Catalytic properties of the eukaryotic exosome

Aleksander Chlebowski, MSc

Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
Institute of Biochemistry and Biophysics, Polish Academy of Sciences
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
olobiolo@gmail.com

Rafał Tomecki, PhD

Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
Institute of Biochemistry and Biophysics, Polish Academy of Sciences
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
rtomecki@yahoo.com

María Eugenia Gas López , PhD

Institute of Genetics and Molecular and Cellular Biology
B.P. 10142, 67404 ILLKIRCH CEDEX, FRANCE
mgas@igbmc.fr

Bertrand Séraphin, PhD

Institute of Genetics and Molecular and Cellular Biology
B.P. 10142, 67404 ILLKIRCH CEDEX, FRANCE
seraphin@igbmc.fr

Andrzej Dziembowski, PhD

Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
Institute of Biochemistry and Biophysics, Polish Academy of Sciences
ul. Pawińskiego 5a, 02-106 Warszawa, POLAND
andrzejd@ibb.waw.pl

Abstract

The eukaryotic exosome complex is built around the backbone of a 9-subunit ring similar to phosporolytic ribonucleases such as RNase PH and polynucleotide phosphorylase (PNPase). Unlike those enzymes, the ring is devoid of any detectable catalytic activities, with the possible exception of the plant version of the complex. Instead, the essential RNA decay capability is supplied by associated hydrolytic ribonucleases belonging to the Dis3 and Rrp6 families. Dis3 proteins are endowed with two different activates: the long known procesive 3'-5' exonucleolytic one and the recently discovered endonucleolytic one. Rrp6 proteins are distributive exonucleases. This chapter will review the current knowledge about the catalytic properties of theses nucleases and their interplay within the exosome holocomplex.

Keywords

exosome, exosome ring, RNA degradation, hydrolytic, phosphorolytic, processive, distributive, ribonuclease, PIN domain, RNB domain, catalytic activity

Introduction

When the exosome was first described, it was as a "multienzyme ribonuclease complex" and three subunits of the yeast complex, Rrp4, Rrp41 and Dis3, were shown to be active ribonucleases in vitro.[1] The study demonstrated three different RNase activities within the complex: a distributive hydrolytic one for Rrp4, a processive phosphorolytic one for Rrp41, and a processive hydrolytic one for Dis3. Notably, the three modes of RNA degradation were ascribed to proteins that are vastly different: Rrp41 is similar to the bacterial phosphorolytic enzyme, RNase PH, Dis3 is a member of the RNR family of hydrolytic ribonucleases, whereas Rrp4 contains no domains with known RNase activity. The latter led to Rrp4 becoming the founding member of a new superfamily of ribonucleases.[2]

As more components of the exosome were discovered, they were assumed to be ribonucleases based on sequence similarity to Rrp4 and Rrp41.[3] Expectations grew for the complex to be a very complicated machine with numerous tools for the destruction of RNAs and speculations began about how the many activities are spatially organized, selected for action, and coordinated. Structurally, the exact protein stoichiometry was still unknown, it was even unclear, whether the exosome population is homogenous or not, that is whether all exosome particles share the same subunit composition. And if not, whether the composition of particular complexes is constant or dynamic.

The first clues to the structure of the exosome came relatively early, from analyses of archaeal genomes.[4] It was revealed that most of them encode only two proteins homologous to RNase PH, while no less than six such proteins are encoded in eukaryotic genomes. It was then proposed that, as in the case of the proteasome, the same multimeric structure is assembled using either multiple copies of two proteins or individual copies of six different ones. This pointed to the exosome consisting of, among others, six proteins with the RNase PH fold and three containing the S1 domain or multiples thereof.

RNase PH and S1 domains with a 6:3 stoichiometry were also found in the trimeric structure of *Streptomyces antibioticus* PNPase[5] and thus the yeast exosome was predicted to assume a related architecture, which was confirmed by electron-microscopy studies.[6] These observations were followed by a steady flow of structural data on archaeal and yeast exosomes, strengthening the notion that the exosome is

homogenously built around a structure of defined composition.[7-9] Finally the crystal structure of the human ring complex was solved, giving the first high-resolution view of a 9-subunit exosome complex.[10] Recently a crystal structure was also obtained for yeast Dis3 bound to some ring subunits.[11]

As the structural characterization of the exosome progressed, knowledge about its catalytic properties was revolutionized. Solving the structure of the human ring complex[10] coincided with the contributions of particular subunits to exosome activity being carefully reevaluated.[12] Both studies showed beyond doubt that the 9-subunit ring complex is catalytically inactive in yeast and human. The current view of the exosome is thus quite different from the initial one: only a few proteins in the complex are active RNases. The precise function of the remaining majority is elusive but their involvement in substrate recruitment is often suggested, and a recent study indicates a role in substrate selection.[11]

Exosome composition

The eukaryotic exosome is composed of a ring-like structure and accessory subunits.[*] The ring consists of nine proteins: six composed of a single domain homologous to the phosphorolytic enzyme RNase PH, and three containing RNA binding domains: S1 and KH. The RNase PH subunits form a toroidal hexamer with the RNA binding subunits residing on one side (considered to be the "top"), and a channel running along the pseudosymmetry axis of the complex.[10] The RNase PH-like subunits are: Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and Mtr3, and the RNA binding subunits are: Rrp4, Rrp40 and Csl4.

Despite extensive similarity to RNase PH all of the ring subunits are catalytically inactive in yeast and human, having lost the key catalytic amino acid residues of the bacterial enzyme.[10, 12] The enzymatic activities detected previously in Rrp4 and Rrp41

---

[*] *The name "exosome core" is used to describe several structural elements. Most commonly, it refers to the 9-subunit complex, archaeal or eukaryotic, but has also been used to mean the hexameric ring of RNase PH subunits of Archaea, as well as the 10-subunit complex of eukaryotes, encompassing the 9-subunit ring and Dis3. To avoid confusion we shall refrain from using "exosome core" throughout the chapter, and instead use the term "exosome ring" to denote the 9-subunit complex of six RNase PH subunits and three S1/KH subunits.*

were apparently artefactual. Of all eukaryotic exosome ring proteins studied so far, only plant Rrp41 homologues seem capable of catalysis. Indeed, Rrp41 of *Arabidopsis thaliana* has been convincingly shown to have phosphorolytic activity,[13] indicating that exosome mediated RNA turnover pathways in plants may differ substantially from those of other organisms. This is covered in detail in the chapter *The plant exosome* by Heike Lange and Dominique Gagliardi.

The exosome ring associates with additional subunits that supply catalytic activity to the complex. In yeast these are the essential Dis3, present both in the nucleus and the cytoplasm and the non-essential Rrp6, which is found only in the nucleus.[14] Thus, the yeast exosome exists in two forms: the 10-subunit cytoplasmic and the 11-subunit nuclear form, both of which interact with more factors and protein complexes in their respective compartments. This model of exosome composition was thought to be ubiquitous among eukaryotes until some recent studies have challenged this assumption, as more catalytic exosome proteins were discovered.

The human genome encodes two proteins of the Dis3 family, named DIS3 and DIS3-like (DIS3L), and the same appears to be true for all vertebrates. DIS3, the closest homologue of yeast Dis3, is localized mostly, if not exclusively, in the nucleus, while DIS3L is strictly cytoplasmic. It would seem the nuclear and cytoplasmic functions of yeast Dis3 have been divided between two proteins in vertebrates. Notably, in human cells DIS3 is absent from nucleoli, unlike the yeast protein, which is enriched in this compartment.[15-16] Interestingly, the single human Rrp6 protein is found not only in the nucleus (with nucleolar enrichment typical of exosome ring proteins) but also in the cytoplasm, unlike its yeast counterpart.[17-18]

In plants on the other hand, while only one protein of the Dis3 family is encoded in currently known plant genomes, three genes of the *RRP6* family are present in *Arabidopsis*, rice and poplar, encoding the proteins AtRRP6L1, AtRRP6L2 and AtRRP6L3 in *Arabidopsis*. Like the vertebrate Dis3 paralogues, the plant Rrp6 proteins are differentially localized in the cell: AtRRP6L1 and AtRRP6L2 were both found to localize in the nucleus, although with different distributions, whereas AtRRP6L3 was exclusively cytoplasmic.[19] The subcellular localization of Dis3 in plants has not been determined. Neither have physical interactions between the plant exosome ring and any of the putative catalytic subunits been observed so far.[20]

Thus the interplay between the catalytic subunits of the exosome may be quite different between eukaryotic taxa. The different isoforms of the exosome are depicted in Figure 1.

Notably, genomes of eukaryotes, including plant and vertebrate, do seem to encode other proteins of the RNR family than Dis3, but these are unlikely to physically interact with the exosome, given their lack of the PIN domain (see below). Their potential role in RNA metabolism notwithstanding, they are not considered to be exosome proteins.

Dis3

Dis3 proteins are highly conserved in all eukaryotes and although there are differences at the sequence level, in terms of domain organization they are all identical. The generic Dis3 protein is a representative of the RNR superfamily of ribonucleases, with similarity to bacterial RNase II and RNase R. Like the bacterial enzymes, it consists of three RNA binding domains (two cold-shock (CSD) and one S1) and one catalytic domain (RNB). However, unlike the bacterial RNases, Dis3 additionally has a long N-terminal region containing a PIN domain that is responsible for stable association of the protein with the exosome ring. This is absolutely crucial for Dis3 to function as part of the exosome complex.[11, 21] The domain composition of Dis3 is shown in Figure 2.

Enzymes of the RNR family all exhibit hydrolytic activity against RNA, which they digest processively in the 3'-5' direction. When presented with a substrate, they release nucleoside 5'-monophosphates and end products: short oligonucleotides, 2-5 nt in length, depending on the particular enzyme.[12, 22] However, they can have quite different substrate specificities. For instance, RNase II acts only on single-stranded RNA and stops a few nucleotides before structured regions, RNase R on the other hand is able to digest through both intra- or intermolecular secondary RNA structures, provided a single-stranded region at the 3' end is available for the protein to bind.[22] All RNR proteins require the presence of divalent metal ions for catalysis, preferably magnesium.

RNB is a well characterized domain. Crystal structures of substrate-bound *Escherichia coli* RNase II and *Saccharomyces cerevisiae* Dis3 reveal the structural basis for magnesium ion-dependent hydrolytic RNA cleavage as well as for the

substrate specificities of the enzymes. The active site lies buried deep within the domain, and is connected to the surface by a narrow channel that can only accommodate single-stranded substrates. Two magnesium ions are coordinated in the active site by conserved aspartic acid residues and mutation of one of those residues results in complete loss of catalytic activity.[23-24]

The PIN domain is found in proteins of different functions and was initially associated with signaling pathways.[25] but some PIN domains were later predicted and shown to have nuclease activity.[26-27] Cleavage of RNA by the PIN domains in vitro requires the presence of divalent metal ions, usually manganese.[28]

Crystal structures of different PIN domains show an RNase H-like fold with the catalytic site placed in a ridge running along the surface of the domain.[26] Metal ions are coordinated in the active site by four conserved acidic amino acid residues and mutation of any of these residues abolishes RNase activity.[26, 28] Interestingly, some PIN domains act as oligomers with different catalytic properties.[26]

Exonuclease activity

Since the first demonstration that yeast Dis3 is an exoribonuclease[1] the exonucleolytic activity of Dis3 has been exhaustively assayed and while the majority of studies concerned the activity of the yeast protein, those of the *Drosophila melanogaster* and human homologues have also been described.[12, 17-18, 29] As expected of an RNR protein, Dis3 is a processive 3'-5' exoribonuclease that requires divalent metal ions, releases 5'-mononucleotides and leaves a short end product.

Maximum activity of Dis3 in vitro is observed for unusually low concentrations of magnesium ions, about 50 μM.[12] This is a little puzzling, given that the intracellular concentrations of this cation are more than ten times higher.[30] Apparently the cell "intentionally" keeps exosome activity low by maintaining a suboptimal environment or perhaps the activity is enhanced by some unidentified cofactors. However, such results and conclusions should be taken with a pinch of salt. One must keep in mind that determining actual amounts of free magnesium ions in relevant environments is difficult: magnesium is bound by free trinucleotides and RNA and these species are very abundant in cells as well as in enzymatic assay reaction solutions. For these reasons we will concentrate more on qualitative results of activity studies than on quantitative ones.

Dis3 is active on various substrates, but does exhibit some specificity when degradation efficiency is considered. The strongest activity was observed for AU-rich substrates, which may reflect a role of the exosome in degrading ARE-containing mRNAs.[31-32] Activity on adenine homopolymers or oligoadenylated heteropolymers was more modest, but they were still degraded completely.[10, 24]

Importantly, Dis3 is able to degrade structured substrates, making it biochemically more similar to RNase R than to RNase II. Like RNase R, Dis3 appears to bind RNA very efficiently and even though the rate of hydrolysis drops significantly on structured regions, the protein does not dissociate and degrades the molecule to end products.[10]

Degradation of structured substrates

Dis3 is one of the very rare exonucleases that can degrade both single- and double-stranded substrates. The enzyme cannot bind blunt-ended dsRNA molecules, but rather needs a single stranded 3' end.[24] Once bound, Dis3 begins processive hydrolysis and even if secondary structures are encountered, the protein does not dissociate, slowly moving through the structured region. These RNA unwinding properties apply to both inter- and intramolecular secondary structures.[10, 12, 24]

Like RNase II, the RNR part of yeast Dis3 has also been crystallized and its structure solved, giving more insight into the *modus operandi* of the RNR family of nucleases. Both proteins were expressed as mutants unable of catalysis (RNase II^{D209N} and Dis3^{D551N}) and crystallized in the presence of ssRNA and both show that the active site of the RNB domain is situated on the bottom of a narrow channel deep enough to accomodate 5 or 6 nucleotides.[23-24] Comparison with RNase II reveals strong structural conservation within all individual domains, but a markedly different spatial arrangement of the RNA binding domains. This results in the substrate strand taking a different route on its way to the inward channel of the RNB domain: in RNase II the substrate is threaded between the CSD2 and S1 domains only (Figure 3A), while in Dis3 it is threaded between the CSD1 and RNB domains, contacting different surfaces (Figure 3B). It has been suggested that this different positioning of domains and the resulting bending of the RNA strand by Dis3 is responsible for the difference between RNase II and RNase R-like activity on structured RNA substrates. Specific point mutations, designed according to the

structure model, impaired Dis3 activity on double-stranded, but not on single-stranded RNA, lending support to this hypothesis.[24] But that is not the end of it. The problem has also been tackled by a comparative study of RNase II and RNase R, in which the contribution of individual domains to enzymatic activity was measured. Curiously, the difference between the two bacterial enzymes lies within the RNB domain itself, whereas the CSD and S1 domains, while important for substrate binding and greatly enhancing the catalytic properties of the proteins, are not essential and have no bearing on substrate specificities.[33] The same authors demonstrated that high-affinity binding of substrate by the RNB domain of RNase R allows the enzyme to take advantage of thermal breathing of RNA duplexes and thus move along and digest structured substrates.[34]

The two models, while not strictly exclusive, seem contradictory. Either one of them is wrong or RNase R and Dis3 have developed completely different strategies to achieve the same result. Such functional convergence would be very interesting since the proteins are closely related, but it seems more likely that the underlying mechanism of structured substrate digestion is the same throughout the RNR superfamily. In an attempt to reconcile the two hypotheses it has been suggested that the positioning of the single RNA strand in the Dis3 crystal is in fact artefactual.[34]

According to this model, in an RNase R-like enzyme there are two paths for RNA to take: one leading between one CSD and the S1 domains toward the channel entrance, and the other leading away from the channel entrance, between one CSD and the RNB domains, and outward. The free 3' end of an otherwise structured substrate would approach the RNB domain by the first path, as observed for RNase II.[23] When the double-stranded region would reach the entrance to the channel, the enzyme would cleave off the last easily accessible nucleotide at the 3'-end and then pause, but not dissociate, due to high affinity of the RNB domain towards RNA.[33] A wedge is then postulated to protrude from the surface of the RNB domain and push into the double-stranded region of the RNA, disrupting the structure and enhancing thermal breathing effects. Upon transient melting of the secondary structure, the enzyme would immediately move along the substrate, driven by the thermodynamically favourable filling of the active site by the 3' terminus of the RNA, and push the other strand into the second path. Thus the double-stranded

region would be passively unwound, split on the edge of the channel in the RNB domain (Figure 3C).

The hypothesis is supported by several observations. First, activity of RNase R against structured substrates depends on temperature to a greater extent than its activity against ssRNA, suggesting involvement of thermal breathing.[34] Second, the RNB domain of RNase R binds RNA much stronger than the RNB domain of RNase II, probably by making strong contacts between the RNA backbone and the walls of the channel.[33] Only two such contacts are seen in the crystal structure of RNase II.[23] Furthermore, an RNase R mutants, in which some residues within the channel are changed to those found in RNase II have catalytic properties similar to those of RNase II, including faster action on unstructured substrates and problems with structured ones.[34] Third, blocking the second path by introducing a bulky amino acid side chain makes it harder for Dis3 to digest structured substrates with no effect on its activity towards unstructured ones. Such blockage occurs naturally in RNase II.[24] Finally, the second path is blocked in the crystal structure of yeast Dis3 bound to ring components (Rrp41 and Rrp45), while the first remains open, and a ring+Dis3 complex is still able to degrade structured substrates.[11] The latter not only is consistent with the thermal breathing-driven mechanism of RNA unwinding by pulling one strand into a narrow channel, but also indicates that secondary structures can be split on other surfaces than the RNB domain itself.

Endonuclease activity

Recently three groups independently discovered an additional endoribonuclease activity of yeast Dis3, which resides in the PIN domain.[21, 35-36] The protein was able to digest circular substrates as well as linear ones, and the latter produced similar patterns of products regardless of which end of the substrate was labeled. No cleavage was observed in double-stranded regions of structured substrates.[21] The endoribonuclease activity was dependent on divalent metal ions, highest with manganese, but also supported by zinc and, to a small extent, magnesium.

The endonuclease activity of the PIN domain is unspecific to substrate sequence, but has been reported to be specific towards phosphorylated 5' ends. When linear substrates were used, molecules with 5' phosphate groups were cleaved very efficiently, whereas ones with 5' hydroxyl groups were not.[36] However, another group

was unable to reproduce this result and observed comparable activity for 5'-phosphorylated and -unphosphorylated substrates (Rafał Tomecki, unpublished data). Phosphorylation of the 3' terminus is irrelevant for the endoribonuclease activity.[36]

Paralogy in the Dis3 family

Vertebrate cells differ from others in that they contain two Dis3 proteins, called Dis3 and Dis3l, most likely resulting from a gene duplication event. The two subfamilies have diverged significantly: in terms of sequence similarity Dis3l proteins are equally distant from their respective Dis3 paralogue and from Dis3 of *Saccharomyces cerevisiae*. Nevertheless, the domain structure has remained the same in both paralogous lines. The human proteins, DIS3 and DIS3L have recently been characterized after purification from human cell cultures.[18] Both were found to be exoribonucleases with properties typical of the RNR family. However, only DIS3 exhibited endonuclease activity like yeast Dis3. This is consistent with Dis3l proteins lacking two of the amino acid residues that were shown to be important for catalysis in PIN domains of other proteins.

Active sites

The active site of the RNB domain is formed by a short stretch of the polypeptide chain containing four aspartic acid residues that coordinate magnesium ions. In all RNR proteins tested so far, exoribonuclease activity can be abolished by a single point mutation targeting one of those residues. This mutation is D209N in RNase II,[37] D280N in RNase R,[38] D551N in yeast Dis3,[12] and for human proteins: D487N in DIS3 and D486N in DIS3L[17-18] (Figure 2).

Endoribonuclease activity of the Dis3 PIN domain is also dependent on coordination of manganese ions by four acidic amino acid residues and mutation of any of these residues results in an inactive protein. These residues are D91, E120, D171, D198 for yeast Dis3,[21, 35-36] and in human D69, E98, D146, D177 for DIS3 and D62, A91, T140, D166 for DIS3L.[18] Notably two of the residues in DIS3L are naturally non-acidic and consequently DIS3L does not have endoribonuclease activity (Figure 2).

The exoribonuclease activity of the RNB domain and the endoribonuclease activity of the PIN domain are independent of each other in that the disruption of one does not affect the other. This was shown by testing mutant proteins, and is also apparent from the properties of DIS3L.[18, 21, 35-36]

Rrp6

Rrp6 proteins are not as strongly conserved as Dis3. They are related to *Escherichia coli* RNase D, but again, the similarity is not as strong as within the eukaryotic and bacterial branches of the RNR superfamily. RNase D is ubiquitous in eukaryotes, but very rare in bacteria, and archaeans lack the enzyme altogether. In fact, it has been suggested that bacteria have acquired RNase D from eukaryotes by horizontal gene transfer. RNase D and Rrp6 belong to the DEDDy family, a subset of the large DEDD superfamily, which contains various nucleases, including the nuclease domain of DNA polymerases and oligoribonuclease. The hallmark of the DEDD family is four acidic amino acid residues (three aspartate and one glutamate) required for catalysis, which is dependent on divalent metal cations. The DEDD residues do not form a continuous motif, but are distributed between three separate motifs.[2]

The domain structure of yeast Rrp6 entails the catalytic DEDD domain with the four conserved residues, and one HRDC domain of unknown function. Notably, these two domains cover only one-third of the polypeptide chain and little is known about the rest of the protein. The crystal structure of Rrp6 shows the conserved amino acid residues, D238, E240, D296, D365 and Y361 coordinating two ions, one zinc and one manganese.[39]

Yeast Rrp6 has a strong tendency to aggregate and is difficult to work with, which may explain why numerous attempts at purification and thorough characterization have failed. Purification of the free protein from yeast is additionally complicated by its very high affinity to the exosome ring[12] and RNA (Andrzej Dziembowski, unpublished data). The few studies that succeeded revealed that the protein releases nucleoside 5'-monophosphates in a distributive manner. It is active only on 3' ends with hydroxyl groups, but not on ones with phosphate groups, and is unable to process structured RNAs. End products are oligonucleotides of 4-5 nt. Human Rrp6 had similar properties to the yeast protein.[10, 14]

Holoenzyme activities

Dis3 exonuclease

The first studies concerning activity of the holoexosome involved purification of the native complex from *Saccharomyces cerevisiae*. This was achieved by copurification with tagged versions of either Dis3 itself or exosome ring proteins used as bait.[12, 24] More or less at the same time the yeast and human exosome complexes were reconstituted in vitro. Ring subunits were expressed in *Escherichia coli*, individually or in pairs, purified and then mixed. Dis3 and Rrp6 were also expressed in bacteria, purified separately and added to the reconstituted exosome ring to obtain 10- and 11-subunit versions of the complex.[10]

In the first study demonstrating holoenzyme activity[12] the native ring+Dis3 complex had essentially the same activity on both unstructured and structured substrates, a synthetic oligonucleotide and an in vitro transcribed tRNA precursor, respectively. Differences were observed but were very modest, and these activities were very similar to those observed for Dis3 alone. Later studies, however, showed that the activity of yeast Dis3 on structured substrates containing duplexes 17 bp long was inhibited by the exosome ring.[24] A different result was obtained even though the same purification procedure was applied. The reason for this inconsistency is discussed below.

For the proteins purified from bacteria, exonuclease activity of yeast Dis3 on single-stranded substrates was slightly attenuated by the ring complex, but apart from slowing down, substrate specificity and extent of degradation were unaffected. However, when structured substrates were applied, hardly any activity of the ring+Dis3 complex was observed.[10] On the other hand, in a later study such a complex did degrade structured substrates, but required a significantly longer single-stranded stretch at the 3' end to do so.[11] These observations are readily explained by RNase protection assays and structural data. The unstructured 3' end that Dis3 requires in order to initiate degradation is threaded into the channel of the RNB domain and reaches the catalytic site. Only when hydrolysis occurs, can secondary structures be unwound. For Dis3 alone an unpaired stretch of 10 nt is perfectly sufficient.[10, 24] When Dis3 is bound to the exosome ring, the substrate must first pass through the ring channel before it can enter the RNB domain, so a longer

unstructured stretch is necessary: the 10-subunit complex is inactive on substrates with a 10 nt overhang, but completely digests those with a 35 nt overhang.[11] See Figure 4 and chapter *Structural components and architectures of RNA exosomes* by Kurt Januszyk and Christopher Lima for more details.

Conflicting results with structured substrates

The exosome ring does not inhibit Dis3 activity on structured substrates as such, but simply changes the minimal length of the 3' overhang that can be bound and digested (Figure 4). Thus it is puzzling why the first study testing holoexosome activity showed that binding to the exosome ring had no significant effect on the activity of Dis3 against structured substrates,[12] especially since a very strong effect was observed in later studies.[24]

One possible explanation is that the two studies used different structured substrates: an in vitro transcribed pre-tRNA was degraded with similar efficiency by yeast Dis3 alone and the 10-subunit complex, whereas an intermolecular duplex with 3' overhangs of 14 nt and less was only degraded by free Dis3. Both these substrates contain secondary structures, but with arguably different stabilities. Notably, the efficiency of hydrolysis by Dis3 is to a certain extent dependent on the length of the free 3' end: highest activity is observed with overhangs of 10 nt and longer, but substrates with overhangs as short as 5 nt can still be degraded.[24] This is because the protein requires a single-stranded binding site in order to thread the 3' terminus into the active site, which is also true for the exosome complex.[11] It is easy to imagine how an unstable secondary structure constitutes a smaller obstacle to binding, thermal breathing events transiently elongating the single-stranded region.

The duplex used in the second study contained a continuous 17 bp region, making for great stability. The pre-tRNA from the first study on the other hand was completely unmodified, which could locally weaken secondary structures and allow for efficient binding by both Dis3 and the complex. Such an effect would be in line with the ability of yeast Dis3 to recognize hypomodified tRNAs, presumably due to their misfolding.[40]

Rrp6 and Dis3 endonuclease

Endonuclease activity of yeast Dis3 was not assayed in the study that first demonstrated complex reconstitution, as it was unknown at the time. It was later shown that binding of the protein to the ring complex has little or no effect on the activity. This suggests that RNA can access the PIN active site directly, without going through the ring channel.[11]

Activity of yeast Rrp6 was also unaffected by binding to the ring. Although activity of a ring+Rrp6 complex was never tested directly, assays were performed for Rrp6, ring+Dis3 and ring+Dis3+Rrp6. Degradation patterns for the 11-subunit complex were the sum of the patterns obtained with Rrp6 alone and with the 10-subunit complex, revealing no modulation of Rrp6 activity.[10]

Cooperation of the three activities

Rrp6 and Dis3 can bind to the ring concomitantly, as evidenced by copurification of the two proteins[3, 12] and successful reconstitution of the 11-subunit complex.[10] The three activities of the catalytic subunits can be brought together in one molecular assembly and possibly act on the same substrate, but whether the cell prefers 11- to 10-subunit complexes (ring+Dis3 or ring+Rrp6) is unknown. Certainly in yeast Dis3 and Rrp6 can act at different steps of one process, as exemplified by processing of the 5.8S rRNA. A precursor molecule is first processed by Dis3, producing an intermediate elongated by 30 nt at the 3' end,[41] which is then trimmed by Rrp6, leaving an 8 nt extention, later removed by the Rex RNases.[42] The length of the tail in the intermediate left by Dis3 is perfectly rationalized by threading of substrates through the exosome ring and into Dis3, although why the complex dissociates is not explained.[11] Processing of precursor RNAs first by processive and then by distributive RNases has also been suggested in tRNA maturation in bacteria[43-44] and is probably a general mechanism of RNA maturation, preventing overzealous processive enzymes from damaging the body of mature molecules.[43]

Sequential cooperative action of the endo- and exonuclease activities may also enhance the degradation potential of the exosome, allowing it to destroy troublesome molecules. It was suggested that the endoribonuclease activity of Dis3 serves to create additional entry points in substrates that cause the exosome to stall.[35] This is strongly supported by the fact that a yeast mutant devoid of the exoribonuclease activity (*dis3^{D551N}*) accumulates decay intermediates resulting from structured molecules

having been cut in predicted single-stranded regions. The same type of cleavage was observed in vitro.[21, 35]

Noncatalytic activities of the exosome ring

The exosome ring has only a modest effect on the enzymatic properties of the catalytic subunits, raising questions about the function of the ring proteins in fungi and animals. So far three explanations have been proposed for the continued existence of the ring even without catalytic functions: substrate recruitment, substrate restriction and interaction with regulatory factors.

Substrate recruitment

The ring is considered too ancient and important to remove. The consensus is that an exosome ring-like molecular machine, a "proto-exosome" so to speak, arose even before differentiation of the three domains of life. Pathways of RNA processing and decay already existed and all were centered around a capped hexameric ring of phosphorolytic RNases. Later the system diverged into the ones we observe today: PNPase and the archaeal and eukaryotic exosomes. They do have different additional proteins and biochemical activities, but are still major knots in the RNA degradation frameworks of their respective cell systems. In bacteria and eukaryotes hydrolytic activities have been attached to the proto-exosome, both endo- and exonuclease, and even though analogous proteins have not been identified in archaeans, their existence can hardly be excluded. All the while the proto-exosome has been a crucial meeting point for RNA decay substrates and effectors. In eukaryotes it lost its catalytic activity, but its other functions were too complex to be easily taken over by active enzymes. The system became trapped with an inactive assembly of proteins that are still required for bringing together enzymes, their substrates and regulatory factors.[12]

Exosome activation

Concordantly, the exosome can, and in vivo may need to, be activated by the action of additional cofactors, such as the TRAMP complex in the nucleus, which marks RNAs for exosomal destruction by polyadenylation,[45-46] the Ski complex in the cytoplasm[47] and possibly even some AU-rich element binding proteins (AUBPs).[31-32] Both TRAMP and Ski complexes contain RNA helicases[46, 48] and the AUBP RHAU

was also shown to bear helicase activity that is necessary for exosome activation by that protein.[49] Moreover, the Ski complex, as well as some AUBPs, have been shown to interact with the exosome ring,[31, 47] and even a direct interaction of the human Rrp45 with AU-rich elements was reported.[32] Thus, the function of the ring as an interaction platform seems relatively well documented.

Substrate filtering

It is clear that the presence of the exosome ring does not alter much the activity of Dis3, but rather acts as a substrate filter, rejecting substrates with free 3' ends too short to reach the catalytic site. This observation sheds a light on how the exosome works in the cell. It was suggested on several occasions that RNAs are targeted to degradation when their structure is compromised.[50] Be it due to a transcription error or hypomodification, deviation from the proper structure would be recognized and the molecule would be removed. Weakening the structure might expose a free 3' end long enough to be bound by Dis3, which was indeed observed: yeast Dis3 preferentially binds to a hypomethylated tRNA$_i$ when presented with total RNA.[40] Still, since Dis3 needs a very short overhang to bind a molecule, such activity would put many normal RNAs in danger of becoming falsely identified as misfolded. Association of Dis3 with the exosome ring prevents the enzyme from attacking RNAs unless they have a long unstructured stretch (unbound by proteins), filtering potential substrates. At the same time the ring provides a way for a substrate RNA to be tagged for degradation by adding an unstructured tail, such as in the case of polyadenylation mediated by the TRAMP complex.[45-46]

This hypothesis also explains why the exosome is activated by complexes containing helicase activity. If normal, undamaged RNAs must be removed, the protective effect of their "healthy" structure and bound proteins may need to be counteracted. In vitro data suggest that Dis3 and the holoexosome can degrade quite stable secondary structures, but this was never tested on large molecules such as rRNA. Neither is there any indication of how quickly marked RNAs should be removed, but intuitively the process should be rapid and efficient. Attaching an RNA helicase activity to the exosome would speed up the process, preventing both accumulation of marked RNAs and titrating the exosome.

The situation of Dis3 in the complex is such that the only entrance to the RNB domain leads through the ring channel, but the active site of the PIN domain is exposed outward and freely accessible. Rrp6 is positioned on top of the ring and appears to be accessible at all times too.[51] So far regulation of these activities by the ring has not been observed.[10-11] For them the function of the ring seems to be limited to substrate recruitment and activation.

Why the exosome would carefully restrict access to one of its active sites, but permit it to two others is difficult to explain. Though perhaps no explanation is needed. It is very unlikely that such a central piece of the cell machinery would be left unchecked. It makes sense for the most aggressive activity (Dis3 exonuclease) to be the most conspicuously restrained and its control mechanism was relatively easily unraveled. The mechanisms controlling the other two activities surely exist, but are for the time being unknown.

References

1. Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell.* Nov 14 1997;91(4):457-466.
2. Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic acids research.* Mar 1 2001;29(5):1017-1026.
3. Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P. The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development.* Aug 15 1999;13(16):2148-2158.
4. Koonin EV, Wolf YI, Aravind L. Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. *Genome research.* Feb 2001;11(2):240-252.
5. Symmons MF, Jones GH, Luisi BF. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure.* Nov 15 2000;8(11):1215-1226.
6. Aloy P, Ciccarelli FD, Leutwein C, et al. A complex prediction: three-dimensional model of the yeast exosome. *EMBO reports.* Jul 2002;3(7):628-635.
7. Buttner K, Wenig K, Hopfner KP. Structural framework for the mechanism of archaeal exosomes in RNA processing. *Molecular cell.* Nov 11 2005;20(3):461-471.
8. Hernandez H, Dziembowski A, Taverner T, Seraphin B, Robinson CV. Subunit architecture of multimeric complexes isolated directly from cells. *EMBO reports.* Jun 2006;7(6):605-610.
9. Lorentzen E, Walter P, Fribourg S, Evguenieva-Hackenberg E, Klug G, Conti E. The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nature structural & molecular biology.* Jul 2005;12(7):575-581.
10. Liu Q, Greimann JC, Lima CD. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell.* Dec 15 2006;127(6):1223-1237.
11. Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E. The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell.* Oct 30 2009;139(3):547-559.
12. Dziembowski A, Lorentzen E, Conti E, Seraphin B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nature structural & molecular biology.* Jan 2007;14(1):15-22.
13. Chekanova JA, Shaw RJ, Wills MA, Belostotsky DA. Poly(A) tail-dependent exonuclease AtRrp41p from Arabidopsis thaliana rescues 5.8 S rRNA processing and mRNA decay defects of the yeast ski6 mutant and is found in an exosome-sized complex in plant and yeast cells. *The Journal of biological chemistry.* Oct 20 2000;275(42):33158-33166.
14. Burkard KT, Butler JS. A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. *Molecular and cellular biology.* Jan 2000;20(2):604-616.
15. Staals RHJ, Bronkhorst AW, Schilders G, Heck AJR, Raijmakers R, Pruijn G. Dis3-like 1: a novel exoribonuclease associated with the human exosome. submitted.
16. Tomecki R, Kristiansen MS, Lykke-Andersen S, et al. The human core exosome utilizes differentially localized processive ribonucleases: hDIS3 and hDIS3L. submitted.
17. Staals RHJ, Bronkhorst AW, Schilders G, Heck AJR, Raijmakers R, Pruijn G. Dis3-like 1: a novel exoribonuclease associated with the human exosome. *The EMBO journal.* 2010.
18. Tomecki R, Kristiansen MS, Lykke-Andersen S, et al. The human core exosome utilizes differentially localized processive ribonucleases: hDIS3 and hDIS3L. *The EMBO journal.* 2010.

19. Lange H, Holec S, Cognat V, et al. Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in Arabidopsis thaliana. *Molecular and cellular biology.* May 2008;28(9):3038-3044.

20. Chekanova JA, Gregory BD, Reverdatto SV, et al. Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell.* Dec 28 2007;131(7):1340-1353.

21. Schneider C, Leung E, Brown J, Tollervey D. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic acids research.* Mar 2009;37(4):1127-1140.

22. Cheng ZF, Deutscher MP. Purification and characterization of the Escherichia coli exoribonuclease RNase R. Comparison with RNase II. *The Journal of biological chemistry.* Jun 14 2002;277(24):21624-21629.

23. Frazao C, McVey CE, Amblar M, et al. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature.* Sep 7 2006;443(7107):110-114.

24. Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Molecular cell.* Mar 28 2008;29(6):717-728.

25. Makarova KS, Aravind L, Galperin MY, et al. Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome research.* Jul 1999;9(7):608-628.

26. Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *The Journal of biological chemistry.* Apr 16 2004;279(16):16471-16478.

27. Clissold PM, Ponting CP. PIN domains in nonsense-mediated mRNA decay and RNAi. *Curr Biol.* Dec 14-28 2000;10(24):R888-890.

28. Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *The EMBO journal.* Nov 1 2006;25(21):5117-5125.

29. Mamolen M, Andrulis ED. Characterization of the Drosophila melanogaster Dis3 ribonuclease. *Biochemical and biophysical research communications.* Dec 18 2009;390(3):529-534.

30. Gunther T. Concentration, compartmentation and metabolic function of intracellular free Mg2+. *Magnes Res.* Dec 2006;19(4):225-236.

31. Chen CY, Gherzi R, Ong SE, et al. AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell.* Nov 16 2001;107(4):451-464.

32. Mukherjee D, Gao M, O'Connor JP, et al. The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *The EMBO journal.* Jan 15 2002;21(1-2):165-174.

33. Vincent HA, Deutscher MP. The roles of individual domains of RNase R in substrate binding and exoribonuclease activity. The nuclease domain is sufficient for digestion of structured RNA. *The Journal of biological chemistry.* Jan 2 2009;284(1):486-494.

34. Vincent HA, Deutscher MP. Insights into how RNase R degrades structured RNA: analysis of the nuclease domain. *Journal of molecular biology.* Apr 3 2009;387(3):570-583.

35. Lebreton A, Tomecki R, Dziembowski A, Seraphin B. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature.* Dec 18 2008;456(7224):993-996.

36. Schaeffer D, Tsanova B, Barbas A, et al. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature structural & molecular biology.* Jan 2009;16(1):56-62.

37. Amblar M, Arraiano CM. A single mutation in Escherichia coli ribonuclease II inactivates the enzyme without affecting RNA binding. *The FEBS journal.* Jan 2005;272(2):363-374.

38. Matos RG, Barbas A, Arraiano CM. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *The Biochemical journal.* Oct 15 2009;423(2):291-301.

39. Midtgaard SF, Assenholt J, Jonstrup AT, Van LB, Jensen TH, Brodersen DE. Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. *Proceedings of the National Academy of Sciences of the United States of America.* Aug 8 2006;103(32):11898-11903.

40. Schneider C, Anderson JT, Tollervey D. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Molecular cell.* Jul 20 2007;27(2):324-331.

41. Briggs MW, Burkard KT, Butler JS. Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. *The Journal of biological chemistry.* May 22 1998;273(21):13255-13263.

42. van Hoof A, Lennertz P, Parker R. Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *The EMBO journal.* Mar 15 2000;19(6):1357-1365.

43. Deutscher MP. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic acids research.* 2006;34(2):659-666.

44. Li Z, Pandit S, Deutscher MP. 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America.* Mar 17 1998;95(6):2856-2861.

45. LaCava J, Houseley J, Saveanu C, et al. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell.* Jun 3 2005;121(5):713-724.

46. Vanacova S, Wolf J, Martin G, et al. A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS biology.* Jun 2005;3(6):e189.

47. Araki Y, Takahashi S, Kobayashi T, Kajiho H, Hoshino S, Katada T. Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. *The EMBO journal.* Sep 3 2001;20(17):4684-4693.

48. Brown JT, Bai X, Johnson AW. The yeast antiviral proteins Ski2p, Ski3p, and Ski8p exist as a complex in vivo. *RNA (New York, N.Y.* Mar 2000;6(3):449-457.

49. Tran H, Schilling M, Wirbelauer C, Hess D, Nagamine Y. Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. *Molecular cell.* Jan 16 2004;13(1):101-111.

50. Reinisch KM, Wolin SL. Emerging themes in non-coding RNA quality control. *Current opinion in structural biology.* Apr 2007;17(2):209-214.

51. Cristodero M, Bottcher B, Diepholz M, Scheffzek K, Clayton C. The Leishmania tarentolae exosome: purification and structural analysis by electron microscopy. *Molecular and biochemical parasitology.* May 2008;159(1):24-29.

Figure legends

Figure 1. Composition of yeast and human exosome complexes and differential localization of catalytic subunits. The exosome ring is localized in the cytoplasm, nucleus and nucleoli in all eukaryotes. In *Saccharomyces cerevisiae* (**left**) the single Dis3 protein is present in all three compartments and the single Rrp6 protein is confined to the nucleus. In *Homo sapiens* (**right**) the single RRP6 homologue is found in all compartments. One Dis3 paralogue – DIS3 – is mostly, if not only, nuclear, but excluded from nucleoli, while the other – DIS3L – is strictly cytoplasmic.

Figure 2. Domain organization of the RNR superfamily. All RNR proteins contain one catalytic domain, RNB, and three RNA binding domains: CSD1, CSD2 and S1. Proteins of the Dis3 subfamily, found in all eukaryotes but never in prokaryotes, also contain a catalytic PIN domain. The latter is inactive in the Dis3l proteins, missing two of the conserved acidic amino acid residues (blue arrows).

Figure 3. Unwinding secondary structures by Dis3. A. Substrate position observed in crystals of *Escherichia coli* RNase II. The RNA approaches the RNB channel between the RNA binding domains, making contacts with CSD2 and S1. **B.** Substrate position observed in crystals of *Saccharomyces cerevisiae* Dis3. The RNA takes a different route, between CSD1 and RNB. **C.** Proposed position of a structured RNA molecule in Dis3. The 3' portion of the substrate reaches the catalytic site and is processively hydrolyzed. The double stranded segment approaches the RNB channel between the RNA binding domains, but is too wide to enter RNB itself. As the 3' end is pulled into the active site with each successive cleavage event, the secondary structure is pulled apart, the 3' nucleotide entering the RNB channel and the 5' nucleotide leaving the protein between RNB and CSD1. Presumably, once the secondary structure is completely unwound, the 5' portion of the RNA is pulled back and threaded into the RNB channel to be completely digested.

Figure 4. Substrate filtering by the exosome ring complex. Any RNA molecule with secondary structure can be digested by Dis3, provided it has a single stranded stretch at the 3' end long enough to reach the active site buried within the RNB domain. **A.** Molecules with free ends shorter than 4 nt are resistant to hydrolysis. **B.** Molecules with free ends of 5 nt and longer can be digested to end products by free Dis3. **C.** Substrates threaded through the exosome ring require a longer free end to reach the active site of Dis3 and thus molecules with single-stranded 3' ends as long as 30 nt are protected. **D.** Free ends longer than 33 nt are sufficient to reach the Dis3 active site through the exosome ring.