MEETING REVIEW

RNA 3’-end formation

Two major areas in the study of gene expression and its regulation are transcriptional events such as initiation and termination and the subsequent processing of the primary transcripts. Each of these fields has expanded to such an extent in recent years that entire meetings are devoted to one or the other topic, with the result that cross-fertilization of interests and expertise at such meetings is increasingly rare. Representatives of these two areas, with shared interests in the RNA-processing reactions that generate 3’ ends and in transcription termination, were reunited recently at the Advanced Workshop on RNA 3’-End Formation. The presence of investigators studying prokaryotic and eukaryotic systems lent additional perspectives to discussions throughout. There were several major recurring themes at this meeting: the importance of both DNA and RNA structure, as well as specific protein–nucleic acid interactions, in termination; the realization that signals for both RNA processing and termination are often more complex than previously thought; the existence of multicomponent complexes that catalyze 3’ processing; and the likelihood of significant interplay between transcription and processing. Space constraints prevent us from discussing all of the interesting findings that were described, and we have concentrated therefore on the talks most relevant to the above themes.

The prokaryotic paradigm

Transcription termination appears to be the general mechanism by which mRNA 3’ ends are formed in bacteria, and, unlike in higher cells, processing events play a less important role. The major focus of talks concerned three areas: (1) ‘simple’ termination brought about by RNA hairpin structures followed by uridine runs, (2) ρ factor-dependent termination, and (3) antitermination. It was agreed that simple (often called ‘factor-independent’) terminators should be termed ‘intrinsic’ termination sites, reflecting their effectiveness with RNA polymerase alone in vitro, while admitting the possibility that in vivo additional factors might play a modulating role.

An important question concerning intrinsic terminators is whether the sequence of a hairpin-U stretch is sufficient to define a termination site. Consistent with this, a thermodynamic model, based on calculating the stabilities of competing DNA–DNA, DNA-RNA, and RNA–RNA structures in the ternary transcription complex, accurately predict both the position and the efficiency of termination for a number of sites from the DNA sequence alone (von Hippel, Eugene). However, terminators of similar predicted stabilities display wide variations in efficiency in vitro as a function of salt, counterions, and Mg2+ concentrations (although at 0.4 mM Mg2+, all were very efficient) and sequences some distance downstream of the hairpin, and thus not yet transcribed, were shown to have substantial effects on termination efficiency (Chamberlin, Berkeley). Because these results would not have been predicted by the above thermodynamic analysis, factors in addition to the sequence of the terminator, including RNA–enzyme and DNA–enzyme interactions, probably are important in determining termination efficiency at many intrinsic sites.

The function of the well-studied termination factor ρ involves binding of the protein to target sites on nascent mRNA, translocation of the ρ protein toward the transcript’s 3’ end (driven by hydrolysis of NTPs), and an intrinsic RNA–DNA helicase activity. Beyond a characteristic lack of secondary structure, specific sequences in the RNA target site have not been identified. It was observed previously, though, that some target sites are cytosine-rich, and elimination of this cytosine-richness by mutagenesis (in the most extreme case converting 26 of 28 Cs to Us), did increase readthrough transcription, but only to about 70% (Platt, Rochester). Analysis of the RNA–DNA helicase activity of the ρ protein in vitro demonstrated that spacer regions up to 450 nucleotides separating the target ρ binding site and the RNA–DNA helix are not detrimental to function. However, a short RNA oligomer hybridized to the spacer can block helicase activity, suggesting that ρ must maintain more-or-less continuous contact with single-stranded regions during its 5’ to 3’ action (Platt). During translocation ρ remains anchored to its target site, and movement involves ATP-dependent interactions between the downstream RNA and secondary binding sites on the protein (Richardson, Indiana). Perhaps to facilitate this translocation, ρ is active as a hexamer, and each subunit contains both RNA- and ATP-binding domains, characterized by the appropriate consensus sequences (von Hippel, Richardson, Platt).

Mechanisms of antitermination are now more clearly defined for both sites and factors. The classical N protein-mediated antitermination in phage λ has been reconstituted in vitro with purified proteins (Greenblatt, Toronto). Several host proteins (the nus proteins), in addition to N, are required to form an antitermination...
transcription complex. Ribosomal protein S10 [nusE] joins polymerase holoenzyme at initiation, as the elongation complex [now also with nusA] proceeds through a nut site [which appears to function as RNA], the addition of N protein catalyzes incorporation of nusB and the newly characterized nusG protein, forming a termination-resistant elongation complex. Transcription of host genes can be antiterminated by a similar mechanism. For example, the ribosomal RNA [rrn] operons each have two copies of a nut-like sequence that allow antitermination of distal p-dependent termination sites [Squires, Columbia]. The host equivalent of N, however, remains unknown. Transcription from the λ late promoter is also subject to antitermination control, by the λ Q protein. Q mediates release from a strong promoter–proximal pause by a mechanism that both requires initiation at the late promoter and also eliminates pausing at distal sites [Roberts, Cornell]. Mutations that affect Q action occur within the promoter itself as well as some distance downstream of the pause site. Promoter-specific pausing and antitermination also occur in eukaryotes [see below]. It will be of interest to learn in the future if the mechanisms employed are similar to those that operate in prokaryotes.

**Termination by eukaryotic RNA polymerases**

New insights into how eukaryotic RNA polymerases I, II, and III terminate transcription emerged, and, indeed, several common themes were apparent. One of these was the identification of protein factors that interact with sites and/or enzyme to facilitate termination. For example, RNA polymerase III [pol III] terminates at simple runs of four or more Ts in a G-C–rich environment. However, completion of the last few nucleotides of the transcript and release of enzyme and RNA from template also requires an additional factor, the 50-kD La protein [Gottlieb, Yale [MRC, Cambridge]].

The tandemly repeated rRNA genes possess multiple terminators that depend on the binding of specific proteins to the DNA template to bring about termination. Each gene copy contains identical sites that lie both in the region encoding the 3' end of the long rRNA precursor as well as in or near the pol I promoters itself. The terminator at the 3' end of the pre-rRNA has been examined most thoroughly in mouse [Grummt, Wurzburg], where an 18-bp sequence is necessary and sufficient for pol I termination, given the binding of a specific factor [TTF I]. This protein has been extensively purified and termination has been reproduced in vitro. Interestingly, the termination signal is orientation dependent and invisible to other polymerases, suggesting that its function is more subtle than simply blocking elongation. The situation in the human rRNA cluster is analogous, and a nearly identical sequence and a DNA-binding protein cause termination just upstream of the rRNA transcription start site [Sollner-Webb, Johns Hopkins]. Such promoter-proximal sites function to prevent transcriptional interference, or promoter occlusion, by readthrough of polymerases from the spacer regions that separate each rDNA repeat. Curiously, an RNA-processing function is associated with the pre-rRNA 3’ site in *Xenopus laevis* [Reeder, Hutchinson]. A single C-to-G change distinguishes this site [T2] from the promoter-proximal 8-bp termination site. Polymerases read through T2 unpended, although RNA 3’ ends are formed efficiently. Has the single-base change uncoupled transcript release from polymerase release? Or is there a distinct RNA-processing signal coincident with the nonfunctional termination site?

RNA polymerase II appears to require two signals to terminate transcription of genes encoding polyadenylated mRNAs—a functional poly[A] addition signal and a sequence located at or near the actual site of termination. This second element can lie hundreds, or even thousands, of bp 3’ to the poly[A] site, and several talks dealt with these downstream sites. In the human α-globin gene, a 50-bp sequence located ~300 bp downstream of the poly[A] site was identified [Proudfoot, Oxford]. This orientation-dependent sequence can facilitate termination not only at its natural position, but also within an intron, so long as a functional poly[A] site is inserted upstream. In adenoavirus, an inverted CCAAT box sequence just upstream of the late promoter functions as a termination sequence, perhaps, by analogy with rRNA termination, to prevent transcriptional interference from upstream genes [Connelly, Columbia [FMI]]. Interestingly, this site works in only one orientation, and analysis of single-base substitutions suggests that binding of a CCAAT protein is required, adding another role in addition to those in transcription initiation and DNA replication for these multifunctional proteins. The ability of the c-myc gene attenuator [see below] to function as a terminator when inserted downstream of the mouse β-major globin poly[A] site was also described [Falck-Pederson, Cornell]. This site induces termination in HeLa cells in which endogenous c-myc expression is repressed, but not in cells expressing high levels of c-myc. Could this indicate the existence of a protein that recognizes this site and blocks polymerase elongation?

Termination of histone mRNA transcription occurs by a mechanism similar to that for polyadenylated mRNAs [Marzluff, Tallahassee]. In the mouse, the two most highly expressed histone H2a and H3 genes are 800 bp apart and are transcribed from the same DNA strand. Termination of H2a transcription occurs close to the H3 promoter, and appears to require not only sequences near the termination site, but also an intact H2a mRNA 3’ processing signal. Although it remains to be seen whether protein binding to the H3 promoter-proximal element is involved in its termination function, this situation is remarkably similar to the adeno virus late promoter terminator described above.

Termination of small nuclear [sn] RNA transcripts synthesized by pol II is unusual in its promoter specificity: Only polymerases that initiate from an snRNA promoter are affected by the so-called 3’ box situated at the end of the gene. In the case of human U1 snRNA, 3’ ends are formed ~20 bp upstream of the 3’ box [Dahl-
Termination of snRNA gene transcription is perhaps more analogous to attenuation or premature termination [see below] than to termination of mRNA-encoding genes. Both attenuation and snRNA gene termination occur after the polymerase has transcribed only a short distance, and both display an element of promoter specificity. In contrast, termination of mRNA transcripts occurs only after the polymerase has traversed considerable distances, and there is no evidence that such termination is promoter specific.

The biochemistry of pre-mRNA 3'-end processing

The 3' ends of eukaryotic mRNAs are not formed by transcription termination, but rather by RNA processing. These reactions require specific signal sequences in the pre-mRNA, and are catalyzed by complex sets of factors. Significant progress toward understanding these reactions in higher eukaryotes as well as in yeast was reported at the meeting. In metazoans, formation of the 3' ends of polyadenylated mRNAs requires the highly conserved AAUAAA sequence, usually located 10–30 nucleotides 5' to the site of poly(A) addition, as well as, in many cases, a GU- or U-rich sequence just downstream of this site (which appears to contribute to the efficiency of the reaction). Unexpectedly, sequences upstream of AAUAAA may also be important in some instances (Alwine, Penn.). The polyadenylation reaction consists of a site-specific endonucleolytic cleavage at the mRNA 3' end, followed by polymerization of a 200 to 300-nucleotide poly(A) tail. These two reactions are tightly coupled in vivo, but can be experimentally uncoupled in vitro, and it has been shown, for example, that both 'half-reactions' require the AAUAAA sequence.

As many as five separable factors are required to carry out the cleavage-polyadenylation reaction, generating a processing complex that is somewhere between 500 and 1000 kD [Manley, Columbia; Keller, Basel; Nevins, Duke]. A key component is a factor of ~250 kD, which is required for both the cleavage and polyadenylation half-reactions. This 'specificity' factor is a strong candidate for the component that recognizes AAUAAA. However, since it has not yet been purified to homogeneity, the identity of this factor is unknown, and the question of whether or not it is an snRNP particle remains unresolved. Three additional proteins appear to be required to catalyze the cleavage reaction. While the precise functions of these factors is unknown, one, a 200-kD, three-subunit protein, has been purified to apparent homogeneity [Manley]. For polyadenylation, the only other component required is an ~50-kD poly(A) polymerase (PAP). This enzyme is apparently identical to the n-specific PAPs first identified over 25 years ago [Wickens, Madison], and it appears that the enzyme becomes AAUAAA-dependent by interaction with the specificity factor. Multiple chromatographic forms of PAP that function in specific polyadenylation were described [Manley, Keller], but their significance is currently unclear. Synthesis of the poly(A) tail occurs in two phases, in which addition of approximately the first 10 A residues requires AAUAAA, while addition of the remainder does not [Wickens]. PAP can also influence the efficiency and accuracy of the cleavage reaction; however, it is unlikely to contain the actual cleavage activity, because at least one pre-RNA can be cleaved efficiently in its absence. Finally, a possible role for hnRNP proteins, in particular the C proteins, in enhancing the efficiency of the cleavage-polyadenylation reaction by interacting with downstream U-rich sequences was suggested [Shenk, Princeton].

The 3' ends of mRNAs in the yeast Saccharomyces cerevisiae are also formed by cleavage-polyadenylation. Several different pre-RNAs can be processed in an in vitro reaction that appears similar to the mammalian reaction [Butler, Rochester]. Moreover, extracts from one mutant temperature-sensitive strain (in the L. Hartwell collection) are defective in polyadenylation but not in cleavage. Genetic approaches to identify this and other genes involved in 3' processing should provide insights into the factors, as well as illuminate the troublesome lack of defined nucleotide sequences that signal 3' end formation in S. cerevisiae. For example, many yeast mRNAs lack the universal metazoan AAUAAA. Indeed, in the ADH2 gene, which does contain this sequence, mutagenesis showed it to be unnecessary in vivo, although a single-base change farther upstream did inhibit 3' end formation [Moore, Tufts].

The formation of nonpolyadenylated histone mRNA 3' ends also occurs by endonucleolytic cleavage, but the signal sequences and factors involved are totally different from those required for 3' cleavage of polyadenylated mRNAs. An upstream conserved hairpin loop is important for efficient 3'-end formation in vitro, but not absolutely required, while a downstream sequence, complementary to a sequence in U7 snRNA, is essential [Steitz, Yale]. In addition to playing a role in processing and termination of histone transcripts see above, the stem-loop also appears important for the transport and stability of the mRNA in the cytoplasm [Marzluff].

Three separable factors are required to catalyze histone 3'-end formation in vitro [Schumperli, Berne, Vassero], IMP]. The U7 snRNP, whose function involves base-pairing to the histone pre-mRNA downstream sequence, may exist in multiple forms. One of these is inactive, perhaps due to the presence of additional protein components that could shield the RNA sequence involved in base-pairing with the pre-mRNA [Vassero]. Two additional proteins are required for the reaction, a heat-labile factor (HLF), and a hairpin-binding factor, which interacts with the upstream stem-loop structure. Although these factors can be separated from U7 snRNP in vitro, it is possible that they may be snRNP-associated in vivo. The activity of HLF is cell-cycle regulated, appearing during G1 and persisting into S phase. This suggests that regulation of the activity of HLF, and thus of histone mRNA 3' end formation, plays an im-
important role in the control of histone synthesis [Schumperli].

Regulation of gene expression by 3'-end formation

A number of examples showing that 3'-end formation can play a role in regulating gene expression have been documented. A λ N protein-mediated antitermination control, described above, is a classical example. Several talks dealt with two other types of regulation involving RNA 3'-end formation: poly[A] site selection in eukaryotes and attenuation, or controlled premature termination of transcription, which occurs in both prokaryotes and eukaryotes.

The best-studied case of differential poly[A] site usage is in immunoglobulin genes, where the membrane-bound and secreted forms of antibody are generated by alternative poly[A] site usage. Nearly every transcriptional, and post-transcriptional, mechanism may be employed in regulating this process. Some B cells terminate transcription between the two poly[A] sites, thereby favoring the weaker upstream secreted-form poly[A] site [pAs] [Nevins, Chen-Kiang, Mt. Sinai]. Alternatively, the two poly[A] sites may compete for factors, so that if these are limiting, that is, in pre B-cells, the stronger membrane-form poly[A] site [pAm] predominates; however, when the factors are produced in higher amounts, the weaker upstream pAs is favored [Nevins]. Another twist to this complex picture is that the upstream pAs is within an intron that must be spliced out to form the membrane-bound antibody. The 5' splice site of this intron is weak, and mutations that make this site more efficient promote increased usage of the downstream pAm by splicing out the upstream site [Perry, Fox Chase]. Thus, changes in polyadenylation, termination, and/or splicing efficiencies may all contribute to this developmentally regulated switch.

Two in vitro studies provide support for the ideas that both trans-acting factors and splicing can influence the efficiency of polyadenylation. An example of regulated poly[A] site selection was described in the herpes simplex virus system, where extracts from infected cells were shown to contain an activity that preferentially increases the use of a viral late mRNA poly[A] site [McLauchlan, MRC, Glasgow]. These findings provide the first evidence that a trans-acting factor can control poly[A] site selection, and they set the stage for its identification. Will the factor be virally encoded? Or will the activity of a host protein be qualitatively or quantitatively altered? In vitro studies also suggest that polyadenylation efficiency may be linked to splicing of upstream introns [Berger, Baylor]. With pre-mRNAs containing both introns and poly[A] sites, each processing mechanism is more efficient when the other is occurring on the same precursor, and mutations in either signal affect both processes. Polyadenylation is also stimulated when just the 3' splice site is present upstream on the same pre-mRNA. Since recognition of the acceptor site is one of the first steps in splicing, it is possible that in some cases polyadenylation may be linked to this initial stage in RNA processing.

An interesting form of regulated polyadenylation occurs during oocyte maturation in which poly[A] tails are added to certain preexisting cytoplasmic mRNAs. Microinjection of RNAs into X. laevis oocytes revealed that this polyadenylation requires two signal sequences, AAUAAA and a [U]₆AU sequence located nearby [Wickens]. Cytoplasmic polyadenylation may be required for translational activation of the relevant mRNAs, and it will be of interest to learn more details concerning the mechanism of this process. For example, although the reaction occurs in the cytoplasm, are some of the same factors that catalyze nuclear polyadenylation involved?

Another mechanism for regulating gene expression by 3'-end formation is the attenuation-controlled premature termination of transcription. In prokaryotes, attenuation is well known to involve a coupling of translation and transcription, a process that requires the elongating polymerase to pause long enough for the ribosomes to initiate translation upstream. The sequences required for such pausing are proving to be complex, and they involve elements both upstream (in an RNA hairpin) and downstream of the pause site [Landick, St. Louis]. Even though eukaryotes cannot regulate premature termination by such coupling, it is apparent that attenuation or pausing occurs in several eukaryotic genes. Purified pol II terminates specifically at three clustered sites within the first intron of the human histone H3.3 gene, the strongest of which is characterized by a critical run of Ts [Kane, Berkeley]. Intriguingly, mutations that prevent termination also destroy a bend in the DNA, suggesting that DNA structure may be important for this termination. Pol II in cell extracts prematurely terminates at sites in the SV40 and adenovirus late leader sequences, and the signals here may involve hairpin loops [Aloni, Weizmann]. Although the histone and viral sites are not obviously related, the transcription factor TFIIIS promotes readthrough of all of them. Could this protein play a role similar to the antiterminators in Escherichia coli?

Transcription of the mouse and human c-myc genes terminates prematurely in the first exon, and regulation of this process is important in controlling c-myc expression. Although the exact sequences responsible remain elusive, a 180-bp DNA fragment containing the attenuation site, as well as multiple protein binding sites, can function in an orientation-dependent manner with a heterologous promoter [Wright, U. California, San Francisco]. However, the attenuator appears to be very particular about which promoter it will work with. For example, fusion of certain enhancers to c-myc sequences leads to readthrough of the attenuator. This does not appear to be due to changes in promoter strength, and it raises the interesting possibility that certain enhancer-mediated initiation events may be qualitatively different from others [Bentley, ICRF]. Also, in cases of Burkitt's lymphoma in which c-myc is expressed at high levels,
transcription appears to initiate from an upstream promoter that does not recognize the attenuator (Groudine, Hutchinson). It will be of interest to determine the difference between ‘terminating’ and ‘readthrough’ polymerases. Could TFIIS or possible E. coli nus-homologs be involved?

Summary and perspectives

The mechanisms by which RNA 3’ ends are formed are both diverse and complex. Even the simplest terminators, such as the intrinsic hairpin loop sites in E. coli and pol III sites in eukaryotes, require more in the way of signal sequences and/or protein factors than was previously realized. Termination by both pol I and pol II can be effected by proteins bound at or near termination sites. However, pol II termination appears to be more complex, in that it undoubtedly involves additional factors and can occur by several distinct pathways. Likewise, intricate mechanisms are involved in promoter-proximal antitermination and attenuation in both prokaryotic and eukaryotic systems. The main factors that process RNA 3’ ends have been identified, and examination of the mechanism and control of these complex, multicomponent reactions with purified factors is foreseeable.

Some of the intriguing questions raised at the meeting include the following: What are the precise roles played by RNA and DNA structures in termination? How do sequence-specific DNA-binding proteins interact with RNA polymerases to mediate termination? Is there a eukaryotic p factor? How and why is 3’-end processing linked to termination? How frequently and by what factors is 3’-end formation regulated? Reports from the meeting make us optimistic that these and other questions will soon be answered.

James L. Manley,¹ Nicholas J. Proudfoot,² and Terry Platt³

¹Department of Biological Sciences, Columbia University, New York, New York 10027 USA; ²Sir William Dunn School of Pathology, University of Oxford, Oxford, OX1 3RE UK; ³Department of Biochemistry, University of Rochester Medical School, Rochester NY 14642 USA.
RNA 3′-end formation

James L. Manley, Nicholas J. Proudfoot and Terry Platt

*Genes Dev.* 1989, 3:
Access the most recent version at doi:10.1101/gad.3.12b.2218

License

**Email Alerting Service**
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).