate ion release and its coupled trigger loop motion from closed to open state in RNA polymerase II. *Journal of the American Chemical Society*. 2012; 134:2399–2406. [PubMed: 22206270]
47. Roghanian M, Yuzenkova Y, Zenkin N. Controlled interplay between trigger loop and Gre factor in the RNA polymerase active centre. *Nucleic acids research*. 2011; 39:4352–4359. [PubMed: 21266474]
48. Furman R, Sevostyanova A, Artsimovitch I. Transcription initiation factor DksA has diverse effects on RNA chain elongation. *Nucleic acids research*. 2012; 40:3392–3402. [PubMed: 22210857]
49. Cheung AC, Cramer P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature*. 2011; 471:249–253. [PubMed: 21346759]
50. 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. [PubMed: 19478184]
51. Sydow JF, Brueckner F, Cheung AC, Damsma GE, Dengl S, Lehmann E, Vassilyev D, Cramer P. Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol Cell*. 2009; 34:710–721. [PubMed: 19560423]
52. Erie DA, Hajiseyedi O, Young MC, von Hippel PH. Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science*. 1993; 262:867–873. [PubMed: 8235608]
53. Larson MH, Zhou J, Kaplan CD, Palangat M, Kornberg RD, Landick R, Block SM. Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:6555–6560. [PubMed: 22493230]
54. Weilbaecher RG, Awrey DE, Edwards AM, Kane CM. Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. *The Journal of biological chemistry*. 2003; 278:24189–24199. [PubMed: 12692127]
55. Izban MG, Luse DS. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'→5' direction in the presence of elongation factor SII. *Genes Dev*. 1992; 6:1342–1356. [PubMed: 1378419]
56. Reines D. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. *J Biol Chem*. 1992; 267:3795–3800. [PubMed: 1371280]

57. Walmacq C, Kireeva ML, Irvin J, Nedialkov Y, Lubkowska L, Malagon F, Strathern JN, Kashlev M. Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. *The Journal of biological chemistry*. 2009; 284:19601–19612. [PubMed: 19439405]
58. Walmacq C, Cheung AC, Kireeva ML, Lubkowska L, Ye C, Gotte D, Strathern JN, Carell T, Cramer P, Kashlev M. Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage. *Molecular cell*. 2012; 46:18–29. [PubMed: 22405652]
59. Kireeva ML, Opron K, Seibold SA, Domecq C, Cukier RI, Coulombe B, Kashlev M, Burton ZF. Molecular dynamics and mutational analysis of the catalytic and translocation cycle of RNA polymerase. *BMC Biophys*. 2012; 5:11. [PubMed: 22676913]
60. Awrey DE, Weilbaecher RG, Hemming SA, Orlicky SM, Kane CM, Edwards AM. Transcription elongation through DNA arrest sites. A multistep process involving both RNA polymerase II subunit RPB9 and TFIIS. *The Journal of biological chemistry*. 1997; 272:14747–14754. [PubMed: 9169440]
61. Koyama H, Ueda T, Ito T, Sekimizu K. Novel RNA polymerase II mutation suppresses transcriptional fidelity and oxidative stress sensitivity in rpb9Delta yeast. *Genes to cells : devoted to molecular & cellular mechanisms*. 2010; 15:151–159. [PubMed: 20088966]
62. Domecq C, Kireeva M, Archambault J, Kashlev M, Coulombe B, Burton ZF. Site-directed mutagenesis, purification and assay of *Saccharomyces cerevisiae* RNA polymerase II. *Protein Expr Purif*. 2010; 69:83–90. [PubMed: 19567268]
63. Kireeva ML, Domecq C, Coulombe B, Burton ZF, Kashlev M. Interaction of RNA polymerase II fork loop 2 with downstream non-template DNA regulates transcription elongation. *The Journal of biological chemistry*. 2011; 286:30898–30910. [PubMed: 21730074]
64. Miropolskaya N, Artsimovitch I, Klimasauskas S, Nikiforov V, Kulbachinskiy A. Allosteric control of catalysis by the F loop of RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:18942–18947. [PubMed: 19855007]
65. Weinzierl RO. The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. *BMC Biol*. 2010; 8:134. [PubMed: 21034443]
66. Nedialkov YA, Gong XQ, Hovde SL, Yamaguchi Y, Handa H, Geiger JH, Yan H, Burton ZF. NTP-driven translocation by human RNA polymerase II. *J Biol Chem*. 2003; 278:18303–18312. [PubMed: 12637520]
67. Feig M, Burton ZF. RNA polymerase II with open and closed trigger loops: active site dynamics and nucleic acid translocation. *Biophys J*. 2010; 99:2577–2586. [PubMed: 20959099]
68. Feig M, Burton ZF. RNA polymerase II flexibility during translocation from normal mode analysis. *Proteins*. 2010; 78:434–446. [PubMed: 19714773]
69. Malagon F, Kireeva ML, Shafer BK, Lubkowska L, Kashlev M, Strathern JN. Mutations in the *Saccharomyces cerevisiae* RPB1 gene conferring hypersensitivity to 6-azauracil. *Genetics*. 2006; 172:2201–2209. [PubMed: 16510790]
70. Yarnell WS, Roberts JW. Mechanism of intrinsic transcription termination and antitermination. *Science*. 1999; 284:611–615. [PubMed: 10213678]
71. Larson MH, Greenleaf WJ, Landick R, Block SM. Applied force reveals mechanistic and energetic details of transcription termination. *Cell*. 2008; 132:971–982. [PubMed: 18358810]
72. Weinzierl RO. Nanomechanical constraints acting on the catalytic site of cellular RNA polymerases. *Biochem Soc Trans*. 2010; 38:428–432. [PubMed: 20298196]
73. Hein PP, Palangat M, Landick R. RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. *Biochemistry*. 2011; 50:7002–7014. [PubMed: 21739957]
74. Bochkareva A, Yuzenkova Y, Tadigotla VR, Zenkin N. Factor-independent transcription pausing caused by recognition of the RNA-DNA hybrid sequence. *The EMBO journal*. 2012; 31:630–639. [PubMed: 22124324]
75. Zhang Z, Dietrich FS. Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE. *Nucleic acids research*. 2005; 33:2838–2851. [PubMed: 15905473]
76. Gleghorn ML, Davydova EK, Basu R, Rothman-Denes LB, Murakami KS. X-ray crystal structures elucidate the nucleotidyl transfer reaction of transcript initiation using two nucleotides. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:3566–3571. [PubMed: 21321236]

77. Shultzaberger RK, Chen Z, Lewis KA, Schneider TD. Anatomy of *Escherichia coli* sigma70 promoters. *Nucleic acids research*. 2007; 35:771–788. [PubMed: 17189297]
78. Surratt CK, Milan SC, Chamberlin MJ. Spontaneous cleavage of RNA in ternary complexes of *Escherichia coli* RNA polymerase and its significance for the mechanism of transcription. *Proceedings of the National Academy of Sciences of the United States of America*. 1991; 88:7983–7987. [PubMed: 1716768]
79. Uptain SM, Kane CM, Chamberlin MJ. Basic mechanisms of transcript elongation and its regulation. *Annual review of biochemistry*. 1997; 66:117–172.
80. Zenkin N, Yuzenkova Y, Severinov K. Transcript-assisted transcriptional proofreading. *Science*. 2006; 313:518–520. [PubMed: 16873663]
81. Yuzenkova Y, Zenkin N. Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. *Proc Natl Acad Sci U S A*. 2010; 107:10878–10883. [PubMed: 20534498]
82. Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol Cell*. 2010; 38:202–210. [PubMed: 20417599]
83. Trinh V, Langelier MF, Archambault J, Coulombe B. Structural perspective on mutations affecting the function of multisubunit RNA polymerases. *Microbiol Mol Biol Rev*. 2006; 70:12–36. [PubMed: 16524917]
84. Archambault J, Friesen JD. Genetics of eukaryotic RNA polymerases I, II, and III. *Microbiol Rev*. 1993; 57:703–724. [PubMed: 8246845]
85. Arndt KT, Styles CA, Fink GR. A suppressor of a *HIS4* transcriptional defect encodes a protein with homology to the catalytic subunit of protein phosphatases. *Cell*. 1989; 56:527–537. [PubMed: 2537149]
86. Archambault J, Jansma DB, Kawasoe JH, Arndt KT, Greenblatt J, Friesen JD. Stimulation of transcription by mutations affecting conserved regions of RNA polymerase II. *J Bacteriol*. 1998; 180:2590–2598. [PubMed: 9573141]
87. Kwapisz M, Wery M, Despres D, Ghavi-Helm Y, Soutourina J, Thuriaux P, Lacroute F. Mutations of RNA polymerase II activate key genes of the nucleoside triphosphate biosynthetic pathways. *The EMBO journal*. 2008; 27:2411–2421. [PubMed: 18716630]
88. Thiebaut M, Colin J, Neil H, Jacquier A, Seraphin B, Lacroute F, Libri D. Futile cycle of transcription initiation and termination modulates the response to nucleotide shortage in *S. cerevisiae*. *Molecular cell*. 2008; 31:671–682. [PubMed: 18775327]
89. Kolodziej PA, Young RA. Mutations in the three largest subunits of yeast RNA polymerase II that affect enzyme assembly. *Molecular and cellular biology*. 1991; 11:4669–4678. [PubMed: 1715023]
90. Martin C, Young RA. KEX2 mutations suppress RNA polymerase II mutants and alter the temperature range of yeast cell growth. *Molecular and cellular biology*. 1989; 9:2341–2349. [PubMed: 2668732]
91. Archambault J, Lacroute F, Ruet A, Friesen JD. Genetic interaction between transcription elongation factor TFIIIS and RNA polymerase II. *Mol Cell Biol*. 1992; 12:4142–4152. [PubMed: 1508210]
92. Hekmatpanah DS, Young RA. Mutations in a conserved region of RNA polymerase II influence the accuracy of mRNA start site selection. *Mol Cell Biol*. 1991; 11:5781–5791. [PubMed: 1922077]
93. Winston F, Chaleff DT, Valent B, Fink GR. Mutations affecting Ty-mediated expression of the *HIS4* gene of *Saccharomyces cerevisiae*. *Genetics*. 1984; 107:179–197. [PubMed: 6329902]
94. Simchen G, Winston F, Styles CA, Fink GR. Ty-mediated gene expression of the *LYS2* and *HIS4* genes of *Saccharomyces cerevisiae* is controlled by the same SPT genes. *Proc Natl Acad Sci U S A*. 1984; 81:2431–2434. [PubMed: 6326126]
95. Rando OJ, Winston F. Chromatin and transcription in yeast. *Genetics*. 2012; 190:351–387. [PubMed: 22345607]
96. Corden JL. Yeast Pol II start-site selection: the long and the short of it. *EMBO Rep*. 2008; 9:1084–1086. [PubMed: 18846104]

97. Berroteran RW, Ware DE, Hampsey M. The sua8 suppressors of *Saccharomyces cerevisiae* encode replacements of conserved residues within the largest subunit of RNA polymerase II and affect transcription start site selection similarly to sua7 (TFIIB) mutations. *Mol Cell Biol.* 1994; 14:226–237. [PubMed: 8264591]
98. Majovski RC, Khapersky DA, Ghazy MA, Ponticelli AS. A functional role for the switch 2 region of yeast RNA polymerase II in transcription start site utilization and abortive initiation. *J Biol Chem.* 2005; 280:34917–34923. [PubMed: 16081422]
99. Chen BS, Hampsey M. Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. *Mol Cell Biol.* 2004; 24:3983–3991. [PubMed: 15082791]
100. Hull MW, McKune K, Woychik NA. RNA polymerase II subunit RPB9 is required for accurate start site selection. *Genes Dev.* 1995; 9:481–490. [PubMed: 7883169]
101. Furter-Graves EM, Hall BD, Furter R. Role of a small RNA pol II subunit in TATA to transcription start site spacing. *Nucleic Acids Res.* 1994; 22:4932–4936. [PubMed: 7800482]
102. Sun ZW, Tessmer A, Hampsey M. Functional interaction between TFIIB and the Rpb9 (Ssu73) subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Nucleic acids research.* 1996; 24:2560–2566. [PubMed: 8692696]
103. Sun ZW, Hampsey M. Identification of the gene (SSU71/TFG1) encoding the largest subunit of transcription factor TFIIF as a suppressor of a TFIIB mutation in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America.* 1995; 92:3127–3131. [PubMed: 7724527]
104. Ghazy MA, Brodie SA, Ammerman ML, Ziegler LM, Ponticelli AS. Amino acid substitutions in yeast TFIIF confer upstream shifts in transcription initiation and altered interaction with RNA polymerase II. *Mol Cell Biol.* 2004; 24:10975–10985. [PubMed: 15572698]
105. Eichner J, Chen HT, Warfield L, Hahn S. Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.* 2010; 29:706–716. [PubMed: 20033062]
106. Pinto I, Ware DE, Hampsey M. The yeast SUA7 gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection in vivo. *Cell.* 1992; 68:977–988. [PubMed: 1547497]
107. Pinto I, Wu WH, Na JG, Hampsey M. Characterization of sua7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. *J Biol Chem.* 1994; 269:30569–30573. [PubMed: 7982976]
108. Bangur CS, Faitar SL, Folster JP, Ponticelli AS. An interaction between the N-terminal region and the core domain of yeast TFIIB promotes the formation of TATA-binding protein-TFIIB-DNA complexes. *The Journal of biological chemistry.* 1999; 274:23203–23209. [PubMed: 10438492]
109. Zhang DY, Carson DJ, Ma J. The role of TFIIB-RNA polymerase II interaction in start site selection in yeast cells. *Nucleic Acids Res.* 2002; 30:3078–3085. [PubMed: 12136090]
110. Pardee TS, Bangur CS, Ponticelli AS. The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. *The Journal of biological chemistry.* 1998; 273:17859–17864. [PubMed: 9651390]
111. Goel S, Krishnamurthy S, Hampsey M. Mechanism of start site selection by RNA polymerase II: interplay between TFIIB and Ssl2/XPB helicase subunit of TFIIF. *The Journal of biological chemistry.* 2012; 287:557–567. [PubMed: 22081613]
112. Hampsey M, Na JG, Pinto I, Ware DE, Berroteran RW. Extragenic suppressors of a translation initiation defect in the *cyc1* gene of *Saccharomyces cerevisiae*. *Biochimie.* 1991; 73:1445–1455. [PubMed: 1666843]
113. Struhl K. Molecular mechanisms of transcriptional regulation in yeast. *Annual review of biochemistry.* 1989; 58:1051–1077.
114. Venters BJ, Pugh BF. A canonical promoter organization of the transcription machinery and its regulators in the *Saccharomyces* genome. *Genome research.* 2009; 19:360–371. [PubMed: 19124666]

115. Huisinga KL, Pugh BF. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Molecular cell*. 2004; 13:573–585. [PubMed: 14992726]
116. Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic preinitiation complexes. *Nature*. 2012
117. Chen HT, Warfield L, Hahn S. The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol*. 2007; 14:696–703. [PubMed: 17632521]
118. Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills JC, Nilges M, Cramer P, Rappsilber J. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *The EMBO journal*. 2010; 29:717–726. [PubMed: 20094031]
119. Yang C, Ponticelli AS. Evidence that RNA polymerase II and not TFIIB is responsible for the difference in transcription initiation patterns between *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Nucleic acids research*. 2012
120. Kuehner JN, Brow DA. Quantitative analysis of in vivo initiator selection by yeast RNA polymerase II supports a scanning model. *The Journal of biological chemistry*. 2006; 281:14119–14128. [PubMed: 16571719]
121. Lue NF, Flanagan PM, Sugimoto K, Kornberg RD. Initiation by yeast RNA polymerase II at the adenoviral major late promoter in vitro. *Science*. 1989; 246:661–664. [PubMed: 2510298]
122. Hubert JC, Guyonvarch A, Kammerer B, Exinger F, Liljelund P, Lacroute F. Complete sequence of a eukaryotic regulatory gene. *The EMBO journal*. 1983; 2:2071–2073. [PubMed: 6139279]
123. Exinger F, Lacroute F. 6-Azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. *Curr Genet*. 1992; 22:9–11. [PubMed: 1611672]
124. Howe KJ, Kane CM, Ares M Jr. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in *Saccharomyces cerevisiae*. *Rna*. 2003; 9:993–1006. [PubMed: 12869710]
125. Kulish D, Struhl K. TFIIS enhances transcriptional elongation through an artificial arrest site in vivo. *Molecular and cellular biology*. 2001; 21:4162–4168. [PubMed: 11390645]
126. Mason PB, Struhl K. Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol Cell*. 2005; 17:831–840. [PubMed: 15780939]
127. Grigull J, Mnaimneh S, Pootoolal J, Robinson MD, Hughes TR. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Molecular and cellular biology*. 2004; 24:5534–5547. [PubMed: 15169913]
128. Riles L, Shaw RJ, Johnston M, Reines D. Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil. *Yeast*. 2004; 21:241–248. [PubMed: 14968429]
129. Desmoucelles C, Pinson B, Saint-Marc C, Daignan-Fornier B. Screening the yeast “disruptome” for mutants affecting resistance to the immunosuppressive drug, mycophenolic acid. *The Journal of biological chemistry*. 2002; 277:27036–27044. [PubMed: 12016207]
130. Jenks MH, O'Rourke TW, Reines D. Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast. *Mol Cell Biol*. 2008; 28:3883–3893. [PubMed: 18426909]
131. Kopcewicz KA, O'Rourke TW, Reines D. Metabolic regulation of IMD2 transcription and an unusual DNA element that generates short transcripts. *Mol Cell Biol*. 2007; 27:2821–2829. [PubMed: 17296737]
132. Jenks MH, Reines D. Dissection of the molecular basis of mycophenolate resistance in *Saccharomyces cerevisiae*. *Yeast*. 2005; 22:1181–1190. [PubMed: 16278936]
133. Hyle JW, Shaw RJ, Reines D. Functional distinctions between IMP dehydrogenase genes in providing mycophenolate resistance and guanine prototrophy to yeast. *J Biol Chem*. 2003; 278:28470–28478. [PubMed: 12746440]
134. Wind-Rotolo M, Reines D. Analysis of gene induction and arrest site transcription in yeast with mutations in the transcription elongation machinery. *J Biol Chem*. 2001; 276:11531–11538. [PubMed: 11278887]

135. Shaw RJ, Wilson JL, Smith KT, Reines D. Regulation of an IMP dehydrogenase gene and its overexpression in drug-sensitive transcription elongation mutants of yeast. *J Biol Chem.* 2001; 276:32905–32916. [PubMed: 11441018]
136. Shaw RJ, Reines D. *Saccharomyces cerevisiae* transcription elongation mutants are defective in PUR5 induction in response to nucleotide depletion. *Mol Cell Biol.* 2000; 20:7427–7437. [PubMed: 11003640]
137. Gomez-Herreros F, de Miguel-Jimenez L, Morillo-Huesca M, Delgado-Ramos L, Munoz-Centeno MC, Chavez S. TFIIS is required for the balanced expression of the genes encoding ribosomal components under transcriptional stress. *Nucleic acids research.* 2012
138. Davis CA, Ares M Jr. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.* 2006; 103:3262–3267. [PubMed: 16484372]
139. Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol Cell.* 2006; 24:735–746. [PubMed: 17157256]
140. Kuehner JN, Brow DA. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol Cell.* 2008; 31:201–211. [PubMed: 18657503]
141. Martens JA, Wu PY, Winston F. Regulation of an intergenic transcript controls adjacent gene transcription in *Saccharomyces cerevisiae*. *Genes Dev.* 2005; 19:2695–2704. [PubMed: 16291644]
142. Martens JA, Laprade L, Winston F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. *Nature.* 2004; 429:571–574. [PubMed: 15175754]
143. Thebault P, Boutin G, Bhat W, Rufiange A, Martens J, Nourani A. Transcription regulation by the noncoding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNA polymerase II. *Molecular and cellular biology.* 2011; 31:1288–1300. [PubMed: 21220514]
144. Pruneski JA, Hainer SJ, Petrov KO, Martens JA. The Paf1 complex represses SER3 transcription in *Saccharomyces cerevisiae* by facilitating intergenic transcription-dependent nucleosome occupancy of the SER3 promoter. *Eukaryotic cell.* 2011; 10:1283–1294. [PubMed: 21873510]
145. Hainer SJ, Pruneski JA, Mitchell RD, Monteverde RM, Martens JA. Intergenic transcription causes repression by directing nucleosome assembly. *Genes & development.* 2011; 25:29–40. [PubMed: 21156811]
146. Kristjuhan A, Svejstrup JQ. Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo. *The EMBO journal.* 2004; 23:4243–4252. [PubMed: 15457216]
147. Pelechano V, Jimeno-Gonzalez S, Rodriguez-Gil A, Garcia-Martinez J, Perez-Ortin JE, Chavez S. Regulon-specific control of transcription elongation across the yeast genome. *PLoS Genet.* 2009; 5:e1000614. [PubMed: 19696888]
148. Rodriguez-Gil A, Garcia-Martinez J, Pelechano V, Munoz-Centeno Mde L, Geli V, Perez-Ortin JE, Chavez S. The distribution of active RNA polymerase II along the transcribed region is gene-specific and controlled by elongation factors. *Nucleic acids research.* 2010; 38:4651–4664. [PubMed: 20385590]
149. Selth LA, Sigurdsson S, Svejstrup JQ. Transcript Elongation by RNA Polymerase II. *Annual review of biochemistry.* 2010; 79:271–293.
150. Kuehner JN, Pearson EL, Moore C. Unravelling the means to an end: RNA polymerase II transcription termination, *Nature reviews. Molecular cell biology.* 2011; 12:283–294. [PubMed: 21487437]
151. Mayer A, Heidemann M, Lidschreiber M, Schrieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science.* 2012; 336:1723–1725. [PubMed: 22745433]
152. Marquardt S, Hazelbaker DZ, Buratowski S. Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *Transcription.* 2011; 2:145–154. [PubMed: 21826286]
153. Loya TJ, O'Rourke TW, Reines D. A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3. *Nucleic acids research.* 2012

154. Izban MG, Luse DS. Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev.* 1991; 5:683–696. [PubMed: 2010092]
155. Duina AA. Histone Chaperones Spt6 and FACT: Similarities and Differences in Modes of Action at Transcribed Genes. *Genet Res Int.* 2011; 2011:625210. [PubMed: 22567361]
156. Schwabish MA, Struhl K. Evidence for eviction and rapid deposition of histones upon transcriptional elongation by RNA polymerase II. *Mol Cell Biol.* 2004; 24:10111–10117. [PubMed: 15542822]
157. Churchman LS, Weissman JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature.* 2011; 469:368–373. [PubMed: 21248844]
158. Bondarenko VA, Steele LM, Ujvari A, Gaykalova DA, Kulaeva OI, Polikanov YS, Luse DS, Studitsky VM. Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol Cell.* 2006; 24:469–479. [PubMed: 17081995]
159. Brogaard K, Xi L, Wang JP, Widom J. A map of nucleosome positions in yeast at base-pair resolution. *Nature.* 2012; 486:496–501. [PubMed: 22722846]
160. Bushnell DA, Cramer P, Kornberg RD. Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 Å resolution. *Proc Natl Acad Sci U S A.* 2002; 99:1218–1222. [PubMed: 11805306]
161. Armache KJ, Kettenberger H, Cramer P. Architecture of initiation-competent 12-subunit RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America.* 2003; 100:6964–6968. [PubMed: 12746495]
162. Bushnell DA, Kornberg RD. Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc Natl Acad Sci U S A.* 2003; 100:6969–6973. [PubMed: 12746498]
163. Kettenberger H, Armache KJ, Cramer P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell.* 2003; 114:347–357. [PubMed: 12914699]
164. Bushnell DA, Westover KD, Davis RE, Kornberg RD. Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Å. *Science.* 2004; 303:983–988. [PubMed: 14963322]
165. Westover KD, Bushnell DA, Kornberg RD. Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science.* 2004; 303:1014–1016. [PubMed: 14963331]
166. Armache KJ, Mitterweger S, Meinhart A, Cramer P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. *The Journal of biological chemistry.* 2005; 280:7131–7134. [PubMed: 15591044]
167. Tuske S, Sarafianos SG, Wang X, Hudson B, Sineva E, Mukhopadhyay J, Birktoft JJ, Leroy O, Ismail S, Clark AD, Dharia C, Napoli A, Laptenko O, Lee J, Borukhov S, Ebright RH, Arnold E. Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell.* 2005; 122:541–552. [PubMed: 16122422]
168. Lehmann E, Brueckner F, Cramer P. Molecular basis of RNA-dependent RNA polymerase II activity. *Nature.* 2007; 450:445–449. [PubMed: 18004386]
169. Damsa GE, Alt A, Brueckner F, Carell T, Cramer P. Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nature structural & molecular biology.* 2007; 14:1127–1133.
170. Brueckner F, Hennecke U, Carell T, Cramer P. CPD damage recognition by transcribing RNA polymerase II. *Science.* 2007; 315:859–862. [PubMed: 17290000]
171. 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. [PubMed: 18552824]
172. Kostrewa D, Zeller ME, Armache KJ, Seizl M, Leike K, Thomm M, Cramer P. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. *Nature.* 2009; 462:323–330. [PubMed: 19820686]
173. Liu X, Bushnell DA, Wang D, Calero G, Kornberg RD. Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. *Science.* 2010; 327:206–209. [PubMed: 19965383]

174. Tagami S, Sekine S, Kumarevel T, Hino N, Murayama Y, Kamegamori S, Yamamoto M, Sakamoto K, Yokoyama S. Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature*. 2010; 468:978–982. [PubMed: 21124318]
175. Wang D, Zhu G, Huang X, Lippard SJ. X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:9584–9589. [PubMed: 20448203]
176. Liu X, Bushnell DA, Silva DA, Huang X, Kornberg RD. Initiation complex structure and promoter proofreading. *Science*. 2011; 333:633–637. [PubMed: 21798951]
177. Ruan W, Lehmann E, Thomm M, Kostrewa D, Cramer P. Evolution of two modes of intrinsic RNA polymerase transcript cleavage. *The Journal of biological chemistry*. 2011; 286:18701–18707. [PubMed: 21454497]
178. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010

Abbreviations

Pol II	RNA Polymerase II
TL	trigger loop
BH	bridge helix
FRET	Förster Resonance Energy Transfer
NTP	nucleoside triphosphate
EC	elongation complex

Highlights

- Overview of basic mechanisms of RNA polymerase II nucleotide addition and translocation cycles
- Summary of structural analyses on multisubunit RNA polymerases
- Historical overview of molecular genetics of RNA polymerase II in the *S. cerevisiae* model system
- Discussion of gene expression defects that arise from altered RNA polymerase II activity in *S. cerevisiae*

Figure 1. Architecture and subunit composition of *S. cerevisiae* Pol II

a) *S. cerevisiae* Pol II comprises 12 highly conserved subunits assembled in a highly globular fashion. The Pol II structure resembles a crab claw with the two largest subunits, Rpb1 and Rpb2, forming each pincer, between which nucleic acids are bound. **b)** Disassembled view of Pol II subunits showing individual subunits and position of nucleic acids within a Pol II EC relative to the largest subunit, Rpb1. This figure is a composite of two Pol II structures, 2E2H [23] and 3HOU [51] (see Table 1) and was prepared using Pymol [178].

Figure 2. Nucleotide addition and translocation cycles of Pol II

a) Closed active site following addition of NMP. Catalytic Mg^{2+} ions (purple filled circles) are labeled A and B. Trigger loop (magenta) is closed over active site and would be expected to interact with both RNA 3'-NMP and pyrophosphate (connected orange balls). **b)** Opening of the TL in (a) allows pyrophosphate release and together with hypothetical BH (purple) kinking allows Pol II translocation. BH appears to unkink at some point from kinked state, but could still have some dynamic properties. Template nucleotide at position $i+1$ can flip into vacated active site while NTP (orange) substrate may bind non-specifically in the "E" site. Mg^{2+} B is recruited along with the NTP substrate. **c)** Basepairing of NTP with $i+1$ template base allows binding in the "A" site. N-terminal TL helix contacts the NTP. BH omitted for clarity in (c)-(f). **d)** Partial folding of the TL brings it into further contact with NTP. **e)** Complete folding of TL into full interaction with NTP substrate and subsequent catalysis. **f)** TL opening allows pyrophosphate and presumably Mg^{2+} B to be released prior or concomitant with translocation to repeat the cycle.

Figure 3. Pol II trigger loop supports multiple conformations and possible stepwise interaction with NTP substrates

a) Overlay of TLs obtained from multiple Pol II crystal structures with TLs in different conformations. Closed TL from substrate bound 2E2H Pol II EC structure (magenta) [23]. TL from *Sch. pombe* 3H0G structure lacking nucleic acids (yellow) [8]. TL from α -amanitin bound Pol II 3CQZ structure lacking nucleic acids (orange) [35] or Pol II EC 2VUM [171] (blue) structures. α -amanitin from 3CQZ structure shown in orange. α -amanitin binds under the BH (blue, opaque, from 2VUM) and interacts with the TL through Rpb1 H1085, trapping it from closing the active site. In structure 2VUM, the TL (blue) is additionally observed in a wedged conformation where Rpb1 L1081 is between a TL helix and the BH. Open TL from structure 1Y1V [25] (pink) where TFIIS (not shown) has inserted into the Pol II active site, restraining the TL from closing. **(b)-(d)** Interactions with GTP substrate (orange, (c) and (d)) and TL (magenta) may proceed in a stepwise fashion. TL from a single Pol II EC structure with closed TL (2E2H, magenta, same as **3a**) is used to depict the different amino acids proximal to substrate binding A-site in putative steps of TL folding/stabilization, as deduced from multiple Pol II crystal structures and genetic experiments. In absence of substrate, TL is mostly unfolded but Rpb1 residue Q1078 of the TL and Rbp1 active site residues N479 and R446 are positioned for substrate interaction. **c)** In presence of substrate, TL can sometimes be observed partially folded where L1081 and N1082 are positioned for substrate interaction. **d)** Upon complete folding of the TL, H1085 is additionally positioned to interact with substrate and proposed to enhance catalysis. This figure was created using Pymol [178].

Figure 4. Pol II trigger loop and bridge helix domains are surrounded by additional regions that may communicate with them or NTP substrates

Two views of the heart of the Pol II structure assembled from EC structures 2E2H (closed TL structure) [23] and 1Y1V (open TL structure, TFIIIS bound (TFIIIS not shown)) [25].

Amino acids (residue numbers in parentheses) shown in surface view for the following Pol II regions or domains: Rpb1 Switch 1 (1385-1407), Rpb1 F Loop (747-810), Rpb1 α 46- α 47 Loop (1354-1370), Rpb1 α 20- α 21 Loop (696-719), Rpb2 β -DII Loop (526-538), Rpb2 Fork Loop 2 (only ends of loop depicted because loop is not visible completely in this structure, 500-502+509-512), Rpb2 Link Domain (757-776), Rpb2 “E-site” Loop contacting NTP (1014-1020), and Rpb9 Loop (86-101). Closed trigger loop shown in cartoon (magenta, Rpb1 1070-1106). Open trigger loop shown in cartoon (pink, Rpb1 1070-1106). Bridge helix shown in cartoon (cyan, Rpb1 810-846).

Figure 5. Many genes sensitized to altered transcription *in vivo* have interacting transcription units

(a) and (b) Reporter loci *lys2-128* (a) and *his4-912* (b) for the *Spt⁻* phenotype. Each of these genes has a Ty1 Δ -element insertion that compromises transcription in *Spt⁺* (wild type) cells. Mutations conferring the *Spt⁻* phenotype allow transcription from alternate start sites or promoters for one or both genes, allowing suppression of auxotrophy for lysine (a) or histidine (b) conferred by the insertions in *LYS2* or *HIS4*, respectively. *Spt⁻* mutations are widely found in Pol II transcription factors as well as some in Pol II subunits themselves. (c) and (d) Genes in the pathways for GTP synthesis, *IMD2* (c) or UTP synthesis, *URA2* (d) are regulated through attenuation by premature termination of transcripts over upstream regions. Under high GTP conditions, *IMD2* transcripts originate from upstream GTP-initiating start sites, but are attenuated by premature termination. Upon GTP starvation (induced by MPA or 6-AU treatment, for example), *IMD2* can be expressed from a downstream ATP-initiating start site that is not attenuated by premature termination. Mutations in a large number of transcription factors, including Pol II subunits are sensitive to nucleotide depleting drugs because their ability to induce *IMD2* is compromised. Additionally, some Pol II mutations, among other transcription factor mutants, constitutively express *IMD2*. An upstream *URA2* transcript is also regulated by premature termination, but upon uracil starvation, a downstream promoter can be used to make productive *URA2* mRNA. Initiating nucleotide does not appear to be sensed during *URA2* regulation, however, the upstream *URA2* transcript is U-rich and transcription defects in this region may allow the downstream productive promoter to be used. Some Pol II mutants constitutively express *URA2*. (e) The *SRG1-SER3* locus involves transcription of a non-coding RNA (*SRG1*) that causes repression of *SER3* through cotranscriptional nucleosome deposition over the *SER3* promoter. *SER3* is upregulated in a large number of chromatin and transcription factor mutants illustrating the sensitivity of this type of arrangement to multiple transcription defects.

Table I

Ref.	Title	Description	Species	Key PDB IDs	Year
[10]	Crystal structure of <i>Thermus aquaticus</i> core RNA polymerase at 3.3 Å resolution	Structure of <i>T. aquaticus</i> RNAP lacking sigma subunit.	<i>T. aquaticus</i>	1HQM (supersedes 1DDQ)	1999
[6]	Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution	First msRNAP EC structure (10 subunit form of Pol II). Shows Pol II in pre-translocated state.	<i>S. cerevisiae</i>	1I3Q, 1I50	2001
[5]	Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution	First Pol II structure at near atomic resolution (10 subunit form of the enzyme).	<i>S. cerevisiae</i>	1I6H	2001
[160]	Structural basis of transcription: α -amanitin-RNA polymerase II cocrystal at 2.8 Å resolution	Atomic description of the binding site of α -amanitin, a Pol II elongation inhibitor. Proposed to alter translocation dynamics by binding near bridge helix (now known that α -amanitin also targets the TL).	<i>S. cerevisiae</i>	1K83	2002
[11]	Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution	-	<i>T. aquaticus</i>	1L9U	2002
[12]	Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex	-	<i>T. aquaticus</i>	1L9Z	2002
[15]	Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution	-	<i>T. thermophilus</i>	1IW7	2002
[161]	Architecture of initiation-competent 12-subunit RNA polymerase II	Structure of 12 subunit Pol II.	<i>S. cerevisiae</i>	1INT9	2003
[162]	Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription	Structure of 12 subunit Pol II.	<i>S. cerevisiae</i>	1NIK	2003
[163]	Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage	Structure of 12 subunit Pol II with TFIIS, shows insertion of TFIIS into Pol II active site through secondary channel.	<i>S. cerevisiae</i>	1PQV	2003
[164]	Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms	Pol II (10 subunit form) costructure with TFIIB in absence of nucleic acids. Reveals apparent insertion of TFIIB domain into RNA-DNA hybrid binding region of Pol II.	<i>S. cerevisiae</i>	1R5U	2004
[165]	Structural basis of transcription: separation of RNA from DNA by RNA polymerase II	Pol II (10 subunit form) EC showing upstream edge of RNA-DNA hybrid and mechanism separation of RNA and template DNA.	<i>S. cerevisiae</i>	1SFO	2004
[24]	Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center	Pol II (10 subunit form) ECs with or without matched and mismatched NTP substrates. Structures are post-translocated allowing active site access for NTP substrates. Two binding modes for substrates described. A-site binding for matched substrates and E-site binding for mismatched NTP. TL not completely folded or observed even in presence of matched substrate.	<i>S. cerevisiae</i>	1R9S (EC with matched NTP), 1R9T (EC with mismatched NTP), 1TW7 (apo-enzyme with UTP at 2.3 Å resolution)	2004

Ref.	Title	Description	Species	Key PDB IDs	Year
[25]	Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIS	Pol II (12 subunit form) ECs with substrate analog or TFIS. Description of possible pre-insertion site. TL not completely folded or observed even in presence of matched substrate.	<i>S. cerevisiae</i>	1Y1V (refined TFIS-Pol II structure), 1Y1Y (TFIS-Pol II EC), 1Y77 (EC with matched GMPCPP), 1Y1W (12 subunit EC)	2004
[166]	Structures of Complete RNA Polymerase II and its Subcomplex, Rpb4/7	Updated, refined structure of 12 subunit Pol II.	<i>S. cerevisiae</i>	1WCM	2005
[167]	Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation	Structure of RNAP with elongation inhibitor Streptolydigin.	<i>T. thermophilus</i>	1ZYR	2005
[40]	Structural basis of transcription inhibition by antibiotic streptolydigin	Structure of RNAP with elongation inhibitor Streptolydigin, shown to require TL for action.	<i>T. thermophilus</i>	2A6H	2005
[23]	Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis	First msRNAP EC structure with matched substrate to show TL folding/substrate contacts.	<i>S. cerevisiae</i>	2E2H (EC with matched GTP)	2006
[14]	Structural basis for transcription elongation by bacterial RNA polymerase	First EC structure of bacterial RNAP.	<i>T. thermophilus</i>	2O51	2007
[13]	Structural basis for substrate loading in bacterial RNA polymerase	RNAP EC structure showing TL in similar apparent substrate-dependent conformation as in Pol II structure from REF. Also shows that antibiotic Streptolydigin restricts this TL conformation.	<i>T. thermophilus</i>	2O51, 2PPB (EC with substrate analog and Streptolydigin)	2007
[168]	Molecular bases of RNA-dependent RNA polymerase II activity	Pol II complexes bound to dsRNA scaffolds, including a hepatitis delta virus derived stem loop	<i>S. cerevisiae</i>	2R92, 2R93	2007
[169]	Mechanism of transcriptional stalling at cisplatin-damaged DNA	Pol II EC structure containing a cisplatin DNA adduct	<i>S. cerevisiae</i>	2R7Z	2007
[170]	CPD damage recognition by transcribing RNA polymerase II	Pol II EC structure stalled at cyclo-pyrimidine dimer (CPD) DNA lesion	<i>S. cerevisiae</i>	2IA5, 2IA6, 2JA7, 2JA8	2007
[35]	The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by α -amanitin	Refined Pol II structure (10 subunit form) showing α -amanitin interaction with TL through conserved Rpb1 His1085 residue along with biochemical proof that TL-interaction is required for inhibition of Pol II elongation by α -amanitin.	<i>S. cerevisiae</i>	3CQZ	2008
[171]	Structural basis of transcription inhibition by α -amanitin and implications for RNA polymerase II translocation	New Pol II EC structure (12 subunit) bound to α -amanitin. Illustrates both α -amanitin-TL interaction and alternative TL conformation with Rpb1 Leu1081 wedged between TL and BH. Discussed as a hypothetical translocation intermediate.	<i>S. cerevisiae</i>	2VUM	2008
[9]	The X-ray crystal structure of RNA polymerase from Archaea	First RNAP structure from archaeal species	<i>Sulfolobus solfataricus</i>	2PMZ	2008
[50]	Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution	Pol II EC complexes made from reconstituted nucleic acid scaffolds containing mismatched 3' RNA ends to examine conformations of mismatched RNA for the purpose of simulating naturally backtracked complexes with RNA 3' ends displaced from Pol II active site.	<i>S. cerevisiae</i>	3GTG (EC with 12-mer RNA), 3GTJ (EC with 13-mer RNA), 3GTK (EC with 18-mer RNA), 3GTQ (EC backtracked due to DNA damage)	2009

Ref.	Title	Description	Species	Key PDB IDs	Year
[51]	Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA	Pol II EC complexes made from reconstituted nucleic acid scaffolds containing mismatched 3' RNA ends to examine conformations of mismatched RNA for the purpose of simulating naturally backtracked complexes with RNA 3' ends displaced from Pol II active site.	<i>S. cerevisiae</i>	3HOU (EC with T-U mismatch), 3HOW (EC with frayed T-U mismatch), 3HOV, 3HOZ (EC with frayed 3'-GMP), also EC strictures 3HOV, 3HOX, 3HOY	2009
[172]	RNA polymerase II-TFIIB structure and mechanism of transcription initiation	Pol II (12 subunit) structure with TFIIB showing alternate conformation of TFIIB than in REF	<i>S. cerevisiae</i>	3KIF	2009
[8]	<i>Schizosaccharomyces pombe</i> RNA polymerase II at 3.6-Å resolution	12 subunit Pol II from <i>Schizosaccharomyces pombe</i> .	<i>Sch. pombe</i>	3HOG	2009
[173]	Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism	Pol II (10 subunit) structure with TFIIB showing alternate conformation of TFIIB than in REF	<i>S. cerevisiae</i>	3K7A	2010
[174]	Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein	RNAP structure with bound Gfh1, a Gre homolog that inhibits transcription by binding the secondary channel. Shows "ratcheted conformation" with Gfh1 capturing a never before observed RNAP conformation with bent bridge helix and domain rotation. Theorized to relate to a possible translocation intermediate.	<i>T. thermophilus</i>	3AOH	2010
[175]	X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct	Pol II EC structure containing a cisplatin-related antineoplastic DNA adduct.	<i>S. cerevisiae</i>	3M3Y, 3M4O	2010
[49]	Structural basis of RNA polymerase II backtracking, arrest and reactivation	Pol II EC with Pol II-synthesized poly-cytosine RNA in 8 nt backtracked conformation showing path of putative native (not reconstituted) backtracks. Includes TL contacts with backtracked RNA. Also shows displacement of backtracked RNA upon TFIIS binding and hypothetical cleavage product upon TFIIS action.	<i>S. cerevisiae</i>	3PO2 (EC with long backtrack), 3PO3 (Backtracked EC with TFIIS)	2011
[176]	Initiation complex structure and promoter proofreading	Series of structures of Pol II with nucleic acids including series of short RNAs, some with chain-terminated RNA and NTP substrate. Shows that RNA-DNA hybrid with short RNAs is distorted in structure, suggesting possible transition to more stable, regular structure upon synthesis of longer RNA.	<i>S. cerevisiae</i>	3RZO, 3S14, 3S15, 3S16, 3S17, 3S1M, 3S1N, 3S1Q, 3S1R, 3S2D, 3S2H	2011
[43]	Structural basis of initial RNA polymerase II transcription	Series of structures of Pol II with nucleic acids including series of short RNAs, with or without non-hydrolysable substrate analog. Shows that RNA-DNA hybrid with short RNAs is distorted in structure, suggesting possible transition to more stable, regular structure upon synthesis of longer RNA.	<i>S. cerevisiae</i>	4A3B, 4A3C, 4A3D, 4A3E, 4A3F, 4A3G, 4A3I, 4A3J, 4A3K, 4A3L, 4A3M	2011
[177]	Evolution of two modes of intrinsic RNA polymerase transcript cleavage	Pol II structure with chimeric Rpb9-C11 (C11 is a Pol III subunit)	<i>S. cerevisiae</i>	3QT1	2011
[58]	Mechanism of Translesion Transcription by RNA Polymerase II and Its Role in Cellular Resistance to DNA Damage	New Pol II EC structure with CPD lesion partially translocated into active site	<i>S. cerevisiae</i>	4A93	2012