**Title**
Unraveling the Molecular Mechanism of Poly(A)-Dependent Transcription Termination



**Author**
Zhang, Huimin

**Publication Date**
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Unraveling the Molecular Mechanism of Poly(A)-Dependent Transcription Termination

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biochemistry and Molecular Biology

by

Huimin Zhang

2014

ABSTRACT OF THE DISSERTATION

Unraveling the Molecular Mechanism

of Poly(A)-Dependent Transcription Termination

by

Huimin Zhang

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2014

Professor Harold G. Martinson, Chair

Termination of pre-mRNA transcription by RNA polymerase II occurs in two steps: a decrease of elongation rate (pause), followed by the dissociation of the polymerase from the DNA template (release). While the pause can be triggered solely by the AAUAAA hexamer in the nascent transcript, the release can only occur in presence of a complete poly(A) signal, which also requires a GU-rich sequence downstream the hexamer. The hexamer and the GU-rich element are specifically recognized by CPSF and CstF respectively. Contradicting views exist about whether CPSF and CstF define the complete poly(A) signal in a sequential recruitment manner or as a preassemble complex. In Chapter 2, taking advantage of an in vitro system in HeLa nuclear extract, in which pre-mRNA 3'-end processing is coupled to transcription, we found that the functional poly(A) signal is defined by a preassembled CPSF-CstF complex, which potentially captures the upcoming GU-rich transcript more efficiently.

The complete poly(A) signal, together with its associated protein apparatus, is capable of releasing the polymerase from the DNA. But it has been controversial for decades what is the prerequisite of the release step: the poly(A) site cleavage and the subsequent degradation of the downstream RNA, or a particular conformational rearrangement of the transcription complex. In Chapter 3, based on the same in vitro system, we developed a transcription termination assay, which can measure the termination and concurrent poly(A) site cleavage simultaneously. Through this assay, we established that the release of the polymerase does not require poly(A) site cleavage. Rather, the transcription complex experiences a conformational rearrangement immediately after crossing the poly(A) signal, followed by a slow dissociation. The dissociation of the transcription complex can be fully inhibited by α-amanitin, presumably by disruption of the poly(A)-induced conformational rearrangement.

The Dissertation of Huimin Zhang is approved.

Michael F. Carey

Steven G. Clarke

Harold G. Martinson, Committee Chair

University of California, Los Angeles

2014

Dedicated to …

My beloved parents

who gave me a physical life

and

my dear doctoral advisor

who gave me a scientific life

# Table of contents


| Section | | pages |
|---|---|---|
| List of Figures | | vii |
| Acknowledgements | | viii |
| Vita | | ix |
| Chapter 1 | Introduction | 1 |
| | Figures | 17 |
| | References | 21 |
| Chapter 2 | Role of the Nascent RNA in the Assembly of the Cleavage and Polyadenylation Apparatus | 25 |
| | Abstract | 26 |
| | Introduction | 27 |
| | Results | 30 |
| | Discussion | 36 |
| | Materials and Methods | 38 |
| | Figures | 43 |
| | References | 51 |
| Chapter 3 | Poly(A)-Dependent Transcription Termination does not Require Poly(A) site Cleavage | 53 |
| | Abstract | 54 |
| | Introduction | 55 |
| | Results | 61 |
| | Discussion | 73 |
| | Materials and Methods | 77 |
| | Figures | 81 |
| | References | 93 |

# List of Figures


| Figures | pages |
|---|---|
| 1-1 | 17 |
| 1-2 | 19 |
| 2-1 | 43 |
| 2-2 | 45 |
| 2-3 | 47 |
| 2-4 | 49 |
| 3-1 | 83 |
| 3-2 | 85 |
| 3-3 | 88 |
| 3-4 | 90 |
| 3-5 | 92 |

# Acknowledgements

I am deeply grateful to my doctoral advisor Dr. Harold Martinson who has led me in the start of my biochemist's life with tremendous patience and passed me his most valuable scientific spirit, which will remain as the guide for my entire career and life.

I would thank all my committee members who have served years on my dissertation committee and especially thank Dr. Michael Carey and Dr. Steven Clarke who have not only offered valuable advice for my dissertation research and but also for my future scientific career.

I would also thank my rotation advisors Dr. Steven Clarke and Dr. Juli Feigon for introducing me their great researches and opening my minds for a variety of interesting and important scientific topics. I would thank my research mentors, Dr. Benson Ngo, Dr. Alexander Patananan and Dr. Qi Zhang for getting my hands on the bench in order to do real research.

I would thank Dr. Guillaume Chanfreau for the great experience being his teaching assistant and also for his valuable suggestions in my research.

I would thank Dr. Gregory Gilmartin for providing me the plasmid with Rex poly(A) signal, which enables my research on the assembly study of 3'-end processing apparatus.

I would thank Ke-Huan Chow for performing the experiments in Figure 2-1A, Figure 2-3A and Figure 2-4. I would thank Dr. Benson Ngo for performing the experiments in Figure 2-1B and Figure 2-3B. I would thank Dr. Harold Martinson for rewriting the entire Chapter 3 as a manuscript for publication.

I would thank the previous lab alumni, Dr. Frank Rigo who generously offered me the specially modified oligonucleotides, which enable many research questions to be investigated.

# VITA

**EDUCATION**

2009~2014 Ph.D. candidate, Department of Chemistry and Biochemistry, UCLA
Biochemistry and Molecular Biology

2005~2009 B.S., Xiangya School of Medicine, CSU, China
Pharmaceutical Sciences

**TEACHING EXPERIENCE**

Teaching assistant, Department of Chemistry and Biochemistry, UCLA
2013 Biochemistry: Metabolism and regulation
2012 Biochemistry: DNA,RNA and protein
2011 Biochemistry: Biochemical methods II
2010 Biochemistry: Introduction to structure, enzyme and metabolism

**HONORS AND AWARDS**

2014 The ASBMB 2014 Thematic Best Poster Award
2014 The ASBMB 2014 Graduate/Postdoctoral Travel Award
2013 Philip Whitcome Pre-doctoral Fellowship
2013 Roberts A. Smith Research Award
2011 Samson H. Cheng Biochemistry Teaching Award
2009 UCLA University Fellowship
2009 State Scholarship by Chinese Scholarship Council
2008 Excellent Student Model
2006 CSU University Grand Prize

**PUBLICATION**

1. Zhang, H., Rigo, F., and Martinson, H.G. (2014). Poly(A)-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. A manuscript in the final stage.

2. Zhang, H., Ngo, B., Chow, K., Rigo, F., and Martinson, H.G. (2014). Role of the nascent RNA in the assembly of the cleavage-polyadenylation apparatus. A manuscript in the final stage.

# *Chapter 1*

Introduction

Transcription termination is marked by the dissociation of the polymerase from the DNA template. The appropriateness of this activity is crucial for correct genome partitioning and polymerase recycling (Kuehner et al., 2011). The process of RNA polymerase (Pol) I and III termination in eukaryotes is relatively simple: Pol I, transcribing most ribosomal RNAs (rRNAs), terminates at a major terminator sequence; Pol III, transcribing transfer RNAs, 5S rRNA and certain small nuclear RNAs (snRNA), terminates at T-rich sequences; both involve a limited number of auxiliary factors (Richard and Manley, 2009). In contrast, Pol II, transcribing precursor messenger RNAs (pre-mRNAs) and many noncoding RNAs, does not terminate at any specific site, which ranges from a few bases to several kilobases downstream the matured 3'-end of RNA (Proudfoot, 1989). Particularly, due to the vast population of protein-coding genes, the process towards Pol II termination of pre-mRNA undergoes very sophisticated yet neat regulation by a series of RNA and protein factors (Kuehner et al., 2011). A majority of these factors also play pivotal roles in 3'-end processing of pre-mRNA, which matures most coding RNA at the poly(A) signal by cleavage and polyadenylation (Figure 1-1).

The kin relation between pre-mRNA transcription termination and 3'-end processing elicited two views regarding the mechanism of this termination: an anti-termination factor proposed by Shenk's group could leave the transcription complex, altering the conformation of the transcription complex to prepare it for termination (Logan et al., 1987); alternatively, the poly(A) site cleavage proposed by Manley's group was required to provide entry for a 5'-3' exonuclease to degrade the downstream RNA and terminate transcription (Connelly and Manley, 1988). The debate persists for decades along with a plethora of evidences in favor of each view. But the essential prerequisites for pre-mRNA transcription to terminate have never been comprehensively defined, which will readily solve the above issues: whether poly(A) site

cleavage is required for termination; whether the transcription complex undergoes a particular conformation rearrangement in order to terminate.

To explore the core mechanism of pre-mRNA transcription termination, we will start with a critical review on the hitherto sketch of the termination process, which has contributed to the identification and functional dissection of several essential termination factors and may potentially lead us to new insights into its core mechanism.

The first factor that was identified to be crucial for pre-mRNA transcription termination is the poly(A) signal, a single point mutation of which in the α2-globin gene could abolish termination (Whitelaw and Proudfoot, 1986). Further investigation established that most pre-mRNA transcription termination is accomplished in two steps: a decrease of the elongation rate (pause), followed by a stochastic release of the polymerase from DNA (termination), both of which are triggered by the poly(A) signal (Orozco et al., 2002; Park et al., 2004). Hence, we designate the pause as poly(A)-dependent pause (PADP) and the termination as poly(A)-dependent transcription termination (PADT).

**PADP is a checkpoint for the commitment to PADT or pre-mRNA 3'-end processing.**

The fact that termination can occur a few kilobases downstream the poly(A) signal (Citron et al., 1984) encouraged the search of other RNA elements that delineate the termination region. This in turn revealed several pause sites, including the α2, the C2/MAZ4 and the μs pause site, residing in the nascent transcript closely downstream the poly(A) signal (Adamson et al., 2005; Ashfield et al., 1991; Enriquez-Harris et al., 1991; Peterson et al., 2002). Although these sites can enhance the termination efficiency, they can neither trigger termination without a poly(A) signal, nor necessitate a poly(A) signal for its pausing activity (Peterson et al., 2002; Yonaha and

Proudfoot, 1999). Clearly, these pause sites are functionally irrelevant to the poly(A) signal and thereby indistinguishable from other pause sites that the polymerase encounters during elongation. This makes it questionable whether a pause is necessarily associated with PADT. However, a transient increase of the polymerase density was robustly observed immediately downstream the poly(A) signal. Further investigation showed that this pause can occur with solely the poly(A) signal in presence without the assistance of any other sequence-specific RNA elements (Orozco et al., 2002). Therefore, the poly(A) signal is intrinsically a negative elongation RNA element, evoking a pause of the polymerase right after its extrusion, which can but not necessarily be facilitated by other pause elements.

Within the time frame that the pause was observed, the 3'-end processing protein apparatus has not been maturely assembled and hence cleavage and polyadenylation has not occur yet (Chao et al., 1999). Also, the release of the polymerase can be detected only downstream the pause (Orozco et al., 2002). Therefore, PADP precedes both termination and 3'-end processing.

Considering its potential importance as a pioneer activity, the mechanism of PADP was probed, first by dissection of the poly(A) signal. The core of the poly(A) signal is composed of two elements, an AAUAAA hexamer and a GU-rich element (Figure 1-1). While the mutation of the hexamer completely abolished the pause, the deletion of the GU-rich element still supported a sustained increase of polymerase density downstream the hexamer (Nag et al., 2006). However, neither termination nor poly(A) site cleavage can occur in absence of the GU-rich element (Connelly and Manley, 1988; Nag et al., 2006). Thus, the pause is independent of both activities, as already indicated in the temporal order of these events.

The fact that the pause can be triggered by the hexamer alone is well consistent with its previously proposed role as a checkpoint to divert the transcription complex to either 3'-end

processing or termination or simply resumption of elongation, which depends on further integrated information from downstream the hexamer (Orozco et al., 2002).

**The protein carrying out the pause is much less complex.**

Since the hexamer is specifically recognized by the 160-kD subunit of cleavage polyadenylation specificity factor (CPSF) (Murthy and Manley, 1995), the role of CPSF in PADP was further investigated. CPSF is a multi-protein complex composed of seven subunits (Figure 1-1). It is required for cleavage and polyadenylation, with its 160-kD subunit recognizing the hexamer and 73-kD subunit as the endonuclease carrying out poly(A) site cleavage (Chan et al., 2011). When CPSF was inactivated by the binding of influenza virus protein NS1A, the transient increase of polymerase density downstream the poly(A) signal disappeared (Nag et al., 2007). CPSF is thus identified to be also a negative elongation factor specifically mediating the pause of the polymerase transmitted from the hexamer. As the pause precedes termination, not surprisingly, the inactivation of CPSF completely impeded termination, establishing that CPSF is also an essential termination factor (Nag et al., 2007). This is consistent with the previous study in yeast that mutation of Yhh1p, the yeast homologue of the mammalian CPSF160, resulted in drastically defective transcription termination (Dichtl et al., 2002).

In contrast to CPSF, the depletion of cleavage stimulatory factor (CstF), which specifically recognizes the GU-rich element, although severely impaired poly(A) site cleavage, had no effect on the pause (Nag et al., 2006). Similarly, deletion of the C-terminal domain (CTD) of Pol II, although impeding both PADT and 3'-end processing, still supported a sustained pause downstream the poly(A) signal (Nag et al., 2007; Park et al., 2004). These results are consistent with the role of PADP as checkpoint preceding and hence independent of both PADT and 3'-end processing.

**How does the pause state turn to a fully poly(A)-dependent state?**

The occurrence of the pause in absence of the polymerase CTD indicates that it is probably through interaction with the polymerase body that CPSF mediates the pause. Indeed, CPSF was found to co-immunoprecipitate (IP) with the body of Pol II whose CTD was truncated (Nag et al., 2007). Interestingly, although CstF also specifically interacts with CPSF160 through its 77-kD subunit (Murthy and Manley, 1995), in the above co-IP complex CstF was absent. This indicates the binding of CPSF to the polymerase body and to CstF is probably mutually exclusive. Therefore, the authors proposed that while CPSF binds to the polymerase body and triggers the pause, CstF is recruited only when its cis binding partner, the downstream GU-rich element, emerges. A functional poly(A) signal is thus defined and then the CPSF-CstF complex is transferred to the polymerase CTD, where 3'-end processing takes place. This model animatedly describes how the transcription complex is transformed from the pause state to the fully poly(A)-dependent state: through the recruitment of CstF. It well coincides with the ordered assembly pathway of the cleavage and polyadenylation apparatus proposed decades ago, when CPSF was found weakly associate with the hexamer and only after the binding of CstF to the downstream GU-rich sequence, the tertiary complex of CPSF, CstF and RNA became stable (Gilmartin and Nevins, 1989). However, when the protein complex associated with CstF was isolated from the nuclear extract, about 80% CPSF160 was found in the CstF-bound fraction (Takagaki and Manley, 2000). This invoked the preassembly model proposing that the complex of CPSF-CstF is preassembled for more efficient poly(A) signal recognition and RNA processing. Although plausible, neither model can be fully justified by its originating data because they are simply based on protein interaction in the nuclear extract while it is unknown whether the immnoprecipitated complex was actively engaged in transcription or not.

To determine how the pause state turns to a fully poly(A)-dependent state, through preassembly or sequential recruitment of CPSF and CstF, is essentially to determine how they become engaged with the active transcription complex. In Chapter 2, using a transcription-coupled processing system, we examined the enlisted time of CPSF and CstF through the length of RNA that was immnoprecipitated with each protein. We established in two independent approaches that both CPSF and CstF join the transcription complex as early as the extrusion of the hexamer, which is only specifically recognized by CPSF. This supports that CPSF-CstF, the core apparatus for poly(A) signal definition, engages with the transcription complex in a preassembled manner rather than sequentially recruited. Clearly the presence of CstF in proximity before the extrusion of the GU-rich transcript makes the recognition between the two more timely and hence the definition of the complete poly(A) signal more efficient.

**Both PADT and 3'-end processing require the continuity of RNA downstream the poly(A) signal.**

Another RNA element, critical to the cotranscriptional 3'-end processing, constitutes the continuity of RNA downstream the poly(A) signal through tethering the poly(A) signal to the polymerase. When the RNA tether is cut by DNA oligo and RNaseH during transcription (Figure 1-2), the poly(A) site cleavage efficiency is drastically reduced (Rigo et al., 2005). Considering PADT is the effect that the poly(A) signal exerts on the polymerase, we wonder whether the RNA tether that physically connects the two is also critical to PADT. Our data showed that when this tether is cut during transcription, PADT is severely impeded (Figure 3-2D), recapitulating the effect on 3'-end processing. In view of its simple nature, the role of the RNA tether in PADT probably resembles that in pre-mRNA 3'-end processing, which was proposed to hold the poly(A)-recruited apparatus in close proximity to the polymerase until it is mature enough to

stably associate with the polymerase (Rigo et al., 2005). The other possibility, which is minor but does exist, is that the RNA tether may constitute a structural component of the apparatus and solely the binding of DNA oligo, which disrupts the RNA single strand structure, can already ruin its function. This possibility can be tested by 2'-O-methoxyethylribose-modified oligonucleotides (MOE). The hybridization of MOE to the tether mimics that of DNA oligo, but differs in that MOE prevents RNaseH from cutting the RNA (Vickers et al., 2001). When MOE, instead of DNA oligo, binds to the RNA tether, it does not affect pre-mRNA 3'-end processing (Figure 2-1A). Thus, this rules out the possibility that the tether may be structurally involved in the processing apparatus and thereby ascertains that the continuity of RNA downstream the poly(A) signal is critical to both PADT and 3'-end processing through holding the poly(A)-recruited apparatus in proximity to the polymerase.

**Different requirements are exerted on the protein machinery to proceed with PADT and/or 3'-end processing.**

CstF is required for the definition of the complete poly(A) signal as shown previously, so it is not surprising that knockdown of the 64-kD subunit of CstF, which specifically binds the GU-rich sequence, severely impaired poly(A) site cleavage (Nag thesis data). This recapitulates the defect on 3'-end processing by CstF depletion in vitro (Nag et al., 2006). In contrast, PADT remained quite efficient when CstF64 was knocked down (Nag thesis data). But PADT certainly requires the GU-rich element (Connelly and Manley, 1988; Nag et al., 2007), which in turn needs CstF64 for recognition. Clearly, for both poly(A) site cleavage and PADT, the function of CstF in recognition of the GU-rich element is crucial. Then what account for the contrary effects of CstF64 knockdown on each activity are probably their different requirements of the cellular level of CstF. CstF is a three-subunit protein complex, which can self-dimerize through its 77- and 50-

kD subunits. So in the latest structure model of the 3'-end processing complex (Figure 1-1), CstF functions as a dimer (Chan et al., 2011). Thus a plausible explanation is that CstF64 knockdown reduces its concentration to the level that it fails to dimerize, which greatly reduces the 3'-end processing efficiency. But the residual CstF can still function as a monomer to recognize the GU-rich element, which in turn transmits the signal to the transcription complex and triggers termination. A similar phenomenon of Rna15, the yeast homolog of human CstF64, was observed (Qu et al., 2007). When the C-terminal domain of Rna15/CstF64 was mutated, poly(A) site cleavage efficiency drastically reduced whereas termination was barely affected. Therefore, the poly(A) signal can be transmitted either through a CstF monomer or through a C-terminus mutant of Rna15/CstF64 to the transcription complex to proceed with PADT rather than 3'-end processing.

The mutation of Pcf11 can also exert distinct effects on PADT and 3'-end processing. Pcf11 is part of the core apparatus assembled on the poly(A) signal, which is recruited towards the 3'-end of transcription (Mayer et al., 2010). The N-terminus of Pcf11 contains a Pol II CTD interaction domain (CID) (Barillà et al., 2001; Meinhart and Cramer, 2004). Three amino acid changes of the CID in the *pcf11-13* mutant disrupts termination, but 3'-end processing remains equally efficient (Sadowski et al., 2003). Thus the CID of Pcf11 is essential for PADT while may not be as important for 3'-end processing, which can be fully reconstituted in vitro by a CID-deleted version of Pcf11 (Mariconti et al., 2010). Consistently, when the CID is intact but mutations spread elsewhere on the protein in the *pcf11-2* mutant, termination is barely affected despite the abolishment of poly(A) site cleavage (Sadowski et al., 2003). Further investigation showed that CID alone, but not CID-deleted Pcf11, can dismantle the purified elongation complex as efficiently as intact Pcf11 (Zhang and Gilmour, 2006; Zhang et al., 2005). Therefore, the dual

role of Pcf11 in PADT and 3'-end processing can be physically demarcated on the protein: N-terminal CID responsible for termination and the other half for 3'-end processing. This indicates the tasks of PADT and 3'-end processing can potentially be executed independently, as already suggested from the various effects of the *pcf11-13* and *pcf11-2* mutants.

**Which is essential to PADT, RNA cleavage or a conformational rearrangement?**

Since PADT initiates at the end of transcription when the 3'-end of pre-mRNA also needs to be processed, it is physiologically advantageous that the two activities share a majority of the RNA and protein factors. Even so, they are presumably independent of each other, as manifested from the distinct requirements they exert on CstF and Pcf11 (Qu et al., 2007; Sadowski et al., 2003).

However, PADT has been believed to be dependent on poly(A) site cleavage and RNA degradation for decades. When this view was first proposed, the only experimental evidences were that they are both dependent on the poly(A) signal and the RNA downstream the poly(A) cleavage site is unstable (Connelly and Manley, 1988). These evidences are clearly insufficient to explain how the transcription complex dissociates. But the study of E. coli transcription termination provided some inspiration for the story in eukaryotes. The E. coli terminator Rho was identified to be also a helicase and thus was believed to destabilize the transcription complex by unwinding the DNA-RNA duplex (Platt, 1986). In eukaryotes, no such helicase activity was reported to be involved in termination. But since the RNA downstream the poly(A) site is unstable, a 5'-3' exonulcease was proposed, similar to the E. coli helicase Rho, to access the downstream RNA, instead of unwinding DNA-RNA, just degrading the RNA to destabilize the transcription complex. Obviously, PADT is dependent on poly(A) site cleavage and RNA degradation according to this proposal. However, not until 2010, Rho was found to associate with the transcription complex in the very beginning of transcription even without the presence

of RNA and later trigger a conformation rearrangement of catalytic center of the polymerase, which is required for the dissociation of the elongation complex (Epshtein et al., 2010). Also, both E. coli and eukaryotic Pol II transcription complex can be stabilized by as short as 8-nucleotide DNA-RNA hybrid in the catalytic center of the polymerase (Kireeva, 2000; Komissarova et al., 2002). Since an exonuclease has no way to access the RNA deeply buried in the polymerase center, simply degradation of peripheral transcript cannot cause the dissociation of transcription complex.

Even so, attempts to claim the dependency of PADT on cleavage have never stopped. For example, when a strong poly(A) signal was used to replace a weak one, not only 3'-end processing became more efficient, but also termination (Edwalds-Gilbert et al., 1993). The correlation of the efficiency between the two was misconceived as that PADT depends on 3'-end processing, while the fundamental and simplest reason was ignored that they are both dependent on the poly(A) signal and thereby both correlate with the strength of the poly(A) signal. Ten years later, the idea of PADT requiring cleavage become even more ingrained due to the identification of the Rat1 (human homolog Xrn2) as a termination factor. Rat1 is a 5'-3' exonuclease, recruited towards 3'-end transcription like Pcf11 and its association with the transcription complex requires an intact Pcf11 (Luo et al., 2006). Crystal structure reveals that Rat1 interacts indirectly with the polymerase CTD through the CID of Rtt103 (Lunde et al., 2010). Mutation of Rat1 and knockdown of Xrn2 were found to cause PADT defect and hence Rat1/Xrn2 was also identified as a transcription termination factor (Kim et al., 2004; West et al., 2004). The dual role of Rat1/Xrn2 as an exonulcease and a termination factor seems to integrate the story that cleavage is required for termination by generating a free 5' phosphate end for the

access of Rat1/Xrn2 to degrade the downstream RNA and terminate transcription (Kim et al., 2004; West et al., 2004).

Despite the seemingly integrity of the story, the question regarding the core mechanism of PADT remains unanswered whether PADT requires cleavage of RNA. If we revisit the data corroborating the role of Rat1/Xrn2 as a termination factor, the Xrn2 knockdown exhibits just a mild defect in PADT (Figure 1C in West et al., 2004). Also, in the *rat1-1* mutant, termination was barely affected at permissive temperature and only at non-permissive temperature, the defect became dramatic (Figure 4C in Kim et al., 2004). This is in contrast with the *rat1-1* mutant effects on the its exonuclease activity, which were equally destroyed under both temperatures (Figure 3B in Kim et al., 2004). It is also in contrast with the mutants of the essential termination factor Yhh1/CPSF160, in which PADT efficiency drastically reduced under both permissive and non-permissive temperature (Dichtl et al., 2002). Clearly, Rat1/Xrn2 has a role in PADT, but either nonessential, or irrelevant to its exonuclease activity, since fairly efficient termination occurs in the *rat1-1* mutant at 23 ℃ without its RNA degradation activity. On the other hand, the severe termination defect of *rat1-1* only at high temperature, far from contradicts, but well complies with the allosteric view, which underlines a conformational change of the transcription complex, rather than RNA cleavage, is the prerequisite of PADT. Since the conformational change is essentially a thermal motion of protein, it is more vulnerable to increased temperature, and so is PADT. Under this circumstance, any contribution to PADT is nontrivial, including that from Rat1. Thus, the malfunction of Rat1, although normally negligible, can still cause severe termination defect at high temperature.

Then what can be the role of Rat1/Xrn2 in termination? It certainly cannot be attributed to RNA degradation, because when the cytoplasmic version of 5'-3' exonuclease, Xrn1, was introduced

into the nucleus, it can compensate for the nuclear RNA degradation activity, but cannot rescue the termination defect of the *rat1-1* mutant (Luo et al., 2006). This is also consistent with the previously presented evidence that the transcription complex can be stabilized by as short as 8-nucleotide DNA-RNA hybrid (Kireeva, 2000; Komissarova et al., 2002).

Even the RNA degradation is carried out specifically by Xrn2, not Xrn1, termination still does not necessarily occur as manifested by the cotranscriptional cleavage (CoTC) element. The CoTC element was first discovered as a piece of highly unstable RNA downstream the human β-globin poly(A) signal, which was later characterized as a ribozyme bearing autocatalytic cleavage activity (Teixeira et al., 2004). The deletion of the CoTC element resulted in defective PADT (Dye and Proudfoot, 2001; West et al., 2004). Considering its auto-cleavage activity, the role of CoTC in termination was proposed to provide a 5' phosphate end for Xrn2 to degrade the downstream RNA and promote termination (West et al., 2004). However, when the upstream poly(A) signal was mutated, read-through transcription persisted (West et al., 2008), even though the CoTC element retains its auto-cleavage activity and is still able to generate a 5' phosphate end for Xrn2 to degrade the downstream RNA (Teixeira et al., 2004). This well demonstrates that RNA degradation specifically by Xrn2 cannot cause dissociation of the transcription complex. Even when the poly(A) signal is not mutated, a 5' phosphate end can be provided to Xrn2 by RNaseH cutting the RNA tether downstream the poly(A) signal, but rather than promoting termination, this turned out to repress it (Figure 3-2D).

On the other hand, when a hammerhead ribozyme was inserted downstream the poly(A) signal, which generates a 5' hydroxy end blocking the access of Xrn2, the efficiency of PADT was not affected (West et al., 2008). This argues that RNA degradation by Xrn2 is not only insufficient but also unnecessary for PADT to occur.

Thus, the contribution of Rat1/Xrn2 to PADT is far more than degrading RNA, but probably through direct or indirect interaction with the transcription complex. Consistent with this view, Rat1, after removal of its nuclease activity by mutation, can still dismantle the Pol II elongation complex just through interaction with Rtt103, which binds to the polymerase CTD (Pearson and Moore, 2013). Interestingly, the effect of Rat1-Rtt103 on the elongation complex exactly recapitulates that of Pcf11 CID (Zhang et al., 2005), implying both may function as allosteric factors to rearrange the transcription complex for termination. In fact, Rat1/Xrn2 is not the first example of a protein with an established function but also acting as an allosteric factor in termination. It well parallels with the E. coli helicase Rho, which was thought to terminate transcription through unwinding DNA-RNA duplex, but turned out to be an allosteric factor triggering a conformational change of the transcription complex (Epshtein et al., 2010).

In recent years, the yeast helicase Sen1 and its human homolog senataxin have been reported to affect transcription termination of some pre-mRNAs (Skourti-Stathaki et al., 2011; Steinmetz et al., 2006). Although the mechanism of Sen1 in pre-mRNA transcription termination is elusive, it has been well characterized as part of the termination complex of small nucleolar RNA (snoRNA) and other short RNAs. Specifically, Sen1 is recruited to the 5' region of genes through interaction with Nrd1, which has a CID, directly binding the polymerase CTD (Vasiljeva et al., 2008). Again, this resembles the CTD interaction of Rat1-Rtt103 and Pcf11, all resulting in dismantling of the transcription complex. Since the termination defect of the Sen1 mutant was only limited to short pre-mRNAs, the transcription process of which is very similar to snoRNA (Vasiljeva et al., 2008), the same termination mechanism may be adopted by both. Thus, Sen1/senataxin is potentially another allosteric factor that rearranges the transcription complex through interaction with polymerase CTD and facilities termination of short pre-mRNAs.

**PADT does not require RNA cleavage.**

To determine whether the core mechanism of PADT requires RNA cleavage and/or a conformational rearrangement, it is essentially to see whether termination can be uncoupled from RNA cleavage. A positive answer has long been suggested by the fact that certain CstF and Pcf11 mutants support termination but impair poly(A) site cleavage (Qu et al., 2007; Sadowski et al., 2003). However, these data have not aroused enough attention primarily due to decades of dominance of the idea that PADT depends on cleavage. Besides, these data themselves are undermined by their experimental design, in which the termination and poly(A) site cleavage activities were not measured concurrently. While PADT can only be reliably measured in vivo, poly(A) site cleavage efficiency is most accurately determined in vitro. Otherwise, the cleaved mRNA can be easily masked by various RNA degradation activities in vivo (Garneau et al., 2007). Since they were measured at entirely different conditions, the conclusions about their correlation are less persuasive. Therefore, the best approach to settle down this issue is to develop an in vitro termination assay that can concurrently measure PADT and poly(A) site cleavage. A few attempts have been made to assay termination in vitro by looking at the abundance of readthrough RNA, which was thought be stand for defective termination (needing reference). However, as previously mentioned pre-mRNA are vulnerable to various degradation activities (Garneau et al., 2007), so the absence of readthrough RNA cannot be just explained as efficient termination, which is also possibly due to degradation.

To examine definitively whether RNA cleavage and/or a conformational change constitute the core mechanism of PADT, we have developed an in vitro termination assay based on HeLa nuclear extract. This assay measures the Pol II occupancy on the DNA template after crossing a poly(A) signal, which is exactly how PADT is defined, avoiding any interference of the result by

RNA degradation. Also, the concurrent poly(A) site cleavage efficiency can be simultaneously determined just by splitting some sample from the termination assay (Chapter 3). Thus, by this approach the concluded correlation between PADT and poly(A) site cleavage will have no ambiguity. Using this assay, we unequivocally established that PADT occurs efficiently without poly(A) site cleavage. Also, our data revealed a conformational change, elicited by the poly(A) signal in the very early stage (within 4 min of transcription initiation), delivers the competence of the transcription complex to terminate. Specifically, PADT can be inhibited by α-amanitin, which locks the catalytic structures of the polymerase (Brueckner and Cramer, 2008) and presumably perturbs the conformational change required by PADT.

# Figures

**Figure 1-1. The latest model of mammalian pre-mRNA 3'-end processing complex assembles on the poly (A) signal in the newly transcribed RNA**

It includes Pol II (RNAP II) large subunit, four multi-subunit protein complex (CPSF, CstF, CF Im and CF IIm) and poly(A) polymerase (PAP). All these protein complexes except CF Im and PAP are characterized as termination factors of pre-mRNA transcription by Pol II. The RNA elements and the specific peptides of each protein complex involved in termination are highlighted in black and will be further discussed in the context. The white star in the RNA in contact with CPSF73 indicates the poly(A) site where cleavage occurs (modified from Figure 6 in Chan et al., 2011).

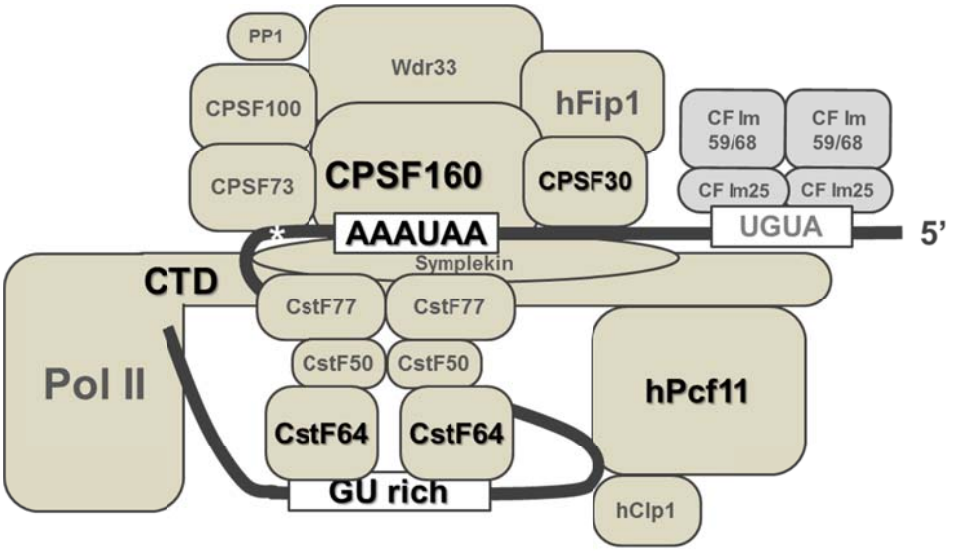
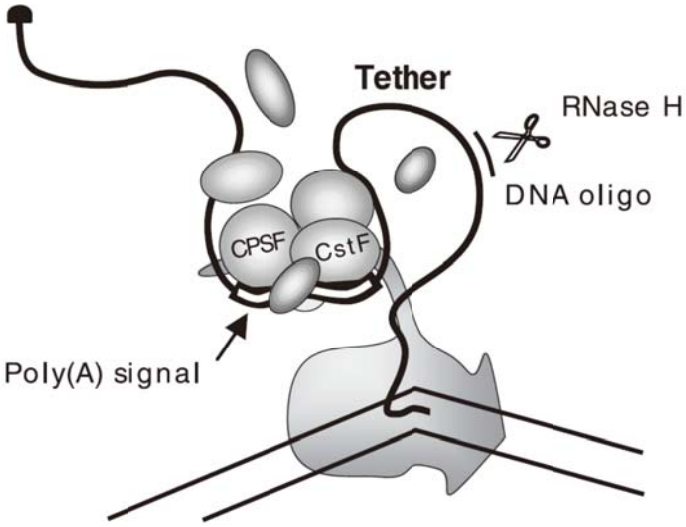

PP1
Wdr33
hFip1
CPSF100
CPSF73
CPSF160
CPSF30
CF Im 59/68
CF Im 59/68
CF Im25
CF Im25
AAAUAA
UGUA
5'
Symplekin
CTD
CstF77
CstF77
CstF50
CstF50
Pol II
hPcf11
CstF64
CstF64
GU rich
hClp1

**Figure 1-2. The RNA tether connects the poly(A) signal to the polymerase**

The cartoon shows a transcription complex in the process of assembling a poly(A)-recruited apparatus. A DNA oligonucleotide hybridizes to its target RNA and directs RNaseH to cut the tether (Rigo et al., 2005).


Tether
RNase H
DNA oligo
CPSF
CstF
Poly(A) signal

## References


Adamson, T.E., Shutt, D.C., and Price, D.H. (2005). Functional coupling of cleavage and polyadenylation with transcription of mRNA. J. Biol. Chem. *280*, 32262–32271.

Ashfield, R., Enriquez-Harris, P., and Proudfoot, N.J. (1991). Transcriptional termination between the closely linked human complement genes C2 and factor B: common termination factor for C2 and c-myc? EMBO J. *10*, 4197–4207.

Barillà, D., Lee, B. a, and Proudfoot, N.J. (2001). Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U. S. A. *98*, 445–450.

Brueckner, F., and Cramer, P. (2008). Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. Nat. Struct. Mol. Biol. *15*, 811–818.

Chan, S., Choi, E.-A., and Shi, Y. (2011). Pre-mRNA 3'-end processing complex assembly and function. Wiley Interdiscip. Rev. RNA *2*, 321–335.

Chao, L.C., Jamil, A., Kim, S.J., Huang, L., and Martinson, H.G. (1999). Assembly of the cleavage and polyadenylation apparatus requires about 10 seconds in vivo and is faster for strong than for weak poly(A) sites. Mol. Cell. Biol. *19*, 5588–5600.

Citron, B., Falck-pedersen, E., Salditt-georgieff, M., and Darnell, J.E. (1984). Nucleic Acids Research. *12*, 8723–8731.

Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. *2*, 440–452.

Dichtl, B., Blank, D., Sadowski, M., Hübner, W., Weiser, S., and Keller, W. (2002). Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. EMBO J. *21*, 4125–4135.

Dye, M.J., and Proudfoot, N.J. (2001). Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. Cell *105*, 669–681.

Edwalds-Gilbert, G., Prescott, J., and Falck-Pedersen, E. (1993). 3' RNA processing efficiency plays a primary role in generating termination-competent RNA polymerase II elongation complexes. Mol. Cell. Biol. *13*, 3472–3480.

Enriquez-Harris, P., Levitt, N., Briggs, D., and Proudfoot, N.J. (1991). A pause site for RNA polymerase II is associated with termination of transcription. EMBO J. *10*, 1833–1842.

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. Nature *463*, 245–249.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. *8*, 113–126.

Gilmartin, G.M., and Nevins, J.R. (1989). An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev. *3*, 2180–2190.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517–522.

Kireeva, M.L. (2000). The 8-Nucleotide-long RNA:DNA Hybrid Is a Primary Stability Determinant of the RNA Polymerase II Elongation Complex. J. Biol. Chem. *275*, 6530–6536.

Komissarova, N., Becker, J., Solter, S., Kireeva, M., and Kashlev, M. (2002). Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol. Cell *10*, 1151–1162.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nat. Rev. Mol. Cell Biol. *12*, 283–294.

Logan, J., Falck-Pedersen, E., Darnell, J.E., and Shenk, T. (1987). A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. Proc. Natl. Acad. Sci. U. S. A. *84*, 8306–8310.

Lunde, B.M., Reichow, S.L., Kim, M., Suh, H., Leeper, T.C., Yang, F., Mutschler, H., Buratowski, S., Meinhart, A., and Varani, G. (2010). Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. *17*, 1195–1201.

Luo, W., Johnson, A.W., Bentley, D.L., Rosonina, E., Kaneko, S., and Manley, J.L. (2006). torpedo model The role of Rat1 in coupling mRNA 3 -end processing to transcription termination : implications for a unified allosteric – torpedo model. Genes Dev. *20*, 954–965.

Mariconti, L., Loll, B., Schlinkmann, K., Wengi, A., Meinhart, A., and Dichtl, B. (2010). Coupled RNA polymerase II transcription and 3’ end formation with yeast whole-cell extracts. RNA *16*, 2205–2217.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat. Struct. Mol. Biol. *17*, 1272–1278.

Meinhart, A., and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3’-RNA-processing factors. Nature *430*, 223–226.

Murthy, K.G., and Manley, J.L. (1995). The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. Genes Dev. *9*, 2672–2683.

Nag, A., Narsinh, K., Kazerouninia, A., and Martinson, H.G. (2006). The conserved AAUAAA hexamer of the poly ( A ) signal can act alone to trigger a stable decrease in RNA polymerase II transcription velocity The conserved AAUAAA hexamer of the poly ( A ) signal can act alone to trigger a stable decrease in RNA polymerase.

Nag, A., Narsinh, K., and Martinson, H.G. (2007). The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat. Struct. Mol. Biol. *14*, 662–669.

Orozco, I.J., Kim, S.J., and Martinson, H.G. (2002). The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. J. Biol. Chem. *277*, 42899–42911.

Park, N.J., Tsao, D.C., and Martinson, H.G. (2004). The Two Steps of Poly ( A ) -Dependent Termination , Pausing and Release , Can Be Uncoupled by Truncation of the RNA Polymerase II Carboxyl-Terminal Repeat Domain The Two Steps of Poly ( A ) -Dependent Termination , Pausing and Release , Can Be Uncoupled .

Pearson, E.L., and Moore, C.L. (2013). Dismantling promoter-driven RNA polymerase II transcription complexes in vitro by the termination factor Rat1. J. Biol. Chem. *288*, 19750–19759.

Peterson, M.L., Bertolino, S., and Davis, F. (2002). An RNA Polymerase Pause Site Is Associated with the Immunoglobulin s Poly ( A ) Site. *22*, 5606–5615.

Platt, T. (1986). AND THE REGULATION OF GENE EXPRESSION.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. *14*, 105–110.

Qu, X., Perez-Canadillas, J.-M., Agrawal, S., De Baecke, J., Cheng, H., Varani, G., and Moore, C. (2007). The C-terminal domains of vertebrate CstF-64 and its yeast orthologue Rna15 form a new structure critical for mRNA 3'-end processing. J. Biol. Chem. *282*, 2101–2115.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases Transcription termination by nuclear RNA polymerases. 1247–1269.

Rigo, F., Kazerouninia, A., Nag, A., and Martinson, H.G. (2005). The RNA tether from the poly(A) signal to the polymerase mediates coupling of transcription to cleavage and polyadenylation. Mol. Cell *20*, 733–745.

Sadowski, M., Dichtl, B., Hübner, W., and Keller, W. (2003). Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. EMBO J. *22*, 2167–2177.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell *42*, 794–805.

Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol. Cell *24*, 735–746.

Takagaki, Y., and Manley, J.L. (2000). Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component. *20*.

Teixeira, A., Tahiri-Alaoui, A., West, S., Thomas, B., Ramadass, A., Martianov, I., Dye, M., James, W., Proudfoot, N.J., and Akoulitchev, A. (2004). Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. Nature *432*, 526–530.

Vasiljeva, L., Kim, M., Mutschler, H., Buratowski, S., and Meinhart, A. (2008). The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat. Struct. Mol. Biol. *15*, 795–804.

Vickers, T. a, Wyatt, J.R., Burckin, T., Bennett, C.F., and Freier, S.M. (2001). Fully modified 2' MOE oligonucleotides redirect polyadenylation. Nucleic Acids Res. *29*, 1293–1299.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522–525.

West, S., Proudfoot, N.J., and Dye, M.J. (2008). Article Molecular Dissection of Mammalian RNA Polymerase II Transcriptional Termination. 600–610.

Whitelaw, E., and Proudfoot, N. (1986). oa-Thalassaemia caused by a poly ( A ) site mutation reveals that transcriptional termination is linked to 3 ' end processing in the. *5*, 2915–2922.

Yonaha, M., and Proudfoot, N.J. (1999). Specific transcriptional pausing activates polyadenylation in a coupled in vitro system. Mol. Cell *3*, 593–600.

Zhang, Z., and Gilmour, D.S. (2006). Pcf11 is a termination factor in Drosophila that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. Mol. Cell *21*, 65–74.

Zhang, Z., Fu, J., and Gilmour, D.S. (2005). CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. Genes Dev. *19*, 1572–1580.

# *Chapter 2*

## Role of the Nascent RNA in the Assembly of the Cleavage and Polyadenylation Apparatus

## Abstract

The continuity of nascent RNA downstream the poly(A) signal is critical to cotranscriptional 3'-end processing of pre-mRNA. The role of the RNA that tethers the poly(A) signal to the polymerase is proposed to keep the 3'-end processing protein machinery in proximity to the polymerase until its maturation. But another possibility has never been ruled out that the RNA tether may structurally constitute the processing apparatus and play an allosteric role in 3'-end processing. Here we took advantage of a specially modified oligonucleotides that only hybridize to the RNA, but not directs RNase H cutting, and established that it is the cutting of the tether but not the hybridization by other oligonucleotides that interferes with 3'-end processing. Also enabled by the specially modified oligonucleotides, we demonstrated the assembly of CstF to the poly(A) signal is as early as the extrusion of the hexamer and its subsequent binding to the downstream GU-rich transcript is fast.

# Introduction

The mature 3'-end of the precursor messenger RNA (pre-mRNA) transcribed by RNA polymerase II (Pol II) is generated by cleavage and polyadenylation at the poly(A) signal. This processing is coupled to transcription both in vivo and in vitro (Bird et al., 2005; Rigo et al., 2005). The coupling phenomenon has been thought to occur through a variety of means such as pre-recruitment of processing factors to the promoter, kinetic coupling, and allosteric activation by the C-terminal domain (CTD) of RNA polymerase II (Bentley, 2002). While several modes of coupling have been proposed, it is not entirely clear to what extent each mechanism contributes to coupling.

In addition to the importance of the CTD in coupling, the RNA tether that connects the poly(A) signal to the polymerase also plays an important role in the coupling of 3'-end processing to transcription in vitro (Rigo et al., 2005). Severing the RNA tether by oligo-directed RNase H cleavage dramatically disrupts the coupled processing. Since the cutting of this same RNA sequence has no effect when 3'-end processing activity is reconstituted in absence of transcription, it was proposed that the role of the RNA tether may be to keep the poly(A) signal close to the polymerase unit the maturation of the cleavage and polyadenylation apparatus during transcription. However, it is not entirely clear whether the disruption of coupling was due to the RNA tether being severed or that the binding of the DNA oligo itself precluded coupling.

The core poly(A) signal is composed of an AAUAAA hexamer sequence and a downstream GU-rich element, which are recognized by core processing factors cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulatory factor (CstF) respectively (Zhao et al., 1999). Besides these two complexes, the full assembly of the cleavage and polyadenylation apparatus

(CPA) may involve greater than 80 proteins according to a recent mass spectrometry study (Shi et al., 2009). While much is known regarding the identity of the processing factors involved in the cleavage and polyadenylation reaction, very little is known about the assembly process of the CPA. However, there is evidence that assembly of the CPA on the poly(A) signal is a stepwise process in vivo (Chao et al., 1999).

More recently, an assembly pathway was proposed based on immunoprecipitation experiments that had identified discrete complexes involving CPSF, CstF and C-terminal domain (CTD)-less RNA polymerase II and CTD-intact RNA polymerase II (Nag et al., 2007). The pathway begins with CPSF riding on the polymerase body before the poly(A) signal is transcribed. CPSF scans the RNA and upon extrusion of the poly(A) hexamer, CPSF binds the AAUAAA sequence and directs the polymerase to pause. Once the GU-rich region is transcribed, CstF joins CPSF at the poly(A) signal, and the two complexes are transferred to the CTD. Additional cleavage and polyadenylation factors such as cleavage factor Im (CF Im), cleavage factor IIm (CF IIm), and poly(A) polymerase (PAP) join sometime later and cleavage and polyadenylation ensues. Based on initial studies showing that the formation of a stable CPSF-CstF complex requires both the AAUAAA hexamer and the downstream GU-rich region (Gilmartin and Nevins, 1989), it has been thought that the two complexes are recruited to their respective binding sites on the poly(A) signal independently and in a step-wise fashion. However, this has not been directly tested.

In recent years, modified oligonucleotides have been utilized to study regulation of polyadenylation site usage. Among them, 2'-O-methoxyethylribose-modified oligonucleotides (MOEs) have been shown to bind RNA without inducing RNase H cutting (Vickers et al., 2001). Another important characteristic is that MOEs bind more strongly to RNA compared to

unmodified DNA oligonucleotides. These two characteristics allowed us to address questions relating to the coupling of RNA 3'-end processing to transcription.

The intent of this study was to study the role of the tether in the coupling of 3'-end processing to transcription and the assembly process of the cleavage and polyadenylation apparatus. We first addressed whether the role of the tether is to keep the poly(A) signal close to the polymerase as proposed in a previous study (Rigo et al., 2005) or whether the tether may act as an allosteric effector. We found that coupling was only disrupted when the RNA tether was cut but not when it was bound by MOEs. Our results are consistent with the previous proposition that the role of the tether may be to keep the poly(A) signal physically close to the transcribing polymerase during the early stages until the processing apparatus is able to remain stably associated with the polymerase on its own. Second, we studied the assembly process of the two essential cleavage and polyadenylation factors, CPSF and CstF. We established in two independent approaches that the recruitment of CstF starts from the AAUAAA hexamer region. Our results suggest that CstF may be initially recruited via protein-protein interactions with CPSF to the hexamer followed by addition stabilization through binding to the GU-rich region. Furthermore, we found that the capture of the subsequently extruded GU-rich elment by CstF is not a slow step, at least faster than the hybridization of MOEs to the RNA.

## Results

**The tether mediates the coupling of 3'-end processing to transcription as a connector rather than an allosteric factor.** The coupling between 3'-end processing and transcription was first demonstrated through the tether-cutting experiment (Rigo et al., 2005). Then the more important question is what the role of the tether is. Since the direct effect of tether-cutting is that the processing apparatus loses its connection with the polymerase and falls apart from the transcription complex, the most favorable proposal is that the tether is a connector that keeps the processing apparatus docking on the transcription complex. But the tether is also part of the transcription complex, so it is equally possible that it is an allosteric factor that has to be kept in a specific conformation to activate the processing apparatus. In this case, simply the binding of the tether by a DNA oligo can disrupt its conformation no matter whether it is cut by RNase H or not. In order to distinguish the two roles, we need to experimentally separate the binding and cutting events. Since MOEs can bind to the tether and meanwhile it is resistant to RNase H cutting (Vickers et al., 2001), it was utilized to mimic the binding of DNA oligo to the tether. The coupled processing assay was carried out to test the role of the tether. To simply put, after the formation of preinitiation complexes, transcription was pulsed with NTPs including [$\alpha$-$^{32}$P] CTP for RNA labeling. During the pulse, either DNA or MOE oligos were added so that they can bind immediately to the targeted region. Then transcription was chased with a high concentration of unlabeled CTP and stopped with $\alpha$-amanitin and 3'-dATP. With another 30min incubation, RNA was gradually cleaved at the poly(A) site (Lane 1 in Figure 2-1A). Consistent with the previous observation (Rigo et al., 2005), poly(A) site cleavage band was mostly gone when transcription was carried out in the presence of complementary DNA oligos that induce RNase H cutting of the transcript in vector sequences 79 nucleotides downstream of the poly(A)

site (Lane 2). In contrast, when MOE oligos complementary to the same region was added, poly(A) site cleavage remained efficient (Lane 3). MOE oligos can hybridize to the tether and disrupt its single strand structure just as DNA oligos, but the hybridization by MOE oligos still supports efficient poly(A) site cleavage. Therefore, the role of the tether in mediating coupled processing cannot be an allosteric factor.

To verify that the above MOE oligos bind to the same region as the DNA oligos, they were added together to see whether the disruption of processing by the DNA oligos can be compromised by competition of MOE. Figure 2-1A shows that processing efficiency was rescued from 5% back to 13% due to addition of MOE (Lane 5 and 6). This suggests the MOE oligos block the binding of the DNA oligos, which compromised the defect of tether-cutting on processing. Consistently, RNase H cutting was very efficient when only the DNA oligos were added, but drastically reduced when the MOE oligos were added together (Lane 5 and 6). This shows that the MOE oligos hybridize to the same region as the DNA oligos.

To test whether the above conclusion also holds for the most basic model of coupled processing, we compared the effects of the DNA and MOE oligos on 3'-end processing that couples to transcription occurring on a promoter-less DNA template. The experiment was carried out similarly as above with a few modifications. A same DNA template but without Pol II promoter was used to exclude any effect from the promoter on the coupling of 3'-end processing to transcription. Also to avoid any effect from the general transcription factors during initiation, the initiation complexes were formed with purified Pol II after a short pulse and then chased in nuclear extract for efficient elongation. Figure 2-1C shows that processing efficiency was sharply reduced with the addition of the DNA oligos cutting 79 nucleotides downstream the poly(A) site (Lane 1 and 2). When the corresponding MOE oligos were added, processing

remained efficient (Lane 3 and 4). Thus, it shows again the disruption of the RNA tether structure does not affect processing. In fact, the MOE oligos that hybridize to the poly(A) signal, instead of the tether, are well capable to disrupt processing. As shown in Lane 6, when anti-GU MOEs, which are complementary to the GU-rich region in SV40L poly(A) signal, were added, processing was completely abolished presumably because the anti-GU MOEs bound to the GU-rich region and blocked its access by the cleavage stimulatory factor (CstF). Therefore, with the advantage of MOE oligos, we demonstrate that in the simplest model, role of the tether in mediating coupled processing is still a connector rather than an allosteric factor.

**CstF joins the processing apparatus from the hexamer instead of the GU-rich region.** It is generally agreed that CPSF is recruited to the transcription complexes from the poly(A) signal, more specifically, from its binding partner on the poly(A) signal, the hexamer. But it is still ambiguous where the recruitment of CstF starts. To address this question, we can pull down the transcription complexes with CstF antibody and identify the RNA associated with CstF, the length of which can indicate where the association of CstF starts. The transcription complexes were generated as those in coupled processing assays. After a few minutes of chase, Anti-CstF immunoprecipitation (IP) was performed to the transcription complexes either containing a wildtype poly(A) signal or a mutant and then the RNA associated with CstF were further analyzed. We found that RNA transcribed beyond the wildtype poly(A) signal were specifically pulled down compared with either shorter ones or the ones bearing the mutant poly(A) signal ( Lane 1 and 2, Figure 2-2A). This is clearly illustrated in the line-graph overlay that RNA start peaking from the poly(A) signal and all the way until the end of transcription. This suggests the recruitment of CstF to the transcription complexes starts from the poly (A) signal. Next, we wanted to know more specifically from the extrusion of which element in the poly (A) signal the

binding of CstF starts. It is known that CPSF is recruited uniquely to the hexamer. If CstF joins from the GU-rich region, the RNA pulled down by CstF should be a little bit longer than that by CPSF. So we compared the IP profiles of CPSF and CstF. However, since the hexamer and GU-rich region are so close to each other (about 20 nucleotide in-between), they are impossible to be resolved on gel. As a result, CstF appeared to be pulled down from the same region as CPSF (Lane 3 and 4). To solve this problem, we inserted a piece of 265-nucleotide RNA between the hexemer and GU-rich region so that they can be resolved from each other (Figure 2-2B). Meanwhile, this piece of RNA can form a stem-loop structure, which will bring the GU-rich region back to proximity to the hexamer and hence resume the 3'-end processing function of the poly (A) signal that is disrupted by the increased distance between the hexamer and the GU-rich region (Ahmed et al., 1991). Since the stem-loop structure was inserted upstream of the poly(A) site, cleavage occurs about 280 nucleotide downstream the hexamer but within 10 nucleotide upstream the GU-rich region (Lane 1 and 2, Figure 2-2B). To know exactly where the recruitment of CstF starts, we carried out the same IP experiment with the transcription complexes containing the above engineered poly(A) signal either in the wildtype form or the mutant form. Line-graph overlay shows that CstF started to pull down RNA from the hexamer instead of the GU-rich region (also comparing Lane 1 and 2 in Figure 2-2C). Since CstF does not bind the hexamer directly, this strongly suggests that CstF is recruited to the hexamer through preassembly with CSPF.

Another reasonable explanation for the above result is that within the 5' region of the stem loop there is a degenerate GU-region which is not efficient enough to activate 3'-end processing but still able to bind CstF. In this case, the recruitment of CstF may not be triggered by the hexamer but instead by this arbitrary GU-rich region. To examine this possibility, we took advantage of

the MOE oligos. Since anti-GU MOEs can bind to the GU-rich region without being cut by RNase H, they can potentially block the binding of CstF so that the GU-rich region cannot contribute to the recruitment of CstF at all (cartoon in Figure 2-3A). With this advantage of anti-GU MOEs, we can use the SV40 late poly(A) signal without any modification. Figure 2-3A shows that anti-GU MOEs can block 3'-end processing efficiently (Lane 1 and 2). This suggests that anti-GU MOEs were successfully directed to its target. Next, we performed anti-CstF IP in presence of anti-GU MOE to see whether CstF is still recruited to the transcription complexes without the access to the GU-rich region. Figure 2-3B shows that CstF specifically pulled down RNA with the wildtype poly (A) signal even in presence of anti-GU MOEs, and the IP efficiency was comparable to that in presence of control MOEs (Lane 5 and 6 vs Lane 1 and 2, Figure 2-3B). Since the access to GU-rich region was blocked by MOE, the only other way to recruit CstF is through CPSF, which binds to the hexamer. To test this, we performed CstF IP in presence of anti-hexamer MOEs, which can also block 3'-end processing (Lane 1 and 3, Figure 2-3A), presumably by blocking the access of CPSF to the hexamer. Figure 2-3B shows that no RNA was pulled down by CstF in presence of anti-hexamer MOEs (Lane 3 and 4). This indicates that CstF is preassembled with CPSF, which specifically binds the hexamer.

**The binding of CstF to the GU-rich region is not a slow step.** We already showed that CstF joins the transcription complex together with CPSF at the hexamer and it is also known that CstF directly binds the GU-rich region. This immediately suggests there is an intermediate step that while CstF locates at the hexamer, it is waiting for the GU-rich region to extrude and catches it once the target shows up. Since the interaction of CstF to the GU-rich region is not strong by itself (Gilmartin and Nevins, 1989), we hypothesize that this intermediate step is a slow one, which may be what limits the rate of 3'-end processing (Figure 2-4A). If the recognition of the

GU-rich region by CstF is slow, then a significant amount of transcription complexes should exist as shown in the left of the cartoon, and thus anti-GU MOEs can inhibit processing while anti-hexamer MOE cannot. To test this hypothesis, we compared the kinetics of the processing inhibition by anti-GU MOEs with that by anti-hexamer MOEs. We profiled the extent of processing in presence of MOEs against the time of MOE addition. To start with a pool of uniform TECs, we first stripped all the processing factors off the TECs and then put them back to the transcription mixture (Figure 2-4B). As expected, the later the MOEs were added, the less inhibition they exerted on processing. This is consistent with the idea that processing factors are gradually assembled on their corresponding poly(A) elements so that MOEs gradually loses the access to their complementary parts. Next, we compared the processing profile in presence of anti-GU MOEs with that in presence of anti-hexamer MOEs. It turned out that anti-GU MOEs did not inhibit processing more than anti-hexamer MOEs regardless of the time of addition (Figure 2-4B). This suggests that the complexes shown in Figure 2-4A is indistinguishable from the ones that already have CstF bind to the GU-rich region; otherwise anti-GU MOEs should block processing to a larger extent than the anti-hexamer MOEs. The right panel shows a repeat of the above experiment with cold transcription. Still, the processing inhibition by the two MOEs was indistinguishable. Therefore, we conclude that the binding of CstF to the GU-rich region is not a slow step.

## Discussion

With the advantage of MOE oligos, we have examined the role of the tether in the coupling of 3'-end processing to transcription. We showed that it is the cutting of the tether rather than the binding of it that severely impairs processing (Lane 2 and 3, Figure 2-1A). This reveals that tether is a connector that keeps the processing apparatus docking on the transcription complex rather than an allosteric factor that has to maintain its structure in the processing apparatus.

Further, we looked into the assembly pathway of processing apparatus in the transcription-coupled manner. We showed that CPSF is recruited to the hexamer region (Lane 3 and 4, Figure 2-2A), which is consistent with the previous CHIP data showing that CPSF is recruited towards the 3'-end of transcription (Glover-Cutter et al., 2008). We further demonstrated that the recruitment of CstF also starts from the hexamer with two independent approaches: one involves the engineering of the poly (A) signal; the other involves the application of MOE oligos (Figure 2-2C and 2-3B). Although this strongly supports that the recruitment of both CPSF and CstF starts from the same region, it does not necessarily imply that they bind the hexamer at the same time as a complex. It is possible that CPSF first recognizes the hexamer, triggers the pause of the polymerase and leaves the polymerase body and then CstF joins CPSF without any assistance of the GU-rich region. In this case, it can be compatible with the previously reported data that CPSF and CstF do not bind the polymerase body at the same time (Nag et al., 2007). But Nag's data was based on IP from the nuclear extract after RNase treatment, which means it does not reflect the behavior of the transcription complex. So it is still open to question whether CPSF and CstF can bind as a complex on the polymerase body. Nevertheless, we unambiguously demonstrated that CstF is recruited to the transcription complex at the hexamer region.

It is known that CstF binds the GU-rich region, but its recruitment starts as early as the hexamer extrudes. Plus we demonstrated that the binding of CstF to the GU-rich region is fast (Figure 2-4B). The above facts collectively are consistent with the following model: CstF is initially recruited by CPSF to the hexamer region so that it can scan the downstream sequences right after their extrusion, which ensures that it captures the GU-rich region fast and efficient.

# Materials and methods

## Oligomers

1: 5'-GTTGGACTCAAGACGATAGTTACC-3'

2: 5'-biotin-CACATTTCCCCGAAAAGTGCCACC-3'

3: 5'-GGGTCTGCGGAGAGGCTGGCAGATTG-3'

4: 5'-CCAAGCTACCGAGCTCTTTTTTGGTACCCCTTGGGAGC-3'

5: 5'-GCTCCCAAGGGGTACCAAAAAAGAGCTCGGTAGCTTGG-3'

6: 5'-CCACTCATGTATGTGATTTTTTAGGACGCAGACGCCAAC-3'

7: 5'-GTTGGCGTCTGCGTCCTAAAAAATCACATACATGAGTGG-3'

8: 5'-TAAGCTGCAATAAACAAGCAGGAGTCTATAAAAGCGTG-3'

9: 5'-CAATTGTTGTTGTTAAGGCAGGGTCAGGCAAAGCGTG-3'

10: 5'- CTTAACAACAACAATTGCATTCATTTT-3'

11: 5'-TTGTTTATTGCAGCTTATAATGGTTAC-3'

12: 5'-GTAACCATTATAAGCTGCAAGTACCAACTTAACAACAACAATTGC-3'

13: 5'-GCAATTGTTGTTGTTAAGTTGGTACTTGCAGCTTATAATGGTTAC-3'

14: 5'-CCATTATAAGCTGCAAGTACCAAGCAGGAGTCTATA-3'

15: 5'-TATAGACTCCTGCTTGGTACTTGCAGCTTATAATGG-3'

Anti-GU MOE: 5'-CTGAAACATAAAATGAATGC-3'

Anti-hexamer MOE: 5'-GTTTATTGCAGCTTA-3'

79 MOE: 5'-GAGATGAGTAGGGAGTATTG-3'

79 DNA : 5'-GAGATGAGTAGGGAGTATTG-3'

Control MOE: 5'-CAAACCACAACTAGAATGCAGTG-3'

Control DNA: 5'-CAAACCACAACTAGAATGCAGTG-3'

**Plasmid templates.** The plasmid pSV40E/L was described previously (Rigo et al., 2005). The plasmid pDsxHIV2/L was identical to Gal5-HIV2dsxΔInt(+ESE) (Rosonina et al., 2003). The plasmid pRexSV40L was generated by cold-fusion cloning (MC101B-1, SBI). The insert of the stem loop was PCR-generated from the pgTAX-R' vector (Ahmed et al., 1991) by oligomer 8 and 9. The linearized pDsxHIV2/L was PCR-generated by oligomer 10 and 11. The mutant pDsxHIV2/L were generated by site-directed mutagenesis from oligomer 12 and 13; the mutant pRexSV40L from oligomer 14 and 15.

**Immobilized templates**. The immobilized templates dsxSV40L were generated by PCRing up the pDsxHIV2/L from -781 bp upstream of the transcription start site down 464 bp past the SV40 Late poly(A) cleavage site using oligomer 1 and 2 as primers. Oligomer 2 is a reverse biotinylated primer which enables bead attachment to the downstream end of the template. The PCR product of correct length was gel-extracted and purified using the QIAquick Gel Extraction Kit. Binding of the template to magnetic Dynal M-280 Streptavidin beads (Invitrogen) was carried out according to manufacturer's protocol. The final concentration of DNA on streptavidin beads was 0.8-3.1 pmol of DNA template per mg of beads. We found that a DNA density greater than 3.1 pmol of 1.7 kb-long DNA per mg of beads gave reduced processing efficiency in a coupled processing assay, presumably due to steric interference. Reducing DNA density on the beads to less than 3.1 pmol of DNA per mg of beads did not have any further effect on processing efficiency but did increase transcription.

Promoterless oligo-dTtaildsx template was synthesized by PCRing up the dsxSV40L template from the +4 position and replacing the first three basepairs with GGG using the forward primer (oligomer 3) and a biotinylated reverse primer (oligomer 2) to amplify up to 464 bp downstream of the SV40 late cleavage site. The templates were then bound to Dynal M-280 Streptavidin beads according to manufacturer's protocol. Final concentration of DNA concentration on beads used was 3.1 pmol of DNA template per mg of beads. The oligo-dT tail was then added by treatment with 40 units of terminal transferase (New England Biolabs) and 7.5 μM TTP for 30 minutes at 37$^{o}$C. The presence of the magnetic beads on the downstream end of the template forced the oligo dT tail to be added only to the upstream end of the template. To discourage RNA pol III transcription, two patches of 6 consecutive thymidine residues (one present beginning at position 174 bp and another at 218 bp downstream of the transcription start site)

were inserted by 2 rounds of site-directed mutagenesis (Strategene) using oligomer 4 and 5 as primers for position 174 and oligomer 6 and 7 as primers for position 218.

**Purification of Gal4-p53 activator.** G4l-p53 construct (gift from James Goodrich) was expressed in the E. coli BL21 strain. T7 promoter expression was induced in these cells with 0.4 mM IPTG for 2 hours. Cells were collected, resuspended in TGED buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 20% glycerol, 0.1 mM EDTA, 1 mM DTT, 1mM PMSF) containing 5 mM $MgCl_2$, and lysed via sonication. Cell lysates were incubated with ANTI-FLAG M2 affinity gel (Sigma, A2220) for 3 hours at 4$^{o}$C. The affinity gel was washed with TGED buffer and TGED buffer containing 1M NaCl. The affinity gel was transferred to a Poly-Prep Chromatography Column (Biorad) and Gal4-p53 proteins were released with TGED buffer containing 0.2 mg/mL FLAG peptide at room temperature. The fractions were collected and immediately stored at -80$^{o}$C. Peak fractions containing Gal4-p53, as determined by SDS-PAGE and sypro ruby staining, were later pooled together, re-aliquoted and stored at -80$^{o}$C.

**Purification of RNA Polymerase II from HeLa nuclear extract.** 8WG16 ascites fluid (Covance) was purified using Protein-A Agarose Beads (Santa Cruz Biotechnology, sc-2001) as described (Thompson et al., 1990). Purified 8WG16 antibodies were conjugated to CNBr-activated Sepharose 4B beads (GE Healthcare Life Sciences) as described, except a final concentration of 2.2 mg of 8WG16 antibodies per mL of beads was used for antibody conjugation and 0.1 M Tris-HCl buffer pH 8.0 was used to block excess unreacted groups on the beads. Nuclear extract was incubated with 8WG16 sepharose beads at a ratio of 3.5 mL of nuclear extract per mL of beads for 3 hours. The beads were washed extensively with 150 bed volumes of TE buffer (50 mM Tris-HCl pH 7.9 and 0.1 mM EDTA) containing 200 mM ammonium sulfate. Polymerase II was eluted with 4 bed volumes of EB buffer containing 0.75

M ammonium sulfate and 40% ethylene glycol at room temperature and dialyzed for 5-6 hours at 4$^{o}$C against TE buffer containing 1 mM DTT, 150 mM ammonium sulfate and 50% glycerol. Purified RNA Polymerase II was stored in liquid nitrogen.

**Coupled processing assay**. Coupled processing assay was carried out essentially as described (Rigo et al., 2005), but in a total volume of 15µl transcription mixture containing 6.7% glycerol, 6.7 mM HEPES (pH 7.9), 33 mM KCl, 66 µM EDTA, 1.7 mM DTT, 33 µM PMSF, 10 U anti-RNase, 4 mM $MgCl_2$, 3.3 mM sodium citrate (pH 6.7), 0.3 µg DNA, 167 µM each of ATP, UTP, and GTP, 16.7 mM creatine phosphate, 1.7 mM CTP and 1.6 µM of [α-$^{32}$P] CTP (20 µCi). For experiments that contained DNA templates other than pSV40E/L, purified Gal4-p53 activator was added during preincubation. The volume of Gal4-p53 addition was selected based on a titration experiment determining the minimum volume required to achieve maximum transcription in a typical coupled processing experiment. In most experiments, oligonucleotides were added during the pulse with a final concentration of 9 ng/µL. α-amanitin and 3′ dATP were added in a 1 µl volume with the final concentration of 32 ng/µl α-amanitin and 340 µM 3'dATP.

**Promoterless template transcription.** Transcription elongation complexes were generated from promoterless oligo-dT tailed templates by incubating 2.4 µL of purified RNA polymerase II with anti-RNase (Ambion), DTT, $MgCl_2$, sodium citrate, beaded DNA template, and water up to 7.8 µL. Volume of RNA polymerase II was selected based on a titration experiment determining the minimum volume required to achieve maximum transcription from 0.16 pmol of oligo-dTtaildsx template in a typical coupled processing experiment. The mixture was preincubated at 30$^{o}$C for 5 min and then pulsed with 3 µL containing 20 µCi of [α-$^{32}$P]-CTP and nucleotide triphosphates without ATP. When TECs were washed, beads were collected on magnet (Dynal MPC-S) for approximately 30-60 seconds and washed with either 200 µL of high salt, high detergent strip

buffer (20 mM HEPES pH 7.9, 1 M KCl, 1% Sarkosyl) or 200 µL of low-salt transcription buffer (3.9 mM HEPES, 20 mM KCl, 39 µM EDTA, 3.9% glycerol, 0.1 mM DTT, 20 µM PMSF). Next, the TECs were resuspended in transcription buffer with nuclear extract to proceed with chase according to steps outlined in coupled processing assay (above). The final concentrations of every component is the same as stated under coupled processing assay, except the final concentrations of ATP, UTP, and GTP is 260 µM.

**Immunoprecipitation.** For CstF pull down, 16 µl of 6A9 and 16 µl 3A7 antibodies [Wallace, 1999] were incubated with 16 µL of Dynabeads Protein G (Invitrogen) at room temperature for about 3 hours. The beads were then rinsed with 32 µL of PBS twice and mixed with transcription mixture for 10min at room temperature. Finally, the beads were separated from the supernatant, washed with 100 µL PBS and exposed to 65 µl of EPK containing 10 mM Tris HCl, 10 mM EDTA, 0.5% SDS, and 100 µg proteinase K (Ambion).

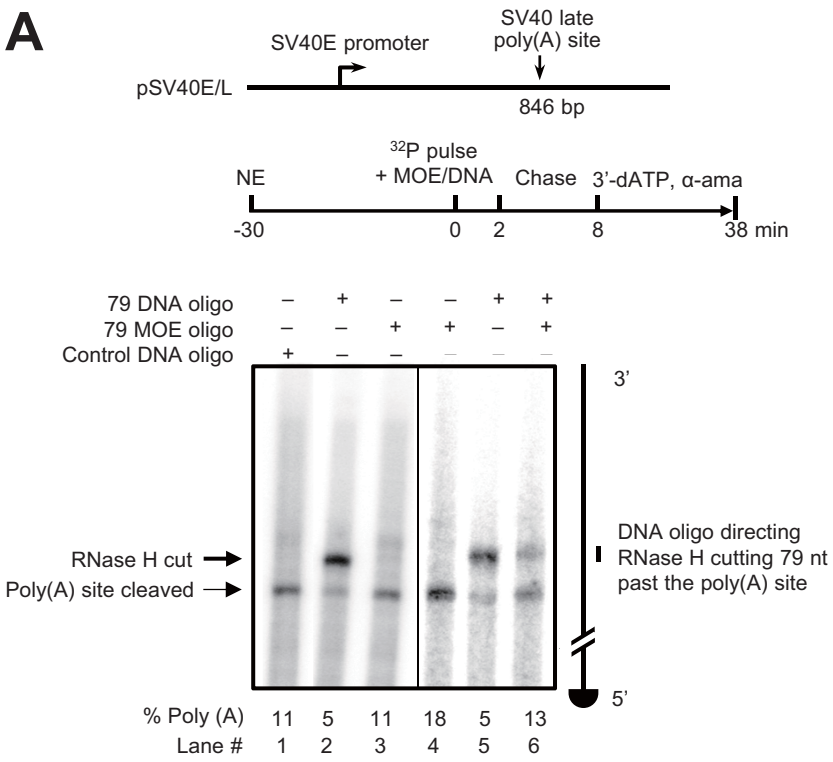
For CPSF pull down, 40 µl of CPSF N20 antibodies (Santa Cruz) were incubated with 20 µl Dynabeads Protein G (Invitrogen) at room temperature for about 5 hours. The following pull down was essentially the same as CstF except that the beads were incubated with transcription mixture for 1 hour and after separation from the supernatant, the beads were washed with 100 µl PBS with 0.02% Tween 20.

## Figures

**Figure 2-1. The tether mediates the coupling of 3'-end processing to transcription as a connector rather than an allosteric factor.**

A. MOE binds the tether but does not interfere with processing. Circular plasmid were transcribed for indicated amount of time. DNA oligo and/or MOE oligo were added together with $^{32}P$. The number in the oligo name refers to the oligo binding position downstream the poly(A) site. Control DNA oligo is complementary to the BGH poly(A) site and does not bind the RNA generated in this experiment. The % poly (A) is the ratio of the cleaved RNA to all RNA past the poly(A) site.

B. The above result is also valid without promoter initiation. Tailed template was used in this experiment. Purified polymerases were used to generate TECs. After 1.5min pulse, TECs were stripped with 1M KCl and 1% sarkosyl and resuspended in nuclear extract with DNA oligo or MOE oligo.


A
SV40E promoter
SV40 late poly(A) site
pSV40E/L
846 bp
NE
32P pulse + MOE/DNA
Chase
3'-dATP, α-ama
-30
0
2
8
38 min
79 DNA oligo
79 MOE oligo
Control DNA oligo
3'
5'
RNase H cut
Poly(A) site cleaved
DNA oligo directing RNase H cutting 79 nt past the poly(A) site
% Poly (A) 11 5 11 18 5 13
Lane # 1 2 3 4 5 6

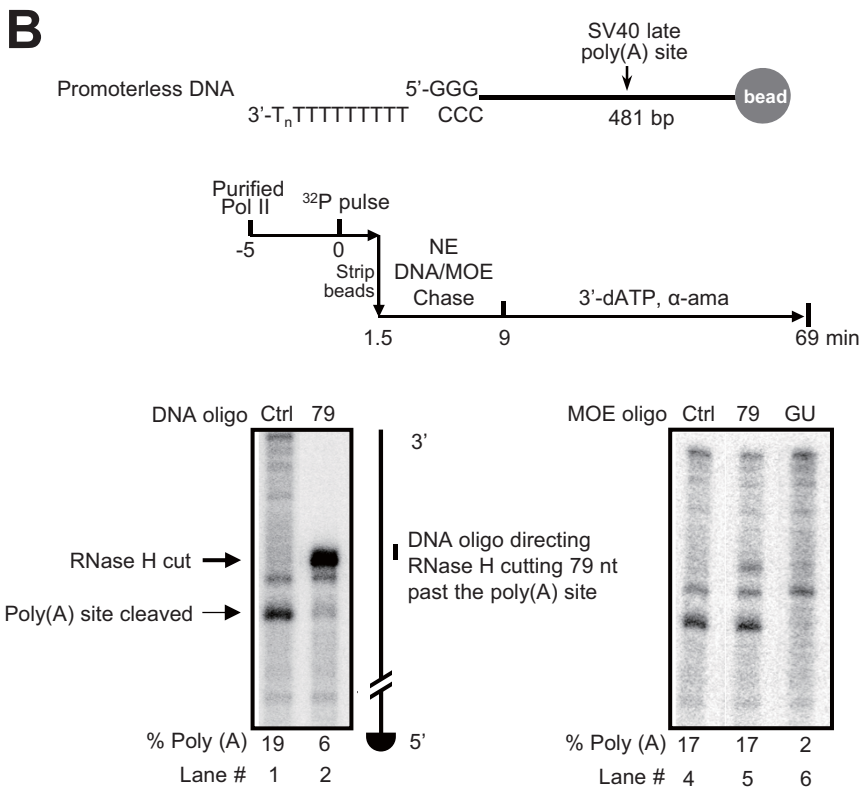

B
Promoterless DNA
5'-GGG
3'-Tn TTTTTTTTT
CCC
SV40 late poly(A) site
bead
481 bp
Purified Pol II
32P pulse
-5
0
Strip beads
NE
DNA/MOE
Chase
3'-dATP, α-ama
1.5
9
69 min
DNA oligo Ctrl 79
MOE oligo Ctrl 79 GU
3'
5'
RNase H cut
Poly(A) site cleaved
DNA oligo directing RNase H cutting 79 nt past the poly(A) site
% Poly (A) 19 6
Lane # 1 2
% Poly (A) 17 17 2
Lane # 4 5 6

**Figure 2-2. CstF joins the processing apparatus as early as the extrusion of the hexamer.**

A. Both CPSF and CstF specifically pull down RNA with poly(A) signal. The plasmid contains either a wildtype (wt) or mutant (mt) hexamer of the SV40 late poly(A) signal. Transcription mixture was subjected to immunoprecipitation (IP) with the indicated antibody. Anti-CstF IP was carried out for 10 min; anti-CPSF IP for 60 min. The RNAs pulled down in the pellet were displayed on gel. On the left of each gel panel is the line-graph overlay of the wildtype and mutant.

B. The processing of the engineered poly(A) signal is efficient. The cartoon illustrates the insertion of a stem loop structure between the hexamer and GU-region of the SV40 late poly(A) signal. Cleavage occurs right upstream the GU-region, and the cleaved RNA is 749 nt.

C. CstF pulls down RNA from the hexamer region.

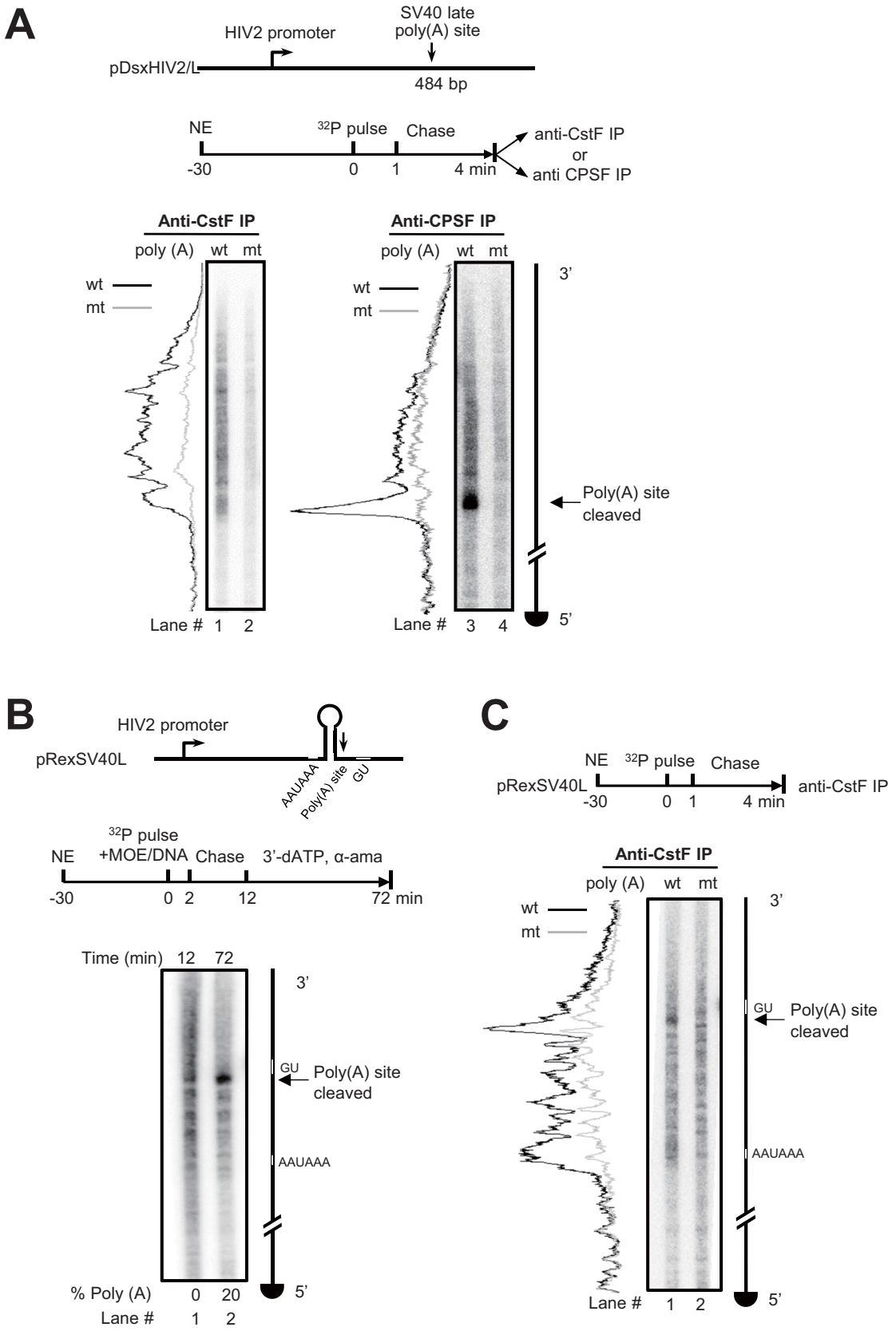

A
HIV2 promoter
SV40 late poly(A) site
pDsxHIV2/L
484 bp
NE
32P pulse
Chase
anti-CstF IP
or
anti CPSF IP
-30
0
1
4 min
Anti-CstF IP
Anti-CPSF IP
poly (A)
wt
mt
wt
mt
3'
Poly(A) site cleaved
5'
Lane #
1
2
3
4
B
HIV2 promoter
pRexSV40L
AAUAAA
Poly(A) site
GU
32P pulse
NE
+MOE/DNA
Chase
3'-dATP, α-ama
-30
0
2
12
72 min
Time (min)
12
72
3'
GU
Poly(A) site cleaved
AAUAAA
5'
% Poly (A)
0
20
Lane #
1
2
C
NE
32P pulse
Chase
pRexSV40L
anti-CstF IP
-30
0
1
4 min
Anti-CstF IP
poly (A)
wt
mt
wt
mt
3'
GU
Poly(A) site cleaved
AAUAAA
5'
Lane #
1
2

**Figure 2-3.** The MOE oligo binding the GU-rich element inhibits poly(A) site cleavage but not the recruitment of CstF.

A. Both ANTI-GU MOE and anti-hexamer MOE inhibit poly(A) site cleavage. The cartoon shows an RNA that has its GU-rich element bound by the MOE oligo so that CstF cannot access its cis-binding partner.

B. CstF recruitment is efficient in presence of ANTI-GU MOE, but blocked by anti-hexamer MOE.

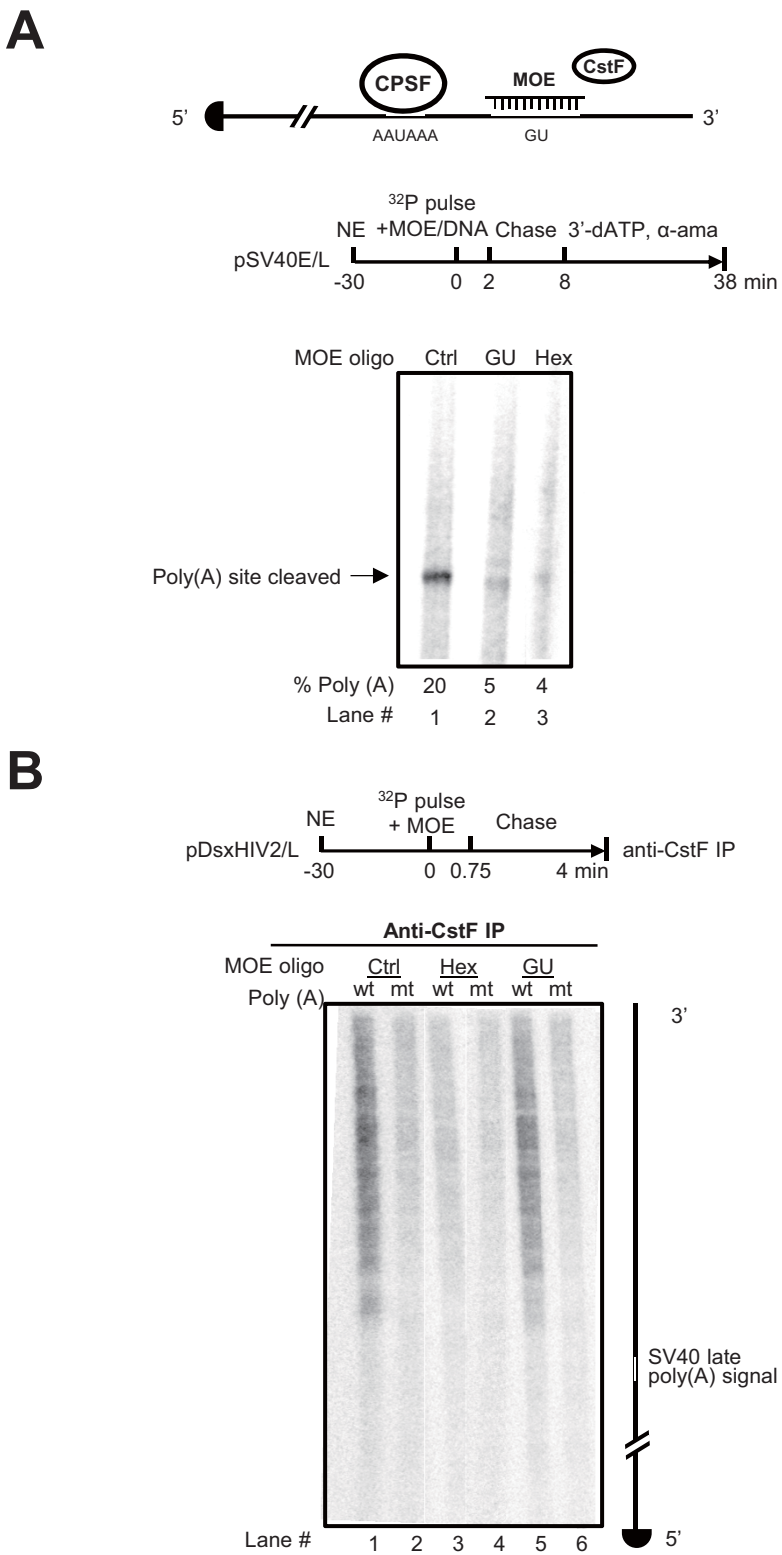
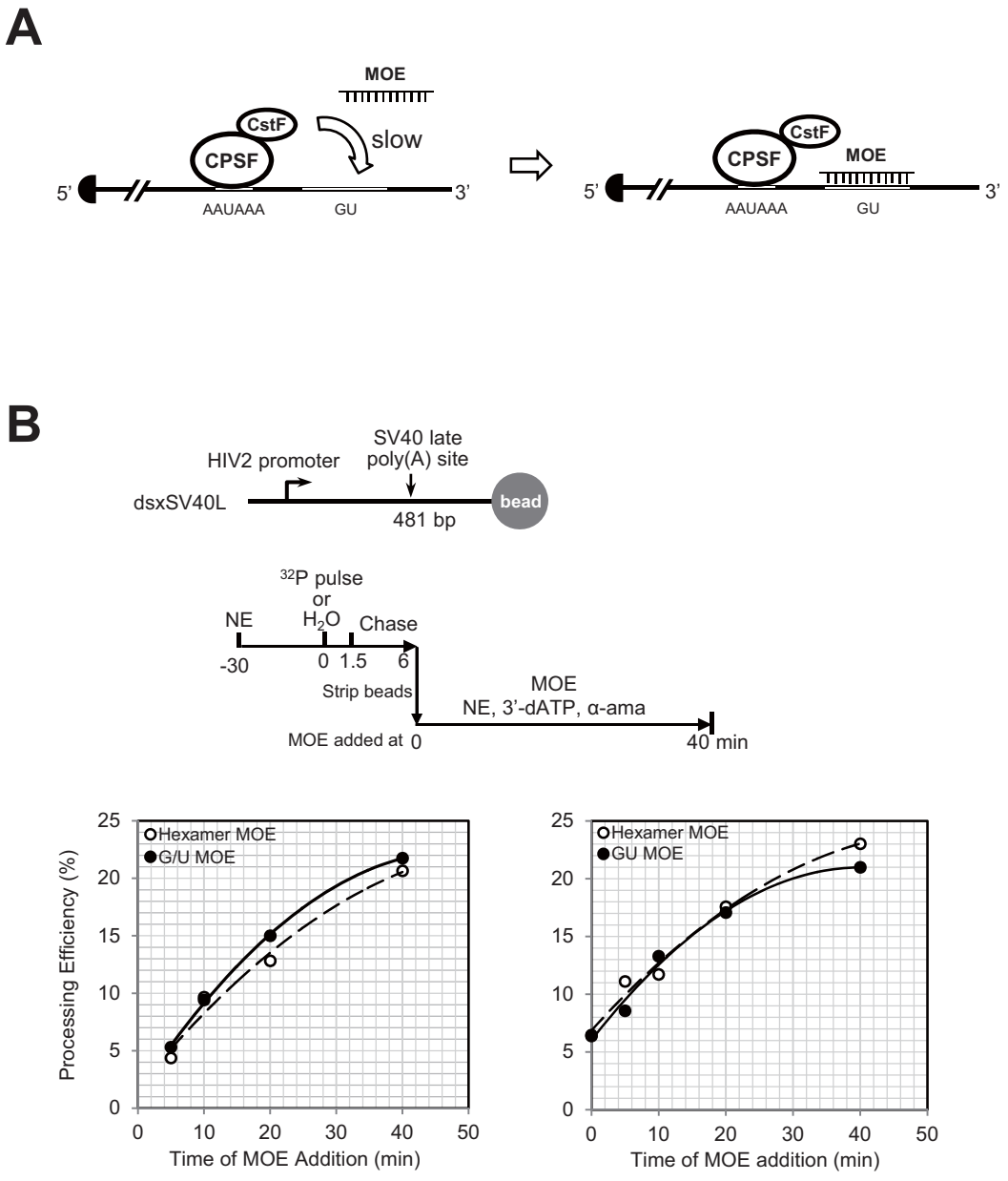

A
CPSF
MOE
CstF
5'
3'
AAUAAA
GU
32P pulse
NE +MOE/DNA Chase 3'-dATP, α-ama
pSV40E/L
-30 0 2 8 38 min
MOE oligo Ctrl GU Hex
Poly(A) site cleaved
% Poly (A) 20 5 4
Lane # 1 2 3
B
32P pulse
NE + MOE Chase
pDsxHIV2/L
anti-CstF IP
-30 0 0.75 4 min
Anti-CstF IP
MOE oligo Ctrl Hex GU
Poly (A) wt mt wt mt wt mt
3'
SV40 late
poly(A) signal
5'
Lane # 1 2 3 4 5 6

**Figure 2-4. The binding of CstF to the GU-rich region is not a slow step.**

A. If the binding of CstF to the GU-region is slow, a large number of transcription complexes should be in the state as shown in the cartoon. For these complexes anti-GU MOE has access to its RNA but anti-hexamer MOE does not, so 3'-end processing can be inhibited more efficiently by anti-GU MOE.

B. The anti-GU MOE does not inhibit processing more efficiently than the anti-hexamer MOE. The indicated MOEs were added after the resuspension of striped TECs. The TECs in the left panel was generated from cold transcription, essentially the same as coupled processing assay except that [$\alpha$-$^{32}$P] CTP was substituted with 2 μL $H_2O$. RNA was analyzed with RNase proection assay. The TECs in the right panel was generated as those in a standard coupled processing assay.


A
MOE
CstF
CPSF
slow
5'
3'
AAUAAA
GU
CstF
CPSF
MOE
5'
3'
AAUAAA
GU
B
HIV2 promoter
SV40 late
poly(A) site
dsxSV40L
bead
481 bp
32P pulse
or
NE
H2O
Chase
-30
0 1.5
6
Strip beads
MOE
NE, 3'-dATP, α-ama
MOE added at 0
40 min
Hexamer MOE
G/U MOE
Processing Efficiency (%)
Time of MOE Addition (min)
Hexamer MOE
GU MOE
Time of MOE addition (min)

# References


Ahmed, Y.F., Gilmartin, G.M., Hanly, S.M., Nevins, J.R., and Greene, W.C. (1991). The HTLV-I Rex response element mediates a novel form of mRNA polyadenylation. Cell 64, 727–737.

Bentley, D. (2002). The mRNA assembly line: transcription and processing machines in the same factory. Curr. Opin. Cell Biol. 14, 336–342.

Bird, G., Fong, N., Gatlin, J.C., Farabaugh, S., and Bentley, D.L. (2005). Ribozyme cleavage reveals connections between mRNA release from the site of transcription and pre-mRNA processing. Mol. Cell 20, 747–758.

Chao, L.C., Jamil, A., Kim, S.J., Huang, L., and Martinson, H.G. (1999). Assembly of the Cleavage and Polyadenylation Apparatus Requires About 10 Seconds In Vivo and Is Faster for Strong than for Weak Poly ( A ) Sites Assembly of the Cleavage and Polyadenylation Apparatus Requires About 10 Seconds In Vivo and Is Faster for Str.

Gilmartin, G.M., and Nevins, J.R. (1989). An ordered pathway of assembly of components required for polyadenylation site recognition and processing. Genes Dev. 3, 2180–2190.

Glover-Cutter, K., Kim, S., Espinosa, J., and Bentley, D.L. (2008). RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat. Struct. Mol. Biol. 15, 71–78.

Nag, A., Narsinh, K., and Martinson, H.G. (2007). The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat. Struct. Mol. Biol. 14, 662–669.

Rigo, F., Kazerouninia, A., Nag, A., and Martinson, H.G. (2005). The RNA tether from the poly(A) signal to the polymerase mediates coupling of transcription to cleavage and polyadenylation. Mol. Cell 20, 733–745.

Rosonina, E., Bakowski, M. a, McCracken, S., and Blencowe, B.J. (2003). Transcriptional activators control splicing and 3'-end cleavage levels. J. Biol. Chem. 278, 43034–43040.

Shi, Y., Di Giammartino, D.C., Taylor, D., Sarkeshik, A., Rice, W.J., Yates, J.R., Frank, J., and Manley, J.L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. Mol. Cell 33, 365–376.

Thompson, N.E., Aronson, D.B., and Burgess, R.R. (1990). Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. J. Biol. Chem. 265, 7069–7077.

Vickers, T. a, Wyatt, J.R., Burckin, T., Bennett, C.F., and Freier, S.M. (2001). Fully modified 2' MOE oligonucleotides redirect polyadenylation. Nucleic Acids Res. 29, 1293–1299.

Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mRNA 3 ′ Ends in Eukaryotes : Mechanism , Regulation , and Interrelationships with Other Steps in mRNA Synthesis Formation of mRNA 3 J Ends in Eukaryotes : Mechanism , Regulation , and Interrelationships with Other Steps in mRNA Synthesis.

# *Chapter 3*

## Poly(A)-Dependent Transcription Termination does not Require Poly(A) site Cleavage

## Abstract

Transcription termination by RNA polymerase II usually requires a functional poly(A) signal in the nascent pre-mRNA. Often called poly(A)-dependent temination, or PADT, it has been widely assumed that the poly(A) signal requirement reflects a poly(A) site cleavage requirement for termination. Cleavage was proposed to provide entry for a 5′ to 3′ exonuclease that would find the polymerase through degradation of the nascent transcript—the torpedo model. To test this model we developed an assay for PADT using Hela nuclear extract. Here we show that cleavage at the poly(A) site is not a prerequisite for PADT, ruling out the torpedo model. Isolated elongation complexes that contain an active poly(A) signal undergo PADT when incubated in buffer, in the absence of extract, nucleotides or cleavage at the poly(A) site. Thus, PADT-proficient complexes undergo a conformational change that leads to termination. PADT is inhibited by α-amanitin, which presumably blocks the required conformational change.

# Introduction

In order to guarantee appropriate termination of transcription at the ends of protein coding genes in eukaryotes, evolution has incorporated the transcription termination mechanism for these genes into their 3′-end processing mechanism (Richard and Manley, 2009). Thus, transcription termination by RNA polymerase II on genes encoding polyadenylated RNAs is dependent on the poly(A) signal (Kuehner et al., 2011; Hsin and Manley, 2012; Kim and Buratowski, 2013; Mischo and Proudfoot, 2013). This was first demonstrated decisively by Whitelaw and Proudfoot (1986) who showed, in the case of a human α-globin gene, that a single point mutation in the poly(A) signal hexamer impaired both 3′-end processing and transcription termination. Subsequently, Logan et al. (1987) offered two potential explanations for the dependence of termination on the poly(A) signal. The first possibility they proposed was that cleavage of the nascent transcript at the poly(A) site may be what leads to destabilization of the transcription complex, perhaps through exonucleolytic degradation of the RNA beginning at the uncapped 5′ end that is produced. The second possibility was that termination may result from an exchange of factors at the poly(A) signal that causes the polymerase to lose processivity. The first of these models was subsequently elaborated in more detail (Connelly and Manley, 1988) and dubbed the "torpedo" model (Osheim et al., 1999), referring to the 5′ to 3′ exonuclease that was proposed to be directly responsible for the termination. The second model is often referred to as the "conformational change" model (although termination by any mechanism obviously must involve conformational changes).

The most fundamental distinction between the torpedo model and the conformational change model is that the torpedo model carries with it an absolute requirement for cleavage at the poly(A) site (or elsewhere) in order to provide entry for the torpedo nuclease. An obvious way to

distinguish experimentally between these two models would be, then, to determine whether it is just the poly(A) signal per se that is needed for termination to occur, or whether actual cleavage at the poly(A) site is also required. Most studies relevant to this question have been done in yeast, where genetic lesions that impair poly(A) site cleavage almost always impair termination as well, consistent with a cleavage requirement for PADT. There are some exceptions (He et al., 2003; Sadowski et al., 2003; Zhelkovsky et al., 2006; Qu et al., 2007), but they suffer from interpretational uncertainties related to the fact that in vitro assays of cleavage must be compared with in vivo assays (i.e. run-on or ChIP) of termination. In metazoans, electron microscopic visualizations of transcription units are consistent with, but do not prove, the converse possibility, that termination may often precede cleavage (Osheim et al., 1999; Osheim et al., 2002). If so, this would indicate the existence of no cleavage requirement for PADT. However, the experimental system was only capable of showing that termination did not follow 3′ end cleavage by a discernable interval, and could not rule out the possibility that cleavage is indeed required, but leads very quickly to termination. Thus, it has not been possible so far to use direct approaches to determine whether poly(A) site cleavage is required for PADT.

The discovery several years ago that the yeast 5′ to 3′ exonuclease Rat1 is a factor that contributes to PADT provided indirect support for the torpedo model through the implication that poly(A) site cleavage would be required to provide this exonuclease access to the nascent RNA (Kim et al., 2004). Similar experiments on the human β-globin termination region revealed that Xrn2, the human homolog of Rat1, likewise contributes to PADT. However, human β-globin PADT does not require cleavage at the poly(A) site, so cleavage instead by a self-cleaving (CoTC) RNA element located in the β-globin termination region was proposed to be required (West et al., 2004; West et al., 2008). Evidence in favor of this possibility included mutations in

the CoTC element that impaired both cleavage in vitro and termination in vivo (Teixeira et al., 2004). However, just as mutations in the poly(A) signal inactivate both cleavage and termination—yet do not establish that cleavage is required for termination, so also, for the CoTC element, it remains unknown whether it is CoTC cleavage per se that contributes to termination, or the proteins that are recruited by the CoTC element in vivo.

Interestingly, nuclease activity per se does not lead to transcription termination (Gu et al., 1996; Luo et al., 2006; Dengl and Cramer, 2009). In contrast, several non-nucleolytic proteins have been shown to be sufficient to induce termination in vitro when they interact with RNA polymerase II (Zhang et al., 2004; Gilmour and Fan, 2008; Pearson and Moore, 2013). In this respect, the parallel with E. coli Rho-dependent transcription temination, which inspired the torpedo model (Connelly and Manley, 1988), is interesting. Since Rho is a termination factor, and is also a robust helicase that tracks along RNA, the predominant view for many years was that Rho causes termination by using this helicase activity to pull nascent RNA out of the polymerase active site (Brennan et al., 1987; Richardson, 2002; Park and Roberts, 2006; Schwartz et al., 2007). However, there was evidence early on that the mechanism of Rho-induced termination is more nuanced than this, because the susceptibility of transcription to termination by Rho was sensitive to mutations in the polymerase (Das et al., 1978). Moreover, whereas yeast RNA polymerase II was shown to be susceptible to Rho-mediated termination in vitro, neither polymerases I nor III of yeast were (Lang et al., 1998). This sensitivity to context would be unusual for a mechanism based primarily on the force of helicase action, and the implication that Rho-mediated termination depends more on conformational changes induced by the Rho-polymerase interaction was eventually established by Epshtein et al. (2010). Through cross-linking, they showed that elements of the RNA polymerase trigger-loop, within the active

site, are functionally involved in the termination mechanism. They also showed that Rho-dependent termination can be blocked by the transcription inhibiter tagetitoxin, which binds to the trigger-loop (Yuzenkova et al., 2013).

To develop a fundamental understanding of eukaryotic PADT it is useful to follow the approach taken so successfully in studying eukaryotic transcription initiation. The initiation studies began with a careful characterization of the core promoter, and the manner in which it was accessed by the general transcription machinery (Thomas and Chiang, 2006). The various enhancing (and inhibitory) elements so prominent in vivo could then be understood in terms of how they modulate the basal initiation mechanism. Accordingly we have focussed for a number of years on the core poly(A) signal, and basal PADT.

Basal PADT requires only a single element, the poly(A) signal (Orozco et al., 2002). No additional downstream elements of any kind are required. The core poly(A) signal alone can direct pausing and then termination of transcription (Orozco et al., 2002; Nag et al., 2007; Kazerouninia et al., 2010). Basal PADT involves a poly(A) signal-mediated drop in polymerase processivity, with the resulting termination-proficient polymerases dissociating from the template according to first order kinetics downstream of the poly(A) signal (Orozco et al., 2002; Kim and Martinson, 2003). Enhancing elements that increase poly(A) signal strength enhance PADT (see discussion in Orozco et al., 2002). Although these enhancers have come to be known as termination elements (Richard and Manley, 2009), it is likely that many or most of them enhance PADT through acceleration of cleavage and polyadenylation apparatus assembly (Orozco et al., 2002). In other words, they function by enhancing the efficiency of the core PADT mechanism.

Strong poly(A) signals assemble the cleavage and polyadenylation apparatus more quickly than weak ones (Chao et al., 1999) and yield higher levels of mature polyadenylated mRNA (see Kazerouninia et al., 2010). Conversely, weak poly(A) signals produce little cleaved and polyadenylated RNA; the RNA not cleaved at the poly(A) site being discarded. The nature of the discard mechanism that efficiently eliminates the uncleaved RNA from the cell, remains unknown, except that it depends on the presence of the very same weak poly(A) signal that failed to get processed in the first place (Kazerouninia et al., 2010). Interestingly, one can account for this dependence of RNA degradation on the presence of an unprocessed poly(A) signal by a very simple model, if it is assumed that PADT is independent of cleavage at the poly(A) site. According to this model, if PADT is cleavage-independent, then the fate of every nascent RNA will depend on a kinetic competition between stabilization by cleavage/polyadenylation or degradation by PADT/exosomal degradation (Kazerouninia et al., 2010). Note, in this model, that both stabilization and degradation depend on the action of an intact poly(A) signal. Accordingly, poly(A) site cleavage is fast for strong poly(A) signals, so that cleavage usually precedes the slower PADT, and the upstream mRNA sequences mostly survive. In contrast, when poly(A) site cleavage is slow, as for weak poly(A) signals, transcription may succomb to PADT before cleavage can take place, and the RNA is lost.

The central requirement of the above kinetic competition model is that PADT be poly(A) signal-dependent but cleavage-independent. Indeed, this begs one of the most fundamental questions about the very mechanism of PADT itself. In the present study, therefore, we ask whether cleavage at the poly(A) site (in the absence of cleavage elsewhere) is required to elicit PADT. Answering this question requires a method to assay both cleavage efficiency and termination efficiency in the same sample. Since we could not envision a reliable in vivo assay for poly(A)

site cleavage, we chose to develop an in vitro assay for PADT. Using such an assay we show here that PADT can proceed efficiently in vitro under conditions which do not permit cleavage at the poly(A) site. Moreover, we show that ternary transcription complexes, generated in nuclear extract, and bearing incompletely assembled cleavage and polyadenylation complexes, undergo dissociation in a poly(A)-dependent manner when resuspended in buffer. Finally, we show that this poly(A)-dependent dissociation can be inhibited by α-amanitin, which binds to the polymerase trigger loop (Martinez-Rucobo and Cramer, 2013). We conclude that the poly(A) signal triggers PADT through a conformational change mechanism that does not require cleavage of the nascent transcript or any 5′ to 3′ exonuclease activity.

## Results

**Poly(A)-dependent transcription termination (PADT) in vitro.** To analyze PADT in vitro we developed the assay shown in Figure 3-1. This assay measures the occupancy of RNA polymerase II on DNA before and after crossing a poly(A) signal. For a typical experiment two DNA templates, wildtype and mutant, differing only by three nucleotides in the poly(A) signal, are transcribed in parallel reactions in HeLa nuclear extract. The templates (attached to magnetic beads) are then stripped of all but the most stably bound proteins by washing in 1% Sarkosyl/1M KCl. There is no crosslinking involved. Only polymerases engaged as genuine transcription elongation complexes (TECs) survive this wash (Figure 3-1A). After restriction enzyme digestion of the stripped DNA (Figure 3-1B) the TECs can be recovered by anti-polymerase immunoprecipitation (Figure 3-1C). Analysis of the immunoprecipitated DNA reveals the relative distribution of polymerases downstream of wildtype and mutant poly(A) signals.

Figure 3-2 shows that PADT is faithfully reproduced in a nuclear extract in vitro. Templates containing either wildtype or mutant versions of the SV40 late poly(A) signal (Figure 3-2A) were transcribed in parallel reactions. At the indicated times samples were withdrawn, stripped, and analyzed as outlined in Figure 3-1. Lanes 5-8 in the gel of Figure 3-2B show how the distribution of ternary complexes evolves over time in the absence of a poly(A) signal [i.e. when the poly(A) signal has been inactivated by mutation]. Under the conditions of our in vitro system there is an initial burst of transcription, which sends a wave of ternary complexes downstream, but reinitiation is slow (unpublished data). In Figure 3-2B, lanes 5-8, this movement of a cohort of polymerases across the template is evident in the evolving co-IP efficiencies of fragments 0, 1 and 2 over time. Thus, fragment 0, the first of the three segments encountered by the ternary complexes, experiences its maximum co-IP efficiency at early times

(4 and 8 min), followed by fragment 1, which peaks between 8 and 15 min in this experiment, and finally by fragment 2 which experiences its maximum co-IP efficiency at 15 and 30 min.

To see the effect of an active poly(A) signal on the distribution of ternary complexes, the co-IP efficiencies for templates having a wildtype poly(A) signal (Figure 3-2B, lanes 1-4) can be compared with those discussed above whose poly(A) signal had been mutated (lanes 5-8). Of course, for fragment 0, there is no effect of an active poly(A) signal on the time-dependence of co-IP efficiency (compare lanes 1-4 with lanes 5-8) because this segment precedes the poly(A) signal. In contrast, template segments downstream of the poly(A) signal (fragments 1 and 2) exhibit reduced co-IP efficiencies at longer times when the poly(A) signal is wildtype (lanes 3 and 4) than when it is mutant (lanes 7 and 8). We interpret this as the stochastic loss of polymerases triggered by an active poly(A) signal, namly, PADT. (Orozco et al., 2002). It is formally possible that the poly(A) signal reduces polymerase occupancy downstream through acceleration of elongation. However, we can rule out this possibility because we have shown previously, using the same in vitro system, that active poly(A) signals, if anything, slow the polymerases down (Kazerouninia et al., 2010), as they do also in vivo (Park et al., 2004; Nag et al., 2007).

The results from three independent PADT experiments are summarized in the graph of Figure 3-2B. In expressing the co-IP data quantitatively for this graph, the signal in the gel for each fragment was assumed to be directly proportional to the degree of occupancy by ternary complexes. This assumption is valid as long as there is no more than one polymerase per fragment, which is likely to be true for fragments of the length used here (Yeung et al., 1998). The plot shows the polymerase occupancy downstream of an active poly(A) signal for both fragment 1 and fragment 2. Each value has been normalized twice: first to fragment 0 as a

loading control (since this fragment is the same in all constructs), and second to the equivalent fragment in the mutant, so that any decrease in co-IP is attributable to poly(A) signal function. The results show that PADT proceeds in vitro with approximately 40% efficiency under our conditions, comparable to the efficiency of cleavage and polyadenylation under similar conditions (Rigo et al., 2005). In addition, the lower polymerase occupancy on fragment 2 than on fragment 1 indicates that unassisted PADT in vitro retains a central feature characteristic of unassisted PADT in vivo, namely the progressive release of polymerases from the template downstream of the poly(A) signal (Orozco et al., 2002). Finally, the poly(A)-dependent pause that precedes termination (Orozco et al., 2002; Nag et al., 2007) is clearly evident in these results. Unexpectedly, polymerase occupancy downstream of the poly(A) signal is often less than 1.0 at very early times, especially for fragment 2. Perhaps this reflects the delay polymerases experience in reaching downstream locations because of the slowing down which starts at the poly(A) signal.

PADT depends on post-transcriptional recognition of the poly(A) signal in vivo (Kim and Martinson, 2003). To confirm that the decrease in polymerase occupancy of fragments 1 and 2 in vitro reflects a process that depends on the extruded poly(A) signal, we took advantage of 2'-O-methoxyethylribose-modified oligonucleotides (MOEs) (Vickers et al., 2001). MOEs have an increased affinity for target sites on the RNA, do not support RNase H cleavage once hybridized, and are much more stable in nuclear extract than RNA oligonucleotides. Fig. 2C shows the results of a PADT experiment essentially like that of Fig. 2B except that inactivation of PADT, for lanes 6-10 in the gel, was by an MOE targeted to the AATAAA hexamer of the poly(A) signal (see Figure 3-2A and 3-2D), rather than by mutation. As expected, for lanes 6-10, whose samples lacked a functioning poly(A) signal, polymerase occupancy peaks early for fragment 0,

later for fragment 1, and later still for fragment 2. This contrasts with lanes 1-5 of Figure 3-2C, where an MOE of unrelated sequence was used, which did not target any element in the transcribed RNA. In these lanes, for which the poly(A) signal in the samples remained active, the polymerase occupancy for fragment 2 did not increase at the later times but, instead, decreased slightly. Evidently, at later times, polymerases are removed by PADT from fragment 2 slightly faster than the newly arriving polymerases from upstream accumulate. These results are summarized graphically beneath the gel of Figure 3-2C.

The graph in Figure 3-2C resembles that of Figure 3-2B in all of its significant features, establishing that the decrease in polymerase occupancy on the DNA template downstream of the poly(A) signal is a poly(A) signal-dependent effect. To substantiate even further that this assay reflects authentic PADT, we used RNase H. Assembly of the cleavage and polyadenylation apparatus is aborted, both in vivo and in vitro, if the downstream RNA that tethers the poly(A) signal to the polymerase is severed before cleavage at the poly(A) site occurs (Bird et al., 2005; Rigo et al., 2005). Releasing the poly(A) signal from the polymerase by cutting the tether should similarly prevent PADT, because in vivo data show that continuous communication of the polymerase with the poly(A) signal is required for PADT (Kim and Martinson, 2003). To test this prediction, an experiment like that of Figure 3-2C was carried out, but in the presence of DNA oligonucleotides targeted to positions in the cloning vector lying roughly one or two hundred nucleotides downstream of the poly(A) site (see Figure 3-2A and 3-2D). As described previously (Rigo et al., 2005), these oligos hybridize to their targets and then recruit RNase H, which cuts the RNA, thereby releasing the poly(A) signal from the polymerase. The results of the experiment are shown in the histogram of Figure 3-2D. Bars 1-3 show the control, in which transcription was carried out for 50 min in the presence of a DNA oligo that was not

complementary to any sequence in the transcribed RNA. PADT occurred normally, with assay fragments 1 and 2 (bars 2 and 3) exhibiting significantly reduced polymerase occupancy compared to fragment 0 (which has an occupancy of 1 by definition because of normalization). In contrast, when transcription was carried out in the presence of complementary oligos that induce RNase H cutting of the transcript in vector sequences 112 or 201 nucleotides downstream of the poly(A) site, PADT is almost completely blocked (Figure 3-2D, bars 4-6 and 7-9). This confirms that the loss of polymerase occupancy measured in our PADT assay reflects a coupling of poly(A) signal function to transcription. Interestingly, cleavage of the RNA downstream of the poly(A) signal in this experiment (to produce a 5' phosphorylated end, Wu et al., 1999) blocked PADT rather than enhancing it—as might be expected according to the torpedo rationale (West et al., 2004).

The results so far establish that the decrease in polymerase occupancy downstream of the SV40 late poly(A) signal in these experiments is the result of authentic PADT in our in vitro system. The PADT can be abrogated by mutating the poly(A) signal sequence in the DNA (Figure 3-2B), by occluding the poly(A) signal in the RNA (Figure 3-2C), and by uncoupling poly(A) signal function from transcription (Figure 3-2D). To determine whether our in vitro system would also support PADT by an unrelated poly(A) signal, we replaced the SV40 late poly(A) signal of dsxSV40L with DNA containing the BGH poly(A) signal. Figure 3-2E shows that PADT is recapitualated also by the BGH poly(A) signal in our in vitro system.

Having now established both the validity and the generality of the PADT assay, we set out to use this in vitro system to address mechanistic questions related to PADT. In the experiments that follow we have used the MOE approach (Figure 3-2C and 3-2D), rather than mutation of the template (Figure 3-2B), to selectively inactivate poly(A) signal function for the control portion of

each experiment. The use of MOEs allows the transcription reactions for the entire experiment to be carried out using the same preparation of DNA, thereby eliminating all experimental error related to differences in DNA template preparations.

**PADT does not require poly(A) site cleavage.** The torpedo model for PADT depends on generating a new 5' phosphorylated end, by cleavage at the poly(A) site, to provide access to Rat1/Xrn2 for degrading the downstream RNA and promoting termination (Kuehner et al., 2011; Kim and Buratowski, 2013). The torpedo model has also been invoked to account for transcription termination in other situations where diverse types of nascent RNA cleavage provide 5'-phosphorylated RNA ends as entry points for Rat1/Xrn2 (Teixeira et al., 2004; El Hage et al., 2008; Kawauchi et al., 2008; Brannan et al., 2012). Yet Figure 3-2D, as we have seen, shows that cutting the nascent RNA with RNase H, which provides a 5' phosphorylated RNA end in the vicinity of both the cleavage/polyadenylation apparatus and the polymerase, far from facilitating termination, instead abrogated it. These results suggest either that the delivery of the new RNA 5' end to Xrn2 during PADT is tightly controlled or, alternatively, that PADT does not require the availability of an RNA 5' end.

To determine whether the generation of a new RNA 5' end is required for PADT we allowed the cleavage and polyadenylation apparatus to assemble, but then blocked its ability to carry out poly(A) site cleavage before proceeding to the PADT assay (Figure 3-3A). For this experiment we took advantage of the fact that, following transcription of the poly(A) signal in vitro, there is a lag of several minutes that precedes cleavage at the poly(A) site (Rigo et al., 2005). During this lag we rinsed the immobilized templates free of creatine phosphate (CP) and then resuspended them in transcription reaction mixture with or without CP. NTPs were not included. CP was originally shown to be required for poly(A) site cleavage when carried out on a pre-

synthesized substrate RNA in vitro (Hirose and Manley, 1997), and we have found that it is similarly required for poly(A) site cleavage that is coupled to transcription (see gel in Figure 3-3A). Thus, RNase protection revealed strong poly(A) site cleavage activity after resuspension in the presence (Figure 3-3A gel, lane 2) but not in the absence (Figure 3-3A gel, lane 3) of CP. Strikingly, as shown by the PADT assays in Figure 3-3A, withdrawal of CP did not block PADT, even though poly(A) site cleavage was completely eliminated under the same conditions. Therefore, PADT does not require cleavage.

Under the conditions of Figure 3-3A, although PADT clearly occurred in the absence of CP, the polymerases terminated, not within fragment 1 (see right hand panel of PADT assay in Figure 3-3A), but only after reaching fragment 2. Thus, PADT was delayed under those conditions. We, therefore, sought to optimize the PADT reaction while maintaining the suppression of cleavage. We found that ATP could substantially increase termination efficiency in the absence of cleavage, provided that a low concentration (0.4 mM) of CP was present (Figure 3-3B). Thus, Figure 3-3B shows that including ATP in the resuspension mixture can restore robust PADT to fragment 1, while increasing its efficiency on fragment 2 to 40%. Yet, RNase protection assays, carried out on a portion of each sample, showed that there is no effect of ATP on cleavage for dsxSV40L, in agreement with previous results showing that ATP, if anything, inhibits cleavage at this poly(A) site (Hirose and Manley, 1997). In the gel of Figure 3-3B there is a faint band at the cleavage position in lanes 3, 4 and 5. However, this band is already present for samples at the 4 min rinse step (lane 6), and may be an unrelated background band because it remains constant at 7% of the total cleaved + uncleaved signal in lanes 3, 4 and 5 , rather than increasing as for PADT. In any case, this band amounts to only a fraction of the 40% termination obtained

in the PADT assay for 250 μM ATP (Figure 3-3B). Thus, these results establish unequivocally that robust PADT does not require poly(A) site cleavage.

We have found that isolated cleavage and polyadenylation complexes, assembled during transcription in vitro, do not carry out poly(A) site cleavage if transferred to buffer in the absence of any extract (unpublished observations). However, the discovery that PADT does not require cleavage prompted us to explore the possibility that these complexes, though they are unable to cleave, may nevertheless be PADT-competent in buffer. The RNase protection assay in Figure 3-4A shows that cleavage and polyadenylation complexes, assembled during the first 4 minutes of transcription (lane 1), can carry out cleavage at the poly(A) site when rinsed and returned to fresh extract (lane 2), but cannot do so if transferred instead to buffer which lacks extract (lane 3). As shown in the timeline for Figure 3-4A, after incubation was complete, samples were split, with 3μl being used for the RNase protection assay just discussed (gel lanes 2 and 3), and with the rest being used for a PADT assay. The histogram of Figure 3-4A shows that substantial PADT occurs without adding any extract back after the rinse (bars 5 and 6). These results, alone, establish two points. First, in confirmation of the conclusion from Figure 3-3, PADT does not require cleavage at the poly(A) site. And second, the apparatus required for PADT is different from the apparatus required for poly(A) site cleavage, the latter perhaps requiring one or more additional components or modifications.

Some time ago we established in vivo that transcription across a functional poly(A) signal led to a dramatic decrease in polymerase processivity, manifested as a first order decrease in polymerase occupancy as a function of distance downstream of the poly(A) signal (Orozco et al., 2002). The in vitro data presented here are similar in that PADT is almost always greater when measured for fragment 2 than for fragment 1. We interpreted our in vivo data as evidence for a

poly(A)-dependent change in the state of the TECs, causing them to terminate transcription stochastically as a function of time after crossing the poly(A) signal (Orozco et al., 2002). This idea can be tested directly using the experimental format of the split sample experiment in Figure 3-4A, because in this procedure rinsed TECs are resuspended in buffer, lacking NTPs, where release from the template can be measured as a function of time only, in the absence of concurrent changes in distance down the template. The results of such an experiment are shown in the time course of Figure 3-4A. Clearly the stationary polymerases are released from the template in a poly(A)-dependent manner as a function of incubation time after resuspension in buffer. Thus, following transcription of the poly(A) signal, the TEC undergoes a change that potentiates termination. Since only the PADT complex itself is present, and there are no added nucleotides and virtually no transcription, and there has been no cleavage at the poly(A) site, this supports the conformational change model as the core mechanism of PADT.

Figure 3-4A shows that the process of PADT is not coupled to ongoing transcription. However, we wondered whether the initial establishment of the PADT-proficient conformation might nevertheless be coupled to the emergence of the poly(A) signal during transcription. To address this question we stripped TECs of auxiliary proteins using high salt and Sarkosyl, and then allowed cleavage and polyadenylation complexes to reassemble on the TECs by resuspending them in extract (Figure 3-4B). Lanes 3 and 4 of the gel in Figure 3-4B show that reassembly of a functional poly(A) site cleavage complex after stripping is quite efficient. Although stripped TECs contain no detectable poly(A) site cleaved RNA (lane 3), they cleave their RNA with 45% efficiency after incubation in extract (in the absence of further transcription, lane 4). Surprisingly, however, incubation in extract after stripping yielded no decrease in polymerase occupancy (i.e. no increment in PADT) compared to the freshly stripped TECs (Figure 3-4B,

compare bars 9 and 12). Thus, reassembled TECs, capable of 45% cleavage at the poly(A) site, yielded no PADT. We conclude that poly(A) site cleavage is neither necessary (Figures 3-3 and 3-4A) nor sufficient (Figure 3-4B) for PADT.

Interestingly, although there was no PADT during the incubation step after stripping (Figure 3-4B, bars 9 and 12), the stripping step itself revealed a significant amount of poly(A)-dependent release of transcription complexes produced during only the first few minutes of transcription (Figure 3-4B, bar 9). In contrast, the time course in Figure 3-4A shows that actual PADT, under more physiological conditions, takes much longer to occur. This suggests that the transcription complex, in its PADT-proficient state, is already bound less tightly to the template as it awaits stochastic release.

**PADT can be inhibited by α-amanitin.** The data presented so far support a conformational change model for RNA polymerase II PADT. A conformational change mechanism has also been proposed for E. coli RNA polymerase Rho-dependent termination (Epshtein et al., 2010). For E. coli this proposal was based in part on the ability of the antibiotic tagetitoxin to inhibit Rho-dependent transcription termination in vitro (Epshtein et al., 2010). The RNA polymerase II inhibitor α-amanitin binds near the active site of the polymerase in a similar location as tagetitoxin, and may act by a similar mechanism (Martinez-Rucobo and Cramer, 2013; Yuzenkova et al., 2013). We therefore set out to ask whether α-amanitin inhibits PADT. Alpha-amanitin inhibits transcription, but active transcription is not required in order for PADT to occur (Figure 3-4A). Accordingly, we allowed transcription to occur for a few minutes, as in Figure 3-4A, and then rinsed and resuspended the transcription complexes in buffer with or without α-amanitin (Figure 3-5A). Bars 2 and 3 in Figure 3-5A confirm, through several repetitions, that

PADT occurs for transcription complexes resuspended simply in buffer. Bars 5 and 6 show that PADT is completely inhibited by α-amanitin.

We have shown in Figure 3-5A that PADT directed by the SV40 late poly(A) signal is inhibited by α-amanitin. We wondered whether this would also be true for an unrelated poly(A) signal. The SV40 early poly(A) signal is quite different from the SV40 late poly(A) signal, and is several fold weaker, both in vivo (Carswell and Alwine, 1989) and in vitro (Kazerouninia et al., 2010). A PADT time course for the SV40 early poly(A) signal (Figure 3-5B) shows that it directs fairly efficient termination in vitro, although at a slightly slower rate than for SV40 late (Figure 3-2B and C), in agreement with measurements in vivo (Orozco et al., 2002). The effect of α-amanitin on PADT by the SV40 early poly(A) signal is shown in Figure 3-5C. Bars 1-3 summarize the 45 min data from Figure 3-5B, for which no α-amanitin was present. Bars 4-6 show the 45 min data from an assay run in parallel, in which α-amanitin had been added 5 min after the start of transcription. The results show that α-amanitin blocks PADT directed by the SV40 early poly(A) signal.

In the experiment of Figure 3-5C, the α-amanitin was added to a sample undergoing active transcription. Thus, the two samples differed not only in containing (bars 4-6) or lacking (bars 1-3) α-amanitin, but also in whether there was ongoing transcription (bars 1-3) or not (bars 4-6). To eliminate this second variable in the SV40 early experiment, we adopted a format similar to that of the SV40 late experiment of Figure 3-5A. In this experiment the beads for both the control and the α-amanitin samples were rinsed free of nucleotides and then resuspended in the presence or absence of α-amanitin. As shown in Figure 3-5D, the results obtained using this format were very similar to those of Figure 3-5C. Control experiments (not shown) confirm that there is no transcription in either the presence or the absence of α-amanitin under the conditions

of Figure 3-5D. Thus, α-amanitin blocks PADT for dissimilar poly(A) signals (SV40 early and SV40 late), and under a variety of conditions (Figure 3-5A, C and D).

## Discussion

Ever since the first suggestion that the poly(A) signal may be responsible for transcription termination (Acheson, 1984), efforts have been directed at understanding the nature of the communication between this RNA processing signal and the polymerase. Because PADT tends to occur at variable distances downstream of the poly(A) signal, and because the many physical associations between processing factors and the transcriptional apparatus were not yet known, early attempts to understand this communication came to focus on how a termination signal might be transmitted along the nascent RNA from the poly(A) site down to the elongating polymerase. This led to the picturesque torpedo model and its requirement for cleavage at the poly(A) site (Connelly and Manley, 1988; Proudfoot, 1989). However, within a decade the close collaboration between the RNA processing factors and the transcriptional apparatus had become clear (Hirose and Manley, 2000; Proudfoot, 2000) and there was no longer a compelling reason to accord primacy to the torpedo model (Kim and Martinson, 2003). It was clear that the processing and transcription factors could communicate directly among themselves. Nevertheless, the torpedo model remains widely favored, because testing it directly has not been possible before now.

Here we have tested the most fundamental assumption of the torpedo model, that cleavage is required for PADT. For these experiments we first developed a reliable assay for PADT in vitro (Figures 3-1 and 3-2). Then, using this assay, we showed that termination occurred in a poly(A)-dependent manner under a variety of conditions that did not permit poly(A) site cleavage of the RNA (Figures 3-3 and 3-4). The most significant of these conditions involved the use of isolated TECs bearing incompletely assembled cleavage/polyadenylation complexes (Figure 3-4). These complexes were isolated simply by removal from the nuclear extract during the lag that occurs

between transcription of the poly(A) signal and cleavage at the poly(A) site. When resuspended in the absence of nuclear extract these partially assembled complexes were unable to carry out cleavage at the poly(A) site, but were nevertheless able to direct progressive polymerase release from the template in a poly(A)-dependent way (Figure 3-4A). This poly(A)-dependent release occurred simply in buffer, in the absence of extract, nucleotides, transcription, or other influences that might contribute to dislodging the complexes from the template. Thus, transcription/3′-processing complexes are intrinsically termination proficient, and do not require cleavage or any additional input to trigger termination.

These results show that there are at least two steps during basal PADT, as previously suggested (Orozco et al., 2002). Both steps are independent of poly(A) site cleavage. First, the transcription complex is converted to a termination-proficient state through at least partial assembly of the cleavage/polyadenylation apparatus. This step presumably coincides with the poly(A)-dependent transcriptional "pause" that is imposed on the transcriptional apparatus during the early stages of cleavage/polyadenylation complex assembly (Nag et al., 2007). The second step is termination itself. This step requires an additional conformational change that apparently involves the trigger loop, deep within the active site of the polymerase. A role for the trigger loop is indicated by the fact that the stalled, termination-proficient TECs, resuspended in buffer, are blocked from template release by the presence of α-amanitin. Thus, the temination-proficient TECs are not simply unstable TECs. Instead, they appear to be stably engaged polymerases awaiting a conformational change that will trigger termination. The α-amanitin sensitivity of PADT resembles the situation for Rho-dependent termination in E. coli, which is similarly inhibited by the trigger loop-binding antibiotic tagetitoxin (Epshtein et al., 2010; Yuzenkova et al., 2013). The convergence of similar conformational change attributes for

mammalian PADT and for Rho-dependent termination in E. coli supports the suggestion by Epshtein et al. (2010) that an allosteric model for transcription termination is likely to be general for all cellular RNA polymerases.

Our focus in the present study has been on the role of the poly(A) signal in basal PADT. Several steps at which the poly(A) signal could influence polymerase function and lead to PADT have been suggested. These include direct recognition of the DNA sequence by the polymerase (Connelly and Manley, 1988; Kim and Martinson, 2003), recognition by TEC-associated proteins of the poly(A) signal as it is extruded (Calvo and Manley, 2001; Nag et al., 2007), recruitment of termination factors by the poly(A) signal (Kim et al., 2004a), and cleavage of the poly(A) signal to provide access to termination factors (Kim et al., 2004b). To date we have explored three of the four junctures mentioned above at which the poly(A) signal could intervene to yield PADT. The last of these was addressed in the present study where we established that cleavage at the poly(A) site is neither necessary nor sufficient for PADT (Figures 3 and 4). Regarding the first possibility mentioned above, we previously found no indication that direct poly(A) signal recognition, at either the DNA or RNA level, prior to extrusion from the polymerase, is involved in PADT (Kim and Martinson, 2003). As for the second possibility, we did find that the mammalian cleavage and polyadenylation factor, CPSF, binds tightly to the body of the polymerase and is required for mediating the poly(A)-dependent transcriptional pause (Nag et al., 2007). Presumably this occurs as the extruding poly(A) signal pushes past the TEC-associated CPSF. However, while this pause normally precedes PADT, it is not sufficient for PADT, and is presumably a precursor to further recruitment of cleavage and polyadenylation factors. Thus, of the four pathways to PADT enumerated above, we are left with the third one, recruitment of termination factors, as the poly(A) signal's principal contribution.

Indeed, recruitment of termination factors is a well-known function of the poly(A) signal (Birse et al., 1998; Kim et al., 2004a; Luo et al., 2006). However, the interdependence of factors involved in the assembly process means that most factors required for efficient assembly of the cleavage/polyadenylation apparatus also contribute to efficient termination. It is therefore useful to search for proteins that contribute directly to the termination mechanism. Two such proteins are Pcf11 (Gilmour and Fan, 2008) and Rat1/Xrn2 (Pearson and Moore, 2013). In both cases these proteins have been shown to have RNA polymerase II termination activity in vitro. Pcf11 is a poly(A) site cleavage factor and RNA polymerase II CTD binding protein that can act alone to dismantle the TEC in vitro through a conformational change mechanism (Gilmour and Fan, 2008). Rat1, when mutated to remove its exonuclease activity, also uses a conformational change mechanism to dismantle TECs, but only when provided an anchor to the CTD through its CTD-binding partner-protein Rtt103 (Pearson and Moore, 2013). Rat1 typically occurs as a complex of three proteins, Rat1•Rai1•Rtt103, as does its mammalian homolog, Xrn2, whose equivalent functional complex in the cleavage/polyadenylation apparatus is Xrn2•Psf•P54$^{nrb}$ (Kaneko et al., 2007). The recruitment of both Pcf11 and Rat1/Xrn2 is facilitated in part by Glc7/PP1 (Schreieck et al., 2014), an RNA polymerase II CTD tyrosine phosphatase/polyadenylation factor that is recruited during assembly of the cleavage and polyadenylation apparatus (Shi et al., 2009).

The termination activities of Pcf11 and Rat1 were produced in the reconstituted in vitro systems in the absence of a poly(A) signal (Gilmour and Fan, 2008; Pearson and Moore, 2013). This is consistent with the in vivo role of the poly(A) signal in PADT being primarily for recruitment. Apparently, in the defined in vitro system, CTD-binding alone is sufficient to orient the Pcf11 and Rat1 appropriately for a conformational change-inducing interaction with the polymerase (Gilmour and Fan, 2008; Pearson and Moore, 2013). Interestingly, with wildtype Rat1 (which

retains its exonuclease activity) CTD binding through Rtt103 is not required for termination. In this case Rat1 (with its partner Rai1) can apparently use the RNA to orient itself for a productive termination-inducing interaction with the polymerase (Pearson and Moore, 2013). Thus, the apparatus recruited by the poly(A) signal has multiply redundant capabilities for triggering termination of transcription. There are at least two unrelated termination-inducing proteins (Pcf11 and Rat1/Xrn2), and one of these (Rat1/Xrn2) can exploit at least two alternative features of the TEC (the CTD and the nascent RNA) to orient itself for a productive polymerase interaction that leads to the termination-inducing conformational change. It is likely that all of these access routes lead to the same ultimate termination-inducing conformational change (Epshtein et al., 2010; Pearson and Moore, 2013).

It is not clear which of the above termination pathways contributed to the PADT we assayed in the experiments described here. It is possible that all of them contributed in competitive fashion for those experiments that allowed poly(A) site processing to occur (Figure 2). For those experiments that did not allow poly(A) site cleavage to occur (Figures 3-5) it is obvious that the last of the possibilities mentioned above (Xrn2 orienting itself via RNA) could not be a contributor to the observed termination. In vivo, therefore, rapid cleavage at strong poly(A) sites (or elsewhere) may favor the Rat1/Xrn2 exonucleolytic pathway, whereas slow cleavage at weak poly(A) sites may favor the Pcf11 and Rtt103/Psf pathways.

The availability of multiple routes to PADT, including ones that do not require cleavage, allows a simple model to explain an old conundrum: Uncleaved RNA bearing a weak poly(A) signal is degraded—but not if the weak poly(A) signal is inactivated completely (see Kazerouninia et al., 2010). Why would the presence of the weak, uncleaved poly(A) signal be required for degradation of the RNA? The likely answer is that slow cleavage at the weak poly(A) signal

(Chao et al., 1999) leaves time for PADT along the cleavage-independent Pcf11 or Rtt103 pathway, whereupon the exosome, presumably associated with the transcription complex (Andrulis et al., 2002; de Almeida et al., 2010), degrades the RNA (Kazerouninia et al., 2010).

## Materials and Methods

The dsxSV40L template was PCR-generated from pDsx, which is identical to Gal5-HIV2dsxΔInt(+ESE) (Rosonina et al., 2003). For mutant dsxSV40L, the SV40 late poly(A) signal AATAAA hexamer of pDsx was converted to AAgtAc by site-directed mutagenesis. For dsxSV40E and dsxBGH, the SV40 late poly(A) signal of pDsx was replaced with either the SV40 early or the BGH poly(A) signal by cold-fusion cloning (MC101B-1, SBI), for which the linearized vector was PCR-amplified from pDsx using oligos 1 and 2 below. The insert for the SV40 early poly(A) signal was produced by annealing oligos 3 and 4, and the insert for the BGH poly(A) signal was PCR-amplified from pCMV/BGH (Rigo et al., 2005) using oligos 5 and 6. The DNA oligos used for generating the above plasmids were as follows:

1: 5′-TTAAAGCAAGTAAAACCCTCGAGTTCTATAGTGTCACC-3′

2: 5′-CAGCTGAAGCTTGCATGCCTGCAGGTCGAC-3′

3: 5′-CATGCAAGCTTCAGCTGAACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTGTTAAAGCAAGTAAAAC-3′

4: 5′-GTTTTACTTGCTTTAACACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTTCAGCTGAAGCTTGCATG-3′

5: 5′-ATGCAAGCTTCAGCTGTGTGCCTTCTAGTTGCC-3′

6: 5′-GGGTTTTACTTGCTTTAAGCCATAGAGCCCACCG-3′

For attaching to magnetic beads all templates were PCR-amplified using the primer pair 5′-CTATGGAAAAACGCCAGCAACGC-3′ and 5′-GCAACGTTGTTGCCATTGCTACAG-3′, in which the latter was biotinylated at the 5′-end. The PCR products were gel-extracted using the QIAquick Gel Extraction Kit (28704, Qiagen), and bound to magnetic Dynal M-280 Streptavidin beads (11206D, Invitrogen) according to the manufacturer's protocol. The final DNA:bead ratio was 1-1.6 pmol per mg.

HeLa nuclear extract was prepared as described (Rigo et al., 2005) except that it was dialyzed additionally, at the end, into several changes of buffer D containing no glycerol. Transcription was carried out essentially as described previously (Rigo et al., 2005) except that nucleotides were unlabeled and Gal4-p53 (a gift from James Goodrich) was required to activate transcription on the pDsx-based templates (Rosonina et al., 2003). Prior to addition of the nucleotides, the templates (1 µl beaded DNA) were preincubated for 30 min with extract (5 µl), Gal4-p53 (0.6-1 µl, amount determined by optimization) and buffer components in a volume of 9.6 µl. The NTPs were then added together with MOEs and/or DNA oligos, in a volume of 5.4 µl, to initiate transcription. For a typical assay, the total final reaction volume was 15 µl containing, in addition to the extract and Gal4-p53, 6.7 mM HEPES (pH 7.9), 33 mM KCl, 66 µM EDTA, 1.7 mM DTT, 33 µM PMSF, 10 U anti-RNase, 3.3~5 mM $MgCl_2$, 3.3~7.5 mM sodium citrate (pH 6.7), 0.2 µg DNA, 167 µM each of the four NTPs, 16.7 mM creatine phosphate and 100 ng/µl MOE. The MOEs and DNA oligos added to the various transcription reactions, were as follows:

Control MOE: 5'- CCACACGCACCATCTTCTGC-3'

MOE complementary to the SV40L hexamer: 5'- GTTTATTGCAGCTTA-3'

MOE complementary to the SV40E poly(A) site: 5'-CAAACCACAACTAGAATGCAGTG-3'

MOE complementary to the BGH poly(A) site: 5'-GACACCTACTCAGACAATGCGATG-3'

Control DNA oligo: 5'-CAAACCACAACTAGAATGCAGTG-3'

DNA oligo 112: 5'-CATAACCTTATGTATCATAC-3'

DNA oligo 201: 5'-GGCTGGCTTAACTATGCGGC-3'

In some experiments, TECs were removed after several minutes of transcription, rinsed with 50 µl of buffer containing 6.7 mM HEPES (pH 7.9), 33 mM KCl, 66 µM EDTA and 0.2 mg/ml BSA, and then resuspended in new, preincubated reaction mixture lacking any DNA template or

Gal4-p53. The resuspension mixture also lacked NTPs, except as noted in the text or figure legends. When α-amanitin was used, it was added in a volume of 0.5 µl to a final concentration of 33 ng/µl.

When the extent of poly(A) site cleavage was to be determined, a 3 µl aliquot of the reaction mixture was processed as before (Rigo et al., 2005) and the RNA subjected to RNase protection, using probes prepared as previously described (Rosonina et al., 2003). The DNA template for generating the SV40L probe was a gift from B. Blencow (McCracken et al., 2002). The DNA templates for generating the SV40E and BGH probes were PCR-amplified from dsxSV40E and dsxBGH respectively, using the primer pair 5′-TAAATTGATGCGCCTGCACTGCTCAAAGGACATAATTTGAATATCTATGACG-3′ and 5′-GGGATCCTAATACGACTCACTATAGGGATCCATACGATTTAGGTGACAC-3′.

To assay for PADT after the desired duration of transcription (Figure 3-1), the TECs on the beaded templates were first stripped of extraneous protein using 80 µl of a stringent wash solution (20 mM HEPES, pH 7.9, 1 M KCl, 1% Sarkosyl) and then rinsed with 100 µl of NEBuffer 4 containing BSA. The rinsed TECs were resuspended in 20 µl NEBuffer 4 containing 7 units of various restriction enzymes together with BSA, and digested for 30 min at 37℃. For dsxSV40L four restriction enzymes were used, as shown in Figure 3-2A. Occassionally, PsiI was replaced with MfeI, which cuts about 31 bp downstream. The dsxSV40E template was digested similarly except that PsiI was replaced with AccI. And for dsxBGH, both EcoRI and PsiI were replaced by BanI, which has cutting sites near both the EcoRI site and the poly(A) site. After restriction digestion, the DNA fragments were prepared for subsequent end-labeling by incubation at 37℃ for 30 min with 20 units (2 µl) of Calf Intestinal Phosphatase (NEB).

To enrich for DNA fragments bearing TECs, immunoprecipitation with an anti-CTD antibody was carried out. The supernatant from restriction enzyme digestion (above) was mixed with anti-Pol II antibody 8WG16 (MMS-126R, Covance), which was pre-bound to protein G dynabeads (10004D, Invitrogen) according to the manufacturer's protocol in a ratio of 4 µl of 8WG16 to 0.6 mg protein G beads. After rotation at room temperature for 30 min, the supernatant was discarded and the protein G beads were washed once with 100 µl PBS containing 0.02% Tween-20. The beads were then digested with 2 µl RNase Cocktail (AM2286, Invitrogen) in 30 µl of 10 mM Tris-HCl (pH7.4), 300 mM NaCl, 5 mM EDTA at 37℃ for 30 min. The RNase digestion was stopped with 80 µl of 0.5% SDS in 10 mM Tris-HCl, 10 mM EDTA.

The DNA was extracted with phenol:chloroform:isoamyl alcohol (AC327115000, Acros Organics), ethanol-precipitated in the presence of glycogen (AM9510, Invitrogen), dissolved in 50 µl T4 polynucleotide kinase reaction buffer (M0201L, NEB), and then labeled for 60 min using 5 units of T4 Polynucleotide Kinase and 1 µl of [$\gamma$-$^{32}$P] ATP (Perkin Elmer, BLU002Z500UC). The DNA was then precipitated, resuspended and loaded onto a 4% denaturing polyacrylamide gel. The gel image was recorded by Molecular Imager FX (Bio-Rad) and analyzed by ImageQuant (Molecular Dynamics).

# Figures

**Figure 3-1. The PADT assay.**

A. Two immobilized DNA templates, one containing a wild type poly(A) signal and the other a mutant, are transcribed in parallel reactions in Hela nuclear extract. At the times indicated in the figures, transcription is stopped and the TECs are stripped of extraneous proteins by washing the beads in 1% sarkosyl, 1 M KCl.

B. The stripped DNA is cut with restriction enzymes, treated alkaline phosphatase, and subjected to immunoprecipitation with anti-Pol II antibody 8WG16.

C. The immunoprecipitated DNA is extracted, 5'-end labeled with [γ-32P] ATP and loaded on denaturing PAGE.

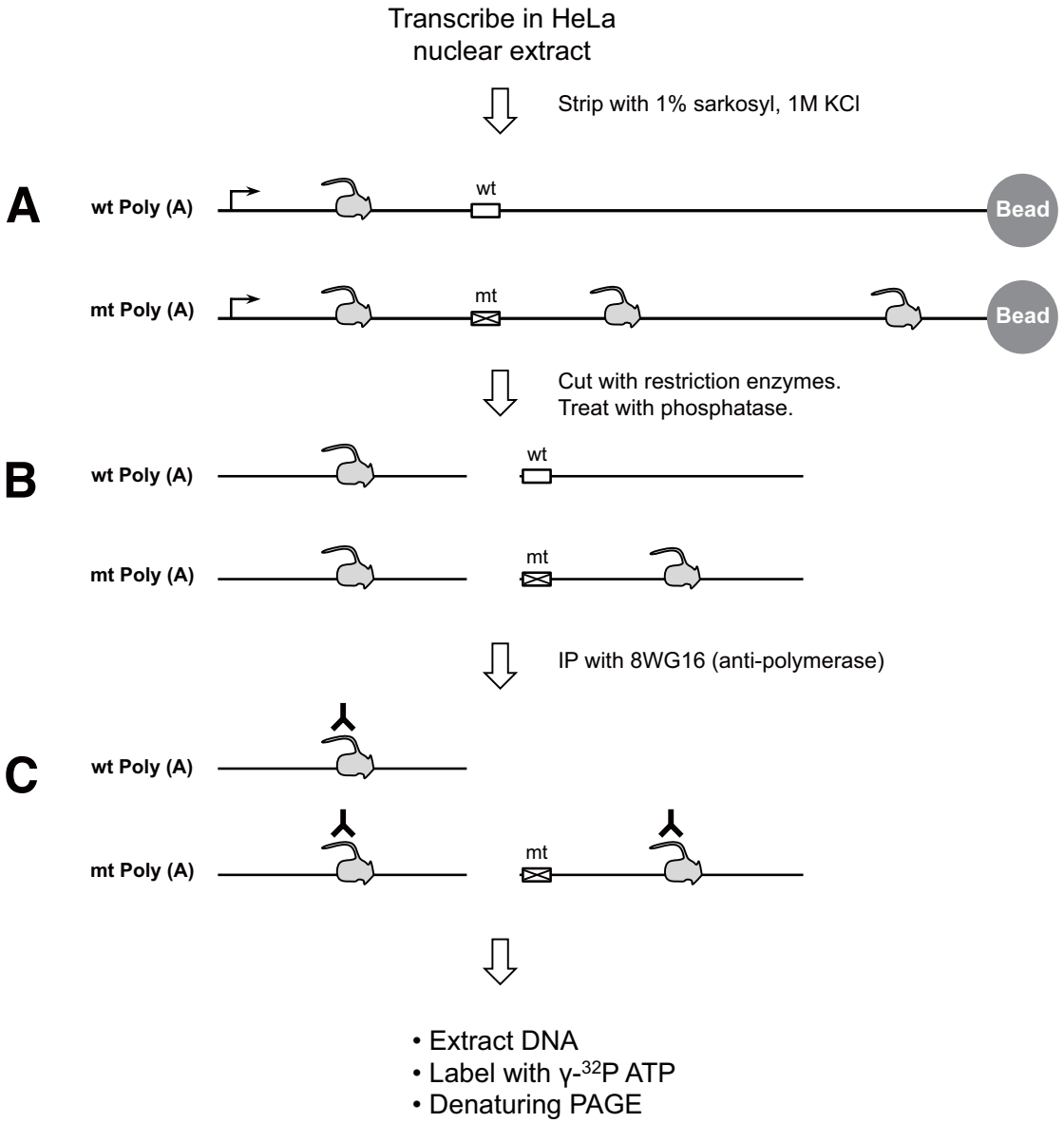

Transcribe in HeLa
nuclear extract
Strip with 1% sarkosyl, 1M KCl
A
wt Poly (A)
wt
Bead
mt Poly (A)
mt
Bead
Cut with restriction enzymes.
Treat with phosphatase.
B
wt Poly (A)
wt
mt Poly (A)
mt
IP with 8WG16 (anti-polymerase)
C
wt Poly (A)
mt Poly (A)
mt
• Extract DNA
• Label with γ-32P ATP
• Denaturing PAGE

**Figure 3-2. PADT is faithfully reproduced in a nuclear extract in vitro.**

A. The dsxSV40L template drawn to scale. The poly(A) signal hexamer (mutated for lanes 5-8 of part B) lies just to the left of the poly(A) cleavage site indicated by the arrow. The size and binding location of the anti-poly(A) MOE, used for parts C and D, and the size and binding location of the DNA oligos used for part D, are shown. The locations of the three restriction fragments quantitated for the PADT assays in parts B, C and D are shown.

B. PADT by the SV40 late poly(A) signal. Transcription was carried out in two pots, one for the wildtype template (pot a), and one for the template whose poly(A) signal had been inactivated by mutation (pot b). At the indicated times, 12 µl aliquots were withdrawn and treated as shown in Figure 3-1. Fragment 0, at the bottom of the gel, shows up as a doublet because one end of the fragment is blunt and the other is a 5'-overhang, which results in the resolution of two single strands on the gel (of 282 and 286 nt) after kinasing and denaturation. A similar situation exists for fragment 1, but it is higher in the gel and the two individual strands are not resolved. The asterisk indicates a band produced by restriction enzyme star activity near one end or the other of fragment 2. The graph below the gel quantitates the results from several experiments. The polymerase occupancy on wildtype fragments 1 and 2, downstream of the poly(A) signal, is defined for each as the efficiency with which it is coimmunoprecipitated relative to that for the mutant. Experimental variations between the wildtype and the mutant transcription reactions are controlled for by expressing all fragment 1 or 2 values as their ratio to the cognate fragment 0. The experiments that were compiled to produce the graph varied somewhat as to the number of time points. Thus, the values at 3, 7, 15 and 30 min are averaged from three experiments (± SD), the value at 56 min is from two experiments (± range), and the value at 50 min is from a single experiment. When the exact time of sampling varied between

experiments, the average was plotted. For example, the value plotted as 56 min is actually an average of samples taken at 50 min in one experiment and 62 min in another.

C. PADT detected using an MOE directed to the poly(A) signal. Transcription was like that for Figure 3-2B except that the poly(A) signal in transcription pot "b" was inactivated by a complementary MOE that blocked the poly(A) signal directly, rather than by mutation of the DNA template. Pot "a" contained a control MOE of irrelevant sequence. Quantitation for the graph below the gel was as for Figure 3-2B except that polymerase occupancy was calculated as the coimmunoprecipitation ratio for transcription in the absence or presence of the complementary MOE (rather than wildtype vs. mutant templates). The values at 2 min are averaged from two experiments (± range); the values at 5, 17 and 30 min are from four experiments (± SD), and the values at 9 and 50 min are from three experiments (± SD). The 6.5 min sample shown in the gel was averaged with one 5 min and two 4 min samples from additional experiments, and is therefore plotted at 5 min on the graph.

D. PADT is blocked when the tether is cut. Transcription was carried out in the presence of either a control MOE or one targeted to the poly(A) signal, as for part C. Also included in each transcription mixture was one of three DNA oligos: a control oligo not complementary to the transcript produced, and two complementary oligos that directed RNase H cutting to positions 112 and 201 nucleotides downstream of the poly(A) site (positions which lie within sequences originating from the cloning vector). The approximate binding locations of the complementary MOE and DNA oligos are shown in the accompanying cartoon (adapted from Rigo and Martinson, 2009) and in part A of this figure.

E. PADT by the BGH poly(A) signal.

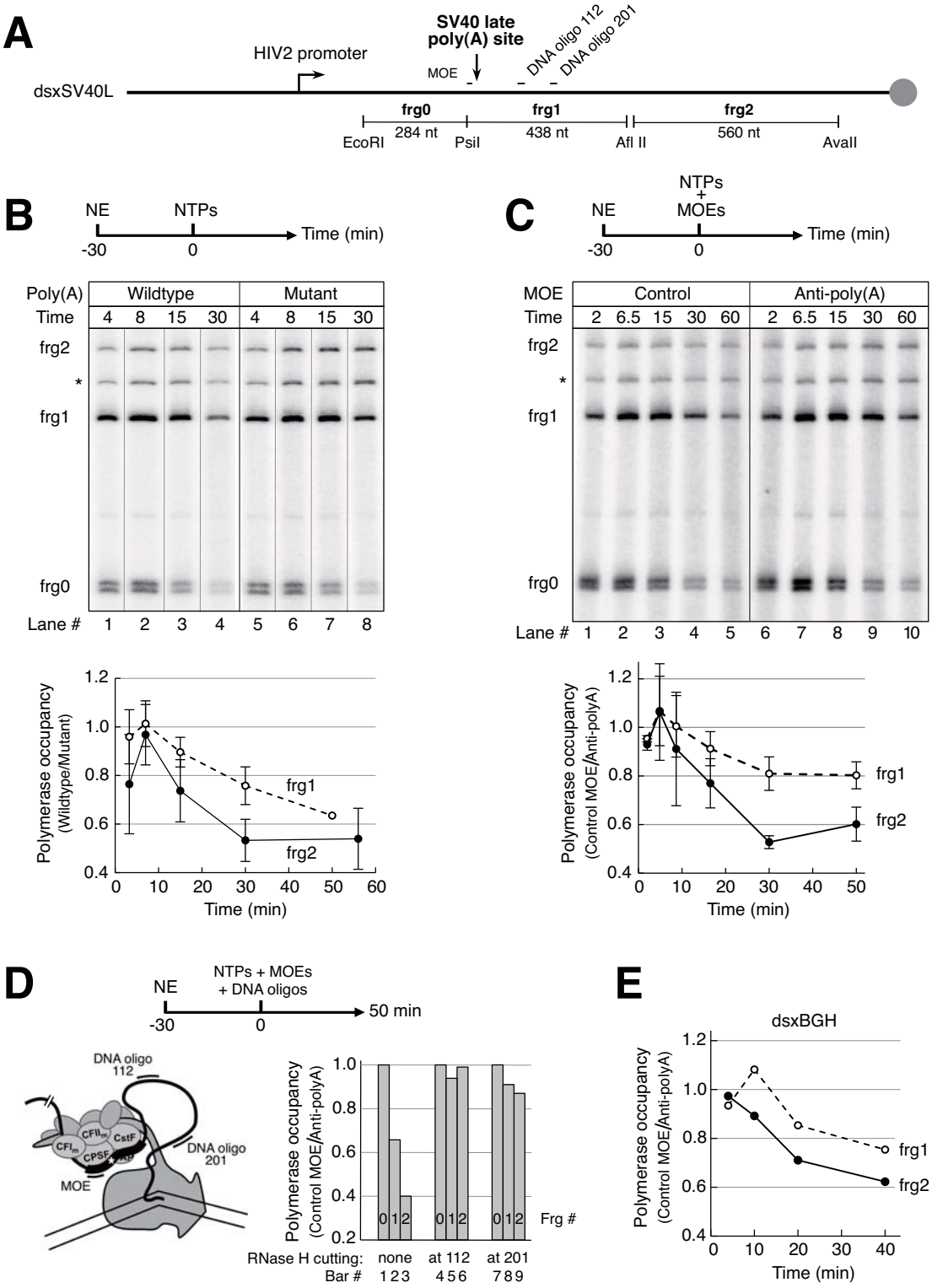
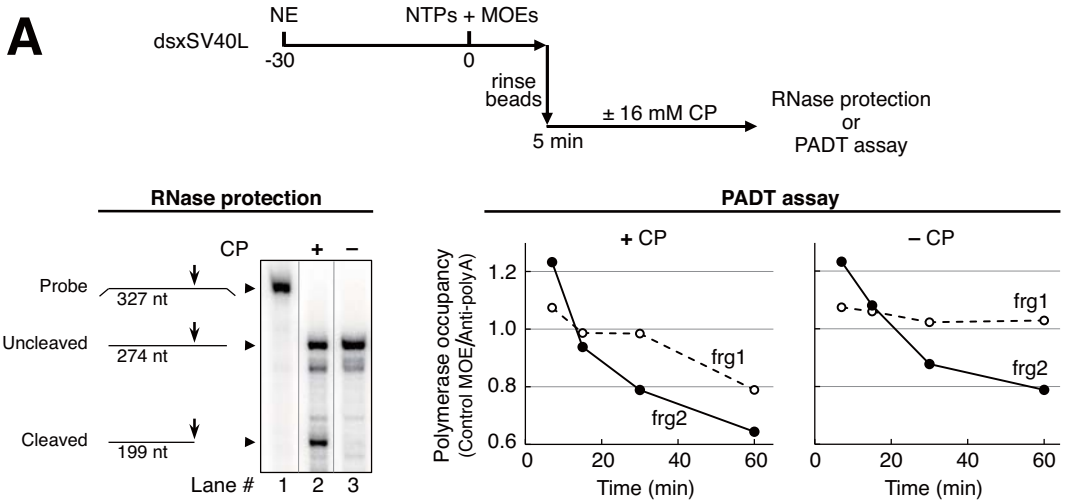

A
SV40 late poly(A) site
HIV2 promoter
MOE
DNA oligo 112
DNA oligo 201
dsxSV40L
frg0
284 nt
frg1
438 nt
frg2
560 nt
EcoRI
PsiI
Afl II
AvaII
B
NE
NTPs
-30
0
Time (min)
Poly(A)
Wildtype
Mutant
Time
4 8 15 30 4 8 15 30
frg2
*
frg1
frg0
Lane # 1 2 3 4 5 6 7 8
Polymerase occupancy (Wildtype/Mutant)
frg1
frg2
Time (min)
C
NE
NTPs + MOEs
-30
0
Time (min)
MOE
Control
Anti-poly(A)
Time
2 6.5 15 30 60 2 6.5 15 30 60
frg2
*
frg1
frg0
Lane # 1 2 3 4 5 6 7 8 9 10
Polymerase occupancy (Control MOE/Anti-polyA)
frg1
frg2
Time (min)
D
NE
NTPs + MOEs + DNA oligos
-30
0
50 min
DNA oligo 112
DNA oligo 201
MOE
Polymerase occupancy (Control MOE/Anti-polyA)
0 1 2
Frg #
RNase H cutting: none at 112 at 201
Bar # 1 2 3 4 5 6 7 8 9
E
dsxBGH
Polymerase occupancy (Control MOE/Anti-polyA)
frg1
frg2
Time (min)

**Figure 3-3. PADT does not require poly(A) site cleavage.**

A. Removal of CP prevents poly(A) site cleavage, but PADT can still occur. After 5 min of trnscription, TECs were rinsed and resuspended in the presence or absence of 16 mM CP. Other components of the resuspension mixture were the same as the original transcription mixture except for lacking NTPs, DNA templates and Gal4-p53. Samples were taken at 50 min for the RNase protection assay.

B. Addition of ATP restores efficient PADT but does not rescue poly(A) site cleavage. At 4 min, TECs were rinsed and resuspended in the same reaction mixture as above which also included 0.4 mM CP and the indicated concentration of ATP. For lane 6 of the RNA gel, the sample was taken at 4 min before the TECs were rinsed.


A
NE
NTPs + MOEs
dsxSV40L
-30
0
rinse beads
± 16 mM CP
5 min
RNase protection or PADT assay
RNase protection
PADT assay
CP
+
–
Probe
327 nt
Uncleaved
274 nt
Cleaved
199 nt
Lane #
1
2
3
+ CP
– CP
Polymerase occupancy (Control MOE/Anti-polyA)
frg1
frg2
Time (min)

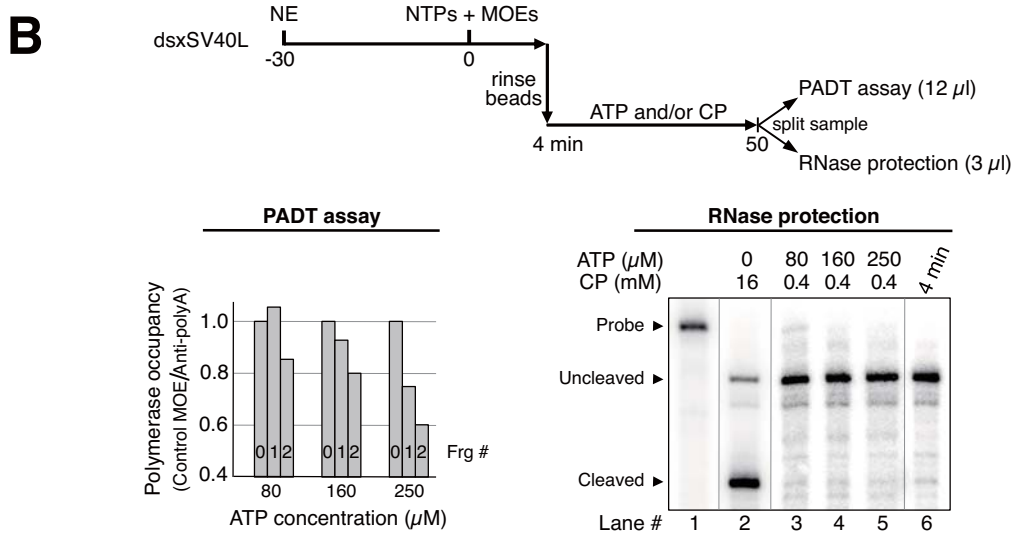
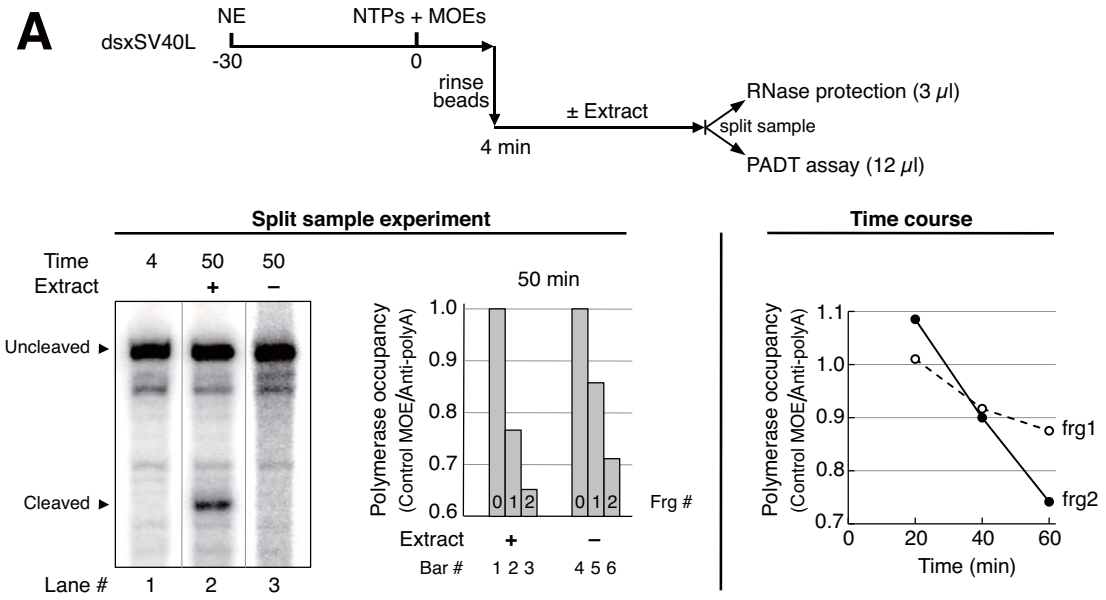

B
NE
NTPs + MOEs
dsxSV40L
-30
0
rinse beads
ATP and/or CP
4 min
50
split sample
PADT assay (12 µl)
RNase protection (3 µl)
PADT assay
RNase protection
Polymerase occupancy (Control MOE/Anti-polyA)
Frg #
ATP concentration (µM)
ATP (µM)
CP (mM)
0
16
80
0.4
160
0.4
250
0.4
4 min
Probe
Uncleaved
Cleaved
Lane #

**Figure 3-4. A conformational change of the TEC drives PADT, not cleavage at the poly(A) site.**

A. PADT, but not cleavage, can occur in the absence of extract. At 4 min, TECs were rinsed and resuspended in the presence or absence of extract. When extract was absent, it was substituted by an equal volume of buffer D (20mM HEPES pH7.9, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 0.1 mM PMSF). Other components of the resuspension mixture were the same as the original transcription mixture except for lacking NTPs, DNA templates and Gal4-p53. For lane 1 of the RNA gel, the sample was taken at 4 min before the TECs were rinsed.

B. Poly(A) site cleavage but not PADT, can be rescued after TECs are stripped with salt and Sarkosyl. After 5 or 8 min of transcription (in different experiments), TECs were either rinsed or stripped (1M KCl and 1% sarkosyl), and then resuspended as usual in the presence of extract. For lanes 1 and 3 of the RNA gel, the samples were taken right after the rinse/strip step. Bars 8, 9, 11 and 12 of the histogram are averaged from two experiments (± range).


A
dsxSV40L
NE
NTPs + MOEs
-30
0
rinse beads
± Extract
4 min
RNase protection (3 μl)
split sample
PADT assay (12 μl)
Split sample experiment
Time
Extract
4
50
+
50
–
Uncleaved
Cleaved
Lane #
1
2
3
50 min
Polymerase occupancy
(Control MOE/Anti-polyA)
Frg #
Extract
Bar #
1 2 3
4 5 6
Time course
frg1
frg2
Time (min)

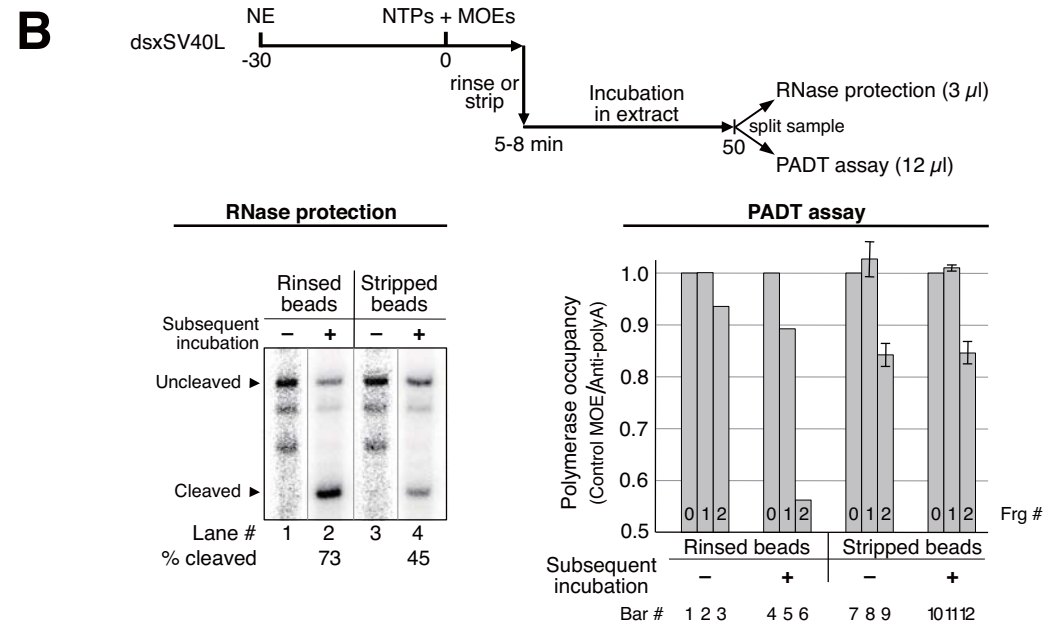

B
dsxSV40L
NE
NTPs + MOEs
-30
0
rinse or strip
5-8 min
Incubation in extract
50
split sample
RNase protection (3 μl)
PADT assay (12 μl)
RNase protection
Rinsed beads
Stripped beads
Subsequent incubation
Uncleaved
Cleaved
Lane #
% cleaved
73
45
PADT assay
Polymerase occupancy
(Control MOE/Anti-polyA)
Frg #
Rinsed beads
Stripped beads
Subsequent incubation
Bar #
1 2 3
4 5 6
7 8 9
10 11 12

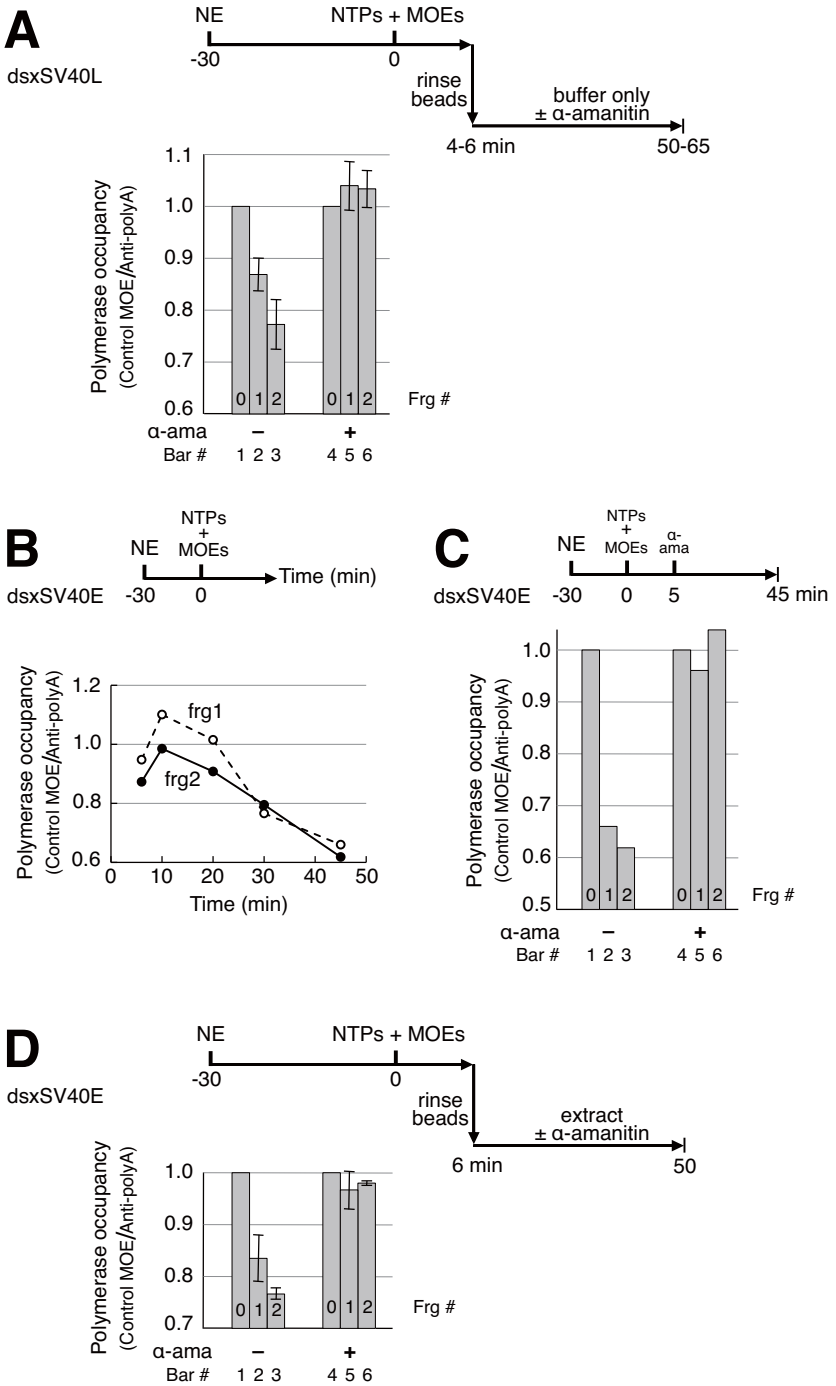
**Figure 3-5. PADT is inhibited by α-amanitin.**

A. PADT by the SV40 late poly(A) signal is inhibited by α-amanitin. Rinsed TECs were resuspended in the absence of extract as for Figure 3-4A, but with or without α-amanitin. Bars 2 and 3 of the histogram are averaged from seven experiments (± SD) and bars 5 and 6 are averaged from four experiments (± SD).

B. PADT by the SV40 early poly(A) signal. The DNA template for this and the following experiments was dsxSV40E.

C. PADT by the SV40 early poly(A) signal is inhibited by α-amanitin (continuous transcription protocol). For this experiment the α-amanitin, when present, was added at 5 min directly to the transcription mixture.

D. PADT by the SV40 early poly(A) signal is inhibited by α-amanitin (rinse and resuspend protocol). Rinsed TECs were resuspended in the presence of extract as for Figure 3-4A, but with or without α-amanitin. In these experiments 25 μM ATP, which enhances PADT by the SV40 early poly(A) signal, was also present. Bars 2, 3, 5 and 6 are averaged from two experiments (± range).


A
dsxSV40L
NE
NTPs + MOEs
-30
0
rinse beads
buffer only ± α-amanitin
4-6 min
50-65
Polymerase occupancy (Control MOE/Anti-polyA)
1.1
1.0
0.9
0.8
0.7
0.6
0 1 2
0 1 2
Frg #
α-ama – +
Bar # 1 2 3 4 5 6
B
dsxSV40E
NE
NTPs + MOEs
-30
0
Time (min)
Polymerase occupancy (Control MOE/Anti-polyA)
1.2
1.0
0.8
0.6
frg1
frg2
0 10 20 30 40 50
Time (min)
C
dsxSV40E
NE
NTPs + MOEs
α-ama
-30
0
5
45 min
Polymerase occupancy (Control MOE/Anti-polyA)
1.0
0.9
0.8
0.7
0.6
0.5
0 1 2
0 1 2
Frg #
α-ama – +
Bar # 1 2 3 4 5 6
D
dsxSV40E
NE
NTPs + MOEs
-30
0
rinse beads
extract ± α-amanitin
6 min
50
Polymerase occupancy (Control MOE/Anti-polyA)
1.0
0.9
0.8
0.7
0 1 2
0 1 2
Frg #
α-ama – +
Bar # 1 2 3 4 5 6

# References


Acheson, N.H. (1984). Kinetics and efficiency of polyadenylation of late polyomavirus nuclear RNA: Generation of oligomeric polyadenylated RNAs and their processing into mRNA. Mol. Cell. Biol. *4*, 722-729.

Bird, G., Fong, N., Gatlin, J.C., Farabaugh, S., and Bentley, D.L. (2005). Ribozyme Cleavage Reveals Connections between mRNA Release from the Site of Transcription and Pre-mRNA Processing. Mol. Cell *20*, 747-758.

Brannan, K., Kim, H., Erickson, B., Glover-Cutter, K., Kim, S., Fong, N., Kiemele, L., Hansen, K., Davis, R., Lykke-Andersen, J., and Bentley, D.L. (2012). mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol. Cell *46*, 311-324.

Brennan, C.A., Dombroski, A.J., and Platt, T. (1987). Transcription termination factor rho is an RNA-DNA helicase. Cell *48*, 945-952.

Carswell, S., and Alwine, J.C. (1989). Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. Mol. Cell. Biol. *9*, 4248-4258.

Chao, L.C., Jamil, A., Kim, S.J., Huang, L., and Martinson, H.G. (1999). Assembly of the cleavage and polyadenylation apparatus requires about 10 seconds in vivo and is faster for strong than for weak poly(A) sites. Mol. Cell. Biol. *19*, 5588-5600.

Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev. *2*, 440-452.

Das, A., Merril, C., and Adhya, S. (1978). Interaction of RNA polymerase and rho in transcription termination: coupled ATPase. Proc. Natl. Acad. Sci. U.S.A. *75*, 4828-4832.

Dengl, S., and Cramer, P. (2009). Torpedo nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. J. Biol. Chem. *284*, 21270-21279.

Dye, M.J., Gromak, N., and Proudfoot, N.J. (2006). Exon tethering in transcription by RNA polymerase II. Mol. Cell *21*, 849-859.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Genes Dev. *22*, 1069-1081.

Epshtein, V., Dutta, D., Wade, J., and Nudler, E. (2010). An allosteric mechanism of Rho-dependent transcription termination. Nature *463*, 245-249.

Gilmour, D.S., and Fan, R. (2008). Derailing the locomotive: transcription termination. J. Biol. Chem. *283*, 661-664.

Gu, W., Wind, M., and Reines, D. (1996). Increased accommodation of nascent RNA in a product site on RNA polymerase II during arrest. Proc. Natl. Acad. Sci. U.S.A. *93*, 6935-6940.

He, X., Khan, A.U., Cheng, H., Pappas, D.L., Jr., Hampsey, M., and Moore, C.L. (2003). Functional interactions between the transcription and mRNA 3′ end processing machineries mediated by Ssu72 and Sub1. Genes Dev. *17*, 1030-1042.

Hirose, Y., and Manley, J.L. (1997). Creatine phosphate, not ATP, is required for 3' end cleavage of mammalian pre-mRNA in vitro. J. Biol. Chem. *272*, 29636-29642.

Hirose, Y., and Manley, J.L. (2000). RNA polymerase II and the integration of nuclear events. Genes Dev. *14*, 1415-1429.

Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119-2137.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev. *22*, 1082-1092.

Kazerouninia, A., Ngo, B., and Martinson, H.G. (2010). Poly(A) signal-dependent degradation of unprocessed nascent transcripts accompanies poly(A) signal-dependent transcriptional pausing in vitro. RNA *16*, 197-210.

Kim, M., and Buratowski, S. (2013). Transcription Termination by RNA Polymerase II. In Posttranscriptional Gene Regulation: RNA Processing in Eukaryotes, J. Wu, ed. (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany), pp. 19-40.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517-522.

Kim, S.J., and Martinson, H.G. (2003). Poly(A)-dependent transcription termination: continued communication of the poly(A) signal with the polymerase is required long after extrusion in vivo. J. Biol. Chem. *278*, 41691-41701.

Kuehner, J.N., Pearson, E.L., and Moore, C. (2011). Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol *12*, 283-294.

Lang, W.H., Platt, T., and Reeder, R.H. (1998). Escherichia coli rho factor induces release of yeast RNA polymerase II but not polymerase I or III. Proc. Natl. Acad. Sci. U.S.A. *95*, 4900-4905.

Logan, J., Falck-Pedersen, E., Darnell, J.E., Jr., and Shenk, T. (1987). A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse $b^{maj}$-globin gene. Proc. Natl. Acad. Sci. U.S.A. *84*, 8306-8310.

Luo, W., Johnson, A.W., and Bentley, D.L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev. *20*, 954-965.

Martinez-Rucobo, F.W., and Cramer, P. (2013). Structural basis of transcription elongation. Biochim. Biophys. Acta *1829*, 9-19.

McCracken, S., Lambermon, M., and Blencowe, B.J. (2002). SRm160 Splicing Coactivator Promotes Transcript 3'-End Cleavage. Mol. Cell. Biol. *22*, 148-160.

Mischo, H.E., and Proudfoot, N.J. (2013). Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim. Biophys. Acta *1829*, 174-185.

Nag, A., Narsinh, K., and Martinson, H.G. (2007). The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. Nat Struct Mol Biol *14*, 662-669.

Orozco, I.J., Kim, S.J., and Martinson, H.G. (2002). The poly(A) signal, without the assistance of any downstream element, directs RNA polymerase II to pause in vivo and then to release stochastically from the template. J. Biol. Chem. *277*, 42899-42911.

Osheim, Y.N., Proudfoot, N.J., and Beyer, A.L. (1999). EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. Mol. Cell *3*, 379-387.

Osheim, Y.N., Sikes, M.L., and Beyer, A.L. (2002). EM visualization of Pol II genes in Drosophila: most genes terminate without prior 3' end cleavage of nascent transcripts. Chromosoma *111*, 1-12.

Park, J.S., and Roberts, J.W. (2006). Role of DNA bubble rewinding in enzymatic transcription termination. Proc. Natl. Acad. Sci. U.S.A. *103*, 4870-4875.

Park, N.J., Tsao, D.C., and Martinson, H.G. (2004). The Two Steps of Poly(A)-Dependent Termination, Pausing and Release, Can Be Uncoupled by Truncation of the RNA Polymerase II CTD. Mol. Cell. Biol. *24*, 4092-4103.

Pearson, E.L., and Moore, C.L. (2013). Dismantling promoter-driven RNA polymerase II transcription complexes in vitro by the termination factor Rat1. J. Biol. Chem. *288*, 19750-19759.

Proudfoot, N. (2000). Connecting transcription to messenger RNA processing. Trends Biochem. Sci. *25*, 290-293.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem. Sci. *14*, 105-110.

Qu, X., Perez-Canadillas, J.M., Agrawal, S., De Baecke, J., Cheng, H., Varani, G., and Moore, C. (2007). The C-terminal domains of vertebrate CstF-64 and its yeast orthologue Rna15 form a new structure critical for mRNA 3'-end processing. J. Biol. Chem. *282*, 2101-2115.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev. *23*, 1247-1269.

Richardson, J.P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim. Biophys. Acta *1577*, 251-260.

Rigo, F., Kazerouninia, A., Nag, A., and Martinson, H.G. (2005). The RNA Tether from the Poly(A) Signal to the Polymerase Mediates Coupling of Transcription to Cleavage and Polyadenylation. Mol. Cell *20*, 733-745.

Rigo, F., and Martinson, H.G. (2009). Polyadenylation releases mRNA from RNA polymerase II in a process that is licensed by splicing. RNA *15*, 823-836.

Rosonina, E., Bakowski, M.A., McCracken, S., and Blencowe, B.J. (2003). Transcriptional activators control splicing and 3'-end cleavage levels. J. Biol. Chem. *278*, 43034-43040.

Sadowski, M., Dichtl, B., Hubner, W., and Keller, W. (2003). Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. EMBO J. *22*, 2167-2177.

Schwartz, A., Margeat, E., Rahmouni, A.R., and Boudvillain, M. (2007). Transcription termination factor rho can displace streptavidin from biotinylated RNA. J. Biol. Chem. *282*, 31469-31476.

Teixeira, A., Tahiri-Alaoui, A., West, S., Thomas, B., Ramadass, A., Martianov, I., Dye, M., James, W., Proudfoot, N.J., and Akoulitchev, A. (2004). Autocatalytic RNA cleavage in the human beta-globin pre-mRNA promotes transcription termination. Nature *432*, 526-530.

Thomas, M.C., and Chiang, C.M. (2006). The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol *41*, 105-178.

Vickers, T.A., Wyatt, J.R., Burckin, T., Bennett, C.F., and Freier, S.M. (2001). Fully modified 2' MOE oligonucleotides redirect polyadenylation. Nucleic Acids Res. *29*, 1293-1299.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522-525.

West, S., Proudfoot, N.J., and Dye, M.J. (2008). Molecular dissection of mammalian RNA polymerase II transcriptional termination. Mol. Cell *29*, 600-610.

Whitelaw, E., and Proudfoot, N. (1986). a-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3' end processing in the human a2 globin gene. EMBO J. *5*, 2915-2922.

Wu, H., Lima, W.F., and Crooke, S.T. (1999). Properties of cloned and expressed human RNase H1. J. Biol. Chem. *274*, 28270-28278.

Yeung, G., Choi, L.M., Chao, L.C., Park, N.J., Liu, D., Jamil, A., and Martinson, H.G. (1998). Poly(A)-driven and poly(A)-assisted termination: two different modes of poly(A)-dependent transcription termination. Mol. Cell. Biol. *18*, 276-289.

Yuzenkova, Y., Roghanian, M., Bochkareva, A., and Zenkin, N. (2013). Tagetitoxin inhibits transcription by stabilizing pre-translocated state of the elongation complex. Nucleic Acids Res. *41*, 9257-9265.

Zhang, Z., Wu, C.H., and Gilmour, D.S. (2004). Analysis of polymerase II elongation complexes by native gel electrophoresis. Evidence for a novel carboxyl-terminal domain-mediated termination mechanism. J. Biol. Chem. *279*, 23223-23228.

Zhelkovsky, A., Tacahashi, Y., Nasser, T., He, X., Sterzer, U., Jensen, T.H., Domdey, H., and Moore, C. (2006). The role of the Brr5/Ysh1 C-terminal domain and its homolog Syc1 in mRNA 3'-end processing in Saccharomyces cerevisiae. RNA *12*, 435-445.