Co-transcriptional monitoring of mRNP formation

Torben Heick Jensen and Michael Rosbash

In eukaryotes, newly manufactured mRNA is subject to strict quality control. If certain standards are not met, offending messages are retained in the nucleus and degraded by the exosome complex. New data suggest that exosome-mediated processes are intimately coupled to the transcription machinery, which makes the mRNA in the first place.

Sydney Brenner, one of three winners of the 2002 Nobel Prize for Medicine or Physiology, liked to describe regulative metazoan development as working on the 'American Plan'. In other words, what matters is neighborhood rather than ancestry. In development, 'neighborhood' would mean interactions and signals shared between nearby cells rather than the lineage of the cells themselves, the 'European plan'. Meanwhile, modern proteomic strategies have allowed us to characterize protein complexes, identifying the friends and neighbors of individual proteins of interest. Indeed, these newly discovered links between proteins have sometimes revealed functional relationships that could not have been anticipated based only on 'European' notions of a given protein's individual identity. One surprising new association was reported in a recent issue of *Nature* by John Lis and colleagues[1], linking mRNA synthesis, or transcription, with RNA degradation wrought by the nuclear exosome.

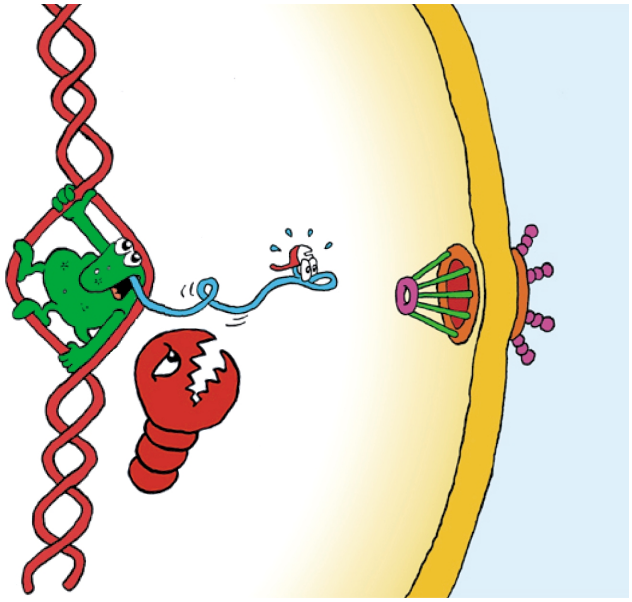


Fig. 1 Cartoon showing quality control of nascent eukaryotic mRNA. The nuclear exosome (red) is shown poised to degrade newly synthesized mRNA (blue) as it is produced by RNA polymerase II (green). This may occur in situations where mRNAs are not assembled into functional mRNP particles. Artwork courtesy of E. Sloth Andersen.

The Spt5 and Spt6 transcription factor proteins are involved in transcription elongation in both yeast and humans. Immunolocalization studies and chromatin immunoprecipitation (ChIP) experiments from the Lis and Winston laboratories have shown that the *Drosophila* versions Spt5 and Spt6 (dSpt5 and dSpt6) are localized at transcriptionally active sites on polytene chromosomes, suggesting that the proteins work the same way in yeast, flies and humans[2,3]. An effort to identify proteins associated with dSpt6 in *Drosophila* nuclear extracts has now led to the co-purification of nine polypeptides at approximately stoichiometric levels with dSpt6. These proteins have been shown by mass spectrometry to be *Drosophila* orthologs of exosome subunits, components of the RNA degrading machine, which has been well characterized in the yeast and human systems. The interaction between dSpt6 and the exosome was resistant to RNase A and therefore is mediated by protein–protein interactions. dSpt5 and the large subunit of RNA polymerase II were also identified in these protein complexes.

Because the exosome is thought to be involved in surveillance of nascent or newly synthesized mRNAs as part of an error-detection mechanism, the authors were prompted to investigate whether the dSpt6–exosome interaction occurred during transcription. Sure enough, antibodies against exosomal components stained transcriptionally active loci on *Drosophila* polytene chromosomes in a pattern very similar to that shown by dSpt6 antibodies. ChIP assays confirmed that antibodies against the exosomal components dRrp6, dRrp4 and dCsl4 all immunoprecipitated chromatin containing the hsp70 and hsp26 heat shock loci, and that these interactions depended on active transcription. Interestingly, signals were obtained from both the 5′ and 3′ ends of the genes, much like the pattern obtained with dSpt6 antibodies. These results indicate that the exosome is recruited to genes early in the transcription process, and suggest that Spt5 and Spt6 are mediators of this recruitment.

Only in the last year or so have other studies hinted that the exosome might function within hailing distance of active chromatin. In the yeast *Saccharomyces cerevisiae*, several types of defective mRNA accumulate at or near the sites of their transcription[4–6]. These include transcripts with defective 3′ ends, as well as transcripts produced in yeast strains in which components of the TREX protein complex (also known as the THO complex) are dysfunctional[6–8]. The TREX

complex is a heterotetrameric protein complex that is recruited to activated genes during transcription and couples the processes of transcription and mRNA export. The TREX complex is probably recruited by RNA polymerase II and either contains or is associated with the mRNA export protein factors Yra1p/REF and Sub2p/UAP56 (Yra1p and Sub2p being the *S. cerevisiae* designations). These mRNA export factors, and perhaps the entire TREX complex, are likely to be deposited on nascent mRNA and contribute to messenger ribonucleoprotein (mRNP) formation and RNA export to the cytoplasm.

Intriguingly, all cases in which mRNAs are known to be retained at the site of transcription in yeast require functional Rrp6 protein as well as other components of the nuclear exosome — in other words, aberrant mRNAs are not retained in yeast mutant for these exosome components[4,6]. This led to the idea that the exosome has an mRNA surveillance role, closely connected to the process of transcription itself. Degradation of aberrant mRNAs was not always observed, suggesting that retention is at least somewhat distinct from the more traditional function of the exosome as an RNA-degrading machine. A direct interaction between yeast Spt6p and histone H3 has also been observed[9], which suggests further links between the exosome–mRNA substrate complex and chromatin.

The exosome was originally shown to participate in the processing of rRNAs, snoRNAs and snRNAs, all of which require its 3′ to 5′ exonuclease activity. Indeed, the bulk of Rrp6p is detected in the nucleolus of both yeast and human cells[10,11]. However, some Rrp6p is found in the nucleoplasm, and the exosome has also been shown to participate in a pre-mRNA degradation pathway[12]. More recently, transcripts with extended 3′ untranslated regions (UTRs) produced in mutant strains defective in 3′ end processing/transcription termination have been shown to be substrates for the nuclear exosome[6,13]. While these studies clearly show that mRNA is degraded in the nucleus, the exact location for this process within the nucleus has remained elusive. The recent *Nature* paper suggests that at least some destruction of mRNA takes place at sites of active transcription[1]. Moreover, dSpt6 immunoprecipitates possess exonucleolytic activity, suggesting that the exosome is capable of degrading mRNA while complexed with chromatin.

Lis and co-workers argue that the exosome surveys and destroys inappropriately made or processed pre-mRNA molecules that would otherwise interfere with normal transcription elongation, RNA processing, or export. But why is the exosome in close proximity to transcription and what exactly is being surveyed? Although the answer is far from certain, co-transcriptional formation of functional mRNPs seems to be a good bet.

It is well established that some mRNA processing factors are recruited to nascent mRNA *via* the C-terminal domain of RNA polymerase II[14]. Recently, co-transcriptional loading of mRNA export factors has also been proposed[7,8,15]. It is possible that coupling of the transcription machinery to mRNA processing is critical for step wise construction of an mRNP competent for export from the nucleus. Difficulties or delays in these early assembly events would result in abnormal mRNPs. In this view, mRNA export from the nucleus requires co-transcriptional mRNP construction.

However, mRNA export factors can also work when they are loaded post-transcriptionally onto mRNA. For example, the mRNA export factors Yra1p/REF and Sub2p/UAP56 appear to be present in the co-transcriptional TREX complex[7], but they are also found in the mammalian exon junction complex, which is deposited onto mammalian mRNAs as a consequence of splicing[16–18]. Although splicing may be co-transcriptional in many cases, there is evidence that mRNA export can be uncoupled from transcription.

In the *Xenopus* oocyte system, synthetic RNAs injected into the nucleus are properly exported to the cytoplasm. This process depends on functional Yra1p/REF and Sub2p/UAP56, indicating that biologically active export proteins can be loaded onto mRNA post-transcriptionally[18,19]. In yeast, transcripts generated by T7 RNA polymerase are localized to the cytoplasm, which circumvents a role of RNA polymerase II and its C-terminal domain in RNA export[20]. Yet nuclear export of these RNAs still requires Yra1p. This export factor is probably loaded post-transcriptionally onto the synthetic mRNA, because the corresponding gene cannot be immunoprecipitated (ChIP'ed) with antibodies against Yra1p (K. Abruzzi and K. Dower, unpublished data). Therefore, co-transcriptional loading of proteins does not appear to be an obligate requirement for mRNA nuclear export.

We suggest that co-transcriptional assembly of mRNPs allows early monitoring of mRNP quality (Fig. 1). Failure to form a proper mRNP would lead to mRNA retention at the site of transcription and degradation by the transcription machinery-associated exosome. Co-transcriptional assembly therefore occurs primarily so that improperly assembled mRNPs can be discarded during or immediately after transcription.

Transcription is also influenced by mRNA processing events[21], and there is recent evidence that U1 snRNA is a component of the transcription factor TFIIH[22]. Although there is as yet no evidence that improper mRNP formation or the exosome send signals upstream to influence the transcriptional machinery, we view the communication between transcription and nascent RNA processing as another aspect of mRNP quality control. These connections may sometimes make it difficult to determine whether a mutant's primary effect is on transcription or on mRNA processing. Components of the TREX complex were originally described as having effects on transcriptional elongation[23], for example, yet the addition of an exosome mutant to a TREX mutant restores proper levels of RNA 3′ ends[6]. This suggests that in a TREX mutant, improper mRNP was being monitored and degraded by the exosome rather than the gene experiencing transcriptional elongation difficulties.

It is, of course, not known in detail how the exosome recognizes aberrant mRNP complexes or what defines correct or incorrect mRNP formation. The answers to these questions will undoubtedly take us into new neighborhoods and identify new friendships among the relevant proteins.

Torben Heick Jensen is in the Department of Molecular and Structural Biology, University of Aarhus, DK-8000 Aarhus C, Denmark. Michael Rosbash is in the Howard Hughes Medical Institute, Department of Biology, Brandeis University, Waltham, Massachusetts 02454, USA. e-mail: thj@mbio.aau.dk or rosbash@brandeis.edu

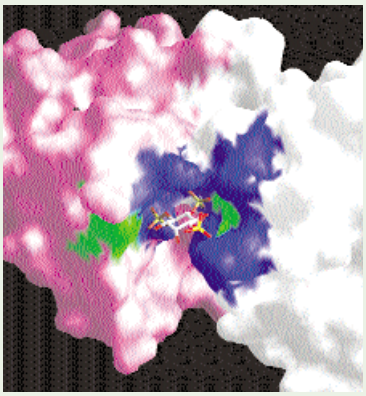
1. Andrulis, E.D., Werner, J., Nazarian, A., Erdjument-Bromage, H., Tempst, P. & Lis, J.T. *Nature* **420**, 837–841 (2002).
2. Andrulis, E.D., Guzman, E., Doring, P., Werner, J. & Lis, J.T. *Genes Dev.* **14**, 2635–2649 (2000).
3. Kaplan, C.D., Morris, J.R., Wu, C. & Winston, F. *Genes Dev.* **14**, 2623–2634 (2000).
4. Hilleren, P., McCarthy, T., Rosbash, M., Parker, R. & Jensen, T.H. *Nature* **413**, 538–542 (2001).
5. Jensen, T.H., Patricio, K., McCarthy, T. & Rosbash, M. *Mol. Cell* **7**, 887–898 (2001).
6. Libri, D., Dower, K., Boulay, J., Thomsen, R., Rosbash, M. & Jensen, T.H. *Mol. Cell. Biol.* **22**, 2254–2266 (2002).
7. Strasser, K. *et al. Nature* **417**, 304–308 (2002).
8. Zenklusen, D., Vinciguerra, P., Wyss, J.C. & Stutz, F. *Mol. Cell. Biol.* **22**, 8241–8253 (2002).
9. Bortvin, A. & Winston, F. *Science* **272**, 1473–1476

(1996).
10. Butler, J.S. *Trends Cell Biol.* **12**, 90–96 (2002).
11. Van Hoof, A. & Parker, R. *Cell* **99**, 347–350 (1999).
12. Bousquet-Antonelli, C., Presutti, C. & Tollervey, D. *Cell* **102**, 765–775 (2000).
13. Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J. & Tollervey, D. *Mol. Cell* **9**, 1285–1296 (2002).
14. Proudfoot, N.J., Furger, A. & Dye, M.J. *Cell* **108**, 501–512 (2002).
15. Lei, E.P. & Silver, P.A. *Genes Dev.* **16**, 2761–2766 (2002).
16. Gatfield, D. *et al. Curr. Biol.* **11**, 1716–1721 (2001).
17. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M.J. *EMBO J.* **20**, 4987–4997 (2001).
18. Luo, M.L. *et al. Nature* **413**, 644–647 (2001).
19. Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M. & Izaurralde, E. *Proc. Natl. Acad. Sci. USA* **98**, 1030–1035 (2001).
20. Dower, K. & Rosbash, M. *RNA* **8**, 686–697 (2002).
21. Manley, J.L. *Nat. Struct. Biol.* **9**, 790–791 (2002).
22. Kwek, K.Y. *et al. Nat. Struct. Biol.* **9**, 800–805 (2002).
23. Chavez, S., Garcia-Rubio, M., Prado, F. & Aguilera, A. *Mol. Cell. Biol.* **21**, 7054–7064 (2001).

From an armadillo to electricity

Electricity drives more than just engines. It is primarily a fundamental feature of life. From single cell organisms to neurons, electrical gradients are used to store energy chemically (as in nucleotide triphosphates, NTPs) or to generate ion fluxes, thus allowing extracellular signals to rapidly transfer from one part of our body (or cell) to another. Ca^{2+} ions play an important role in these forms of cellular signaling and in regulating thousands of downstream targets. One of the molecules that links Ca^{2+} to the function of NTPs in the transduction of extracellular signals, such as growth factors and neurotransmitters, is the intracellular messenger inositol 1,4,5-triphosphate ($InsP_3$). Generated by the enzymatic action of an activated phospholipase on a precursor, $InsP_3$ regulates the release of intracellular Ca^{2+} stores by binding the $InsP_3$ receptor on the surface of the endoplasmic reticulum (ER). Despites decades of research on the ligand and its receptor, a structure had remained elusive.

In a recent issue of *Nature*, Bosanac *et al.* (*Nature* **420**, 696–700; 2002) report the crystal structure of a fragment of the $InsP_3$ receptor corresponding to the $InsP_3$-binding domain. The 381-residue ligand-binding domain is located toward the N terminus of the protein and constitutes a small portion of the 2,749-residue $InsP_3$ receptor. Other structural elements in the receptor moderate the signal transduction process, form a channel for Ca^{2+} release from the ER and reduce $InsP_3$ binding affinity from ~90 pM to ~45 nM. The structure of the high-affinity, core $InsP_3$-binding domain in complex with $InsP_3$ is bipartite (pink and white surface), with an N-terminal region consisting of 12 β-strands and 2 short helices and a C-terminal domain with 8 α-helices. Structural comparisons reveal that the N-terminal region adopts the β-trefoil fold (pink), found in some interleukins and in FGF, whereas the C-terminal domain displays high homology to the armadillo repeat (white) of β-catenin and importin, among others. At the interface of the armadillo and β-trefoil folds lies an $InsP_3$-binding cleft lined with basic residues (blue). The conformation of the $InsP_3$ molecule (stick model) differs in other phosphoinositide-binding proteins, such as the pleckstrin homology and Tubby domains, suggesting that this ligand can be recognized differently.



Nine Arg/Lys residues (blue) coordinate the $InsP_3$ ligand and mediate critical interdomain interactions in the $InsP_3$-binding domain. An additional group of uncharged threonine, glycine and tyrosine residues (green) also engages in ligand recognition. Further, the structure identifies the location of two likely Ca^{2+} ion coordination sites on the β-trefoil–armadillo repeat $InsP_3$-binding domain. Receptor point mutations that eliminate Ca^{2+} binding map to two distinct positions on the domain. The authors hypothesize that the binding of $InsP_3$ to the receptor may facilitate Ca^{2+} binding by relieving the inhibitory function of its N-terminal suppression domain.

This study represents an important milestone in our understanding of these key signal transduction molecules. What it does not yet reveal is precisely how $InsP_3$ causes Ca^{2+} to be released and how ligand binding in the $InsP_3$ receptor appears to be regulated by other receptor domains. No doubt, these questions will soon be answered and increase our molecular understanding of this crucial electrical switch. ***Andreas G. Ladurner***