
RNA surveillance and the exosome

DAVID TOLLERVEY

Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BH, United Kingdom

The discoveries that have had the greatest effect on my work over the past 20 years relate to the nuclear RNA surveillance systems, particularly the exosome complex and its cofactors in budding yeast. These insights were made possible, in large part, by technical progress; in the fields of proteomics, genetics, structural biology, high-throughput sequencing, and RNA-protein crosslinking—and this is typical of many recent advances in RNA biology.

The exosome story started from genetic and biochemical analyses of the yeast ribosome synthesis pathway. Prior to 1994, a small number of ribosome synthesis factors had been identified, but detailed functions were known for very few of these, and none of the processing enzymes had been characterized. In 1994, our genetic and biochemical analyses revealed that the endonuclease RNase MRP acts to provide an entry site for the 5' exonuclease Rat1 (Xrn2), which generates the 5' end of the 5.8S rRNA. Following on from this work, a genetic screen performed in 1995 identified a mutant *rrp4.1* (rRNA processing) that was defective in generating the 3' end of the 5.8S rRNA. The heterogeneous nature of the intermediates seen in the *rrp4.1* strain suggested that, similar to 5' processing, the 3' end of 5.8S was generated by a 3' exonuclease acting from a downstream endonuclease cleavage site. Protein A tags had recently been developed and tagged Rrp4 was shown by Phil Mitchell to co-purify with a 3' exonuclease, supporting this model.

Subsequent work revealed that Rrp4 was the first identified component on the large RNA degradation complex—termed the exosome. Around 1996 several groups were working to develop mass-spectrometry (MS) for protein identification. This key technical advance, allowed the group of Matthias Mann at EMBL to identify five other exosome proteins associated with Rrp4; these were designated as Rrp41, 42, 43, and 44 (also known as Dis3), as well as Mtr3 (mRNA transport) which had been identified by the Tartakof lab via a mutant showing nuclear poly(A) accumulation. With continued improvements in MS technology, further components of the exosome had been identified by 1999; Rrp40, Rrp45, Rrp46, and Csl4. In addition, specific exosome-associated proteins were identified in the nuclear and cytoplasmic complexes,

Rrp6 and Ski7, respectively, which had initially been identified by the Butler and Wickner labs. Together this indicated that a “core exosome” consisting of ten proteins (Rrp4, Rrp40–46, Mtr3, Csl4) was present in both the nucleus and cytoplasm, with Rrp6 and Ski7 specifically associated with the nuclear and cytoplasmic complexes, respectively.

Many exosome subunits appeared to be related to other known nucleases. Biochemical and sequence analyses identified Rrp44 and Rrp6 as hydrolytic 3' exonucleases, that are homologous to *E. coli* RNase R and RNase D, respectively. Subsequently it emerged that Rrp44 also has an additional endonuclease cleavage activity. Six of the exosome subunits were related in sequence to the phosphorolytic 3' exonuclease RNase PH, but it currently appears that none of these are active in the yeast or human exosome. However, the best conserved RNase PH homolog Rrp41 retained its phosphorolytic nuclease activity in *Archaea* and plants.

Biochemical analyses indicated that the exosome had only limited in vitro RNase activity, suggesting that its apparently strong in vivo activity might depend on the presence of activating cofactors. In the late 1990s, the Parker, Johnson, and Takada labs characterized the SKI complex, comprising the DEAH box RNA helicase Ski2 together with the tetratricopeptide repeat (TPR) protein Ski3 and the WD40 repeat protein Ski8. This complex acts together with Ski7 as a cofactor for the cytoplasmic exosome in mRNA turnover and surveillance. Yeast Ski7 shows homology to GTPases that function in translation, suggesting functionally important interaction with the ribosome, but is not clearly conserved to humans.

Around the same time Mtr4, an RNA helicase related to Ski2 was identified as a nuclear exosome cofactor, by the Linder group and ourselves. The identification of these closely related nuclear and cytoplasmic helicases provided a satisfying symmetry. Unexpectedly, however, analyses in the mid 2000s by the Keller, Jacquier, and Anderson groups and ourselves, showed that Mtr4 is not present in a complex similar to SKI. Rather it forms the TRAMP (Trf, Air, Mtr4) polyadenylation complexes, together with a small Zn-finger protein, Air1 or Air2, and a poly(A) polymerase, Trf4 or Trf5, which add short oligo(A) tails to target RNAs. This activity presumably provides a single-stranded “landing pad” that helps

Corresponding author: d.tollervey@ed.ac.uk

Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.050989.115>. Freely available online through the RNA Open Access option.

© 2015 Tollervey This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

initiate RNA degradation by the exosome—as previously shown for *E. coli* poly(A) polymerase. Analyses, by the Jansen and Vanacova groups among others, indicate that nuclear degradation in human cells is even more complicated, since Mtr4 associates with several small RNA binding proteins to generate further exosome cofactor complexes, in addition to TRAMP.

The initial substrates for the exosome were all identified individually by northern hybridization in strains with genetic defects in exosome components or cofactors. These analyses revealed roles in 3' processing of many stable RNAs, functions in nuclear and cytoplasmic RNA turnover, and quality control functions for stable RNA and mRNA precursors. Subsequent analyses, starting in the mid 2000's, made use of tiling microarrays, and then next generation sequencing, to examine transcriptome-wide phenotypes. This work, initially by the Steinmetz, Jansen, Ares, and Jacquier labs, greatly expanded the range of known substrates, identifying very large numbers of short-lived, long noncoding RNAs (lncRNAs) as targets for the TRAMP/exosome system. These can be generated from independent transcription units, or associated with the promoter and enhancer regions of other genes. The TRAMP/exosome surveillance system has therefore emerged as the guardian of the transcriptome, constantly clearing many of the thousands of lncRNAs that are generated in all eukaryotes tested to date. In addition, our transcriptome-wide RNA-protein crosslinking, and analyses of exosome mutants by the Libri lab both indicate high levels of degradation of precursors to stable RNAs, particularly pre-tRNAs. These appear to have surprisingly high rates of rejection by the nuclear quality control systems.

In recent years, structural analyses have transformed our understanding of the exosome and its cofactors. Van Hoof and Parker first proposed that the exosome might resemble the proteasome in having a barrel structure with a central substrate channel and internal degradative active sites. Crystallography, particularly by the Conti and Lima groups, has demonstrated that the exosome core is indeed a barrel through which many RNA substrates pass. In the archaeal exosome, and probably also the plant complex, an internal exonuclease active site lies in Rrp41. Surprisingly, as first reported by the Séraphin lab in 2007, this seems not to be the case for the yeast and human exosomes, which appear to have acquired inactivating point mutations. Despite this, the ge-

ometry of the yeast RNA channel, and the pathway followed by the substrate RNA, is notably similar to the archaeal complex; even including the location of the (former) active site of the Rrp41 component. The exosome core might appear to simply be a tube to conduct RNA from the cofactors to the exonuclease active site of Rrp44 that is localized at the end of the ~31 nt long central channel. However, the high structural conservation strongly suggests that there is more to discover about the functional interactions between the exosome and its many substrate RNAs.

Recent structural analyses have been reported from the Conti group for the nuclear TRAMP and cytoplasmic Ski complexes. These indicate elegant arrangements in which the RNA substrate passes through a loop structure in the helicase in the complex (Mtr4 or Ski2, respectively), before being channeled into the lumen of the exosome. These helicase presumably help remove secondary structure and bound proteins from substrate RNAs. This is particularly important because only single stranded RNA molecules can enter the lumen of the exosome.

Concluding remarks

During the past 20 years it has become apparent that RNA quality control and degradation are pervasive in Eukaryotes. All classes of transcripts examined are subject to surveillance systems, and an extraordinary range of unstable, non-protein coding RNAs are synthesized and degraded. In many of these pathways, the exosome and its cofactors are key players. Work by many groups, only a few of which can be mentioned here, have given us a good idea of the geometry and substrate range of the yeast and human exosome complexes. However, there remain many outstanding questions, which largely relate to the difficulty of determining how the hugely complex range of highly diverse substrates interact with the exosome and its cofactors in living cells. The exosome has at least three nuclease activities and multiple pathways to the corresponding active sites, but it is unclear what substrates normally follow each pathway. In addition, the consequence of exosome activity on its diverse substrates can be quite different—frequently resulting in either complete and very rapid degradation, or precise processing to generate a well defined, stable product—and it remains unclear what features confers this specificity. So there is still plenty of work to be done.