Musing on the structural organization of the exosome complex

Philip Mitchell and David Tollervey

The exosome complex of 3′→5′ exoribonucleases functions in both the precise processing of 3′ extended precursor molecules to mature stable RNAs and the complete degradation of other RNAs. Both processing and degradative activities of the exosome depend on additional cofactors, notably the putative RNA helicases Mtr4p and Ski2p. It is not known how these factors regulate exosome function or how the exosome distinguishes RNAs destined for processing events from substrates that are to be completely degraded. Here we review the available data concerning the modes of action of the exosome and relate these to possible structural arrangements for the complex. As no detailed structural data are yet available for the exosome complex, or any of its constituent enzymes, this discussion will rely heavily on rather speculative models.

A universal feature of stable RNAs is that they are transcribed as 3′ extended precursor molecules that undergo subsequent processing events, generally exonuclease digestion, to generate the 3′ end of the mature molecule. Conversely, RNAs such as cytoplasmic mRNAs, nuclear pre-mRNA and the RNA spacer fragments generated during processing reactions can also be shortened from the 3′ end but these molecules are rapidly and completely degraded. Notably, under circumstances such as the formation of incorrectly assembled ribonucleoprotein (RNP) particles, a precursor RNA that is usually processed to the mature RNA may become a substrate for rapid degradation. Recent data from genetic analyses in the yeast *Saccharomyces cerevisiae* have shown that many of the enzymes that function in RNA 3′ end maturation processes also function in RNA degradation pathways. The regulation of the activities of these enzymes is therefore of great importance.

A major group of RNA processing/degradative enzymes consists of the 3′→5′ exoribonucleases, ribonucleases that progressively digest RNAs specifically from the 3′ end. The genome of *S. cerevisiae* is predicted to encode at least 19 3′→5′ exoribonucleases[1–4], most of which have one or more assigned functions in RNA processing. In contrast to *Escherichia coli*, in which the eight known 3′→5′ exoribonucleases function independently of one another[5], many of the eukaryotic 3′→5′ exoribonucleases form a multienzyme complex known as the exosome[3,4] (Table 1). In yeast, this complex contains all of the 3′→5′ exoribonucleases that are known to be essential for viability, suggesting that it plays a key role in RNA 3′ end maturation and turnover.

Structural models for the exosome

Two related exosome complexes have been identified that are believed to represent cytoplasmic and nuclear forms[4]. In yeast, 10 proteins are present in both the nuclear and cytoplasmic forms of the complex. Three of these have been demonstrated to have 3′→5′ exoribonuclease activity *in vitro*[3] and a further six have high sequence homology to characterized 3′→5′ exoribonucleases (Table 1). Although not formally demonstrated, it therefore seems highly plausible that 9 of the 10 proteins are exoribonucleases. The exception is Csl4p/Ski4p, which is not known to have exonuclease activity and is not clearly homologous to any characterized exonuclease, but does not contain a predicted RNA binding domain similar to that present in the *E. coli* ribosomal protein S1 (S1 RBD). The nuclear exosome includes an additional component, the 3′→5′ exoribonuclease Rrp6p, which is homologous to *E. coli* RNase D[4,6]. Strikingly, neither complex has obvious structural or regulatory subunits. They do, however, include enzymes with distinctly different activities (Table 1). Recombinant Rrp44p and Rrp6p, like their *E.coli* homologs RNase R and RNase D, are hydrolytic enzymes, as is Rrp4p. These nucleases use water as the attacking group, releasing nucleotide 5′ monophosphates. In contrast, Rrp41p is a phosphorolytic enzyme, as is *E.coli* RNase PH, using inorganic phosphate as the attacking group and releasing nucleotide 5′ diphosphates. Moreover, Rrp4p is distributive, whereas Rrp41p and Rrp44p are processive[3]. Once bound, processive nucleases remain associated with the RNA and digest the substrate to completion. Conversely, distributive nucleases repeatedly bind and dissociate from the substrate, removing nucleotides one at a time.

The yeast cytoplasmic complex (lacking Rrp6p) has a predicted mass of 401 kDa based on the known components and has an experimentally determined sedimentation coefficient of 14S (ref. 3). Fractionation experiments suggest that the exosome has a very stable 'core' consisting of the six RNase PH homologs and the three S1 RBD-containing proteins Rrp4p, Rrp40p and Csl4p. These proteins remain stably associated with one another at Mg^{2+} concentrations of up to 1.6 M upon Mg^{2+} concentration gradient elution of immunoaffinity-purified exosome preparations, independent of the epitope-tagged protein[4] (P.M. & D.T., unpublished results). Rrp44p is less tightly bound and can be eluted at a moderate Mg^{2+} concentration of ~0.6 M.

The yeast nuclear exosome processes 3′ extended precursors of the 5.8S ribosomal RNA (rRNA), the U1, U2, U4 and U5 small nuclear RNAs (snRNAs) and many small nucleolar RNAs (snoRNAs), as well as degrading unspliced pre-mRNAs and pre-rRNA spacer regions[4,7–12]. The cytoplasmic exosome degrades mRNAs in a 3′→5′ direction following deadenylation[13]. Genetic depletion of any individual exosome component inhibits all nuclear activities of the complex. Furthermore, cytoplasmic 3′→5′ mRNA degradation is inhibited by mutation of any tested component of the cytoplasmic exosome but not by the absence of Rrp6p, which is specific for the nuclear complex[11] (P.M. & D.T.,

Wellcome Trust Centre for Cell Biology, ICMB, University of Edinburgh, Kings' Buildings, Edinburgh EH9 3JR, UK.

Correspondence should be sent to D.T. *email: d.tollervey@ed.ac.uk* or P.M. *email: pmitch@holyrood.ed.ac.uk*

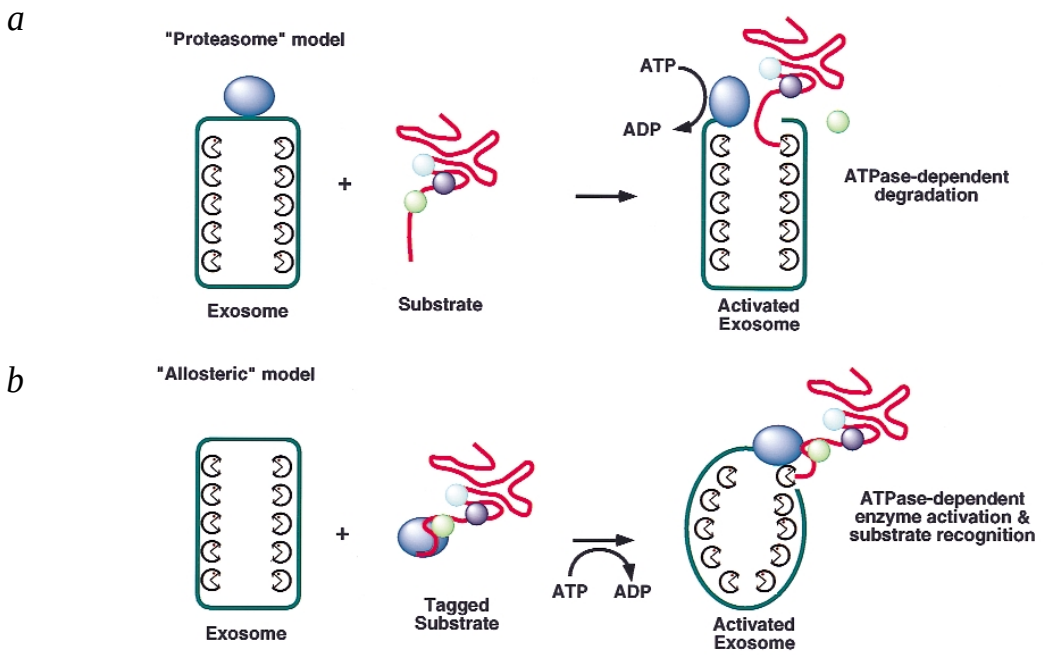

Fig. 1 Models for the activation of the exosome. ***a***, Proteasome model. Access to the active sites of the exonucleases in the exosome is regulated by the RNA helicase Mtr4p or Ski2p (blue oval). Displacement or reorganization of the helicase occurs upon interaction with the RNP substrate. ATP hydrolysis by the activated RNA helicase unfolds the RNA structure and channels the free 3′ end into the lumen of the exosome. ***b***, Allosteric model. The RNA helicase interacts directly with the RNP substrate, recognizing specific marker proteins. The ATPase activity of the RNA helicase allows the exosome to be remodeled into an appropriately active form. The helicase may interact with the RNP substrate, either associated with the exosome or in its absence; for simplicity only the latter case is shown.

unpublished results). These results strongly suggest that all the substrates are processed by intact exosome complexes, rather than by individual enzymes. Notably, all of the characterized exosome substrates are RNP complexes and not naked RNAs.

Clear homologs of the yeast exosome proteins are present in other eukaryotes, including humans[4,14] (Table 1) and are organized in a comparable complex that also has associated exonuclease activity (R. Brouwer, C. Allmang, R. Raijmakers, Y. van Aarssen, W. Vree Egberts, E. Petfalski, W.J. van Venrooij, D. T. & G. Pruij, unpublished results). All tested human homologs are present in nuclear and cytoplasmic complexes that are recognized by autoimmune antibodies present in patients with polymyositis-scleroderma overlap syndrome (PM-Scl). As in yeast, human nuclear and cytoplasmic forms exist which differ by the presence of PM-Scl100 (the homolog of Rrp6p) specifically in the nuclear complex. Four human homologs tested have been shown to be functional when expressed in yeast, complementing mutations in the corresponding yeast genes[3,15,16] (R. Brouwer, C. Allmang, R. Raijmakers, Y. van Aarssen, W. Vree Egberts, E. Petfalski, W.J. van Venrooij, D. T. & G. Pruij, unpublished results; Table 1). The structural organization of the complex must therefore be sufficiently conserved to allow incorporation of the human homologs into the yeast exosome.

Two characteristics of the exosome are of particular significance for potential structural models. Firstly, although the individual recombinant proteins are active *in vitro*, the exosome complex purified from yeast lysates exhibits little activity[3]. This suggests that the assembly of the ribonucleases into the exosome complex largely represses their enzymatic activities. Activation of the exosome would therefore be expected to require additional factors that are not stably associated with the complex. Consistent with this model, genetic analyses have identified proteins that are required for exosome-mediated RNA processing reactions but that are not found in purified exosome preparations. These potential exosome cofactors include the putative ATP-dependent RNA helicases Mtr4p and Ski2p, the putative GTPase Ski7p and the Ski3p and Ski8p proteins[12,13] (A. van Hoof & R. Parker, pers. comm.; Table 1). Secondly, different exosome substrates have very different fates. Precursors to mature RNAs, such as rRNAs, snRNAs and snoRNAs, undergo limited processing reactions that remove nucleotides from the end of the RNA, thereby generating products with discrete 3′ ends. Other substrates, including pre-rRNA spacer fragments, aberrant pre-rRNA processing intermediates, nuclear pre-mRNAs and cytoplasmic mRNAs, are rapidly and processively degraded to completion. Moreover, at least for pre-mRNAs, this degradative activity is regulated[10].

Any model for the organization of the exosome should therefore explain: (i) how 10 exonucleases can be maintained simultaneously in an inactive state; (ii) how each can be activated and gain access to macromolecular substrates; and (iii) how different substrates (or indeed the same substrate under different circumstances) can be subjected to either limited, accurate digestion or complete, processive degradation.

The proteasome analogy

One model recently proposed for the exosome[14] is analogous to the structural arrangement of the 20S proteasome complex (recently reviewed by Baumeister *et al.*[17] and DeMartino and Slaughter[18]). In this model, the active sites of the enzymes are directed towards the hollow center of a tubular structure and are therefore only accessible to substrates that enter this structure (Fig. 1*a*). The function of the Mtr4p and Ski2p helicases would then be directly analogous to the ATPases of the proteasome regulatory subunit, which are thought to unfold the substrate and channel the polypeptide chain into the degradative complex[19].

Table 1 Exosome components and cofactors

Protein	Gene	Activity[1]	Motifs/homologs	Deletion phenotype	Human homolog[2]	Comments
Rrp4p	YHR069c	Hydrolytic, distributive 3′→5′ exonuclease	S1 RBD	Essential	hRrp4p (43%)	hRrp4p complements rrp4-1
Rrp40p	YOL142w	(Hydrolytic, distributive 3′→5′ exonuclease)	S1 RBD	Essential	hRrp40p (30%)	Homologous to Rrp4p
Csl4p/Ski4p csl4-1	YNL232w	?	S1 RBD	Essential	hCsl4p (48%)	hCsl4p complements
Rrp41p/Ski6p	YGR195w	Phosphorolytic, processive 3′→5′ exonuclease	RNase PH	Essential	hRrp41p (35%)	hRrp41p complements GAL:rrp41
Rrp42p	YDL111c	(Phosphorolytic, processive) 3′→5′ exonuclease	RNase PH	Essential		
Rrp43p	YCR035c	(Phosphorolytic, processive 3′→5′ exonuclease)	RNase PH	Essential		
Rrp45p	YDR280w	(Phosphorolytic, processive 3′→5′ exonuclease)	RNase PH	Essential	PM-Scl 75 (38%)	Human KIAA0116 and OIP2 also homologous
Rrp46p	YGR095c	(Phosphorolytic, processive 3′→5′ exonuclease)	RNase PH	Essential	hRrp46p (26%)	
Mtr3p	YGR158c	(Phosphorolytic, processive)	RNase PH	Essential		
Rrp44p/Dis3p	YOL021c	Hydrolytic, processive 3′→5′ exonuclease	RNase R (vacB)	Essential	hDis3p (45%)	hDis3p complements dis3-81
Rrp6p	YOR001w	hydrolytic 3′→5′ exonuclease	RNase D	ts-lethal	PM-Scl 100 (32%)	Component only of nuclear complex
Cofactors						
Mtr4p/Dob1p	YJL050W	(ATP-dependant helicase)	DEAD box	Essential		
Ski2p	YLR398C	(ATP-dependant helicase)	DEAD box	Nonessential	SKIV2L (38%)	
Ski3p	YPR189W		TPR repeat	Nonessential	KIAA0372 (20%)	
Ski7p	YOR076c	(GTPase)		Nonessential		High similarity to Hbs1p
Ski8p	YGL213C		WD repeat	Nonessential		

[1]Demonstrated enzymatic activities are listed; activities in parentheses are predicted from sequence homology.
[2]For the human homologs, the numbers given indicate the percentage identity along the entire length of the protein.

There are, however, several significant differences between the two complexes. The proteasome can be isolated from cell extracts in association with the regulatory ATPase complex, whereas the exosome is not found as a stable complex with either Mtr4p or Ski2p. Furthermore, a key functional difference between the two complexes is that while both degrade many diverse substrates in an apparently processive manner, only the exosome generates precise ends of substrates *via* a clearly exolytic processing mechanism. The proteasome is required for the synthesis of at least one polypeptide, the p50 product of the NF-κB1 gene[20], which corresponds to the N-terminal portion of a functionally distinct protein, p105. However, the mechanism of NF-κB p50 production is unclear[21,22] and convincing evidence exists, at least in mammalian cells, for an initial endoproteolytic cleavage event[23]. Thus, while the structural analogy of the exosome to the proteasome works well for RNA degradation pathways, it is clearly less compatible with accurate 3′ end formation in RNA processing events. Furthermore, it is hard to envisage how such an arrangement would allow the exosome to degrade RNAs in a distributive manner, as is observed *in vitro* with the exosome complex purified from cell lysates or with recombinant Rrp4p[3,24].

An allosteric alternative

An alternative model is that the enzymes are regulated by allosteric mechanisms. In this case, the degradation or processing of a given RNA substrate would be achieved by the selective activation of one or more distinct enzymes upon substrate recognition (Fig. 1*b*). This concept envisages a more fluid structure for the exosome, with the potential to adopt a number of different activated forms that could perform distinct activities, depending upon which enzymes are 'exposed'.

Such an allosteric mechanism would permit both the tight regulation of ribonuclease activity and the versatility of nucleolytic activities suggested by *in vivo* analyses. Data supporting an allosteric mode of activation for the exosome come from *in vitro* studies on the purified complex (P.M. & D.T., unpublished results). Enzymatic analysis of fractionated cell extracts revealed that the majority of exosome activity is associated with only a small fraction (~1%) of the total complex population. This highly active fraction is itself quite heterogeneous; several distinct exonucleolytic activities were observed, even on a short unstructured substrate, which could be partially resolved using biochemical fractionation procedures. No clear differences, however, were observed in the protein compositions of these active fractions and the bulk exosome population, strongly suggesting that the active forms correspond to complexes that are arranged in distinct structural conformations. These active fractions presumably represent complexes that were trapped in a given conformation upon cell lysis. It is unclear whether they represent complexes that had been activated but had not yet initiated processing, complexes that had terminated processing but had not yet returned to the latent state, or complexes that were actively engaged in processing.

Energetics and substrate recognition

Considerable free energy is released during RNA folding and RNP particle assembly processes, and the unfolding of these structures presumably entails a substantial energetic cost. What sources of energy are available to the exosome to facili-

tate the unfolding and processing or degradation of its substrates?

The equilibrium constants for RNA degradation by the phosphorolytic enzymes RNase PH or polynucleotide phosphorylase (PNPase) is close to unity[22,23], the reaction being freely reversible *in vitro*. The low ΔG for this reaction is because the free energy released upon disruption of the phosphodiester linkage is largely recouped by generating the nucleoside diphosphate product. Little free energy is therefore available for phosphorolytic nucleases to unfold the RNA substrate. In the *E. coli* degradosome, PNPase is associated with the endonuclease RNase E and the ATP-dependent RNA helicase RhlB[27,28]. Consistent with the predicted energetics, phosphorolytic degradation of structured RNA substrates by the degradosome is dependent upon the activity of RhlB[27]. In contrast, substantial free energy is released by RNA hydrolysis and this energy can be utilized by any hydrolytic exonuclease that is able to couple RNA degradation to unfolding of the RNA structure. Notably, the Rrp44p-related 3′→5′ exoribonuclease RNase R (also termed vacB) and the hydrolytic 5′→3′ exoribonucleases Xrn1p and Rat1p exhibit substantial activity *in vitro* on structured RNAs such as rRNA and tRNA in the absence of an associated RNA helicase[29–31]. In fact, the energy made available to an exonuclease upon hydrolysis of the phosphodiester backbone within RNA is comparable to that generated by an RNA helicase utilizing one molecule of ATP per nucleotide.

It follows that the ability of the exosome to penetrate regions of secondary structure and displace bound proteins may be strongly dependent on which of the nuclease activities has been activated. Complexes with an activated processive nuclease, such as Rrp44p, may be able to degrade even highly structured substrates, as would be necessary for the rapid degradation of aberrant pre-ribosomes or pre-mRNAs. In contrast, complexes in which a phosphorolytic enzyme is activated would be predicted to degrade relatively unstructured regions but would be blocked by stem loop structures, particularly if these were stabilized by bound proteins. Such an activity is needed for the processing of the 3′ extended precursors to mature snRNAs and snoRNAs. In the case of distributive enzymes, such as Rrp4p, there is no obvious mechanism for coupling hydrolysis to unfolding RNA structure and these activities may be limited to trimming short unstructured 3′ extensions. The fate of exosome substrates would therefore be expected to depend on which component has been activated, a structure-sensitive phosphorolytic enzyme for RNA processing or a structure-insensitive hydrolytic enzyme for RNA degradation.

A key feature of allosteric models for exosome function is that the recognition of a particular substrate induces activation of the appropriate nuclease activity(s). One possible mechanism envisages a modular system in which specific cofactors function as exosome 'adaptors', allowing the complex to dock onto a given class of substrate[13]. A 'one substrate, one adaptor' rule would in principle allow highly specific substrate processing/degradation but would require substantially more cofactors than have been identified to date. Moreover, Mtr4p, the only known nuclear cofactor, appears to be required for both processing and degradation of diverse substrates.

Alternatively, the exosome may interact directly with the RNP structure of its substrates. All exosome substrates characterized so far have undergone prior RNA processing events and they may well be 'tagged' with marker proteins as a result of such processing. This would be similar to current models for the degradation of mRNAs with retained introns by RNA surveillance mechanisms, where proteins that remain associated with the mRNA after pre-mRNA processing are thought to be subsequently recognized by the surveillance complex and trigger rapid degradation of the RNA[32–34]. One possibility is that Mtr4p (or Ski2p in the case of cytoplasmic mRNAs) directly recognizes both processing and degradation substrates, either before or after association with the exosome. Interactions between Mtr4p or Ski2p and the RNP proteins of the substrate would then determine which exosome component, or components, are activated (Fig. 1*b*). Such an allosteric activation could conceivably be driven by the ATPase activities of Mtr4p or Ski2p.

Perspectives

The RNase complement of the exosome constitutes a remarkably powerful RNA degradation machine. The genetic analysis of mutant alleles in yeast has been central in defining the functions of the complex and identifying its cofactors in RNA processing pathways. A major limitation to this approach lies in the identification of the specific roles of individual components in cooperative complexes. Similar difficulties attend other such analyses, for example in identifying the endonuclease component of the pre-mRNA cleavage and polyadenylation machinery. Future work involving a combination of structural studies on the exosome and the development of suitable *in vitro* RNA processing systems should reveal details of the molecular mechanisms of exosome mediated RNA degradation.

Acknowledgments

We would like to thank A. van Hoof and R. Parker for sharing data prior to publication. This work is funded by the Wellcome Trust.

Received 28 July, 2000; accepted 23 August, 2000.

1. Mian, S. *Nucleic Acids Res.* **25**, 3187–3195 (1997).
2. Moser, M.J., Holley, W.R., Chatterjee, A. & Mian, S. *Nucleic Acids Res.* **25**, 5110–5118 (1997).
3. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. *Cell* **91**, 457–466 (1997).
4. Allmang, C. *et al. Genes Dev.* **13**, 2148–2158 (1999).
5. Deutscher, M.P. *J. Bacteriol.* **175**, 4577–4583 (1993).
6. Burkard, K.T.D. & Butler, J.S. *Mol. Cell. Biol.* **20**, 604–616 (2000).
7. Allmang, C., *et al. EMBO J.* **18**, 5399–5410 (1999).
8. Allmang, C., Mitchell, P., Petfalski, E. & Tollervey, D. *Nucleic Acids Res.* **28**, 1684–1691 (2000).
9. Kufel, J. *et al. Mol. Cell. Biol.* **20**, 5415–5424 (2000).
10. Bousquet-Antonelli, C., Presutti, C. & Tollervey, D. *Cell*, **in the press** (2000).
11. van Hoof, A., Lennertz, P. & Parker, R. *Mol. Cell. Biol.* **20**, 441–452 (2000).
12. de la Cruz, J., Kressler, D., Tollervey, D. & Linder, P. *EMBO J.* **17**, 1128–1140 (1998).
13. Jacobs Anderson, J.S. & Parker, R. *EMBO J.* **17**, 1497–1506 (1998).
14. van Hoof, A. & Parker, R. *Cell* **99**, 347–350 (1999).
15. Baker, R., Harris, K. & Zhang, K. *Genetics* **149**, 73–85 (1998).
16. Shiomi, T. *et al. J. Biochem.* **123**, 883–890 (1998).
17. Baumeister, W., Walz, J., Zühl, F. & Seemüller, E. *Cell* **92**, 367–380 (1998).
18. DeMartino, G.N. & Slaughter, C.A. *J. Biol. Chem.* **274**, 22123–22126 (1999).
19. Rubin, D.M., Glickman, M.H., Larsen, C.N., Dhruvakumar, S. & Finley, D. *EMBO J.* **17**, 4909–4919 (1998).
20. Palombella, V.J., Rando, O.J., Goldberg, A.L. & Maniatis, T. *Cell* **78**, 773–785 (1994).
21. Sears, C., Olesen, J., Rubin, D., Finley, D. & Maniatis, T. *J. Biol. Chem.* **273**, 1409–1419 (1998).
22. Lin, L., DeMartino, G.N. & Greene, W.C. *Cell* **92**, 819–828 (1998).
23. Lin, L. & Ghosh, S. *Mol. Cell. Biol.* **16**, 2248–2254 (1996).
24. Mitchell, P., Petfalski, E. & Tollervey, D. *Genes Dev.* **10**, 502–513 (1996).
25. Ost Kelly, K. & Deutscher, M.P. *J. Biol. Chem.* **267**, 17153–17158 (1992).
26. Grunberg-Manago, M., Ortiz, P.J. & Ochoa, S. *Science* **122**, 907–910 (1955).
27. Py, B., Higgins, C.F., Krisch, H.M. & Carpousis, A.J. *Nature* **381**, 169–172 (1996).
28. Miczak, A., Kaberdin, V. R., Wei, C.-L. & Lin-Chao, S. *Proc. Natl. Acad. Sci. USA* **93**, 3865–3869 (1996).
29. Cheng, Z.-F., Zuo, Y., Li, Z., Rudd, K.E. & Deutscher, M.P. *J. Biol. Chem.* **272**, 14077–14080 (1998).
30. Stevens, A. *J. Biol. Chem.* **255**, 3080–3085 (1979).
31. Kenna, M., Stevens, A., McCammon, M. & Douglas, M.G. *Mol. Cell. Biol.* **13**, 341–350 (1993).
32. Hilleren, P. & Parker, R. *RNA* **5**, 711–719 (1999).
33. González, C.I., Ruiz-Echevarria, M.J., Vasudevan, S., Henry, M.F. & Peltz, S.W. *Mol. Cell* **5**, 489–499 (2000).
34. Le Hir, H., Moore, M.J. & Maquat, L.E. *Genes Dev.* **14**, 1098–1108 (2000).