RNA Synthesis

Introductory article

Ronny I Drapkin, *Dana Farber Cancer Institute, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA*

Danny F Reinberg, *Howard Hughes Medical Institute, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, Piscataway, New Jersey, USA*

Article Contents

Transcription is the process of RNA synthesis in which the information stored in DNA is converted to RNA by an enzyme called RNA polymerase. Transcription constitutes a complex reaction whose essential features have been evolutionarily preserved from bacteria to mammals.

Introduction

Gene expression can be regulated at multiple levels, including transcription, RNA processing, messenger RNA (mRNA) stability, translation and posttranslation. In this processional hierarchy, transcription provides the first opportunity for regulation. In fact, the initial step in regulating gene expression is deciding whether or not to transcribe a gene. Transcription is the process of RNA synthesis in which the genetic information that is stored in the nucleus in the form of double-stranded DNA is converted into a single-stranded RNA chain by an enzyme called RNA polymerase (RNAP). The RNA chain is identical to one strand of the DNA, called the coding strand, and complementary to the other strand which provides the template for its synthesis. In prokaryotes, a single form of RNAP shoulders the responsibility of transcribing the genomic DNA. Eukaryotic cells have evolved to accommodate their increased genetic complexity by delegating this function to three nuclear DNA-dependent RNAPs, each responsible for the synthesis of different classes of RNA. Despite the partitioning of the transcription workload in eukaryotic cells, the fundamental mechanics of transcription are conserved from bacteria to mammals.

RNA Polymerase

Most RNA polymerases (RNAPs), with the exception of those encoded by bacteriophages, are multisubunit enzymes. Analysis of subunit composition reveals a trend that mirrors the evolutionary divergence of prokaryotes and eukaryotes. The prokaryotes are composed of two distinct groups: Bacteria and Archaea. Consistent with the fact that Bacteria diverged before Archaea and Eukarya split, the subunit composition of archaeal RNAP is more reminiscent of eukaryotic RNAP than of bacterial RNAP (reviewed in Bell and Jackson, 1998).

The core bacterial RNAP consists of five proteins (two α subunits, one β', one β, and one ω subunit). *In vitro*, this five-subunit enzyme is capable of RNA synthesis from nonpromoter DNA in the absence of additional factors. *In vivo*, the nonspecific activity of core RNAP is modulated by sigma (σ) factors, which are essential for the enzyme's function. Sigma factors are integral components of the complete bacterial RNAP or holoenzyme and function to recognize specific DNA sequences, called promoter elements, located immediately upstream of the transcription start site. In the context of sigma factor, the α subunit is capable of recognizing promoter elements and responding to various regulatory factors. The β and β' subunits together make up the catalytic centre of RNAP.

The pioneering work of R.G. Roeder and P. Chambon identified three mammalian RNAPs (RNAP I, II and III) by their elution profiles on ion-exchange chromatography (DEAE-Sephadex) and later by their differential sensitivity to the bicyclic octapeptide α-amanitin. In all systems studied, RNAP II is rapidly inhibited by low concentrations of α-amanitin; RNAP I is the most resistant; and RNAP III is inhibited at intermediate concentrations. Since these landmark studies, researchers have been able to define further the precise role of each mammalian RNAP. RNAP I transcribes the multicopy genes encoding large ribosomal RNAs (28S, 18S and 5.8S); RNAP II transcribes all protein-coding genes, as well as some small nuclear RNAs (snRNA); and RNAP III transcribes the genes for transfer RNAs, 5S ribosomal RNAs and some snRNAs.

The archaeal RNAP and the three eukaryotic RNAPs are complex enzymes, consisting of 8–14 different subunits. Although the three eukaryotic RNAPs recognize different promoters and transcribe different classes of genes, they share several common features. The two largest subunits of all three eukaryotic RNAPs are related to the B and A′ subunits of archaeal RNAP and to the β and β' subunits of bacterial RNAP. Interestingly, the ω subunit of bacterial RNAP was recently shown to be the homologue of Rbp6, an essential subunit shared by eukaryotic RNAP I, II and III (Minakhin *et al.*, 2001). Rbp6 and ω appear to

promote RNAP assembly and stability. In addition, five subunits of the eukaryotic RNAPs are common to all three eukaryotic RNAP enzymes. Like the core bacterial RNAP, the purified eukaryotic RNAPs can undertake template-dependent transcription of RNA, but are not able to initiate selectively at promoters. However, unlike the bacterial holoenzyme, eukaryotic RNAPs require additional factors to accurately initiate transcription from promoter start sites. This is also true in archaeal RNA synthesis. Though the specifics for each RNAP may differ, the general function of the accessory protein factors is to deliver RNAP to the promoter region and position it over the transcription start site.

The Transcription Cycle

The transcription reaction is highly conserved between prokaryotes and eukaryotes. It begins with the binding of RNAP to specific promoter elements. Although RNAP can interact with DNA, it cannot recognize specific promoter DNA elements on its own. The delivery of RNAP to promoter elements is facilitated by auxiliary factors unique to each RNAP. Once at the promoter, RNAP forms a tight stable complex with the DNA. In the presence of nucleotides, this stable complex can initiate RNA synthesis. RNA synthesis occurs within a 'transcription bubble' in which the DNA duplex is temporarily separated into single strands, and one strand serves as the template for synthesis of RNA. As RNAP moves along the DNA, it unwinds the DNA in front of it and rewinds the DNA behind it. In this way, the size of the 'transcription bubble' remains relatively constant while RNAP elongates the RNA chain. As a consequence, the RNA : DNA hybrid within the 'transcription bubble' is short.

The length of the RNA : DNA hybrid has been a topic of intense debate. Most studies are consistent with a hybrid that is approximately 9–12 nucleotides in length. This short RNA : DNA hybrid is important when one considers the processivity of RNAP. The processivity of an enzyme is defined as the ability of the enzyme to continue to act on the substrate and not dissociate between repetitions of the catalytic event. If RNAP dissociates from the template during transcription, the short hybrid would be too unstable and the RNA would be released. Under these circumstances, there is no mechanism for the polymerase to reassociate and continue transcribing where it left off. This is in contrast to DNA synthesis where the newly synthesized DNA is base paired to the template. In principle, if DNA polymerase dissociated from the template, it could conceivably reassociate and resume synthesis where it left off. Therefore, processivity is extremely important during RNA synthesis.

As RNAP translocates along DNA, the RNA chain is extended until the polymerase reaches a terminator sequence. The terminator sequence signifies the end of the gene and defines a transcription unit as the RNA chain that is synthesized from DNA sequences that begin at the promoter start site and end at the terminator. At this point, the ternary complex of RNAP, RNA, and DNA dissociates and RNAP is able to recycle for another round of transcription.

The details of the RNA synthesis reaction can be divided into six discrete steps: (1) promoter engagement, (2) transition from a closed to opened RNAP : promoter complex, (3) synthesis of initial phosphodiester bond and abortive initiation, (4) promoter clearance, (5) elongation and (6) termination and RNAP recycling (**Figure 1**). Except where noted, these steps are common to all RNAPs.

Promoter engagement

In bacteria, transcription is a process that is largely unregulated. Nevertheless, there are a number of bacterial genes whose expression is governed by regulatory proteins. One of the simplest and most well-studied bacterial regulators is the *Escherichia coli* catabolite activator protein (CAP). In the presence of the allosteric effector cyclic adenosine monophosphate (cAMP), CAP functions by binding to specific DNA sites near target promoters and enhancing the ability of RNAP holoenzyme to bind and initiate transcription. CAP-dependent promoters can be grouped into three classes. Class I promoters require only CAP for transcription activation. In these promoters, the DNA-binding site for CAP is upstream of the binding site for RNAP. Class II promoters also only need CAP for transcription activation but in these promoters the CAP-binding site overlaps with the binding site for RNAP. Class III promoters require multiple regulatory proteins including two or more CAP molecules. In all cases, the mechanism is conserved. CAP binds its cognate DNA element and makes specific protein–protein contacts with the C-terminus of the α subunit of RNAP. This interaction enables the α subunit to make specific contacts with the promoter DNA and position the catalytic subunits of RNAP over the transcription start site.

In eukaryotic systems, the mechanism is conserved but the specifics are different. Since eukaryotic RNAPs cannot initiate RNA synthesis at promoter DNA elements on their own, they need the assistance of accessory factors. These factors are called general transcription factors (GTFs). Their function is to make specific protein–DNA and protein–protein contacts to escort RNAP to the promoter. In many ways, these factors serve a similar function to the sigma factors of the bacterial holoenzyme. RNAP II requires five GTFs to accurately initiate RNA synthesis from most promoters of protein coding genes: TFIIB, TFIID, TFIIE, TFIIF and TFIIH. TFIID is a multi-subunit complex composed of the TATA-binding protein (TBP) and TBP-associated factors (TAFs). TBP is required

for transcription of all genes, including RNAP I and RNAP III genes. TAFs are required at two levels: (1) for general basal levels of transcription and (2) to mediate activated transcription. In a manner analogous to the recruitment of bacterial RNAP to the promoter by CAP, transcriptional regulators can target GTFs to facilitate or hinder the delivery of RNAP II to the promoter. Specific interactions have been reported between various regulators and TBP, TAFs, TFIIB, and TFIIH. The importance of these GTF–activator interactions is emphasized by *in vivo* studies showing that artificially tethering TBP to a promoter overcomes the requirement for an activator to achieve stimulated levels of transcription. The essence of these findings is that the delivery of RNAP to the promoter is an important and regulated step in the transcription cycle.

Transition from a closed to open RNAP–promoter complex

The delivery of RNAP to the promoter results in the binding of RNAP to DNA sequences upstream and around the transcription start site. This complex is referred to as the RNAP–promoter closed (RPc) complex. The RPc complex is very stable as measured by *in vitro* criteria. Specifically, it is stable under nondenaturing gel electrophoresis and is resistant to challenge by nonspecific competitor DNA. Subsequent to the formation of this closed complex, RNAP wraps promoter DNA around its circumference, capturing and interacting with DNA sequences downstream of the transcription start site. As this occurs, there is a conformational change within RNAP that enables it to clamp tightly on to DNA, resulting in an RNAP–promoter intermediate (RPi) complex. RNAP then 'melts' the promoter DNA surrounding the transcription start site to form the 'transcription bubble'. The formation of the bubble renders accessible the genetic information in the template DNA strand to yield an RNAP–promoter open (RPo) complex.

All RNAPs are inherently capable of melting the DNA around the transcription start site with one exception: eukaryotic RNAP II. In the prokaryotic as well as the eukaryotic RNAP I and RNAP III systems, the delivery of RNAP to the promoter signals the completion of promoter engagement. The presence of nucleotide triphosphates (NTPs) triggers the onset of transcription. RNAP II diverges from this paradigm in two ways. First, after the association of the polymerase with the promoter, the complex is not competent to initiate transcription. This system requires the association of two GTFs called TFIIE and TFIIH. The second distinction is that RNAP II requires the input of energy in the form of adenosine triphosphate (ATP) hydrolysis for initiation. Why this particular RNAP requires energy is unclear. What is clear is that the energy derived from ATP is not required for elongation but rather for the formation of a stable RPo complex and for a step subsequent to initiation called promoter clearance. How is this energy utilized? The RNAP II-specific GTF TFIIH is a multisubunit factor that contains as its largest component a protein called XPB/ERCC3. XPB/ERCC3 is an ATP-dependent DNA helicase that appears to function as a 'molecular wrench'. It utilizes the energy stored in the β-γ bond of ATP to rotate downstream DNA relative to fixed upstream protein–DNA interactions. This torsion on the DNA helix presumably results in strand separation around the start site of transcription and the establishment of an RPo complex.

Synthesis of initial phosphodiester bond and abortive initiation

Once formed, the RPo complex is stabilized by interaction between single-stranded DNA and RNAP. The binding of the initiating NTP confers further stability to the open complex without requiring its hydrolysis or formation of a phosphodiester bond. In the RPo complex, the template strand is accessible and directs base pairing with NTPs to initiate phosphodiester bond formation and RNA synthesis. This generates a ternary complex that contains RNA as well as DNA and RNAP. However, the formation of a ternary complex does not predict productive RNA synthesis. In fact, a significant fraction of complexes engage in abortive RNA synthesis in which short (2–8 nucleotides) RNA transcripts are made and released in a repetitive manner (**Figure 1**). Escape from this nonproductive cycle occurs when the RNA transcript reaches a critical length (about 10 nucleotides). In the bacterial system, when an RNA chain of this length is synthesized, sigma factor is released. At this point, RNAP leaves the promoter and becomes committed to productive chain elongation. The physical movement of RNAP away from the promoter is referred to as 'promoter clearance' and signals a vacancy at the promoter that can be filled by another RNAP.

Promoter clearance and elongation

During promoter clearance, RNAP breaks contact with DNA and with protein components at the promoter and becomes stably associated with the RNA and DNA chains. How does this happen? One possible explanation stems from electron crystallographic, electron microscopic and photocrosslinking data that have been used to deduce low-resolution models of RNAP open complexes and RNAP II elongating complexes. These models reveal common features likely shared by all RNAPs. In an RNAP complex that has initiated RNA synthesis, the 8–10 most recently synthesized nucleotides of the nascent RNA are engaged in Watson–Crick hydrogen bonding with the DNA template strand as an RNA–DNA hybrid. Importantly, the 5–8

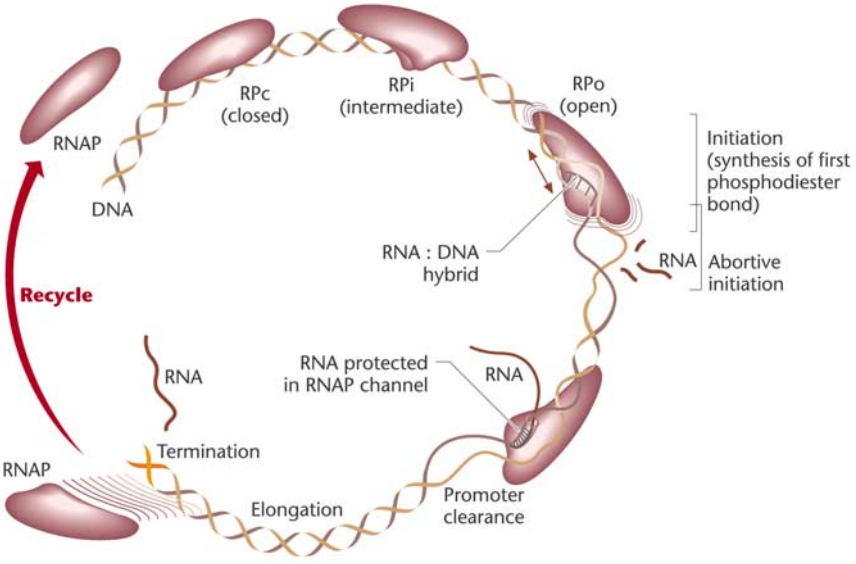

Figure 1 The transcription cycle. RNA polymerases (RNAPs) have an intrinsic affinity for DNA. Promoter-specific DNA binding by RNAPs is facilitated by recruitment factors unique to each RNAP. Once RNAP binds promoter DNA it forms a stable complex with the closed duplex DNA (RPc). The formation of RPc promotes a conformational change between RNAP and the DNA that results in an intermediate complex (RPi) that rapidly converts to an open complex (RPo) by the melting of the DNA around the transcription start site by RNAP. The formation of RPo results in a ternary complex competent to catalyse RNA synthesis. However, most RNAPs engage in nonproductive cycles of abortive RNA synthesis. Transition into a productive cycle occurs when the RNA chain reaches a critical length. At this point, called promoter clearance, RNAP disengages from the promoter and elongates the RNA chain. During the elongation phase, RNAP is subject to DNA and protein influences that can effect the kinetics of RNA synthesis. Once RNAP reaches termination sequences in the DNA, the ternary complex dissociates liberating the newly formed RNA chain. The released RNAP can then recycle for a subsequent round of RNA synthesis.

next most recently synthesized nucleotides of the nascent RNA are engaged in interactions with regions of RNAP that resemble a tunnel or channel. RNA within this tunnel is protected from enzymatic manipulations. The entry of nascent RNA into an exit tunnel appears to confer stability to the elongating RNAP complex. The formation of a stable and productive RNAP ternary complex correlates with the departure of sigma factor. Perhaps it is the entry of RNA transcripts into the exit tunnel that confers stability to the RNAP ternary complex resulting in the release of sigma factor. Although this model is supported by data from all RNAP systems, the situation for RNAP II is more intricate.

The CTD of RNAP II

For RNAP II, the transition from initiation to elongation is accompanied by covalent modifications of an unusual structure at the C-terminal domain (CTD) of its largest subunit. This structure consists of multiple tandem repeats of a heptapeptide (Tyr-Ser-Pro-Thr-Ser-Pro-Ser) that is conserved between fungi and vertebrates. Although the largest subunits of bacterial, archaeal, and eukaryotic RNAPs are members of the same family, the CTD is unique to eukaryotic RNAP II. The length of the heptapeptide repeats seems to correlate with increased genomic complexity; *Saccharomyces cerevisiae* has 26–27 repeats, *Caenorhabditis elegans* 34 repeats, *Drosophila* 43 repeats, and mouse and human 52 repeats. The importance of the CTD is well established, as deletion of the mouse, *Drosophila* or *S. cerevisiae* CTD is lethal.

Owing to the high content of serine and threonine residues in the heptapeptide, the CTD can be found in two states *in vivo*: highly phosphorylated and unphosphorylated. A number of studies have shown that the phosphorylation state of the CTD dictates the activity of RNAP II during the transcription cycle. Specifically, it has been shown that while the nonphosphorylated form of the CTD associates with promoter-bound components prior to initiation of RNA synthesis, it is the hyperphosphorylated form that catalyses RNA chain elongation. These observations point to a critical role for CTD phosphorylation in the disengagement of RNAP from the promoter during the transition from initiation of RNA synthesis to elongation. Phosphorylation of the CTD presumably induces conformational changes within the initiation complex that disrupt certain protein–protein interactions. This disruption may concomitantly stabilize the nascent RNA chain within the RNAP exit tunnel, thereby triggering clearance of the polymerase from the promoter. The above model implies a cycling of the CTD from an unphosphorylated form to a hyperphosphorylated form during rounds of RNA synthesis, and predicts that one of

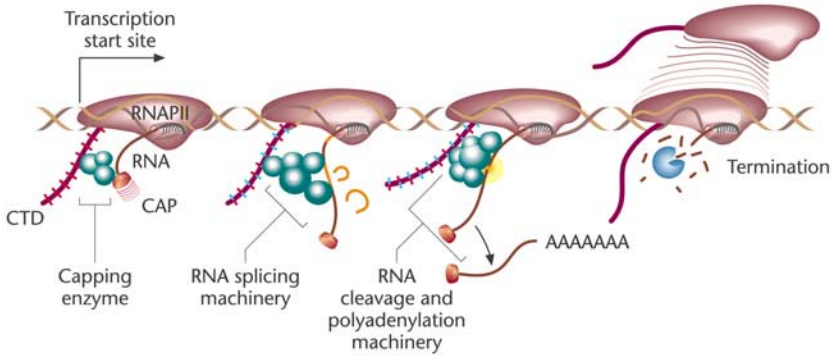

Figure 2 The CTD platform. The unique C-terminal domain (CTD) of RNAP II serves as a platform for the sequential loading of RNA-processing factors. Once the RNA chain exits from its protected environment within RNAP, it becomes a substrate for modification by RNA capping enzyme, splicing factors and RNA cleavage and polyadenylation reactions. The phosphorylation state of the CTD dictates which RNA-processing factors can bind the CTD and modify the RNA. Specifically, phosphorylation of serine 5 of CTD during transcription initiation enables capping enzyme to bind the CTD and catalyse the 'capping' reaction. Subsequent to promoter clearance there is a change in the phosphorylation state of the CTD with a shift of phosphorylation to serine 2. This change probably triggers the release of capping enzyme and the concomitant association of other RNA-processing enzymes. RNA polyadenylation is coupled to termination of RNA synthesis. Following cleavage and polyadenylation of the RNA chain, RNAP II continues to transcribe. However, the newly synthesized RNA molecule does not have a cap. It is possible that the lack of a cap makes it susceptible to degradation by an exonuclease which eventually disrupts the ternary complex. The subsequent or concomitant dephosphorylation of the CTD enables RNAP II to recycle and initiate another round of RNA synthesis.

the RNAP II-specific GTFs would be capable of phosphorylating the CTD. This prediction was validated by the identification of a CTD-specific kinase activity in TFIIH. The TFIIH kinase phosphorylates the CTD after the formation of the first phosphodiester bond and is probably important for promoter clearance.

The function of the RNAP II CTD does not end with promoter engagement and promoter clearance. In the RNAP II system, nascent RNA molecules undergo a number of modifications prior to being exported to the cytoplasm for translation. The transcripts are capped, spliced, cleaved and polyadenylated. These reactions do not proceed independently of one another. *In vivo*, these reactions occur cotranscriptionally; meaning that as the nascent RNA chain protrudes out of the RNAP exit tunnel it becomes a substrate for modification (**Figure 2**).

RNA 'capping' involves three reactions in which the 5′-triphosphate terminus of the RNA is cleaved to a diphosphate by RNA triphosphatase, then 'capped' with guanosine monophosphate (GMP) by RNA guanylyltransferase, and methylated at the N7 position of guanine by RNA methyltransferase. How is capping enzyme targeted to elongating RNA transcripts? Recently it was shown that components of capping enzyme directly interact with the phosphorylated form of the RNAP II CTD. Specifically, the guanylyltransferase component of the capping apparatus binds the CTD containing phosphoserine at either position 2 or 5 of the heptad repeat. Phosphoserine at position 5 stimulates the guanylyltransferase activity of capping enzyme. *In vivo* only the TFIIH-associated CTD kinase is necessary for proper capping enzyme recruitment.

Interestingly, shortly after promoter clearance, TFIIH dissociates from the elongating RNAP II complex. The departure of TFIIH is followed by the departure of capping enzyme. What triggers the release of capping enzyme? Phosphorylation of serine 5 in the heptad by TFIIH promotes the interaction of capping enzyme with the CTD. The release of capping enzyme shortly after promoter clearance invokes a change in the phosphorylation state of the CTD. Current evidence suggests that there is a wave of phosphorylation on serine 2 after promoter clearance. This change in phosphorylation may trigger the release of capping enzyme and recruit factors that are involved in splicing and polyadenylation of RNA. The recruitment of RNA-processing machinery to the CTD of RNAP II provides a mechanism of targeting capping, splicing, and cleavage and polyadenylation of the proper RNA substrate (**Figure 2**). In this system, the CTD serves as a unique platform from which to coordinate these activities.

The elongation phase of RNAP II is also subject to extrinsic influences. Specifically, RNAP II elongation can respond to *cis* (DNA) and *trans* (protein)-acting factors. *Cis*-acting factors constitute natural DNA sequences that cause RNAP II to pause during transcription. The paused RNAP II is acted upon by factors that promote transit beyond these sites. One example is TFIIS. This factor caused the paused RNAP II to back up and then proceed forward past the pause site. Another class of factors, including TFIIF and the elongins, influence the rate of RNAP II elongation. Together, these two classes of factors govern the kinetics of RNAP II elongation.

Termination

Once RNAP has cleared the promoter, it synthesizes RNA until it encounters a terminator sequence. At this point, RNA synthesis stops, RNAP releases the completed RNA product and dissociates from the DNA template. How does this happen? Many terminators require a hairpin to form in the secondary structure of the RNA being transcribed. This suggests that termination depends on the structure of the RNA product and is not simply determined by specific DNA sequences encountered during transcription.

Termination is best understood in the prokaryotic system. *E. coli* RNAP can terminate RNA synthesis at two types of terminators. One class is called intrinsic terminators because RNAP can terminate at these sequences without the assistance of other factors. The other class of terminators is called Rho-dependent terminators; termination at these sites requires the assistance of rho factor.

Two structural features characterize intrinsic terminators: a hairpin in the secondary structure and a stretch of 7–9 uracil residues at the end of the RNA chain. The uracil residues produce a particularly unstable RNA–DNA hybrid. Typically, when RNAP encounters a weak hybrid it pauses and backtracks to form a more stable hybrid. However, when RNAP encounters a terminator sequence, the resulting hairpin in the RNA prevents RNAP from backtracking and also disrupts critical RNA contacts in the exit channel of RNAP.

Rho-dependent termination requires a stretch of 50–90 nucleotides preceding the site of termination that is rich in C and poor in G residues. Rho has an ATPase activity that is dependent on RNA. As RNAP translocates along the DNA template, rho translocates along the RNA transcript. When RNAP pauses at a termination sequence, the secondary structure in the transcript enables rho to disrupt the RNA–DNA hybrid and release the RNA.

In the highly regulated RNAP II system, termination of RNA synthesis is coupled to RNA polyadenylation. Once again, the CTD is invoked to execute this final step in RNA synthesis. Once RNAP II has transcribed through the polyadenylation site (poly(A)), the resulting RNA chain is acted upon by factors that cleave the RNA chain from the moving RNAP II and catalyse the poly(A) addition to the end of the RNA molecule. This modification completes the synthesis and processing of messenger RNA.

Studies on the mouse β-globin gene have shown that transcription terminates within a region 1400 nucleotides downstream of the poly(A) site. What is the nature of the signal that finally releases RNAP II from the template? Two general models can be proposed to address this question. The first suggests that upon the release of the nascent RNA chain, RNAP II continues to synthesize RNA lacking a proper cap at the new 5′ end. The absence of a 5′ cap makes this RNA chain susceptible to degradation. The model invokes the recognition of the unprotected 5′ end of the RNA by an exonuclease that degrades the RNA at a faster rate than RNA synthesis by RNAP II. Eventually, the exonuclease catches up with RNAP II, disrupts the ternary complex, and releases RNAP II from the template (**Figure 2**). The other model suggests that once RNAP II transits past the poly(A) site, a signal is sent to RNAP II, perhaps by processivity factors, causing it to dissociate from the template. One example that illustrates the latter mechanism is transcription termination factor 2 (HuF2). HuF2 is identical to the *Drosophila* lodestar protein. HuF2/lodestar possesses DNA-dependent ATPase activity and can efficiently dissociate RNAP II from the template. Earlier observations showed that lodestar is cytoplasmic during interphase but translocates to the nucleus during mitosis. Perhaps HuF2/lodestar functions to ensure transcriptional silencing during mitosis. Support for this notion comes from analysis of the lodestar mutant in *Drosophila*. Specifically, mutations in lodestar result in chromatin bridges, chromosome tangling and breakage during anaphase. This observation suggests that proper chromosome segregation is hindered by the presence of active transcription complexes during mitosis and that HuF2/lodestar preserves the integrity of the chromosomes by displacing active RNAP II complexes from the template.

RNAP recycling

Once RNAP dissociates from the DNA template, it is free to reassociate with promoter elements and initiate another round of RNA synthesis. Once again, the RNAP II system affords a higher level of regulation. After cleavage of the nascent RNA chain, most of the factors associated with the RNAP II CTD platform have executed their function. As a result, the CTD does not need to remain phosphorylated. Moreover, in order for RNAP II to re-engage the promoter it must be in the dephosphorylated form. This requisite dephosphorylation event is facilitated by a specific phosphatase called FCP1. FCP1 associates with the elongation RNAP II complex and is required for the dissociation of capping enzyme from the complex, presumably by altering the phosphorylation state of the CTD from predominantly serine 5 phosphorylation to mostly serine 2 phosphorylation. When RNAP II encounters certain DNA sequences FCP1 can catalyse the dephosphorylation of the CTD. Importantly, the conversion of the elongating RNAP II to the dephosphorylated form does not promote the release of RNAP II from the template DNA. The ability of FCP1 to dephosphorylate the CTD in the elongation complex suggests that RNAP II is converted to the nonphosphorylated form prior to, or concomitant with, its release from the DNA template. The catalytic activity of FCP1 enables RNAP II to recycle and engage in multiple rounds of productive RNA synthesis.

The Chromatin Challenge

In eukaryotic cells, the genetic blueprint, which extends over a meter if unravelled, is stored in a nucleus that is less than 10^{-5} m in diameter. Chromosomes represent the largest and most visible structures of genetic information. This degree of compaction has severe consequences for processes that require access to DNA. The cell has developed compensatory mechanisms that facilitate access to the DNA during the processes of RNA synthesis, DNA synthesis and DNA repair. The interplay of chromatin remodelling activities with the basal transcription machinery undoubtedly has a profound impact on the overall level and regulation of gene expression.

Further Reading

Bell SD and Jackson SP (1998) Transcription in Archaea. *Cold Spring Harbor Symposia on Quantitative Biology* **63**: 41–51.

Bell SD and Jackson SP (2001) Mechanism and regulation of transcription in archaea. *Current Opinion in Microbiology* **4**: 208–213.

Busby S and Ebright RH (1999) Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology* **293**: 199–213.

Cho H, Kim TK, Mancebo H *et al*.(1999) A protein phosphatase functions to recycle RNA polymerase II. *Genes and Development* **13**: 1540–1552.

Coin F and Egly J-M (1998) Ten years of TFIIH. *Cold Spring Harbor Symposia on Quantitative Biology* **63**: 105–110.

Ebright RH (1998) RNA polymerase–DNA interaction: structures of intermediates, open, and elongation complexes. *Cold Spring Harbor Symposia on Quantitative Biology* **63**: 11–20.

Falck-Pedersen E, Logan J, Shenk T and Darnell JE Jr (1985) Transcription termination within the E1A gene of adenovirus induced by insertion of the mouse beta-major globin terminator element. *Cell* **40**: 897–905.

Hampsey M and Reinberg D (1999) RNA polymerase II as a control panel for multiple coactivator complexes. *Current Opinion in Genetic Development* **9**: 132–139.

deHaseth PL, Zupancic ML and Record MT (1998) RNA polymerase–promoter interactions: the comings and goings of RNA polymerase. *Journal of Bacteriology* **180**: 3019–3025

Kim TK, Ebright RH and Reinberg D (2000) Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**: 1418–1421.

Kim TK, Lagrange T, Wang Y-H *et al*. (1997) Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proceedings of the National Academy of Sciences of the USA* **94**: 12268–12273.

Komarnitsky P, Cho E-J and Buratowski B (2000) Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes and Development* **14**: 2452–2460.

Minakhin L, Bhagat S, Brunning A *et al*. (2001) Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RBP6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. *Proceedings of the National Academy of Sciences of the USA* **98**: 892–897.

Mooney RA, Artsimovitch I and Landick R (1998) Information processing by RNA polymerase: recognition of regulatory signals during RNA chain elongation. *Journal of Bacteriology* **180**: 3265–3275.

Mullen GP and Wilson SH (1997) DNA polymerase beta in abasic site repair: a structurally conserved helix–hairpin–helix motif in lesion detection by base excision repair enzymes. *Biochemistry* **36**: 4713-4717.

de Murcia JM, Niedergang C, Trucco C *et al*. (1997) Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells. *Proceedings of the National Academy of Sciences of the USA* **94**: 7303–7307.

Naryshkin N, Revyakin A, Kim Y *et al*. (2000) Structural organization of the RNA polymerase–promoter open complex. *Cell* **101**: 601–611.

Orphanides G and Reinberg D (2000) RNA polymerase II elongation through chromatin. *Nature* **407**: 471–475.

Orphanides G, Legrange T and Reinberg D (1996) The general transcription factors of RNA polymerase II. *Genes and Development* **10**: 2657–2683.

Pelletier H, Sawaya MR, Wolfle W *et al*. (1996) Crystal structures of human DNA polymerase beta complexed with DNA: implications for catalytic mechanism, processivity, and fidelity. *Biochemistry* **35**: 12742–12761.

Peng J, Liu M, Marion J *et al*.(1998) RNA polymerase II elongation control. *Cold Spring Harbor Symposia on Quantitative Biology* **63**: 365–370.

Reinberg D, Orphanides G, Ebright RH *et al*. (1998) The RNA polymerase II general transcription factors: past, present, and future. *Cold Spring Harbor Symposia on Quantitative Biology* **63**: 83–103.

Schroeder SC, Schwer B, Shuman S and Bentley D (2000) Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes and Development* **14**: 2435–2440.

Shatkin AJ and Manley JL (2000) The ends of the affair: capping and polyadenylation. *Nature Structural Biology* **7**: 838–842.

Zawel L and Reinberg D (1993) Initiation of transcription by RNA polymerase II: A multi-step process. *Progress in Nucleic Acid Research and Molecular Biology* **44**: 67–108.

Zawel L and Reinberg D (1995) Common themes in assembly and function of eukaryotic transcription complexes. *Annual Review in Biochemistry* **64**: 533–561.