NMR studies of the archaeal exosome complex

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Maxime Audin

aus La teste de Buch

Frankreich

Tübingen

2016

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	03.05.2016
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Dr. Remco Sprangers
2. Berichterstatter:	Prof. Dr. Thilo Stehle

SUMMARY

The 3' to 5' exo-ribonucleolytic cleavage of RNA is involved in many processes, including precursor-RNAs trimming and mRNA degradation. In eukaryotes and archaea, the exosome is the major 3' to 5' exo-ribonucleolytic protein complex. Although structural information is available for different exosome complexes, the mechanism that regulates the degradation process remains unclear. Here, we used novel methods in NMR spectroscopy to study the exosome complex from the archaea *Sulfolobus solfataricus*, which has a relatively simple quaternary architecture. The archaeal exosome is composed of three Rrp41: Rrp42 dimers that form a hexameric ring, the catalytic core. Two additional cap proteins (Rrp4 and Csl4) bind this catalytic core and thereby provide substrate specificity. In this thesis, we addressed how the catalytic core interacts with the cap proteins and with substrate RNA.

Molecular motions detected by nuclear magnetic resonance spectroscopy

A combination of methyl group labeling and methyl-TROSY-based NMR experiments allows us to observe dynamic processes in the 180 KDa exosome core. First, we observed that the N-terminal helix of the Rrp42 subunit exchanges between two conformations, reflecting unpredicted flexibility of the exosome core. The exchange rate between the two conformations is in the order of tens per second. We determined that one conformation is selected upon Rrp4 and RNA binding. In addition, we determined that the 3’end of the RNA inside the exosome core is highly mobile and jumps between the three active sites with rates in the order of thousands per second.

Efficient degradation requires the exosome quaternary structure

We demonstrated that the oligomerization of the exosome into a hexameric ring is important for the efficiency of the enzyme. First, this results

in the formation of a neck that ensures processive degradation of the substrate RNA. Second, oligomerization creates a lumen with a high active site and RNA concentration. As a result, the low affinity between the RNA and the active site is compensated by local high concentration. Using a combination of biochemical and NMR experiments we could furthermore determine that about 100 RNA: active site encounters are required for one cleavage event to occur.

In summary, this thesis provides unique and novel insights into the mode of action of the exosome complex.

ZUSAMMENFASSUNG

Der exonucleolytische Abbau von RNA in 3'-5' Richtung ist in zahlreiche zelluläre Prozesse, wie z.B. den mRNA-Abbau oder die Prozessierung von vorläufigen RNAs, involviert. Der Exosomkomplex ist dabei die wichtigste 3'-5' Exoribonuclease in Eukaryoten und Archaeen. Obwohl die Struktur zahlreicher Exosomkomplexe aufgeklärt wurde, ist der Regulationsmechanismus des RNA-Abbaus weitgehend unverstanden. In dieser Arbeit wurden daher neuartige Methoden der NMR-Spektroskopie eingesetzt, um den Exosomkomplex aus dem Archaeon Sulfolobus solfataricus zu untersuchen, welcher eine vergleichsweise einfache Quartärstruktur aufweist. Der katalytische Kern archaeler Exosomkomplexe besteht aus drei Rrp41:Rrp42 Protein Dimeren, die einen hexameren Ring bilden. Substratspezifität wird durch zwei zusätzliche Proteine vermittelt (Rrp4 und Csl4), die mit dem Exosomkern eine Verbindung eingehen. In der vorliegenden Arbeit wurde die Wechselwirkung des katalytischen Kerns mit Substrat- RNA sowie den Rrp4/Csl4 Proteinen untersucht.

Untersuchung der Exosomdynamik mittels NMR

Zur Untersuchung der Moleküldynamik des 180 kDa Exosomkerns wurde eine Kombination aus methylgruppenspezifischer Isotopenmarkierung und Methyl-TROSY-basierten NMR-Experimenten verwendet. Die Ergebnisse zeigen, dass die N-terminale Helix der Rrp42 Untereinheit in zwei unterschiedlichen Konformationen vorliegt, zwischen denen ein schneller Austausch im ms-Bereich erfolgt. Eine der beiden Konformationen wird dabei durch die Bindung von Rrp4 und RNA selektiert. Zudem konnten wir nachweisen, dass die Substrat-RNA im Exosomkern eine hohe Mobilität aufweist, und mit einer Rate von einigen tausend pro Sekunde zwischen den drei aktiven Zentren des Exosomkerns wechselt.

Die Quartärstruktur des Exosomes als Grundlage des effizienten RNA-Abbaus

Im Laufe dieser Arbeit konnte des Weiteren die Bedeutung der Oligomeren Struktur des Exosomes für die effiziente Katalyse gezeigt werden. Die Oligomerisierung führt zur Bildung einer Engstelle auf dem Weg in den katalytischen Kern des Exosomes, die für die Prozessivität des RNA-Abbaus entscheidend ist. Zudem entsteht durch die Oligomerisierung ein Hohlraum im inneren des Exosomes, der zu einer hohen lokalen Konzentration von RNA und katalytisch aktiven Zentren führt, wodurch die niedrige Affinität der katalytischen Zentren für RNA kompensiert wird. Durch die Kombination von biochemischen und NMR-Experimenten konnte außerdem gezeigt werden, dass ein RNA-Molekül in etwa 100-mal mit einem aktiven Zentrum in Verbindung treten muss, bevor es zur Katalyse kommt.

Im Verlauf der Arbeit konnten somit zahlreiche neue Einblicke in die Funktionsweise des Exosomkomplexes gewonnen werden.

TABLE OF CONTENTS

LIST OF PUBLICATIONS

Peer-reviewed articles

Paper 1:

Audin MJ, Dorn G, Fromm SA, Reiss K, Schütz S, Vorländer MK, Sprangers R. (2013).
The archaeal exosome: identification and quantification of site-specific motions that correlate with cap and RNA binding.
Angew Chem Int Ed Engl. 2013 Aug 5;52(32):8312-6

Paper 2:

Maxime J. Audin, Jan Philip Wurm, Milos A. Cvetkovic and Remco Sprangers (2016)
The oligomeric architecture of the archaeal exosome is important for processive and efficient RNA degradation.
Nucleic acids research, accepted for publication.

ABBREVIATIONS

AA	Acrylamide
ADP	Adenosine diphosphate
B. subtilis	*Bacillus subtilis*
CCR4	Carbon catabolite repressor 4
CEST	Chemical exchange saturation transfer
CPMG	Carr-Purcell-Meiboom-Gill
CPSF	Cleavage and polyadenylation specificity factor
Csl4	Cep1 synthetic lethal 4
CSP	Chemical shift perturbation
FPLC	Fast performance liquid chromatography
HMQC	Heteronuclear multiple quantum coherence
HPLC	High pressure liquid chromatography
hSuv3	Human suppressor of var1 3
INEPT	Insensitive nuclei enhanced by polarization transfer
ITC	Isothermal titration calorimetry
KH domain	K-homology domain
MQ	Multiple quantum
mRNA	Messenger ribonucleic acid
mtRNA	Mitochondrial ribonucleic acid
Ni-NTA	Nickel-Nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
NOT	Negative on TATA
p.p.pm.	Parts per million
PABP	Poly(A) binding protein
PAN2	Poly(A) nuclease 2
PAN3	Poly(A) nuclease 3
PAP	Poly(A) polymerase
PDB	Protein data bank
Pi	Inorganic phosphate
PNPase	Polynucleotide phosphorylase
PriSL	Primase small and large catalytic subunit

RF	Radiofrequency
RhlB	RNA helicase B
RNA	Ribonucleic acid
RNAse	Ribonuclease
rRNA	Ribosomal ribonucleic acid
Rrp4	Ribosomal RNA-processing protein 4
Rrp41	Ribosomal RNA-processing protein 41
Rrp42	Ribosomal RNA-processing protein 42
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SPR	Surface plasmon Resonance
SQ	Single quantum
TFE	Trifluoroethanol
TRAMP	Trf4/Air2/Mtr4p Polyadenylation complex
tRNA	Transfer ribonucleic acid
TROSY	Transverse relaxation optimized spectroscopy

I. INTRODUCTION

RNA degradation plays a central role in the regulation of RNAs levels. Indeed, defective RNAs and RNAs whose expression is no longer required are targeted to degradation. All types of RNA, from the stable functional RNA (such as ribosomal RNA or transfer RNA) to the more ephemeral messenger RNA are eventually subjected to degradation.

The difference of stability between the various types of RNAs resides in their capacity to resist to the action of ribonucleases. Surprisingly, the maturation of many non-coding ribosomal RNAs involves the processing of precursors RNAs by endo- and/or exo-ribonucleases that are also involved in RNA decay (1). One notable example for such an enzyme is the exosome, the major 3' to 5' RNA degradation complex in eukaryotes. However, the regulatory mechanism that discriminates the partial degradation (trimming) from complete degradation remains unclear, despite the increasing amount of structures that are available of the complex. Structural studies of the exosome have benefitted from the discovery of the complex in Archaea (2) that harbors a simpler architecture than the eukaryotic one and, at the same time, shares similarities with the polynucleotide phosphorylase (PNPase), the counterpart of the exosome in bacteria (also found in mitochondria and chlorplasts).

1. The exosome-like complexes

The PNPase, the archaeal and eukaryotic exosome are protein complexes composed of subunits that are evolutionary related to the bacterial RNase PH (figure 1). All these complexes form a barrel-like structure (3-7) that enables only single stranded RNA to enter in the ring chamber. Except for the eukaryotic exosome, these complexes harbor several active sites due to the intrinsic activity endowed in the subunits that form the catalytic core. As a result, the RNase PH possesses 6 active sites, where the PNPase and the archaeal exosome possess 3 active sites. Except for the eukaryotic exosome, these protein complexes also share a phospholytic activity (8) and are the

only enzymes that possess 3' to 5' phosphorolytic exoribonucleases that are known so far (6). In contrast to the RNase PH, the PNPase and the archaeal exosome also exhibit a polymerase activity, which is suspected to be the primary function from an evolutional point of view (9). The eukaryotic exosome evolved into a similar but asymmetric structure due to the fact that every polypeptide chain is different (5,10). Over evolution, none of its subunits conserved an active site and the eukaryotic exosome thus requires an additional subunit for its activity (11). The exosome-like complexes are involved in the majority of the 3' to 5' degradation RNA degradation pathways and are found in all domains of life. However, their implications in the RNA 3' end degradation vary. The different functions of the PNPase, of the archaeal and eukaryotic exosome are described in the following paragraphs.

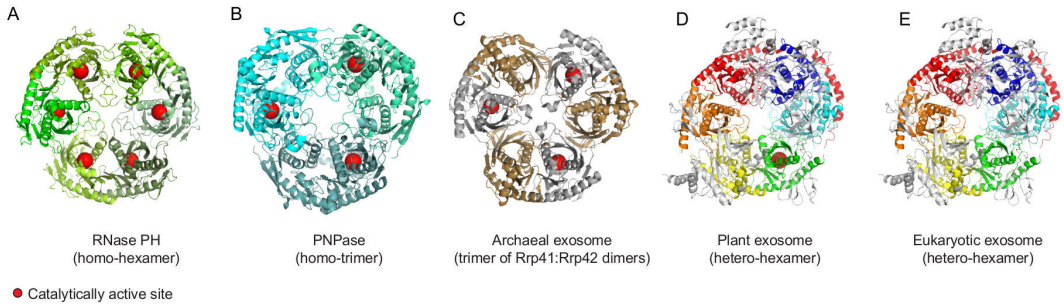

Figure 1: Comparison of the exosome-like structures

For each protein complex, the active sites are depicted with red spheres. The exosome and exosome-like complexes are sorted to visualize the progressive loss of active sites.

A) The homo-hexameric RNase PH from *Aquifex aeolicus* (12). The six identical protein chains are colored in different shades of green (PDB: 1UDN). B) The home-trimeric PNPase from *Streptomyces antibioticus* (13). The three identical protein chains are colored in different shades of cyan (PDB: 1E3H). C) The archaeal exosome core complex from *Sulfolobus solfataricus* (4) that contains three Rrp41: Rrp42 dimers, colored grey and sand, respectively (PDB: 2BR2). D) and E) The eukaryotic exosome core structure (5) that contains six different protein chains (PDB: 2NN6). The three different proteins that form the cap of the complex are colored grey. The exosome from *Arabidopsis thaliana* seems to possess a single active site in Rrp41 (14) (D), in contrast to all the other eukaryotic exosomes that are inactive (E).

2. The eukaryotic exosome

The exosome is a protein complex involved in various RNA pathways, from maturation to decay (15,16). The exosome is the main 3' to 5' exoribonuclease in the nucleus and in the cytoplasm of all eukaryotes (animals, plants, yeasts and trypanosomes). In eukaryotes, the role of the exosome varies depending on its recruiting/assisting co-factors (17). The exosome participates in the maturation of rRNAs (18) by trimming the 3' end part of the pre-rRNA and by degrading the by-products during rRNA maturation. It is also involved in the maturation of other non-coding RNAs, such as snRNAs, snoRNAs or rRNAs (19). Abnormally folded non-coding RNAs (such as tRNAs) can be adenylated by the TRAMP complex prior to its degradation by the exosome. For degradation of RNA with secondary structures, the exosome requires the helicase activity of the ski complex.

The exosome is also a key component in the mRNA surveillance. It degrades aberrant mRNAs (no-go decay, nonsense-mediated decay, non-stop decay) and participates in the turnover of normal mRNAs as well. In eukaryotes, the poly(A) tail is essential for the integrity of the mRNA and ensures mRNA stability. It is bound to Poly(A) Binding Proteins (PABP) that offer protection against non-specific exonuclease activities. The poly(A) degradation is not initiated by the exosome but by the PAN2/PAN3 complex, followed by the CCR4/NOT complex. Then, the exosome removes the last nucleotides of the poly(A) tail and further degrade the mRNA (20).

3. The eukaryotic PNPase

The exosome family also includes the polynucleotide phosphorylase (PNPase) that is present in bacteria, but also in chloroplasts and mitochondria of many eukaryotes. The PNPase seems to be important for the mammalian mitochondria homeostasis but the reason why is still unknown. Indeed, the PNPase is located in the mitochondrial intermembrane space where no RNA substrate has been detected yet (21). However, recent studies suggest that the PNPase, together with the RNA helicase hSuv3, are responsible for mtRNA decay in human (22). In Plant, the PNPase is found in the stroma of

the chloroplast where the RNA degradation is polyadenylation-dependent. There, the PNPase is able to add a heteropolymeric tail to the truncated transcripts in order to target them to degradation (23). In contrast to the mammalian mitochondria, the PNPase is located in the matrix of the plant mitochondria. Its function is very similar to one observed in chloroplast where it can behave as an exoribonuclease or as a polymerase for polyadenylation-dependent RNA degradation (24).

4. The bacterial PNPase

Despite a 5' to 3' degradation observed in B. subtilis (25), the main RNA degradation pathway in bacteria starts with the endonucleolytic cleavage of RNA which is facilitated by the 5' pyrophosphate removal (26). The resulting fragments then undergo a 3' polymerization that targets the RNA to a rapid 3' to 5' exoribonucleolytic degradation by either the PNPase, the RNase R and /or the RNase II. Although the polyadenylation is mainly carried out by the Poly(A) polymerase (PAP), the PNPase can also add heteroplymeric tails. This dual function seems to be regulated by the local concentration of ADP and Pi (27). The 3' end degradation of the RNAs is thus a balance between the action of PAP and the activity of PNPase, RNase II and RNase R (28).

Alone, the PNPase can hardly degrade fragments with secondary structure. Nevertheless, in few cases, the PNPase is capable of nibbling the bottom of an hairpin and after several cycles of polydenylation followed by further nibbling by the PNPase, the hairpin could be completely degraded (29). The PNPase is often found associated with other proteins in the so-called degradosome. The degradosome is composed of the Rnase E, the RNA helicase B (RhIB), the enolase and the PNPase (30) together with other components in a sub stoichiometric amounts that modulate its activity (31).

The importance of the PNPase varies between bacterial organisms. In cyanobacteria, PNPase appears to be the only polymerase and its deletion is lethal (32). But in E.coli, the deletion of PNPase is not lethal and results in a slower elimination of fragments resulting from mRNA decay (33). It implies that other enzymes with similar activities can compensate the lack of PNPase.

PNPase is shown to be also involved in degradation of stable RNA (34) but has not been observed to play a role in RNA maturation in bacteria.

5. The archaeal exosome

In general, archaea show higher similarities to bacteria in term of morphology and higher similarities to eukaryotes at the molecular level. However, mRNA degradation in Archaea seems to be closer to that of bacteria. Like in bacteria, several RNase J homologs (CPSFs) have been found to exhibit either a 5' to 3' degradation or an endonucleolytic activity (35,36). In the exosome-containing archaea, the exosome appears to be involved in the majority of the 3' to 5' RNA degradation (mRNA, rRNAs) and is potentially involved in the rRNA maturation (37). Like in bacteria, a tail added post-transcriptionnaly to RNAs target them to rapid degradation. So far, the exosome is the only one polymerase detected (38) in archaea and appears to be the major 3' to 5' exoribonuclease. Concerning the exosome-less archaea, even fewer is known. In these organisms, no polyadenylation is observed and the degradation is predicted to be either performed by an Rnase R homolog (like in Halophilic archaea)(39) or would solely depend on the putative Rnase J (as no other 3' to 5' exoribonuclease has been found yet in Methanogenic archaea)(35).

In archaea, the exosome can be formed by up to 10 subunits (2). First, there is the catalytic core, which consists of a trimer of a Rrp41:Rrp42 dimer that assembles to form a barrel-like structure (4)(figure 1C). Rrp41 and Rrp42 are two RNase PH-like proteins where only Rrp41 possess an active site. The catalytic core interacts with a trimeric complex composed of the RNA binding proteins Rrp4 (40) and Csl4 (41). Rrp4 has been shown to increase the degradation of poly(A) rich RNAs whereas Csl4 doesn't change the substrate specificity of the catalytic hexamer (42). The Csl4 subunit main role is to bind the 10^{th} subunit, DnaG (43). DnaG is a bacterial type primase that coexist in archaea with the eukaryotic type primase PriSL protein complex (44). DnaG increases the affinity of the Csl4-exosome for poly(A) rich RNAs (43) and appears to be responsible for the membrane localization of the exosome (45).

DnaG also seems to be involved in the polynucleotidylation and the degradation of rRNA by the exosome (46).

6. In-vivo regulation of the S. solfataricus exosome

The presence of an exosome in archaea was accurately predicted in 2001, with Rrp41, Rrp42 and Rrp4 in the same operon (37), whereas Csl4 is located in another operon. The presence of an exosome in archaea was confirmed several years later and revealed that the DnaG is strongly associated to the exosome (2). As the co-immunoprecipitation always leads to the detection of Rrp4 and to a lower extent Csl4 and DnaG, it is believed that the hexameric core is not present in-vivo without these RNA-binding proteins (41). The presence of RNA binding-domains (zinc finger, S1 and KH domains) in the cap proteins allows a stronger binding of the RNA to the exosome and thus improves degradation (42). In addition, the cap proteins differ in their substrate specificity suggesting that the composition of the cap affects the type of substrate RNA recruited and consequently degraded by the hexamer (42). The heteromeric caps are found in-vivo but the exact stoichiometry of Rrp4/Csl4 remains unclear. Csl4-rich exosomes are mainly found in the insoluble fractions whereas Rrp4-rich exosomes are mostly present in the soluble fraction (41). The presence of the Csl4-rich exosomes in the insoluble fractions is correlated with the co-localization of Csl4-exosome with DnaG to the membrane (45). DnaG is also suspected to facilitate the addition of heteropolymeric tail to native rRNAs which enhances their degradation (46). Taken together, the different composition of the cap proteins together with DnaG seems to affect its localization and could regulate its availability.

II. AIM OF THE PROJECT

As described above, the exosome is a major component of the RNA degradation in the majority of the archaea species. The archaeal exosome exhibits a phosphorolytic and polymerase activity that is simply regulated by local concentration of ADP and inorganic phosphate. It also harbors a three-fold axis symmetry, which results in the presence of three identical active sites.

The aim of my thesis is to obtain more information on the mechanistic by which the archaeal exosome degrades RNA in solution. Indeed, a better comprehension of this mechanism is necessary to elucidate the regulation that discriminate between partial trimming from complete degradation or polymerization. In addition, understanding this mechanism would be helpful to explain the progressive loss of active sites in the exosome-like complexes throughout evolution.

Due to its relatively simple structural architecture, the archaeal exosome could be efficiently purified. It is also stable in isolation (especially the catalytic core), which considerably increases the range of applicable techniques. We decided then to focus on the hexameric core of the exosome from *Sulfolobus solfataricus* to clarify how it interacts with RNA during the catalytic process.

Among the different techniques used in this project, Nuclear Magnetic Resonance (NMR) experiments were decisive to detect the molecular motions, essential for comprehending protein functions. The importance of protein dynamics and the description of the main NMR experiments used to detect these dynamics are emphasized in the next paragraph, followed by the results section. Results from the work that forms the basis of this thesis have been published in international peer reviewed journals. In the results section, I will summarize these results and elaborate on experiments that were not part of the publications.

III. PROTEIN DYNAMICS

1. Structural studies: importance of the dynamics

Many of the discoveries mentioned before have been obtained with the help of crystal structures that brought a considerable amount of information. However, the paradigm structure ↔ function might not be sufficient to explain in detail the mechanisms that regulate these proteins.

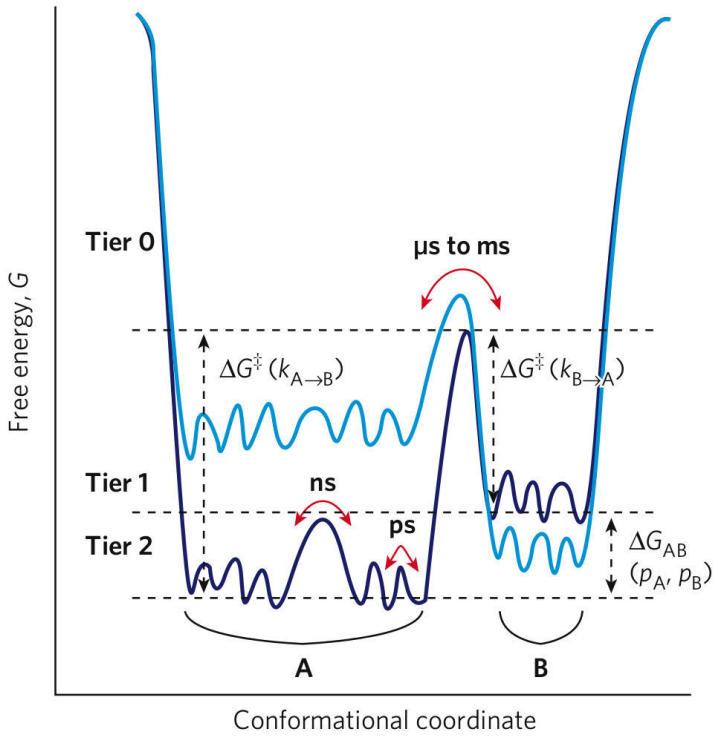

Figure 2: The conformational energy landscape of proteins.

A protein can be present in solution in a multitude of different conformations, separated from each other by energy barriers. The substrates and equilibrium fluctuations possess a hierarchy that can be separated in tiers, as described by Ansari & al. (47). Tier 0 dynamics corresponds to slow fluctuations between kinetically distinct states (here state A and state B). Within these tier-0 states, faster fluctuations between closely related states are observed and are classified in the lower tiers. A change in the system (ligand binding, mutation…) could affect the energy landscape (represented in light and dark blue). (Figure taken from Henzler-Wildman & Kern (48).)

In recent years, dynamics has been shown to play a crucial role in protein functions and led to the notion of energy landscape of proteins. The energy landscape of a protein defines the relative probabilities of the conformational states (thermodynamics) and the energy barriers between them (kinetics) (47)(Figure 2). From this emerges the concept of "conformational selection and population shift" (49) where the ligand does not induce the formation of a new structure (like in the induce fit model (50)) but, instead, selects a pre-existing structure. But the distinction between the two models is not absolute as an increasing number of cases show that conformational selection is often followed by conformational adjustment (51). The conformational selection, or in other words stabilization of a high-energy conformation, has been found for various types of interaction (protein-ligand, protein-protein, RNA and DNA interactions), regulation (where phosphorylation stabilize one conformer), and allostery (where ligand binding induces population redistribution) (reviewed by Boehr & al. (52)).

2. Detection of protein dynamics

The transient high-energy conformations, in dynamic with lower-enegy ground state conformations, are often populated to only low level and are therefore difficult to observe. The structures determined by X-ray crystallography only show one possible conformation and could hide the highest-energy conformers that we are interested in. However, these invisible states begin to be observable thanks to new advances in Nuclear Magnetism Resonance spectroscopy (NMR) techniques (53,54) and in single-molecule fluorescence energy transfer (FRET)(55). Whether they are directly observable or not, the different conformations are exchanging with different rates, depending on the type of structural changes (Figure 2 and 3).

Among the various techniques that could be used to detect dynamics (Molecular Dynamics simulations, FRET, time-resolved X-ray crystallography, NMR, Structural Mass Spectrometry) we decided to use Nuclear Magnetic Resonance spectroscopy (NMR), since it gives atomic resolution over a wide range of exchange process frequencies (56) (Figure 3). Until recently, the NMR applications were limited to rather small proteins but new methodologies

now enable the study of protein complexes that could reach up to 1.1 MDa (57).

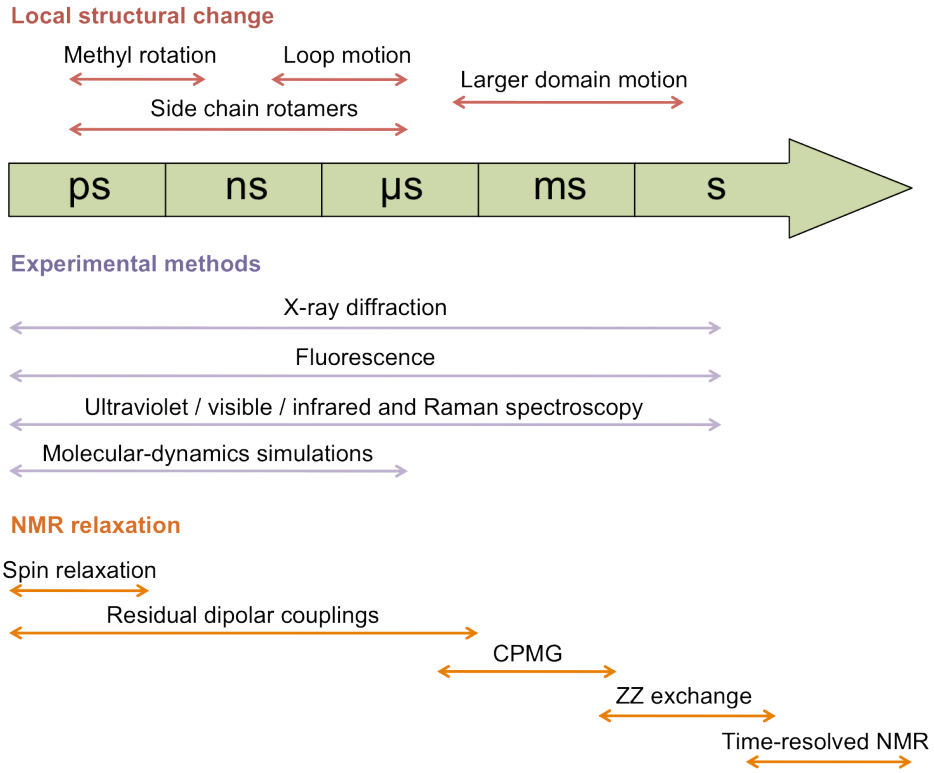

Figure 3: Time scales of protein dynamics in proteins and the different techniques to detect these fluctuations.

(Figure adapted from Henzler-Wildman & Kern (48) and from Bothe J.R & al. (58).)

IV. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

1. Brief overview of the NMR technique

NMR is a method that exploits the magnetic properties of certain nuclei; especially the nuclei with a spin ½ because they only possess a high and a low energy state. The most frequently spin ½ used in biomolecular NMR are the ^{1}H, ^{13}C, ^{15}N and ^{31}P that enable the proteins to be “NMR visible”. The sample is placed in a strong magnetic field (B_0) that, on average, align the nuclear spins in its direction. The low-energy state of the nuclei can undergo a transition to the high-energy state by the absorption of a photon that matches exactly the energy difference between the two states. This energy is provided to the sample by applying radio frequency pulses that are perpendicular to B_0. As a result, the nuclear spins are not aligned anymore with B_0 and the precession of the nuclear spins can be detected. The frequency of precession, called Larmor Frequency, is detected as a time-dependent current (sinusoid signal) named the Free Induction Decay (FID). The FID time domain signal (sinusoid signal) is finally converted, via the Fourier transformation, into a frequency domain (spectrum) with the different nuclear spin frequencies (chemical shifts).

The power of NMR spectroscopy results from the fact that each nucleus precesses at a unique frequency depending on the type of nucleus and on its local chemical environment. Solvent exposure, secondary structure, aromatic ring currents, bond torsion angles and hydrogen bonds are several examples of factors that determine the local magnetic field for each spin and thus make them distinguishable from each other.

The NMR experiments differ from each other by their pulse sequences, which defines the type of nucleus to excite, the magnetization transfer pathway from one nucleus to another, the suppression of unwanted signal and the selection of a specific quantum coherence. As an example, the pulse sequence of a Heteronuclear Multiple-Quantum Correlation (HMQC) experiment is described in the following section.

2. HMQC experiment

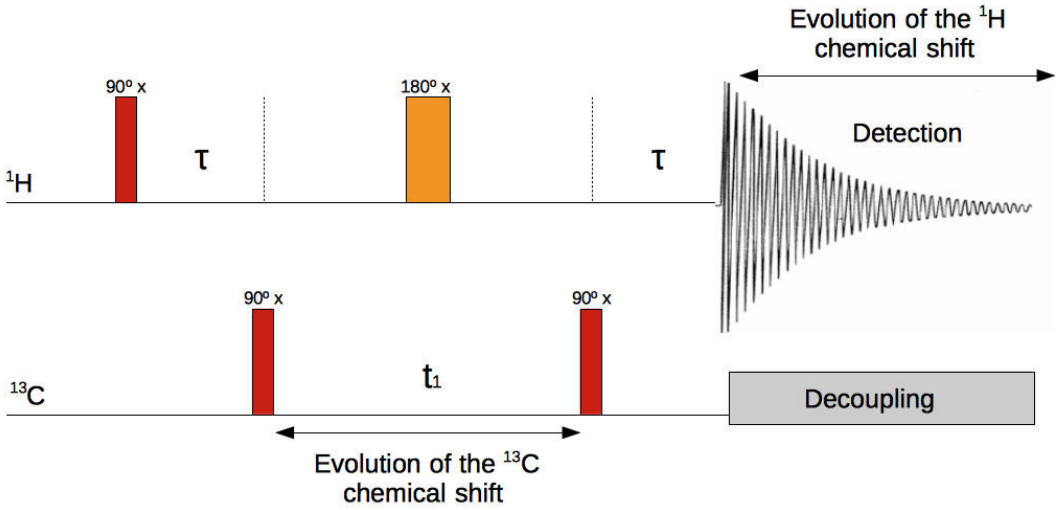

Figure 4: Pulse sequence of a HMQC experiment

In this pulse sequence, the homonuclear J-couplings between ^{1}H-^{1}H and ^{13}C-^{13}C are not suppressed and would lead to peak broadening. However, the samples we used only contain ^{1}H and ^{13}C nuclei in the methyl groups of selected amino acids (see below) in an otherwise fully ^{12}C and ^{2}H labeled background, therefore ^{1}H-^{1}H and ^{13}C-^{13}C J-couplings are not present. During detection, the ^{1}H-^{13}C J-couplings are suppressed with a decoupling pulse on the carbons to avoid splitting of the signal.

Let's assume that an HMQC pulse sequence (figure 4) is applied on an isoleucine. First, all the protons are excited with a 90 degree pulse that creates transverse magnetization. Each proton - H_N, Hα, Hβ, Hγ12, Hγ13, Hγ2, Hδ (illustrated in figure 5A) - evolves at their Larmor frequencies during the first period τ and under the effect of heteronuclear scalar coupling (connection of two spins through chemical bonds) from N, Cα, Cβ, Cγ1, Cγ2, and Cδ. Note that the first pulse creates coherence between the coupled ^{1}H and ^{13}C, which becomes maximal after the period τ. Then, the polarization is transferred from the ^{1}H to their coupled ^{13}C with a 90 degree pulse on the carbons which creates heteronuclear multiple-quantum coherences (giving its name to the pulse sequence) that doesn't allow evolution of the heteronuclear couplings. During the period t_1, we let the carbon nuclei evolving at their corresponding Larmor frequencies. The 180 degree pulse on the protons in

the middle of t_1 ensures that no proton chemical shifts are evolving during t_1. Then, the magnetization is transferred back to the proton with a 90 degree pulse on carbon, followed by a second period τ to get the optimal polarization transfer. Furthermore, the second period τ refocuses the chemical shift of the protons that evolved during the first period τ. During the acquisition time, the different chemical shifts of the protons are recorded, Fourier transformed and stored as a 1D spectrum. The same pulse sequence is applied to the sample with different periods of t_1 and stored separately. Since the modulation of the amplitude of the signals in the different 1D spectra is due to the evolution of carbon chemical shifts during the evolution time t_1, a second Fourier transformation is performed in the orthogonal dimension to extract the chemical shift of the coupled carbons.

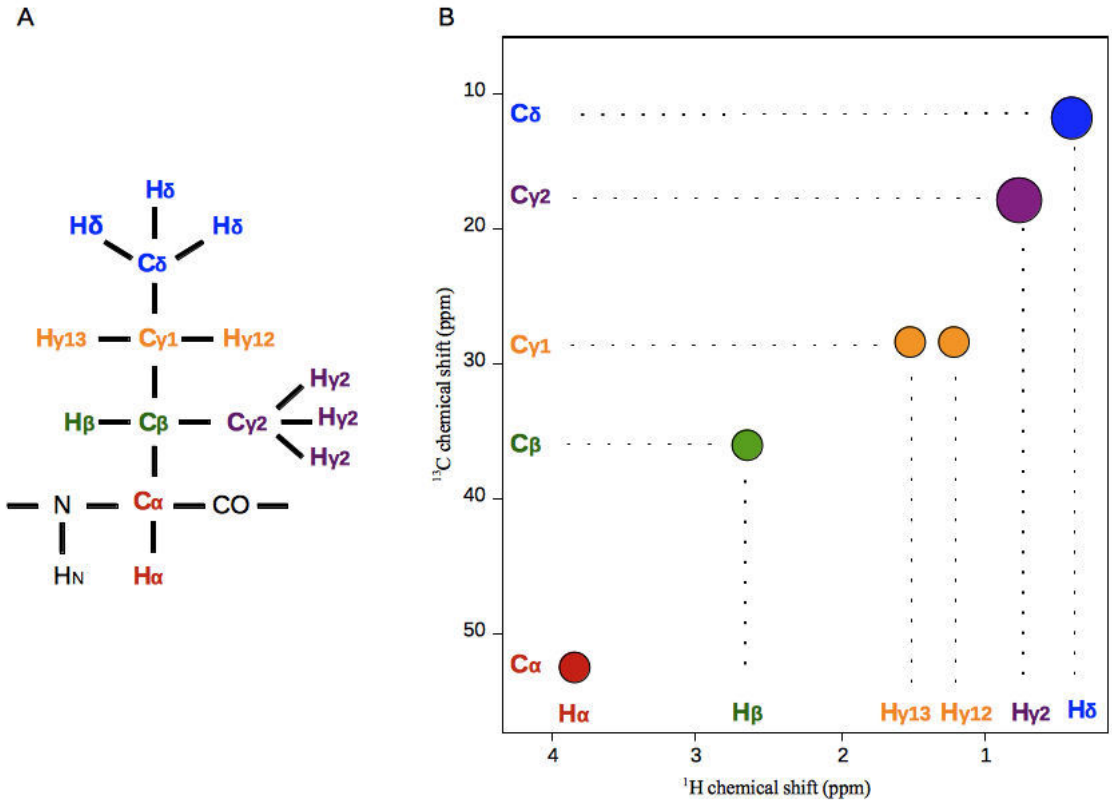

Figure 5: Theoretical HMQC spectrum of an isoleucine.

A) Schematic representation of an isoleucine involved in a peptide bond.

B) Corresponding spectrum of the isoleucine represented in A).

The spectrum of one isoleucine could, in principle, looks like the one depicted in figure 5B. The red peak corresponds to the proton Hα that is coupled to the Cα. The peak arises at the intersection of the chemical shift of the proton Hα and the chemical shift of the Cα. The same will be observed for the coupled Cβ and Hβ, Cγ2 and Hγ2, Cδ and Hδ. Note that the peak intensity of the two latter ones are more intense due to the fact that the three protons of the methyl groups are rotating so fast around the methyl symmetry axis that there are behaving as three identical protons. Unlike the methyl group protons, the two protons Hγ12 and Hγ13 give two different peaks. As they are coupled to the same carbon Cγ1, they appear at the same carbon chemical shift. However, the proximity of the carbon Cγ2 will influence the local magnetic field of the proton Hγ12 and leads to a slightly different proton chemical shift than Hγ13.

3. Relaxation in NMR

Relaxation refers to the process by which the spins in the sample return to equilibrium due to interactions with the surrounding. After the RF pulse, the loss of coherence among nuclei that rotate at their own Larmor frequency is called the transverse relaxation. In addition, the high-energy state of the nuclei converts back to the low energy state by transfer of energy to the surrounding; this phenomenon is referred to the spin-lattice relaxation. These two types of relaxation vary from one protein to another and influence the quality of the signal. Fast relaxation is one major cause of the decay of the FID, which eventually causes peak broadening in the spectrum. The size of the protein is the main limitation for NMR investigations because larger proteins induce faster relaxation. However, new methodologies have been developed to slow down the relaxation and extend the NMR applications to larger proteins.

4. Dipolar interaction

Due to its high sensitivity and its ubiquity in biomolecules, the proton is an essential nucleus in biomolecular NMR. However, the ubiquitous protons

also possess other properties that become a disadvantage for NMR. Indeed, because of its high magnetic dipole, the proton engenders dipolar interaction between spins especially the ^{1}H-^{1}H and the heteronuclear ^{1}H-^{13}C or ^{1}H-^{15}N. These dipolar interactions are the most common cause of relaxation of protons. Depending on their relative orientations with respect to the external filed B_0, the dipole-dipole interaction will slightly increase or decrease the local magnetic field. For small proteins, the dipolar coupling is close to zero due to their fast tumbling rates that tends to average out the relative orientation in the strong magnetic field. However, the slow tumbling rates of large molecules poorly average out these dipole-dipole interactions and some equivalent nuclei (nuclei with the same chemical shift) from different proteins will experience small differences in their local magnetic field. As a result, the precession of these nuclei will not be perfectly identical at all times and will lead to transverse relaxation.

There are two approaches to overcome the relaxation induced by the dipole-dipole interactions. First, the tumbling rate of large molecules could be increased by simply performing NMR experiments at higher temperatures. Second, the dipole-dipole interaction could be diminished by replacing the proton ^{1}H by deuterium ^{2}H (59,60). The interesting characteristic of deuterium lies in its gyromagnetic ratio that is 6.5 times lower than that of the proton. Since the dipolar interaction between ^{13}C or ^{15}N to an attached proton is proportional to the square of their gyromagnetic ratios, the benefits of replacing proton to deuterium in a protein appears obvious. In addition, the presence of deuterium in the solvent suppresses spin diffusion that takes place between protons from the protein and protons from the solvent (59).

5. Transverse Relaxation Optimized Spectroscopy (TROSY)

Another method to overcome the high relaxation of large molecules is the Transverse Relaxation Optimized Spectroscopy (TROSY)(61). As merely broached in the description of the HMQC pulse sequence, a pulse sequence has the aim to create coherence between coupled spins in order to correlate them in the spectrum. The type of coherence that is created depends on the pulse sequence and several types of coherences could be present at the

same time. The TROSY idea is to select only the coherences that relax slowly and to isolate (not mix) these from those that relax faster. In the HMQC, the coherence that is selected is the slow relaxing multiple quantum coherence. But at the same time, the HMQC experiment creates ^{1}H-^{1}H couplings that enhance relaxation, which is poorly averaged out for large molecules (see above). Therefore, for large molecule the HMQC experiment is a balance between slow relaxing coherence and relaxation induced by J couplings. It turns out that for large molecules, the benefits of the multiple quantum coherence is higher than the relaxation induced by the ^{1}H-^{1}H dipole relaxation, especially for methyl groups. Although the gain of sensitivity induced by the TROSY enlarged the range of proteins amenable for NMR investigation (62), large asymmetric protein complexes will also suffer from a significant amount of overlapping peaks.

6. Peak overlap

High molecular weight proteins or protein complexes means a high number of amino acids. Even by choosing one group per residue (such as the amide group), the number of unique resonances can be very large (100s), which can result in significant resonance overlap in two dimensional spectra. Reducing the number of peaks appears then as a necessity to study large molecular complexes by NMR. This could be achieved by labeling the protein with few selected amino acids. The methyl group labeling is an interesting option as several amino acids contain a methyl group (methionine, isoleucine, leucine, valine, threonine and alanine). Because of their hydrophobicity, these residues are found in the hydrophobic core of proteins and are therefore excellent probes of the structure (63). Conveniently, they are also widely spread over the primary sequence of the protein. In addition, the three-fold symmetry, together with the rapid rotation around the methyl symmetry axis, generates favorable relaxation properties (64), used by the HMQC (see above). Finally, the methyl groups are well dispersed in ^{1}H–^{13}C correlation map which lower even more the probability of peak overlap.

7. Methyl-TROSY-HMQC

Among the different NMR experiments developed for large proteins (62,65), the best approach to study very large protein complexes by NMR appears to be a combination of the strategies mentioned above. First, the protein is deuterated and combined with the methyl group labeling. The NMR visible nuclei (^{1}H and ^{13}C) are therefore only present in methyl group of selected amino acids. Deuteration also cancels the ^{1}H-^{1}H dipole relaxation that could arise from the use of HMQC. Relaxation is further reduced with the use of TROSY-HMQC (63) that select the slow relaxing multiple quantum coherence. Various biological assemblies have been studied using methyl TROSY spectroscopy (64), extending the range of NMR applications to very large molecular complexes.

All the NMR experiments presented in this thesis derive from the TROSY-HMQC and involve the methyl group labeling strategy. It enabled us to study the structural dynamics of the exosome from the archaea *Sulfolobus solfataricus* and, together with biochemical methods, helped us to characterize more precisely its mode of action.

V. RESULTS

1. The N-terminal helix of Rrp42 exhibits two conformations

A. Detection of the two conformations

The methyl group assignment of the isoleucines (δ1), leucines and valines (ILV) of Rrp42 in the monomeric form was previously obtained thanks to the work of Georg Dorn. The assignment of ILV residues from the monomeric Rrp42 was then used to facilitate the assignment of the ILV residues of Rrp42 in the hexameric complex (the “devide and conquer” approach, paper 1 figure S2A). It reveals that in the hexameric complex, five residues (Ile 10, Ile 13, Ile 19, Ile 27 and Ile 220) display two sets of peaks (figure 7; paper 1 figure 1) in contrast to the monomeric form of Rrp42 where these five residues display only one set of peaks. Except for the Ile 10, the peak intensity ratio between the two sets of peaks (refer to state A and state B) is about 3. Interestingly, these five residues cluster in the N-terminal helix of Rrp42 and could reflect a dynamics process in this region (figure 8; paper 1 figure 1). Since the detection of two peaks for a single residue is a hallmark of slow exchange on the NMR scale, we choose to perform longitudinal exchange experiment to characterize the dynamics of the N-terminal helix of Rrp42.

B. Quantification of the exchange rate

To extract the exchange rate between the two states, we used two independent NMR experiments: the longitudinal exchange (explained in figure 6 and the multiple quantum dispersion (explained in figure 7). We show that the N-terminal helix of Rrp42 experience a global conformational exchange (paper 1 figure 2) which is induced by the formation of the hexameric core. The two experiments give similar exchange rates which are 36 (± 1.9) per second for the longitudinal exchange and 44 (± 26.8) per second for the multiple quantum dispersion (paper 1 figure 2).

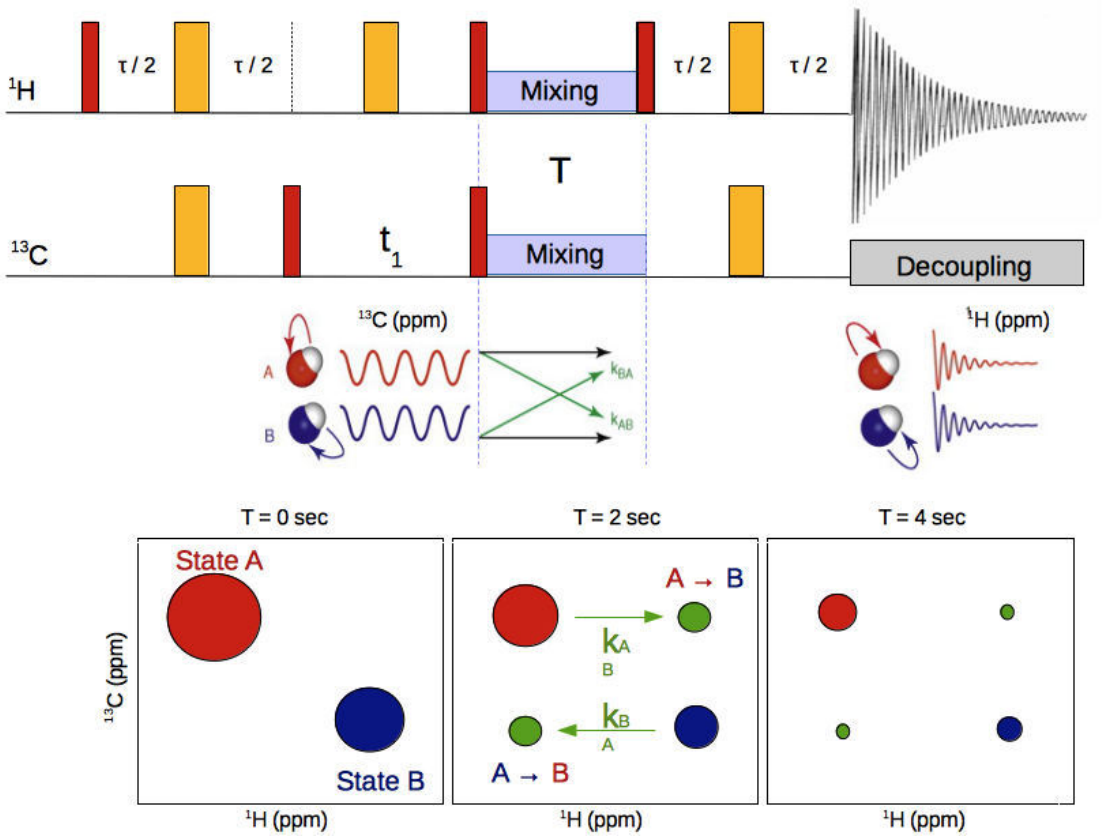

Figure 6: Principle of the longitudinal exchange experiment.

(Note that this experiment is also called magnetization exchange spectroscopy (EXSY) or zz-exchange.) The z-magnetization is not subject to the fast transverse relaxation but only to the slower spin-lattice relaxation. This allows the "stabilization" of the magnetization for up to ~5 sec and therefore the detection of slow exchange processes during the mixing time T.

Top) Because of the introduction of a mixing time period in the pulse sequence, the protons could not be efficiently refocused before the detection in a standard HMQC. Therefore, the heteronuclear polarization transfer is here performed by an INEPT that allows optimal refocusing of the protons before detection. The 90 degree proton pulse at the end of the INEPT is removed to create multiple quantum (MQ) coherence instead of single quantum (SQ) coherence. During the mixing time, the ^{13}C and ^{1}H spins are both in the z-axis, which allows exchange between the two states without chemical shift evolution or J-couplings. Bottom) During the mixing time T, the inter-conversion between the state A and the state B continues to occur and gives rise to a cross-peak at the ^{13}C chemical shift of A and the ^{1}H chemical shift of B, and vice versa. HMQC spectra are then recorded for different values of the mixing period T. The intensities of the auto and cross peaks are plotted as a function of the mixing times. The intensities of the cross-peaks (AB and BA) initially increase as a function of T, as the molecules are given more opportunity to exchange. For longer delays, the intensities of all peaks decrease, due to spin relaxation.

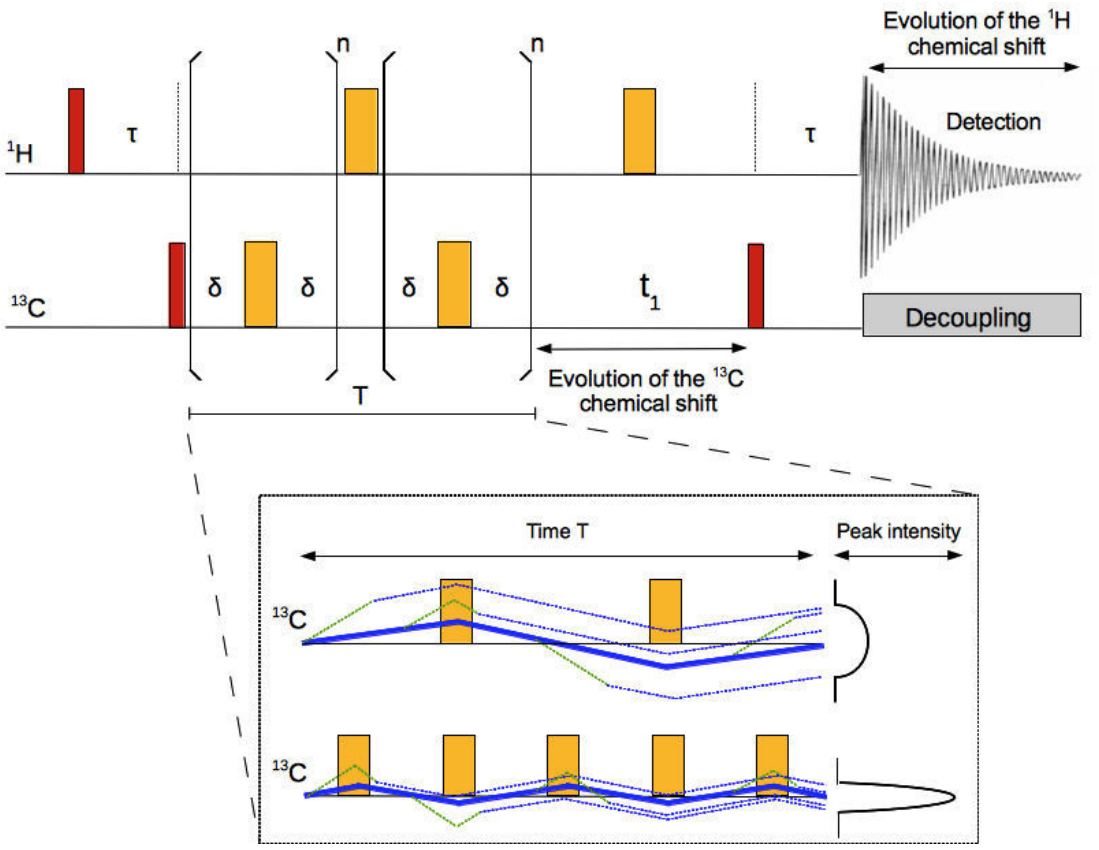

Figure 7: Multiple quantum dispersion experiment.

Top) Simplified pulse sequence of MQ dispersion experiment, also called Carr-Purcell-Meiboom-Gill (CPMG) dispersion experiment. During the period T, which stays constant, an increasing number of refocusing spin echo pulses (δ-180°-δ) is applied on carbons, with a 180 degree pulse on protons in the middle of the CPMG pulse train to refocus the chemical shift evolution of the protons during the period T. Bottom) Effect of different CPMG frequencies on the peak intensity. The blue lines correspond to the chemical shift evolution of the carbon in the state A, whereas the green dotted line corresponds to the chemical shift evolution of the carbon in the state B. The bold and solid blue line corresponds to the chemical shift evolution of the state A that, in absence of exchange into state B, is perfectly refocus regardless the number of applied spin echo pulses. Exchange into state B induces a different chemical shift evolution (dotted blue and green lines) that is not perfectly refocused anymore and leads to peak broadening. Increasing the number of spin echo pulses (CPMG frequency) diminishes the refocusing efficiency and eventually leads to stronger peak broadening. The exchange rates are extracted from the relation between the CPMG frequencies and the peak intensities.

In order to determine the importance of each state for the interaction with cap proteins, for the interactions with RNA and for it degradation, we first need to find the structural features that stabilize these two conformations. Indeed, understanding the structural origins that lead to the two conformations would enable the design of mutants where only one state is conserved; the role of the two states could be thus compared.

C. Finding the structural origins of the two Rrp42 conformations

The comparison of the different crystal structures obtained for the *S. solfataricus* exosome (figure 8) shows slight structural difference for only one region: the N-terminal tail and helix of Rrp42. Indeed, in the presence of Rrp4, a kink in the helix is present at the residue 21, which is not observed in the absence of Rrp4. Several mutations have been undertaken to obtain a mutant that stabilizes the state B conformer. Unfortunately, the HMQC spectra of these mutants show that all these mutations failed to eliminate the state A.

In the crystal structure, the residue that exhibits the kink of the N-terminal helix is the Ser 21. As it is not involved in any intra-molecular interaction, we decided to mutate it into an alanine, which is the residue with the highest α helical propensity. Unexpectedly, this mutation induces the destabilization of the state B (or stabilization of the state A). Similar results have been obtained for residues with a lower α helical propensity (S21G and S21R). More intriguing, the mutation of the Ser 21 into a proline hardly affected the population distribution. Position 21 thus does not seem to play an important role in determining the structural state of the Rrp42 helix.

To address the stability of the Rrp42 helix in more general terms, we used trifluoroethanol (TFE), an agent that favors and stabilizes α helices. TFE, was added to the exosome, however, no major effect on the populations were detected until the presence of 30% TFE, after which the exosome started to precipitate. Since the attempt of stabilizing the N-terminal helix was not successful, we decided to add a chaotropic agent to destabilize the helix. However, the addition of urea to the exosome affected the integrity of the whole Rrp42 preventing the extraction of any reliable information.

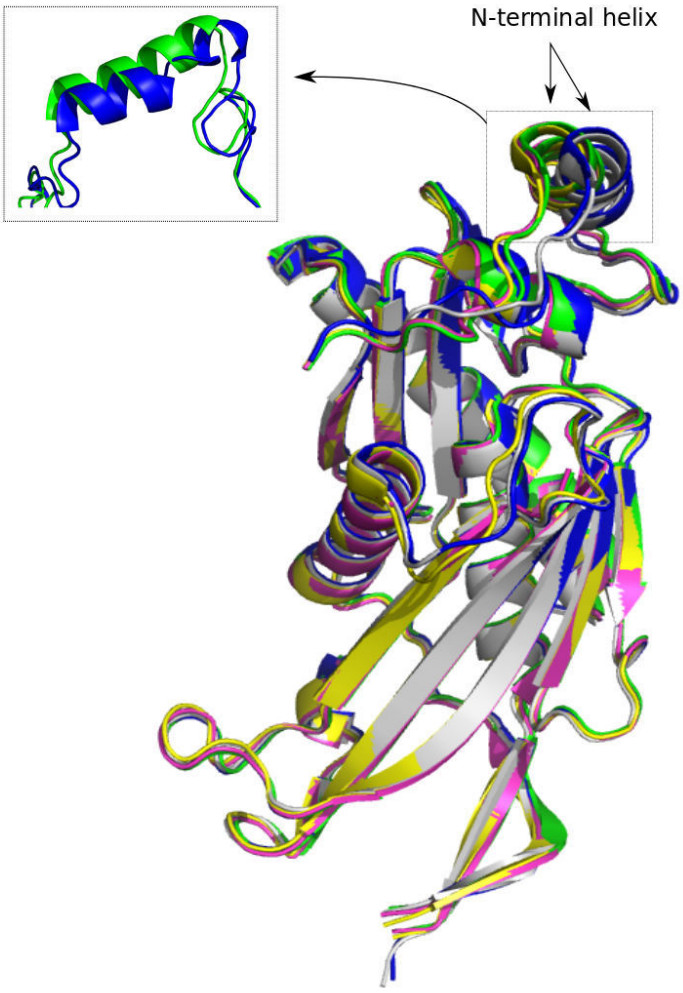

Figure 8: The N-terminal helix of Rrp42 with and without cap proteins.

Structure alignment of exosomes crystallized without Rrp4 (blue and grey; PDB numbers 2JE6 and 2JEA, respectively) and crystallized with Rrp4 (green, pink and yellow; PDB numbers 2BR2, 2C37 and 2C38, respectively).

In addition to the N-terminal helix, the N-terminal tail slightly differs in the crystal structures (Figure 8 and 9). However, all structures display an extended β-sheet between the Met 1, Ser 2 and Ser 3 of Rrp42 with the Leu 77, Arg 78 and Val 79 of Rrp41 (Figure 7). These interactions are absent in the monomeric form of Rrp42 and might potentially participate in the stabilization of conformation A or B. We thus decided to realize several truncations of the N-terminal tail of Rrp42. Interestingly, the removal of two or four amino acids (Δ1-2 and Δ1-4 mutants) resulted in a small shift of the population towards the state B but far from “locking” Rrp42 in the state B. The Δ1-6 mutant showed similar effect but to a lower extend. On the other hand,

the Δ1-8 and Δ1-10 mutants led to the complete loss of the state B. The different truncated versions of Rrp42 suggest that the first four N-terminal residues are implicated in the stabilization of the state A. Then, progressive removal of additional residues eventually leads to the loss of the state B meaning that the residues involved in the stabilization of state B are missing.

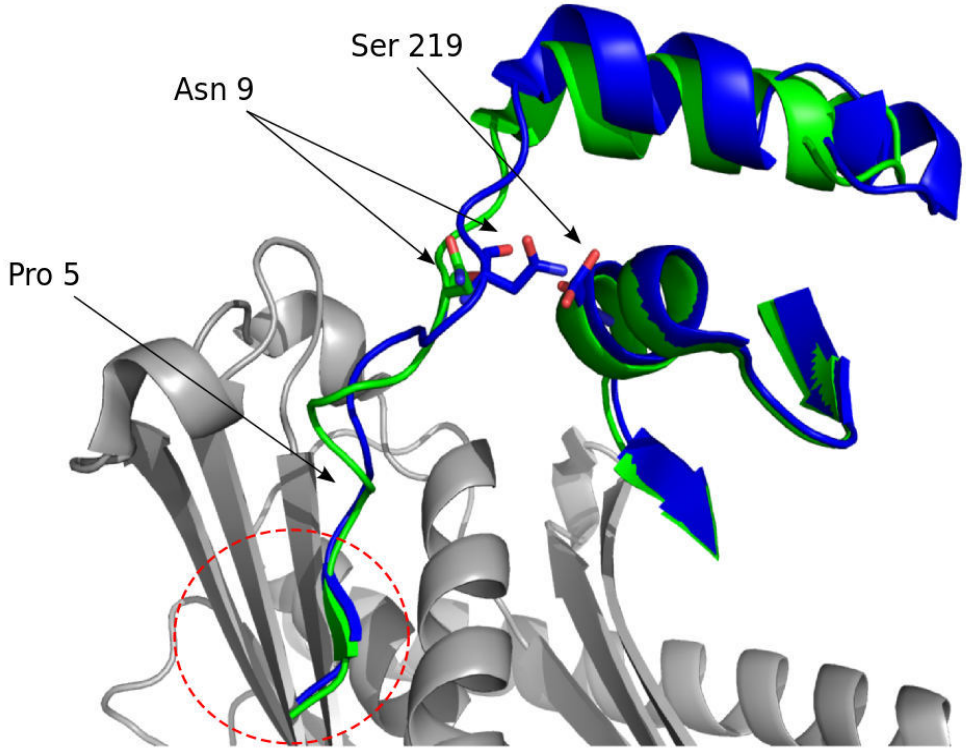

Figure 9: The N-terminal tail of Rrp42

Structure alignment of an exosome crystallized without Rrp4 (blue; PDB numbers 2JE6) and an exosome crystallized with Rrp4 (green; PDB numbers 2BR2). Rrp41 is colored in grey. The dashed circle indicates the region where the Met 1, Ser 2 and Ser 3 are involved in an extended b-sheet with Rrp41.

As we observe a destabilization of the state B already in the Δ1-6 mutant, it is possible that either the pro 5 and/or the Ser 6 are responsible for it. Because these two residues are not involved in any interaction in the crystal structure, one could speculate that the Pro 5 could participate to the state B formation through isomerization. The reason why this hypothesis is attractive is the fact that the energy barrier between the two conformations of Rrp42 has been extracted from the NMR experiments (paper 1 figure 2) and gave an activation enthalpy of 26.2 (± 0.8) kcal/mol, which is in range of

values observed for proline isomerization. The Pro 5 appears in the trans configuration in all the crystal structures. Nevertheless, the electron density of the Pro 5 in several structures could easily allow another conformation and does not guarantee that the Pro 5 is always in this configuration. Furthermore, a different orientation/configuration of Pro 5 would explain the different orientations of the Ser 6 displayed in these structures. On the other hand, the different orientations of the Ser 6 could be simply due to the flexibility of the N-terminal tail. The mutation of the Pro 5 would certainly provides experimental insight into the influence of the N-terminal tail in the stabilization of state B.

Another difference between the structures of the exosome is found in the vicinity of the Asn 9. Indeed, in the cap-free exosome, the side-chain amide groups of the Asn 9 are 2.6 Å and 3.9 Å away from the side-chain hydroxyl group of the Ser 219 of Rrp42, creating favorable conditions for hydrogen bonding. In contrast, the Asn 9 and the Ser 219 side-chain groups are 8 Å away in the Rrp4-exosome. The mutation of the Asn 9 into an alanine (N9A) provoked the loss of the state B conformer, probably due to the disruption of the interactions involved between the Asn 9 and the Ser 219. It thus suggests that these interactions are important to stabilize the state B conformer.

The mutagenesis experiments described above thus show that the two conformations in the N-terminal helix of Rrp42 don't result from a simple absence or presence of specific interactions. Nevertheless, we have been able to design a mutant where only the state A is present (N9A or state A mutant). Despite our intense efforts, we have not been able to "lock" Rrp42 at 100% into the state B conformation.

D. Chemical shift prediction

Assignment of an NMR spectrum just by using information available from a crystal structure could avoid the need to perform several time consuming NMR experiments (such as HNCOCA/HNCA, HNCOCB/HNCB or NOESY). MAP-XS (66) is a chemical shift prediction program developed to that end. We tested the program MAP-XS to assigned state A or state B in the spectra to the structure of Rrp42 with or without a kink in the helix. The correlation

between the experimental and the predicted methyl group chemical shifts of Rrp42 in the exosome were not conclusive (paper 1 figure S4). Since Yingqi Xu *& al.* obtained good results on the proteasome, the high molecular size of the exosome complex should not be the main limitation. Still, we decided to decrease the complexity of the system and we re-oriented our MAP-XS analysis to the monomeric form of Rrp42. Figure 10 shows a comparison of the predicted HMQC with the experimental one. Also for the Rrp42 monomer, the chemical shift prediction was poor. The reasons why the chemical shift prediction gives accurate data for the proteasome and not for the exosome remains unclear. For the current work, we were thus not able to use the program MAP-XS to link the resonances in our HMQC spectra directly to the structure of the complex.

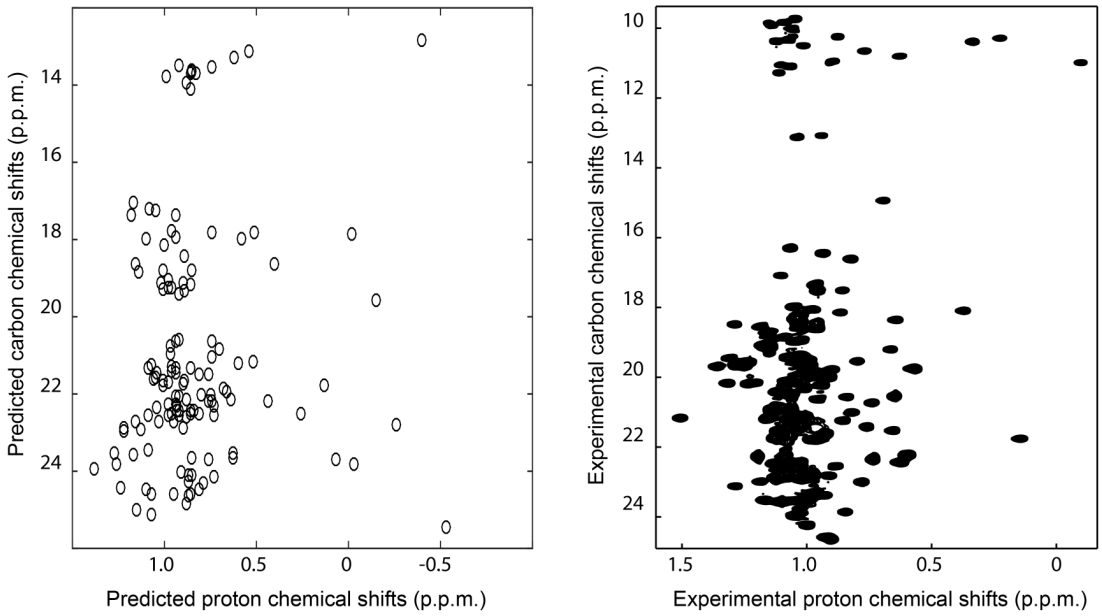

Figure 10: Chemical shift prediction

Comparison between the predicted (left) and the experimental (right) HMQC of the monomeric Rrp42. The methyl groups of isoleucines, leucines and valines were used for the comparison.

2. Importance of the different conformations

As mentioned above, the mutant N9A only exhibits the state A of Rrp42. This mutant is used to address the effect of the absence of the state B on cap binding (Rrp4 and Csl4), RNA binding and RNA catalysis.

A. Cap binding

As already mentioned in the introduction, we cannot determine whether the main conformer observed in crystallography corresponds to the low-energy conformer or not. Nonetheless, it is tempting to think that, in absence of the cap protein, the low-energy conformer is the one without the kink in the helix (state A).

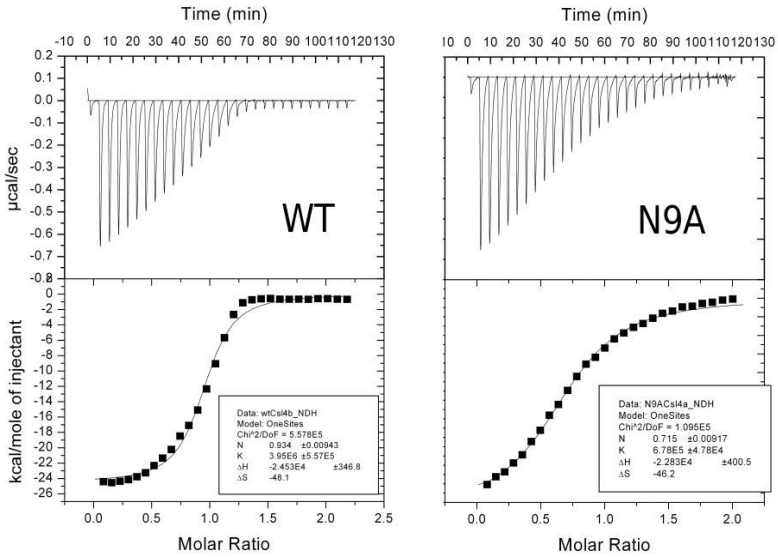

Figure 11: Interaction of the hexameric core with the cap protein Csl4.

ITC results obtained by the titration of 50 uM of Csl4 into 5 uM exosome . The extracted binding constant (kd) are 250 nM for the wild type (left) and 1.4 uM for the N9A mutant.

Our NMR experiments indicate that the high-energy conformer, which would correspond to the one with the kink in the helix (state B), pre-exists in solution. The crystal structures of the Rrp4-containing exosomes would then suggest that the state B becomes the low-energy conformer upon binding of Rrp4 by affecting the conformational energy landscape (Figure 2) of the hexameric core.

To test this hypothesis, we first recorded HMQC spectra of WT_exosome mixed with Rrp4 and Csl4, separately (paper 1 figure 4A, 4B). We observed that upon cap binding, only one state is present. The large chemical shift changes prevented us to determine which conformer is selected. The presence of only one conformation has been confirmed by the lack of dispersion of the N-terminal helix of Rrp42 in presence of the cap

proteins (paper 1 figure S3). However, the addition of Rrp4 to the N9A mutant results in the same chemical shift changes observed for the wild-type exosome (paper 1 figure 4C). It indicates that the transient state B is not required for binding Rrp4 and suggests that the selected conformer is the state A. Surface Plasmon Resonance (SPR) experiments (undertaken by Kerstin Reiss) indicate that both WT_exosome and N9A_exosome interacts very strongly to Rrp4 (Kd < 10 pM) and no difference in affinity could be observed. The state A is therefore the conformer selected upon Rrp4 binding.

Due to its tendency to aggregate over time, SPR experiments with Csl4 were not possible. Therefore, we decided to use Isothermal Titration Calorimetry (ITC) to determine the affinity of the exosomes for Csl4. We observed that the Csl4 binds slightly stronger to the N9A_exosome than to the WT_exosome with a Kd of 250 nM and 1.4 uM, respectively (figure 11). Albeit the state B is not absolutely required for binding Csl4, it does appear to facilitate its binding.

Unfortunately, the comparison with an exosome mutant that would stabilize the state B was not possible as the design of such mutant remained unsuccessful. So far, the function of the state B remains unclear. One possible rationale is that the state B is the conformer involved in the binding of a yet unknown partner of the exosome. The state B could also be the conformation of the monomeric Rrp42; the formation of the hexameric ring might shift Rrp42 towards the biological relevant state A and leads to the presence of these two states in equilibrium, in the absence of cap proteins. The last explanation is that the state B is the result of the flexibility required by the exosome for being able to bind different partners.

B. RNA also selects the state A conformer for binding

To address the effect of substrate binding to the exosome we recorded HMQC experiments in the presence of RNA. To prevent the RNA degradation while recording the NMR experiments, we used RNA with a 5' stem-loop followed by 20 As with a cyclic phosphate at the terminal 3' nucleotide. Indeed, the stem-loop cannot enter the catalytic chamber and the cyclic phosphate prevents RNA cleavage. As the exosome activity is phosphorolytic, HEPES

buffer was used. To establish that the RNA is indeed not degraded and is still properly interacting with the exosome, the NMR sample was subjected to an FPLC analysis after recording NMR experiments at 50°C (Figure 12).

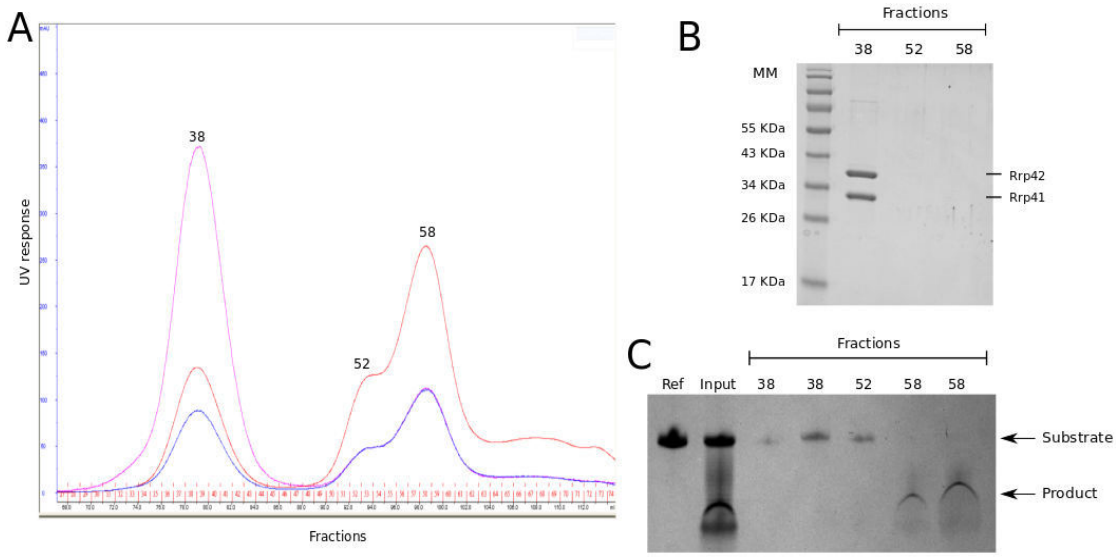

Figure 12: The RNA-exosome sample integrety.

A) Size Exclusion Chromatogram of the NMR sample. The blue, red and purple curves corresponds to the detection at different wavelengths, respectively 280nm, 260nm and 230nm. The 260/280 ratios < 2 at the fraction 38 confirms a mixture of protein and RNA. The 260/280 and 260/230 ratios of ~ 2 at the fractions 52 and 58 indicates that these fractions contains only RNA. However, the low 260/230 ratio at the fraction 38 is not really informative and could potentially indicates co-elution of salts with the exosome. B) and C) correspond respectively to 12% acrylamide (AA) SDS gel and denaturing 8M urea gel 8% AA of the fractions indicated in the chromatogram. Two different volumes (5 uL and 10 uL) have been loaded for the fractions 38 and 58.

It is worth noting that the exosome and the RNA were mixed in a respective 1:2 ratio. Fractions 38, 52 and 58 were run onto an SDS-PAGE gel to detect the exosome and also run onto a denaturing urea gel to detect the RNA. The totality of the exosome is found in fraction 38. The full length RNA is also found in the fraction 38, as well as in the fraction 52. Because the exosome is present in only one peak in the chromatogram, we can affirm that all the exosome molecules are bound to RNA. The fraction 52 corresponds then to an excess of RNA. Truncated species of the RNA were detected in the fraction 58, indicating residual degradation activity. The fractions corresponding to the first peak were collected and a HQMC was applied on

this sample; the spectrum is identical to the one recorded before the FPLC. We can conclude that chemical shift perturbations observed in Rrp42 are only induced by the substrate RNA bound to the exosome.

In the NMR spectra of the exosome in complex with RNA, the peaks corresponding to the state B were not detected anymore whereas the peaks corresponding to the state A display almost no chemical shift perturbations (paper 1 figure 4D). This indicates that the N-terminal helix of Rrp42 is locked in the state A upon RNA binding. State A is thus not only responsible for the interaction with the cap proteins, but also for the interaction with substrate.

C. RNA degradation

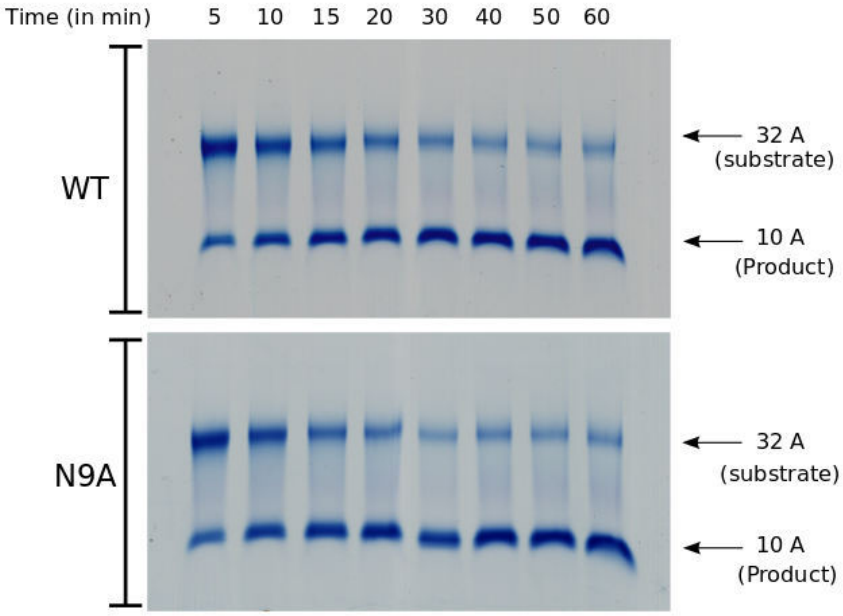

Figure 13: RNA degradation by the wild-type exosome and the N9A mutant.

The degradation reaction was followed by analyzing the band intensities of the RNA substrate and product, at different times, for the wild-type (top) and the state A mutant (N9A) (bottom). The substrate was separated from the product with a denaturing 8M urea 8% acrylamide (AA) gel. Similar degradation rates were obtained for the exosome wild-type and the state A mutant.

To address the effect of the state A mutation on the activity of the exosome we performed RNA degradation experiments. There, the RNA degradation by the state A mutant shows a similar pattern than the wild-type

exosome (Figure 13), reinforcing the idea that the state B conformer is not required for efficient RNA degradation.

3. Path of the RNA inside the hexameric ring

A. The RNA is mobile inside the catalytic core

Upon RNA binding, in addition to the selection of the state A conformer, we observed that the peak intensity of the Ile 85 of Rrp42 was dramatically decreased. This Ile 85 is located next to the nucleotide N4 of the substrate and is very likely to interact directly with the nucleotide N5. Therefore, Ile 85 is a perfect reporter of the interaction between the substrate RNA and the active sites. Due to the length of the substrate RNA used to study the interaction with the exosome, only one RNA could interact with an exosome unit. Normally, we should detect one peak for the Ile 85 free and one peak for the Ile 85 bound to the RNA, with a ratio of 2/3 and 1/3 respectively. Here, only one broad peak is detected, suggesting that the Ile 85 exchanges between the free and the RNA-bound state in the ms range. Indeed, when the exchange rate is similar to the difference of chemical shift between two states (called intermediate state or coalescence), only one broad peak appears close to the one with the highest populated state. To confirm this hypothesis, we performed CPMG Relaxation dispersion experiments (67,68) where we show that the 3' end of the substrate RNA is highly mobile and moves from one active site to another at a frequency of ~ 1700 per second at 50ºC (paper 2 figure 5A). The chemical shift difference between the Ile 85 free and the Ile 85 bound to the RNA was also extracted from the CPMG Relaxation experiment and gave Δ = 0.25 ppm. Using Chemical Exchange Saturation Transfer (described in Figure 14) experiments allowed us to “visualize” the invisible peak of Ile 85 bound to the RNA, 0.25 ppm downstream from the resonance frequency of the peak of the Ile 85 free (paper 2 figure 5B). The two independent experiments are in excellent agreement and confirm the reliability of our data. In order to saturate all the active sites and therefore induce the appearance of the Ile 85 bound peak, HMQC spectra were recorded with an excess of 7-mer RNAs, which are small enough to reach a

stoichiometry of 3:3 with the active sites per exosome unit. Unfortunately, the peak broadening was still too intense to observe the chemical shift of the Ile 85 bound. This shows that the small RNA does not stably bind to the active sites and is also exchanging rapidly between a free form and an exosome bound form.

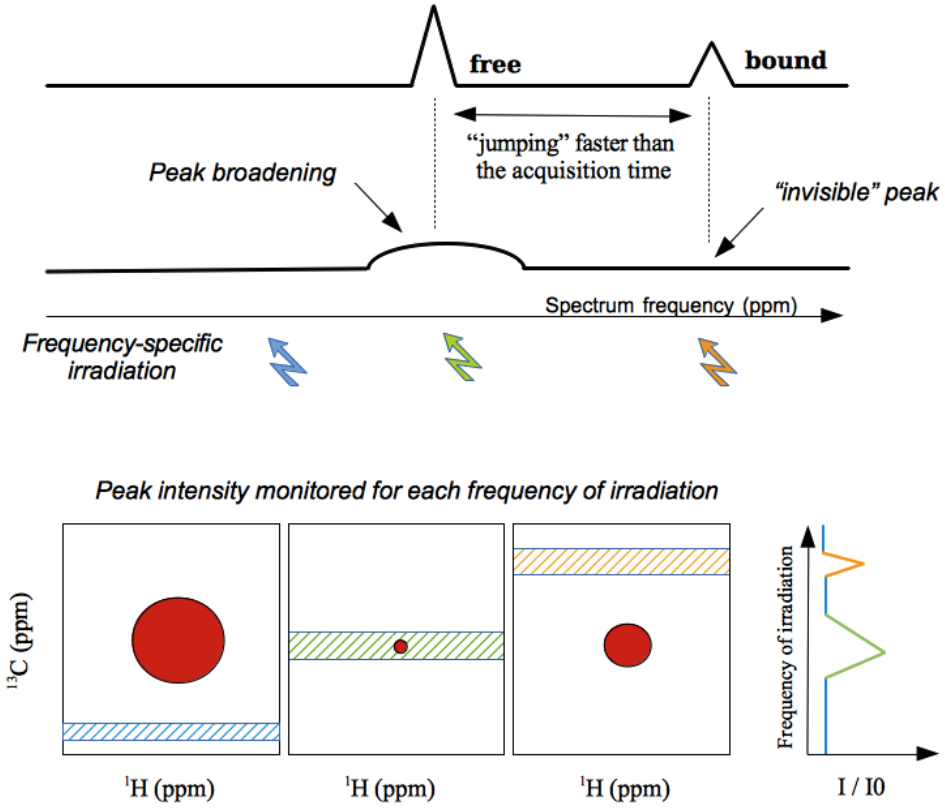

Figure 14: Principle of the CEST experiment

A) Because the exchange process is intermediate, only one broad peak is observed instead of one for the free and one for the bound state. The CEST experiment relies on the saturation of specific nuclei, achieved by irradiating at different frequencies of the spectrum. B) Three scenarios could result from the frequency-specific irradiation. HMQC spectra are recorded with different frequencies of irradiation and the effect on the peak intensities are monitored (right). If the frequency does not match the chemical shift (blue), then the peak intensity will be unaffected. On the contrary, if it does match the chemical shift (green), the peak intensity will severely drop. The interesting case is when we observe a decrease in the peak intensity by saturating another frequency than its chemical shift (orange). In that case, it means that another state (invisible) linked to the visible peak has been irradiated. When exchanging from the invisible to the visible state, the nuclei will still be partially saturated and will create a decrease in the peak intensity of the visible peak.

B. Additional probes in the catalytic chamber

The lack of electron density observed for the RNA between the active site and the neck region has been explained by the presence of motions inside the hexameric core. Indeed, based on Ile 85, we were able to observe and quantify the motions of the RNA from one active site to another. The fluorescence anisotropy experiments (paper 2 figure 2A and 3B) allowed us to state that the neck is important for the strong binding of the substrate RNA and for its processive degradation. Nevertheless, we are still missing information to clearly define the path of the RNA inside the barrel. The co-expression of Rrp41 and Rpr42 permitted the methyl group labeling of isoleucines and methionines of both Rrp41 and Rrp42. The formation of the hexameric core, while being over-expressed, significantly increased the yield of soluble Rrp41 after purification. The resulting spectrum contains thus the isoleucine and methionine peaks from both Rrp41 and Rrp42 (see figure 15).

Upon addition of RNA, we observe the same chemical shifts that are detected when only Rrp42 were labeled, that is to say the loss of the state B conformer as well as the peak broadening of the Ile 85 (figure 15, green and purple dashed box respectively). But peaks that belong to Rrp41 also display chemical shift perturbations (CSP). The Rrp41 Ile and Met methyl groups that seemed to be the best candidate to interact with RNA inside the lumen were mutated in order to assign them. Among all isoleucine and the methionine residues of Rrp41 that display CSPs, we were able to assign two: the Ile 101 and the Ile 131 (Figure 15; paper 2 figure 4C). Interestingly, upon RNA binding, theses two residues shift rather than broaden, indicating a fast binding-unbinding process. The Ile 131 is close to the nucleotide N1 and probably senses the RNA only when the RNA is dislocated from the active site. The Ile 101 is located between the neck and the active sites but could be away from the RNA at a distance that could reach up to 9 Å (see figure 3C paper 2). Altogether, these CSP suggests that, even bound to an active site, the RNA might remain flexible between the active sites and the neck.

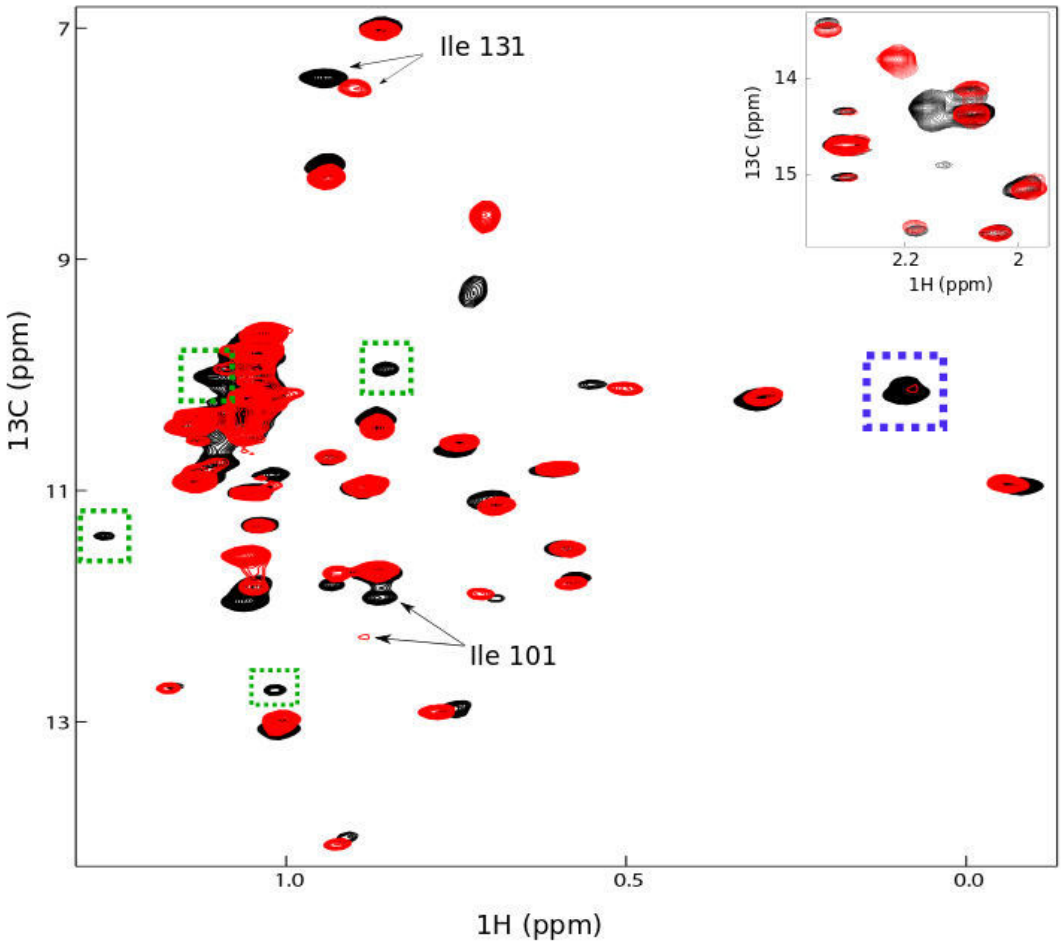

Figure 15: Isoleucine and methionine HMQC Spectrum of Rrp41 and Rrp42.
HMQC spectra in the absence (black) and in the presence of RNA (red) of the exosome where the methyl groups of isoleucines (main spectrum) and methionines (top-right spectrum) of Rrp41 and Rrp42 are both labeled. In presence of RNA, the peak corresponding to minor state of Ile 10, Ile 14, Ile 19 and Ile 220 disappear (dashed green boxes) and the Ile 85 peak undergo intense broadening (dashed purple box).

4. All active sites equally participate in the RNA degradation

To determine whether the RNA motions observed by NMR (see above) are important for RNA degradation or not, we decided to perform RNA degradation assays with exosomes that contain different amount of active sites. This implies the reconstitution of exosomes with active and inactive Rrp41 subunits. Prior to perform degradation assays, the dual activity of the exosome has to be checked to confirm that the exosomes that we reconstitute are functional.

A. Balance between polyadenylation and degradation

The exosome is able to perform RNA polynucleotidylation in addition to RNA degradation. To check that the exosome that we reconstitute was able to perform the polyadenylation reaction, we incubated RNA with the hexameric core of the exosome. As shown in the figure 16A, we can observe that in absence of ADP, the exosome degrades RNA. Addition of 15mM of ADP shifted the reaction towards polyadenylation. Since the buffer contained 10 mM of inorganic phosphate, more polyadenylated RNA products were detected before reaching the equilibrium. At equilibrium, all lengths of RNAs were detected but long RNA species (>150 nucleotides) and RNAs smaller than the original substrate were predominant. The dual function of the exosome is mainly regulated by the local concentrations of ADP and inorganic phosphate (69). As phosphorolytic degradation increases the concentration of ADP and decreases the concentration of inorganic phosphate over time, we carried out our degradation assays (see below) at times where these changes in concentration are negligible.

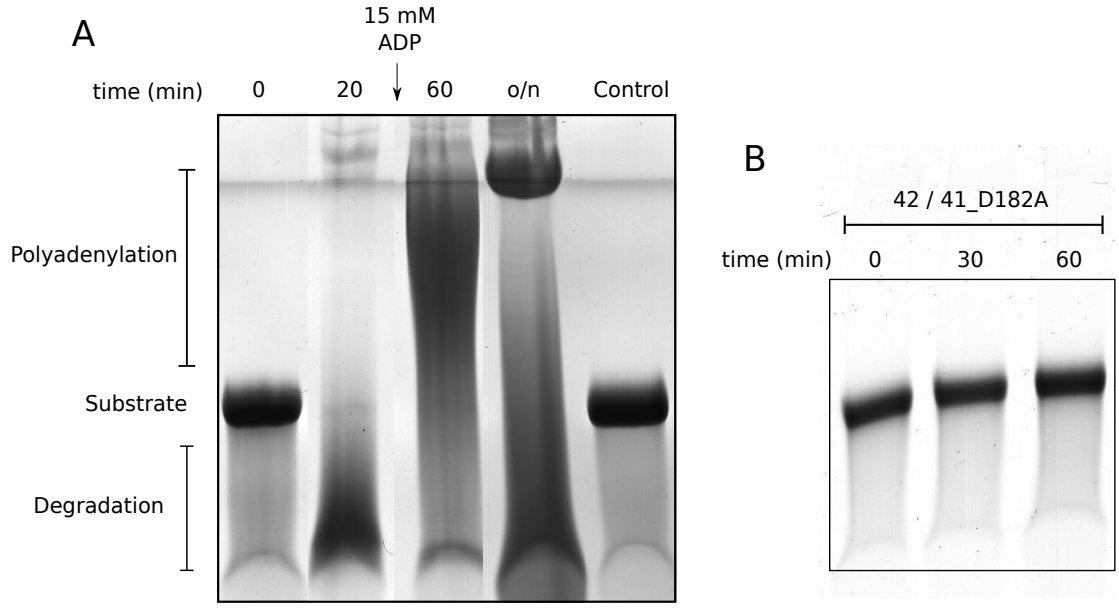

Figure 16: Balance between degradation and polymerization

A) Denaturing 8M urea 12% acrylamide (AA) gels. After 20 min, 15 mM of ADP was added to the degradation reaction. The reaction was followed after 60 min and overnight. The control corresponds to a sample without exosome and taken at the end of the reaction. B) Degradation assays performed with the inactive Rrp41.

B. The inactive exosome also binds RNA

To obtain exosomes with different amount of active sites, we reconstituted exosomes with variable ratios of inactive Rrp41. The inactivation of Rrp41 is achieved by the mutation of the Asp 182 into an alanine (4). The Asp 182 was originally thought to serve as the acid involved in the protonation of the new 3' end of the RNA, after the cleavage of the phosphodiester bond between the N1 and N2 nucleotides by the inorganic phosphate (70). However, recent studies indicate that the Asp 182 function would rather be the coordination of a divalent magnesium ion required for catalysis (71). As a result, the Rrp41_D182A mutant is completely inactive (Figure 15B). Indeed, HMQC experiments of the exosome containing Rrp41_D182A show selection of the Rrp42 state A, together with the peak broadening of the Ile 85 from Rrp42; two hallmarks of RNA binding by the exosome (paper 1 figure 4; figure 15). The observed peak broadening of Ile 85 suggests that the RNA binds to the active sites in the uM range. This is in agreement with experiments performed on the wild-type exosome (72). Therefore, the inactive D182A Rrp41 binds the substrate RNA with a similar affinity than the active Rrp41.

C. Degradation with different ratios of active Rrp41s

Exosomes with different amounts of inactive Rrp41 subunits (from 0 to 100%) were reconstituted to analyze the impact of the number of active sites on the degradation rates. The exosome degrades completely RNA in a processive manner, which results in the absence of intermediate products. In other words, degradation assays lead to the detection of only the substrate and the final degradation product (figure 12). The RNAs used in our degradation assays contain a 5’ hairpin that prevents the RNA from being completely degraded by the exosome. It also ensures that the substrate RNA long enough to interact with the neck and the active sites at the same time and hence maintain the degradation in the processive regime (73). The degradation rates were first analyzed by measuring the band intensity of the substrate and the product (illustrated in paper 2 figure S4). However, the data were not accurate enough to extract precise degradation rates. Consequently, we decided to

separate and quantify the RNA substrate and product by anion exchange combined with High Pressure Liquid Chromatography (HPLC). The later approach (illustrated in paper 2 figure S5) allowed the acquisition of precise data that unambiguously shows (73) that the degradation rates are linearly correlated with the number of active sites (paper 2 figure 6A, 6B). It reveals that each active site is equally used during RNA degradation.

D. Exosome with discrete number of active sites

To independently confirm these results, we designed a purification strategy to obtain exosomes with a discrete number of active sites (explained in figure 17), that is to say that exosomes with either 1, 2 or 3 active sites are in three separate samples. The degradation assays obtained with these exosomes gave very similar results than the ones with different ratios of inactive Rrp41s (paper 2 figure 6C, 6D). In both experiments, the exosome degrades about 10 nucleotides per second, which is approximately 100 times lower than the number of RNA: active site encounters (see above). Since the inactive form of Rrp41 is still able to bind the substrate RNA, the linear correlation between the degradation rates and the number of active sites is a consequence of the RNA motions inside the exosome core. The degradation assays with different number of active sites therefore confirm that the RNA motions detected by NMR does occur during catalysis. However, it does not necessarily mean that having three active sites is beneficial for the exosome. Indeed, the RNA concentration inside the hexameric core could be estimated to be around 100 mM and is therefore predicted to be in contact to one active site for >98% of the time.

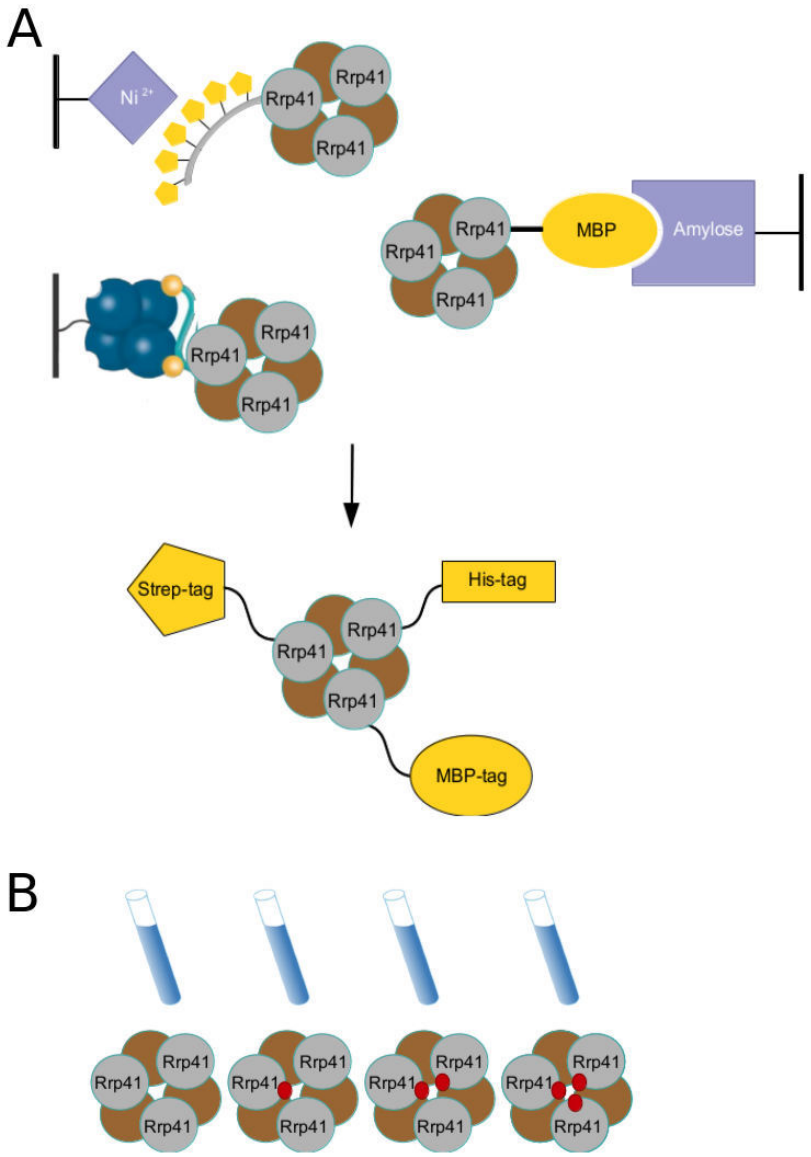

Figure 17: three tags strategy

A) Constructs co-expressing Rrp42 and three differently tagged Rrp41s are purified through three successive affinity columns (Ni-NTA, amylose and streptavidine resins). At the end, the purified exosomes will contain three differently tagged Rrp41 subunits. B) The combination of active/inactive Rrp41s in the construct enables to collect exosomes with 0, 1, 2 or 3 discrete active sites in separate samples.

VI. OLIGOMERIZATION AND EVOLUTION

Like the archaeal exosome, the Rnase PH forms a hexameric ring involved in tRNA maturation. It has been shown that the oligomerization of the Rnase PH is required for its activity in *Pseudomonas aeruginosa*, as the

dimeric form is not binding to the tRNA substrate (74). The Rnase PH is probably recognizing its substrate in a structure-specific manner (75), explaining why the dimeric form is completely inactive. Indeed, the oligomerization of the RNase PH induces the formation of a neck, absent in the dimer. The trimming of tRNA, whose structure consists of a stem-loop followed by a single-stranded tail, is then limited by the ring architecture of the exosome. For all exosome-like complexes, the neck is interacting with the single-stranded substrate RNA (40,70,76,77) and mutations located in the neck significantly decrease the activity of the E.coli PNPase (3) and of the S. solfataricus exosome (paper 2 figure 3A).

Another consequence of the oligomerization is the formation of a catalytic chamber, a lumen where the high concentration of active sites is sequestered from the environment. When threaded to the active sites, the substrate RNA is restraint into the catalytic chamber. The concentration of the RNA inside the catalytic chamber becomes very high and ensures a quasi-permanent binding of the RNA to an active site, even if only one active site would be present in the chamber. This might explain the progressive loss of active sites over evolution; from the RNase PH with 6 active sites, the PNPase and archaeal exosome with 3 active sites, to the eukaryotic exosome that only act as a scaffold and requires an additional subunit that possess endo-exo hydrolytic activity.

As for the eukaryotic exosome, the reason to lose the polynucleotidylation, and consequently the phosphorolytic activity, might be the appearance in eukaryotes of poly(A) polymerases that took over the role of the exosome for the polyadenylation-dependent degradation (9). Another reason for further evolution into inactive scaffold could be the higher degree of regulation of its activity. Indeed, in eukaryotic cells, the exosome is required in many diverse pathway (17) and two additional active subunits allow, in addition to the various co-factors, a better control of the catalytic potential over time and space (17,78).

REFERENCES

1. Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A. and Gleizes, P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA*, **6**, 225-242.
2. Evguenieva-Hackenberg, E., Walter, P., Hochleitner, E., Lottspeich, F. and Klug, G. (2003) An exosome-like complex in Sulfolobus solfataricus. *EMBO Rep*, **4**, 889-893.
3. Shi, Z., Yang, W.Z., Lin-Chao, S., Chak, K.F. and Yuan, H.S. (2008) Crystal structure of Escherichia coli PNPase: central channel residues are involved in processive RNA degradation. *RNA*, **14**, 2361-2371.
4. Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G. and Conti, E. (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat Struct Mol Biol*, **12**, 575-581.
5. Liu, Q., Greimann, J.C. and Lima, C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223-1237.
6. Harlow, L.S., Kadziola, A., Jensen, K.F. and Larsen, S. (2004) Crystal structure of the phosphorolytic exoribonuclease RNase PH from Bacillus subtilis and implications for its quaternary structure and tRNA binding. *Protein Sci*, **13**, 668-677.
7. Hartung, S. and Hopfner, K.P. (2009) Lessons from structural and biochemical studies on the archaeal exosome. *Biochem Soc Trans*, **37**, 83-87.
8. Januszyk, K. and Lima, C.D. (2010) Structural components and architectures of RNA exosomes. *Adv Exp Med Biol*, **702**, 9-28.
9. Slomovic, S., Portnoy, V., Yehudai-Resheff, S., Bronshtein, E. and Schuster, G. (2008) Polynucleotide phosphorylase and the archaeal exosome as poly(A)-polymerases. *Biochim Biophys Acta*, **1779**, 247-255.
10. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. and Tollervey, D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell*, **91**, 457-466.
11. Dziembowski, A., Lorentzen, E., Conti, E. and Seraphin, B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat Struct Mol Biol*, **14**, 15-22.
12. Ishii, R., Nureki, O. and Yokoyama, S. (2003) Crystal structure of the tRNA processing enzyme RNase PH from Aquifex aeolicus. *J Biol Chem*, **278**, 32397-32404.
13. Symmons, M.F., Jones, G.H. and Luisi, B.F. (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure*, **8**, 1215-1226.
14. Chekanova, J.A., Shaw, R.J., Wills, M.A. and Belostotsky, D.A. (2000) 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. *J Biol Chem*, **275**, 33158-33166.
15. Schmid, M. and Jensen, T.H. (2008) The exosome: a multipurpose RNA-decay machine. *Trends Biochem Sci*, **33**, 501-510.

16. Butler, J.S. (2002) The yin and yang of the exosome. *Trends Cell Biol*, **12**, 90-96.
17. Houseley, J., LaCava, J. and Tollervey, D. (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol*, **7**, 529-539.
18. Allmang, C., Mitchell, P., Petfalski, E. and Tollervey, D. (2000) Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res*, **28**, 1684-1691.
19. Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. and Tollervey, D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J*, **18**, 5399-5410.
20. Norbury, C.J. (2013) Cytoplasmic RNA: a case of the tail wagging the dog. *Nat Rev Mol Cell Biol*, **14**, 643-653.
21. Chen, H.W., Rainey, R.N., Balatoni, C.E., Dawson, D.W., Troke, J.J., Wasiak, S., Hong, J.S., McBride, H.M., Koehler, C.M., Teitell, M.A. *et al.* (2006) Mammalian polynucleotide phosphorylase is an intermembrane space RNase that maintains mitochondrial homeostasis. *Mol Cell Biol*, **26**, 8475-8487.
22. Borowski, L.S., Dziembowski, A., Hejnowicz, M.S., Stepien, P.P. and Szczesny, R.J. (2013) Human mitochondrial RNA decay mediated by PNPase-hSuv3 complex takes place in distinct foci. *Nucleic Acids Res*, **41**, 1223-1240.
23. Yehudai-Resheff, S., Hirsh, M. and Schuster, G. (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. *Mol Cell Biol*, **21**, 5408-5416.
24. Gagliardi, D., Stepien, P.P., Temperley, R.J., Lightowlers, R.N. and Chrzanowska-Lightowlers, Z.M. (2004) Messenger RNA stability in mitochondria: different means to an end. *Trends Genet*, **20**, 260-267.
25. Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T. and Condon, C. (2007) 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell*, **129**, 681-692.
26. Celesnik, H., Deana, A. and Belasco, J.G. (2007) Initiation of RNA decay in Escherichia coli by 5' pyrophosphate removal. *Mol Cell*, **27**, 79-90.
27. Mohanty, B.K. and Kushner, S.R. (2000) Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in Escherichia coli. *Proc Natl Acad Sci U S A*, **97**, 11966-11971.
28. Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S. and Arraiano, C.M. (2009) The role of 3'-5' exoribonucleases in RNA degradation. *Prog Mol Biol Transl Sci*, **85**, 187-229.
29. Coburn, G.A. and Mackie, G.A. (1998) Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. *J Mol Biol*, **279**, 1061-1074.
30. Carpousis, A.J. (2007) The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol*, **61**, 71-87.
31. Kaberdin, V.R. and Lin-Chao, S. (2009) Unraveling new roles for minor components of the E. coli RNA degradosome. *RNA Biol*, **6**, 402-405.
32. Rott, R., Zipor, G., Portnoy, V., Liveanu, V. and Schuster, G. (2003) RNA polyadenylation and degradation in cyanobacteria are similar to the

chloroplast but different from Escherichia coli. *J Biol Chem*, **278**, 15771-15777.
33. Deutscher, M.P. (2006) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Res*, **34**, 659-666.
34. Cheng, Z.F. and Deutscher, M.P. (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. *Proc Natl Acad Sci U S A*, **100**, 6388-6393.
35. Levy, S., Portnoy, V., Admon, J. and Schuster, G. (2011) Distinct activities of several RNase J proteins in methanogenic archaea. *RNA Biol*, **8**, 1073-1083.
36. Evguenieva-Hackenberg, E., Hou, L., Glaeser, S. and Klug, G. (2014) Structure and function of the archaeal exosome. *Wiley Interdiscip Rev RNA*, **5**, 623-635.
37. Koonin, E.V., Wolf, Y.I. and Aravind, L. (2001) Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. *Genome Res*, **11**, 240-252.
38. Portnoy, V., Evguenieva-Hackenberg, E., Klein, F., Walter, P., Lorentzen, E., Klug, G. and Schuster, G. (2005) RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polynucleotidylates RNA in Sulfolobus. *EMBO Rep*, **6**, 1188-1193.
39. Portnoy, V. and Schuster, G. (2006) RNA polyadenylation and degradation in different Archaea; roles of the exosome and RNase R. *Nucleic Acids Res*, **34**, 5923-5931.
40. Lorentzen, E., Dziembowski, A., Lindner, D., Seraphin, B. and Conti, E. (2007) RNA channelling by the archaeal exosome. *EMBO Rep*, **8**, 470-476.
41. Witharana, C., Roppelt, V., Lochnit, G., Klug, G. and Evguenieva-Hackenberg, E. (2012) Heterogeneous complexes of the RNA exosome in Sulfolobus solfataricus. *Biochimie*, **94**, 1578-1587.
42. Roppelt, V., Klug, G. and Evguenieva-Hackenberg, E. (2010) The evolutionarily conserved subunits Rrp4 and Csl4 confer different substrate specificities to the archaeal exosome. *FEBS Lett*, **584**, 2931-2936.
43. Hou, L., Klug, G. and Evguenieva-Hackenberg, E. (2013) The archaeal DnaG protein needs Csl4 for binding to the exosome and enhances its interaction with adenine-rich RNAs. *RNA Biol*, **10**, 415-424.
44. Zuo, Z., Rodgers, C.J., Mikheikin, A.L. and Trakselis, M.A. (2010) Characterization of a functional DnaG-type primase in archaea: implications for a dual-primase system. *J Mol Biol*, **397**, 664-676.
45. Roppelt, V., Hobel, C.F., Albers, S.V., Lassek, C., Schwarz, H., Klug, G. and Evguenieva-Hackenberg, E. (2010) The archaeal exosome localizes to the membrane. *FEBS Lett*, **584**, 2791-2795.
46. Hou, L., Klug, G. and Evguenieva-Hackenberg, E. (2014) Archaeal DnaG contains a conserved N-terminal RNA-binding domain and enables tailing of rRNA by the exosome. *Nucleic Acids Res*, **42**, 12691-12706.
47. Ansari, A., Berendzen, J., Bowne, S.F., Frauenfelder, H., Iben, I.E., Sauke, T.B., Shyamsunder, E. and Young, R.D. (1985) Protein states and proteinquakes. *Proc Natl Acad Sci U S A*, **82**, 5000-5004.
48. Henzler-Wildman, K. and Kern, D. (2007) Dynamic personalities of proteins. *Nature*, **450**, 964-972.

49. Ma, B., Kumar, S., Tsai, C.J. and Nussinov, R. (1999) Folding funnels and binding mechanisms. *Protein Eng*, **12**, 713-720.
50. Koshland, D.E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci U S A*, **44**, 98-104.
51. Csermely, P., Palotai, R. and Nussinov, R. (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci*, **35**, 539-546.
52. Boehr, D.D., Nussinov, R. and Wright, P.E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol*, **5**, 789-796.
53. Hansen, D.F., Vallurupalli, P. and Kay, L.E. (2008) Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J Biomol NMR*, **41**, 113-120.
54. Bouvignies, G. and Kay, L.E. (2012) A 2D (1)(3)C-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding. *J Biomol NMR*, **53**, 303-310.
55. Diez, M., Zimmermann, B., Borsch, M., Konig, M., Schweinberger, E., Steigmiller, S., Reuter, R., Felekyan, S., Kudryavtsev, V., Seidel, C.A. *et al.* (2004) Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. *Nat Struct Mol Biol*, **11**, 135-141.
56. Mittermaier, A. and Kay, L.E. (2006) New tools provide new insights in NMR studies of protein dynamics. *Science*, **312**, 224-228.
57. Sprangers, R. and Kay, L.E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, **445**, 618-622.
58. Bothe, J.R., Nikolova, E.N., Eichhorn, C.D., Chugh, J., Hansen, A.L. and Al-Hashimi, H.M. (2011) Characterizing RNA dynamics at atomic resolution using solution-state NMR spectroscopy. *Nat Methods*, **8**, 919-931.
59. Sattler, M. and Fesik, S.W. (1996) Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure*, **4**, 1245-1249.
60. Gardner, K.H. and Kay, L.E. (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct*, **27**, 357-406.
61. Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A*, **94**, 12366-12371.
62. Tzakos, A.G., Grace, C.R., Lukavsky, P.J. and Riek, R. (2006) NMR techniques for very large proteins and rnas in solution. *Annu Rev Biophys Biomol Struct*, **35**, 319-342.
63. Tugarinov, V., Hwang, P.M., Ollerenshaw, J.E. and Kay, L.E. (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J Am Chem Soc*, **125**, 10420-10428.
64. Wiesner, S. and Sprangers, R. (2015) Methyl groups as NMR probes for biomolecular interactions. *Curr Opin Struct Biol*, **35**, 60-67.
65. Foster, M.P., McElroy, C.A. and Amero, C.D. (2007) Solution NMR of large molecules and assemblies. *Biochemistry*, **46**, 331-340.

66. Xu, Y., Liu, M., Simpson, P.J., Isaacson, R., Cota, E., Marchant, J., Yang, D., Zhang, X., Freemont, P. and Matthews, S. (2009) Automated assignment in selectively methyl-labeled proteins. *J Am Chem Soc*, **131**, 9480-9481.
67. Lundstrom, P., Vallurupalli, P., Religa, T.L., Dahlquist, F.W. and Kay, L.E. (2007) A single-quantum methyl 13C-relaxation dispersion experiment with improved sensitivity. *J Biomol NMR*, **38**, 79-88.
68. Korzhnev, D.M., Kloiber, K., Kanelis, V., Tugarinov, V. and Kay, L.E. (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J Am Chem Soc*, **126**, 3964-3973.
69. Evguenieva-Hackenberg, E., Roppelt, V., Finsterseifer, P. and Klug, G. (2008) Rrp4 and Csl4 are needed for efficient degradation but not for polyadenylation of synthetic and natural RNA by the archaeal exosome. *Biochemistry*, **47**, 13158-13168.
70. Lorentzen, E. and Conti, E. (2005) Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol Cell*, **20**, 473-481.
71. Lorentzen, E. and Conti, E. (2012) Crystal structure of a 9-subunit archaeal exosome in pre-catalytic states of the phosphorolytic reaction. *Archaea*, **2012**, 721869.
72. Oddone, A., Lorentzen, E., Basquin, J., Gasch, A., Rybin, V., Conti, E. and Sattler, M. (2007) Structural and biochemical characterization of the yeast exosome component Rrp40. *EMBO Rep*, **8**, 63-69.
73. Hartung, S., Niederberger, T., Hartung, M., Tresch, A. and Hopfner, K.P. (2010) Quantitative analysis of processive RNA degradation by the archaeal RNA exosome. *Nucleic Acids Res*, **38**, 5166-5176.
74. Choi, J.M., Park, E.Y., Kim, J.H., Chang, S.K. and Cho, Y. (2004) Probing the functional importance of the hexameric ring structure of RNase PH. *J Biol Chem*, **279**, 755-764.
75. Li, Z., Pandit, S. and Deutscher, M.P. (1998) 3' exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in Escherichia coli. *Proc Natl Acad Sci U S A*, **95**, 2856-2861.
76. Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E. and Conti, E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell*, **139**, 547-559.
77. Navarro, M.V., Oliveira, C.C., Zanchin, N.I. and Guimaraes, B.G. (2008) Insights into the mechanism of progressive RNA degradation by the archaeal exosome. *J Biol Chem*, **283**, 14120-14131.
78. Lykke-Andersen, S., Tomecki, R., Jensen, T.H. and Dziembowski, A. (2011) The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. *RNA Biol*, **8**, 61-66.

ACKNOWLEDGMENTS

First, I will be very original and thank my supervisor for giving me the opportunity to work in such a fantastic atmosphere. I consider myself lucky to have worked in the Sprangers' lab where I have enjoyed doing my Ph.D. I particularly thank Georg Braach (mostly known as Georg Dorn in Switzerland) for starting the whole archaeal exosome project. I also thank Kerstin Reiss for her collaboration and all the members of the Wiesner's lab for the nice (non)scientific talks. Enough of work-related acknowledgment! Now, I would like to mention the people that made my Ph.D. a very pleasant experience:

- The Heidelberg pre-doc trio:
Nik, the “Walter tiger” who stays “exactly similar”
Markus, who was nice enough to always let me win our badminton games.
Simon, for all the rugby related activities and for being my host in Heidelberg.

- Miloš for joining me in the exosome studies and for the Porto tasting.

- Nina, for keeping the lab organized and a relaxed place to work in.

- Ancilla, for the English tea, the countless advices and our skiing trips in Parpan.

- Philip, the hiccup master, for his German translations.

- Stefan, for surviving next to me in the office and for sharing the excitement of the 'tour de France'.

- Lisa and Anna-Lisa, the best impact players of our MPI football team.

- Magnus, for arriving 5 minutes after me in the Alb-Gold cycling race.

- Erik, Jan and Daniel, the main protagonists of the senseless talks in the lab!

- Thanks as well to our neighbors, the plant people, for making the Christmas bowling exciting, already two weeks before it happens.

- Of course to my parents, for all their support during my studies.

PEER-REVIEWED ARTICLES

Paper 1:

Audin MJ, Dorn G, Fromm SA, Reiss K, Schütz S, Vorländer MK, Sprangers R. (2013).
The archaeal exosome: identification and quantification of site-specific motions that correlate with cap and RNA binding.
Angew. Chem. Int. Ed. Engl. 2013 Aug 5;52(32):8312-6

Paper 2:

Maxime J. Audin, Jan Philip Wurm, Milos A. Cvetkovic and Remco Sprangers (2016)
The oligomeric architecture of the archaeal exosome is important for processive and efficient RNA degradation.
Nucleic Acids Res. 2016 Apr 7;44(6):2962-73

Molecular Machines

DOI: 10.1002/anie.201302811

The Archaeal Exosome: Identification and Quantification of Site-Specific Motions That Correlate with Cap and RNA Binding**

*Maxime J. C. Audin, Georg Dorn, Simon A. Fromm, Kerstin Reiss, Stefan Schütz, Matthias K. Vorländer, and Remco Sprangers**

Large molecular machines perform many cellular processes and it is of fundamental interest to understand how these enzyme complexes work in detail. In this regard, not only an accurate description of the static three-dimensional (3D) structure is required, but also a description of how these machines change their structure over time. These internal protein motions have been shown to be important for, for example, biomolecular recognition,[1] allostery,[2] protein stability,[3] and enzymatic activity.[4] NMR spectroscopy is especially suited to study internal motions and when combined with methyl TROSY techniques[5] is able to address this aspect in very large molecular assemblies.[6]

The exosome complex is a large molecular machine that degrades or trims different RNA substrates in the 3′ to 5′ direction.[7] The archaeal and eukaryotic exosome complexes consist of nine subunits arranged in a hexameric ring structure (the exosome core) that interacts with a trimeric cap structure.[8] In the archaeal exosome complex the core contains three Rrp41/Rrp42 heterodimers[9] and the cap contains three copies of Rrp4 or Csl4 or a mixture thereof[10] (Figure S1 in the Supporting Information). A number of crystal structures of the archaeal exosome have been solved that show that the RNA is funneled through a hole at the top of the exosome core.[10a,11] In isolation the hexameric 173 kDa exosome core is catalytically active, where binding of the trimeric Rrp4 (or Csl4) cap modulates both catalytic activity and substrate selectivity.[12]

Herein, we present methyl TROSY NMR experiments that address the potential dynamics of the exosome complex in solution. A prerequisite for detailed NMR studies of biomolecular structures, interactions, and dynamics is that high-quality spectra can be recorded. For large molecular machines this is challenging in several ways. First, significant signal overlap arises due to the high number of unique resonances. To simplify such spectra, we prepared exosome complexes that contain NMR-active Rrp42 in an otherwise NMR-inactive background. Second, fast relaxation rates in large complexes lead to substantial broadening of the NMR signals. To overcome this, we made use of methyl TROSY spectroscopy[5] on samples that contained NMR-active Ile-δ1, Leu-δ, and Val-γ [^{1}H,^{13}C] methyl groups in an otherwise fully deuterated background (Figure 1 A and Figure S2).

To assign the Ile-δ1, Leu-δ,and Val-γ methyl resonances of Rrp42 in the exosome core and exosome–cap complexes, we followed the divide-and-conquer approach (Figure S2A).[6a] To validate and complete the Rrp42 assignments in the exosome complexes, we took a mutational approach[6c,13] (Figure S2B, Table S1). In summary, we have assigned more than 70 % of the methyl groups in the exosome core complex. It should, however, be mentioned that 100 % of the fully resolved resonances that can be used to study interactions and dynamics were assigned.

The assignment of the methyl groups in the exosome core complex revealed that a subset of the residues gives rise to two sets of resonances (Figure 1 A), which indicates that the complex adopts two structurally different conformations in solution. In the following we will refer to the set of peaks that are more intense as state A and the other set of peaks as state B. It should be noted that the isolated Rrp42 (Figure S2A) adopts only a single conformation, implying that Rrp41 is required to induce the different states in the exosome complex. In the exosome core, the residues that display two states are clustered at the top of the core complex (Figure 1 B), a region that contains the entrance pore for the substrate and that is responsible for the interaction with the Rrp4/Csl4 cap complex (Figure S1).

There is no evidence for conformational variability in the free exosome core based on the high-resolution crystal structure of the complex.[9] There, the unit cell contains four exosome complexes and all independent copies of Rrp42 superimpose with a backbone root-mean-square deviation (RMSD) of less than 0.13 Å. The only evidence for structural

[*] M. J. C. Audin,[+] G. Dorn,[+] S. A. Fromm, S. Schütz, M. K. Vorländer, Dr. R. Sprangers
Max Planck Institute for Developmental Biology
Spemannstrasse 35, 72076 Tübingen (Germany)
E-mail: remco.sprangers@tuebingen.mpg.de
K. Reiss
Interfaculty Institute of Biochemistry
University of Tübingen (Germany)
[+] These authors contributed equally to this work.

[**] We thank E. Conti (MPI Munich) for DNA of the exosome proteins, O. Weichenrieder (MPI Tübingen) for the plasmid used for in vitro transcription, and Thilo Stehle (University of Tübingen) for providing access to the SPR machine. We acknowledge Lewis Kay (University of Toronto) and Silke Wiesner (MPI Tübingen) for critically reading the manuscript. We thank all group members for discussions and Janina Ullmann for excellent technical assistance. This study was supported by the Max Planck Society and a Marie Curie Reintegration Grant (FP7/2007–2013, grant agreement no. 239164, to R.S.).
Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201302811.

plasticity in the exosome core can be derived from the static structure of the exosome–cap complex. There, Rrp4 induces a small structural change in Rrp42[10a] in the region that we observe to be structurally inhomogeneous in the free exosome core (Figure 1 B,C).

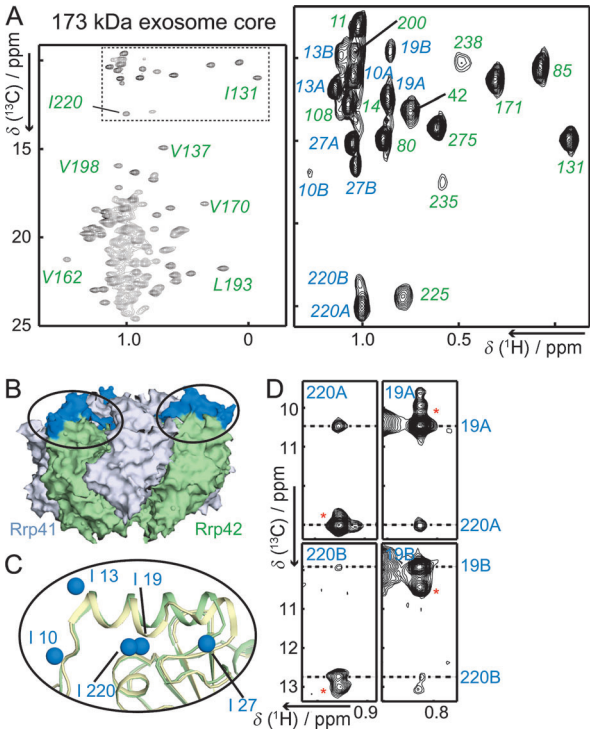

Figure 1. A) Methyl TROSY spectrum of the 173 kDa exosome core that contains NMR-active Rrp42. The boxed Ile-δ1 region is enlarged and residues that show two sets of peaks are labeled "*A*" and "*B*". B) The region that adopts two conformations is colored blue on the surface of the exosome core. This region is part of the cap-binding interface (see also Figure S1). C) Enlargement of the indicated region in (B), where Rrp42 as found in the exosome core (green; PDB: 2BR2) is superimposed on Rrp42 and as found in the exosome-Rrp4 complex (olive; PDB: 2JE6). The helix in the Rrp42 cap-binding region adopts a different conformation upon interaction with the Rrp4 cap. The blue spheres show the positions of the exchanging Ile residues. Note the short distance between Ile 19 and Ile 220. D) Planes from a 3D (H)-C-C-H NOESY spectrum that displays interproton NOE contacts between Ile 19 and Ile 220 in both states, indicating that the structures of state A and state B in the free exosome complex are similar. Cross peaks indicated with a red asterisk result from chemical exchange between the two states (see Figure 2).

exchange rate and one set of populations (Figure 2 A). From this, we determined that the exosome core exchanges between state A and state B with a rate of 36.1 ± 1.9 Hz and that the population of state A is 0.76 ± 0.01 at 60 °C. To independently

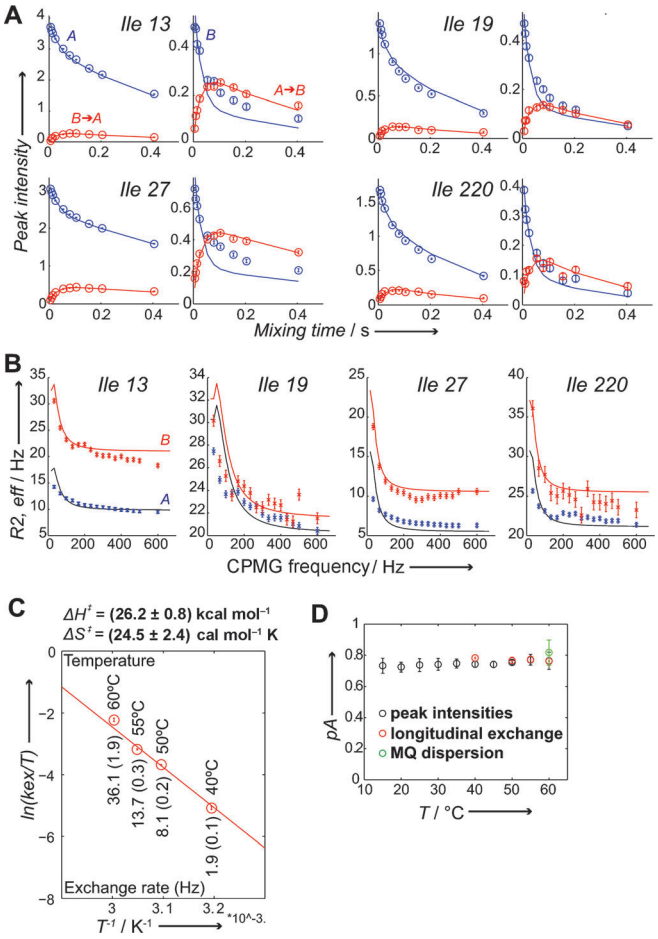

Figure 2. Quantification of the exchange process in the exosome core. A) Longitudinal exchange experiments. Auto and cross peaks are colored blue and red, respectively. Experimental data points are indicated with open circles, and the solid lines correspond to the best fits of the data. (All data are fitted simultaneously to one global exchange rate and one set of populations.) B) MQ dispersion data. The blue and red curves correspond to the states A and B, respectively. (All data are fitted simultaneously to one global exchange rate and one set of populations.) C) Eyring plot derived from kinetic data determined at four different temperatures. D) The population of state A does not change significantly over a large temperature range.

To further characterize the two observed Rrp42 states we recorded methyl–methyl NOESY spectra. Interestingly, we observe the same methyl–methyl contacts for state A and state B (Figure 1 D), indicating that the two states have similar structures. To determine the rates and populations that are connected with the exchange process between states A and B, we used longitudinal exchange NMR experiments.[6c,14] In the exosome core, we could reliably extract exchange rates and populations for four isoleucine residues (13, 19, 27, and 220). As all four residues have very similar exchange rates we used one global fit to extract a single

validate the extracted exchange parameters we used multiple quantum (MQ) relaxation dispersion experiments,[15] from which we extracted an exchange rate of 44.24 ± 26.8 Hz and a population of state A (pA) of 0.82 ± 0.08 at 60 °C (Figure 2 B). These parameters are in agreement with the values obtained from the longitudinal exchange experiments and thus independently validate these data.

To access thermodynamic properties associated with the exchange process we measured exchange rates and populations at temperatures between 40 and 60 °C. Based on that, we derived an activation enthalpy of 26.2 ± 0.8 kcal mol^{-1} and an

activation entropy of $24.5 \pm 2.4\ cal\,mol^{-1}\,K^{-1}$ for the conformational exchange (Figure 2C). The positive value for the activation entropy indicates that the disorder in the transition state is greater than in the two ground states. If one assumes that the hydration does not change significantly during the exchange process, this implies that the Rrp42 N-terminal helix partially unfolds upon changing between states A and B. The magnitude of the activation entropy is similar to that observed for the aromatic ring flips in the core proteins[16] and implies that the transition state is structurally distinct from the two ground states. Between 15 and 60 °C the populations of states A and B are almost invariable at $pA = 0.75$ (Figure 2D). This indicates that the two states have similar enthalpy, which is consistent with state A and state B being structurally similar (Figure 1D). The small entropy difference also implies that state A (which is populated to a larger extent) is favored over state B due to increased entropy. In agreement with that, we observe lower order parameters (high flexibility) for two of four methyl groups in state A (Figure 3). Taken together, we have shown that the cap-binding region in the isolated exosome core complex exchanges between two states in solution. From our NOE and thermodynamic data we can conclude that the two conformations are structurally similar.

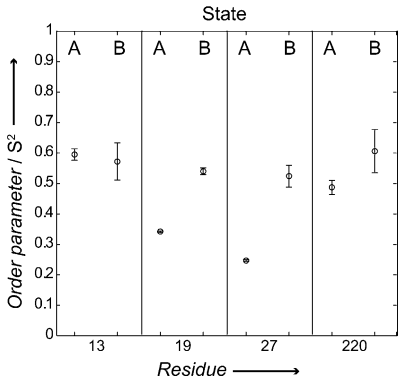

Figure 3. Order parameters for states A and B of Ile13, 19, 27, and 220. State A is significantly more flexible for residues 19 and 27, suggesting that state A is more flexible than state B; this is in agreement with a higher expected entropy of that state.

To determine whether the motions in the free exosome are directly correlated with the cap-binding process, we prepared exosome samples in complex with NMR-inactive Rrp4 and Csl4. Interestingly, we noticed that the residues that show two conformations in the free exosome core complex only display a single set of resonances in the exosome–cap complex (Figure 4A,B). This shows that the dynamics in the exosome core are significantly affected or even diminished upon interaction with the cap. To validate this observation, we recorded MQ dispersion experiments on the exosome–Rrp4 and exosome–Csl4 complexes that show that the residues that are dynamic in the free exosome core complex do not show micro- to millisecond time-scale dynamics in the cap-bound form (Figure S3). These observations are compatible with a scenario where one of the two states in the free exosome recruits Csl4 or Rrp4. It is, however, not straightforward to determine whether state A or state B interacts with the cap proteins. First, due to the chemical shift changes that Rrp42 experiences upon cap binding (Figure 4A,B), we cannot directly deduce which state the exosome adopts when in complex with the cap. Second, the interaction between the cap and the exosome core is slow on the NMR chemical shift timescale, which prevents the determination of which of the two states is selected for by the cap in a titration experiment. Finally, chemical shifts that are predicted[17] based on the two structures (free and cap-bound) do not correlate with the observed chemical shifts (A or B) in any of the possible

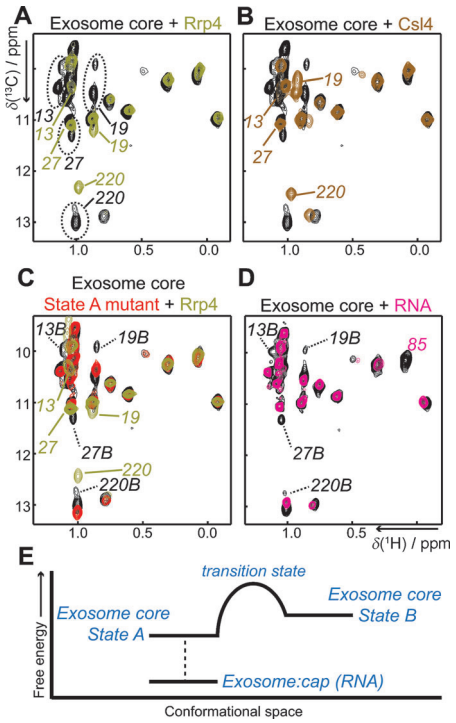

Figure 4. State A in the exosome core is stabilized upon interaction with cap proteins or with substrate RNA. (A, B) Spectra of the exosome core in the absence (black, dashed oval, two states) and presence of Rrp4 (A, olive) or Csl4 (B, brown). Upon cap interaction only a single state is observed in the exosome. The cap complexes interact with state A (see text), although this cannot be derived from the spectra directly due to the chemical shift changes induced by the cap. The signal to noise ratio for isoleucines 13, 19, 27, and 220 is larger than 10; thus, a potential minor state with a population of 25% would have been observed in the spectra. C) Superposition of the WT exosome (black), the Rrp42 state A mutant where state B is no longer visible in the absence (red) and presence of Rrp4 (olive). In complex with Rrp4, Rrp42 is structured identically in the state A mutant exosome (olive) and in the WT exosome (Figure 1A, olive spectrum) (Figure 4A). D) Exosome core in the absence (black) and presence (pink) of an equimolar amount of RNA substrate (one RNA molecule per hexameric exosome complex). State B in the exosome core is no longer visible upon formation of the substrate–enzyme complex. In addition, residues close to the active site experience chemical shift perturbations (e.g Ile85). E) Schematic representation of the reaction coordinates of the exosome core. States A and B have similar enthalpy; state A has a lower free energy due to increased entropy. The transition state is more disordered than the ground states A and B. State A interacts with the cap structure through a mechanism of conformational selection.

combinations (Figure S4), likely reflecting difficulties in the accurate prediction of methyl groups' chemical shifts for large protein complexes.

To experimentally address which state in the free exosome complex interacts with the cap, we designed a mutant exosome core in which the equilibrium between the two states is significantly altered. Interestingly, we found a point mutant that is remote from the cap-binding interface (Figure S5) that displays only state A in $^{1}H,^{13}C$-HMQC spectra (N9A in Rrp42, referred to as "state A mutant") (Figure 4C, red). We then used this state A mutant to probe binding for the cap proteins; if this mutant efficiently interacts with Cls4 and Rrp4, this is strong evidence that state A is selected by the cap proteins. Interestingly, this is exactly what we observe. First, an NMR spectrum of the state A mutant in complex with Rrp4 (Figure 4C, olive) is identical with the spectrum of the wild-type (WT) exosome in complex with Rrp4 (Figure 4A, olive). Secondly, surface plasmon resonance (SPR) experiments confirm that the state A mutant still interacts strongly with Rrp4 (Figure S6). Thirdly, a reduced form of Rrp4, where one of the three protein domains is deleted, interacts stronger with the exosome of the state A mutant than with the WT exosome in NMR titration experiments (Figure S7). In summary, all binding experiments show that the exosome complex of the state A mutant is fully capable of interacting with the cap proteins. We thus propose that state A in the exosome complex corresponds to the cap-bound conformation. Our data are thus compatible with a model where the exosome–cap interaction occurs through a process where only one of the possible conformations is selected upon complex formation.

Upon addition of substrate RNA to the exosome core complex (Figure 4D)[9,11a] we observed significant chemical shift perturbations for Rrp42 residues Ile85 and Val86. Based on the structure of the *Pyrococcus abyssi* exosome in complex with RNA,[18] these residues are indeed close to the substrate. In addition, we observe a significant shift towards state A for the residues at the cap-binding region. This indicates that substrate RNA, like the cap structure, interacts with state A in the free exosome complex (Figure 4E).

In summary, we have complemented the known static crystal structure of the archaeal exosome core with quantitative information regarding unanticipated internal dynamics. Our data show that molecular motions often remain undetected in protein structures and underscore the importance of studies that address the localization, quantification, and interpretation of these functionally important aspects of large molecular machines. We anticipate that future work of the sort presented here will be able to address the relation between dynamics and function in many biomolecular assemblies.

Experimental Section

NMR spectra were recorded between 15 and 60°C on Bruker AVIII-600 and AVIII-800 spectrometers. All spectra displayed in the figures were recorded at 50°C. Longitudinal exchange experiments were recorded as a series of 3D (C-C-H) data sets with mixing times between 0 and 800 ms. Exchange parameters were extracted as described.[6c] Errors in the measured data were determined based on the noise level in the spectra. The error in the extracted parameters is based on Monte Carlo simulations, where back-calculated data were randomly changed according to the experimental error. Methyl TROSY relaxation dispersion experiments[15] were recorded as a series of 2D data sets using constant-time relaxation periods of 50 ms and CPMG (Carr–Purcell–Meiboom–Gill) frequencies ranging from 33 to 600 Hz. The dispersion data were fitted numerically as described, where the chemical shift differences were extracted directly from the spectra.[6c] Errors in the parameters were based on Monte Carlo simulations and on duplicate measurements. Uncertainties in the extracted chemical shift differences were accounted for by varying $\Delta\omega$H and $\Delta\omega$C by 0.005 and 0.01 ppm, respectively. Methyl group order parameters were determined as described,[19] using a rotational correlation time of 86 ns, as derived from the program HYDRONMR.[20] All NMR data were processed with the nmrpipe/nmrdraw suite of programs.[21] Figures displaying NMR spectra were prepared using NMRview (onemoonscientific.com), molecular structures were drawn using PyMol (pymol.org).

Received: April 4, 2013
Published online: June 26, 2013

Keywords: exosome · molecular machines · NMR spectroscopy · protein dynamics

[1] a) D. D. Boehr, R. Nussinov, P. E. Wright, *Nat. Chem. Biol.* **2009**, *5*, 789–796; b) S. R. Tzeng, C. G. Kalodimos, *Nature* **2012**, *488*, 236–240.
[2] a) S. Bruschweiler, P. Schanda, K. Kloiber, B. Brutscher, G. Kontaxis, R. Konrat, M. Tollinger, *J. Am. Chem. Soc.* **2009**, *131*, 3063–3068; b) S. R. Tzeng, C. G. Kalodimos, *Nature* **2009**, *462*, 368–372.
[3] A. L. Lee, A. J. Wand, *Nature* **2001**, *411*, 501–504.
[4] E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, D. Kern, *Nature* **2005**, *438*, 117–121.
[5] V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, *J. Am. Chem. Soc.* **2003**, *125*, 10420–10428.
[6] a) R. Sprangers, L. E. Kay, *Nature* **2007**, *445*, 618–622; b) R. Sprangers, A. Velyvis, L. E. Kay, *Nat. Methods* **2007**, *4*, 697–703; c) R. Sprangers, A. Gribun, P. M. Hwang, W. A. Houry, L. E. Kay, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16678–16683; d) I. Gelis, A. M. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou, C. G. Kalodimos, *Cell* **2007**, *131*, 756–769.
[7] P. Mitchell, E. Petfalski, A. Shevchenko, M. Mann, D. Tollervey, *Cell* **1997**, *91*, 457–466.
[8] a) Q. Liu, J. C. Greimann, C. D. Lima, *Cell* **2006**, *127*, 1223–1237; b) E. Lorentzen, J. Basquin, E. Conti, *Curr. Opin. Struct. Biol.* **2008**, *18*, 709–713.
[9] E. Lorentzen, P. Walter, S. Fribourg, E. Evguenieva-Hackenberg, G. Klug, E. Conti, *Nat. Struct. Mol. Biol.* **2005**, *12*, 575–581.
[10] a) E. Lorentzen, A. Dziembowski, D. Lindner, B. Seraphin, E. Conti, *EMBO Rep.* **2007**, *8*, 470–476; b) C. Witharana, V. Roppelt, G. Lochnit, G. Klug, E. Evguenieva-Hackenberg, *Biochimie* **2012**, *94*, 1578–1587.
[11] a) E. Lorentzen, E. Conti, *Mol. Cell* **2005**, *20*, 473–481; b) E. Lorentzen, E. Conti, *Archaea* **2012**, 721869.
[12] P. Walter, F. Klein, E. Lorentzen, A. Ilchmann, G. Klug, E. Evguenieva-Hackenberg, *Mol. Microbiol.* **2006**, *62*, 1076–1089.
[13] C. Amero, M. Asuncion Dura, M. Noirclerc-Savoye, A. Perollier, B. Gallet, M. J. Plevin, T. Vernet, B. Franzetti, J. Boisbouvier, *J. Biomol. NMR* **2011**, *50*, 229–236.
[14] N. A. Farrow, O. Zhang, J. D. Forman-Kay, L. E. Kay, *J. Biomol. NMR* **1994**, *4*, 727–734.

[15] D. M. Korzhnev, K. Kloiber, V. Kanelis, V. Tugarinov, L. E. Kay, *J. Am. Chem. Soc.* **2004**, *126*, 3964 – 3973.
[16] G. Wagner, A. DeMarco, K. Wuthrich, *Biophys. Struct. Mech.* **1976**, *2*, 139 – 158.
[17] B. Han, Y. Liu, S. W. Ginzinger, D. S. Wishart, *J. Biomol. NMR* **2011**, *50*, 43 – 57.
[18] M. V. Navarro, C. C. Oliveira, N. I. Zanchin, B. G. Guimaraes, *J. Biol. Chem.* **2008**, *283*, 14120 – 14131.
[19] V. Tugarinov, R. Sprangers, L. E. Kay, *J. Am. Chem. Soc.* **2007**, *129*, 1743 – 1750.
[20] J. Garcia de La Torre, M. L. Huertas, B. Carrasco, *J. Magn. Reson.* **2000**, *147*, 138 – 146.
[21] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* **1995**, *6*, 277 – 293.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

The Archaeal Exosome: Identification and Quantification of Site-Specific Motions That Correlate with Cap and RNA Binding**

*Maxime J. C. Audin, Georg Dorn, Simon A. Fromm, Kerstin Reiss, Stefan Schütz, Matthias K. Vorländer, and Remco Sprangers**

anie_201302811_sm_miscellaneous_information.pdf

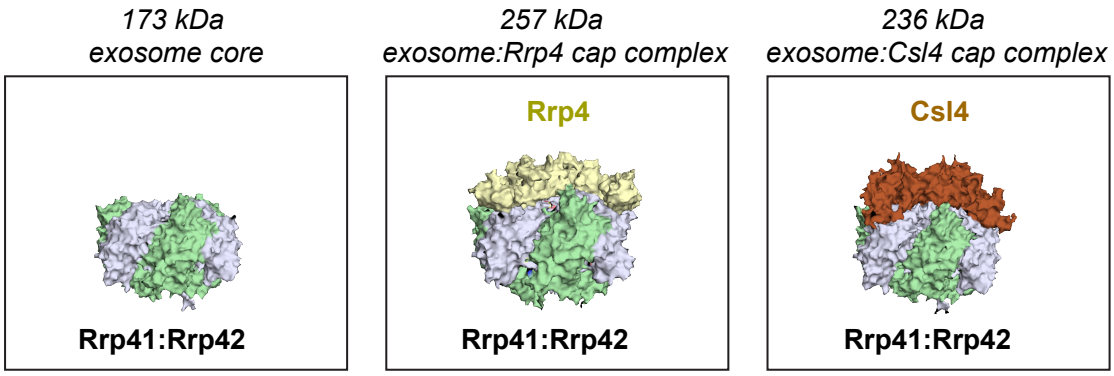

Figure S1. Architecture of the exosome complex. Rrp42 is displayed as a cartoon (green), Rrp41 is shown in gray, Rrp4 in yellow and Csl4 in brown. The displayed structures are based on PDB-entries 2BR2 (exosome core)[1], 2JEA (exosome Rrp4 complex)[2] and 3M7N (exosome Csl4 complex, displayed structure is a homology model based on the *archaeoglobus fulgidus* structure of the complex).[3]

The exosome core (Rrp41:Rrp42) can interact with substrate RNA and degrade this in a processive manner in the 3' to 5' direction. During this process, the exosome does not release the substrate. We thus expect that cap proteins will not be recruited to the processing exosome core: substrate RNA complex. It should be noted that the amounts of cap-free exosome (Rrp41:Rrp42) is expected to be very low in a cellular context as cap proteins, Rrp41 and Rrp42 are present in similar relative amounts.[4]
The interaction between the free exosome and the cap proteins (Csl4 or Rrp4) is very tight. After formation of the exosome:cap complex RNA substrate can be recruited and degraded; the cap proteins will remain bound to the exosome core during this process.

To prepare the samples used in the current study, the *Sulfolobus solfataricus* Rrp41, Rrp42, Rrp4 and Csl4 DNA (a kind gift from E. Conti, MPI Munich) was cloned into pET vectors carrying a TEV cleavable N-terminal His6-tag. Point mutations were introduced using the Quikchange approach (Stratagene). U-[^{2}H, ^{15}N] Ile-δ1, Leu-δ, Val-γ [^{1}H, ^{13}C] labeled Rrp42 proteins were obtained by overexpression of the corresponding gene in BL21(DE3) Codon Plus RIL (Stratagene) cells in 100 % D2O minimal medium, as previously described.[5] Purification of all constructs was achieved by using Ni affinity chromatography followed by cleavage of the histidine tag and size-exclusion chromatography. If required, amide protons of Rrp42 were back-exchanged by refolding the GuHCl denatured protein in H_2O based buffer. Exosome core complexes were reconstituted by combining separately purified components. Exosome-cap complexes were obtained by addition of purified Rrp4 or Csl4 to the exosome core complex. NMR samples contained between 0.05 and 1.5 mM protein (monomer concentration) in 30 mM KPO4 pH 6.8 (or 25 mM Hepes pH 7.5), 150 mM NaCl, 1 mM DTT in 100 % D2O or in 95:5 H2O:D2O. For the exosome RNA complex, excess of RNA was removed by size exclusion chromatography such that one RNA molecule was present per hexameric exosome core complex.

Substrate RNA (20 adenines linked to a hairpin structure) was produced by *in-vitro* transcription using a linearized pSP64 plasmid that contains the substrate RNA followed by a 3' HDV ribozyme that auto-cleaves the RNA cotranscriptionally. The RNA was purified over a Dionex DNAPac PA-100 column at 75 °C using a NaCl gradient in 5 M urea. Substrate RNA was not degraded during NMR experiments due to lack of phosphate in the buffer, the 2',3'-cyclic phosphate at the 3' end of the RNA and the hairpin structure at the 5' end of the RNA.

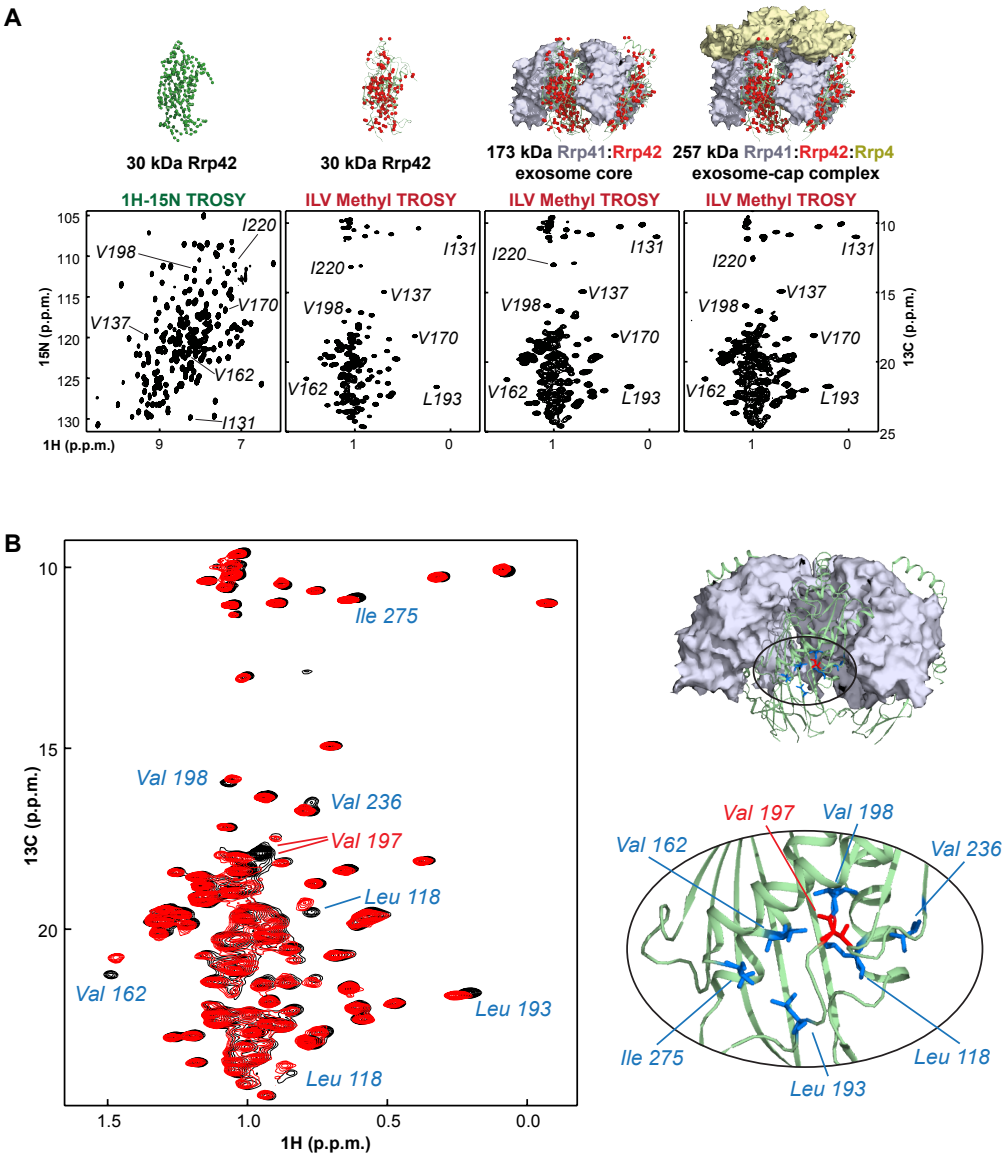

Figure S2. Methyl group assignment strategy. (A) Building blocks used in the divide and conquer approach. First panel: ^{1}H-^{15}N TROSY spectrum of [^{2}H, ^{13}C, ^{15}N] labeled Rrp42 monomer. H-N groups are shown as green spheres. Second panel: ^{1}H-^{13}C methyl TROSY spectrum of U-[^{2}H, ^{15}N] Ile-δ1, Leu-δ, Val-γ [^{1}H, ^{13}C] Rrp42 as monomer. Third panel: Rrp42 within the exosome core complex. Last panel: Rrp42 within the exosome-Rrp4 complex. Labeled methyl groups are shown as red spheres. Exemplary assignments are indicated. Rrp42 backbone sequential assignments were completed using TROSY versions of HNCACB/HNCOCACB experiments. Methyl groups in the Rrp42 monomer were assigned using C(C)(CO)NH TOCSY, H-N-H and H-N-C NOESY spectra. C-C-H HMQC-NOE-HMQC and H-C-H NOE-HMQC spectra were used to assign methyl groups in the exosome-core and exosome-cap complexes. (B) Assignments by point mutations. Left: spectra of WT (black) and V197A (red) exosome complexes. The assignment for V197 is indicated in red. Residues that are in the vicinity of the mutation and that thus experience secondary chemical shift changes are labeled in blue. Right: Location of V197 and the residues that experience secondary chemical shifts on the crystal structure of the exosome core complex.

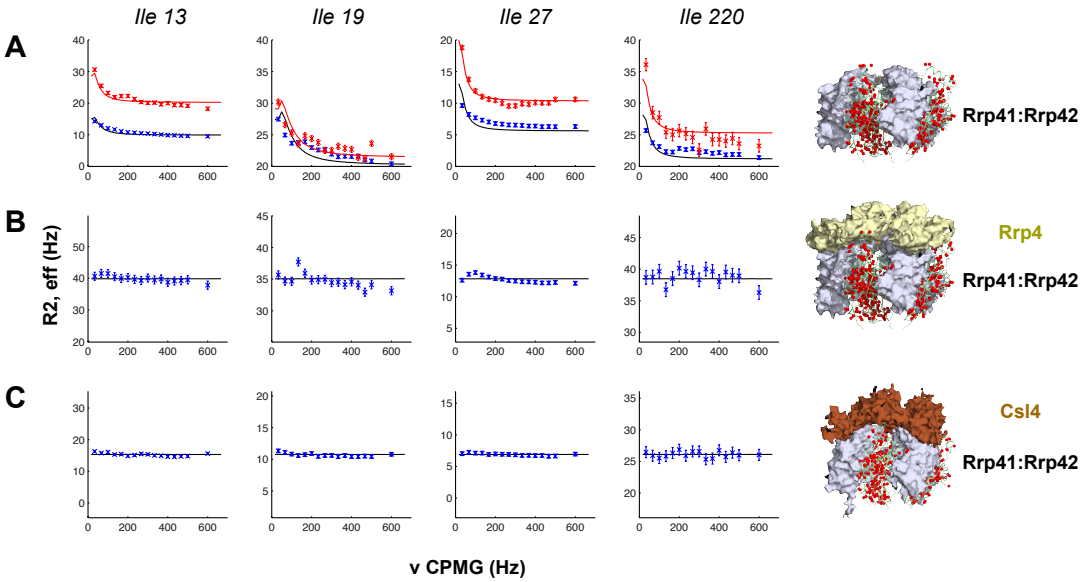

Figure S3. The cap proteins change the dynamics in the exosome core. MQ dispersion profiles observed for Ile 13, 19, 27 and 220. Note that the y-axis has the same range for all graphs for a specific residue to allow for direct comparison of the data. (A) Profiles in the exosome core (identical to Fig 2B in the main text). Blue and red correspond to state A and B respectively. See main text for details. The structure of the exosome core is indicated on the right. (B) Profiles in the exosome-Rrp4 complex. Note that only one state is present in the spectra. The structure of the exosome-Rrp4 complex is shown on the right. (C) Profiles in the exosome-Csl4 complex. Note that only one state is present in the spectra.

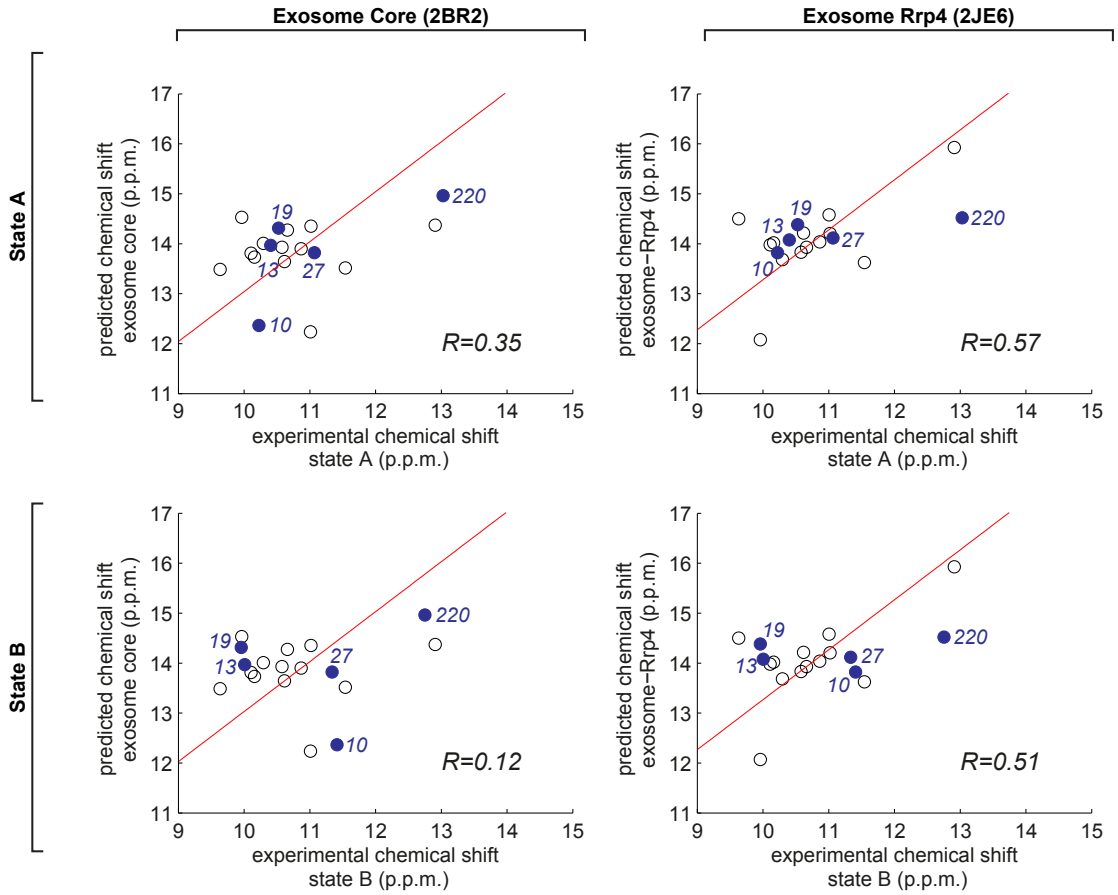

Figure S4. Predicted versus measured chemical shifts for all assigned isoleucine residues. The chemical shifts were predicted using shiftx2[6] using the free exosome complex (2BR2) or the exosome-Rrp4 complex (2JE6) from which Rrp4 was removed as input. The methyl groups that show two conformations have been labeled; the Pearson R correlation coefficient is indicated. A red drawn line indicates the best fit between the predicted and measured shifts (y=x+A), where A corrects for an (potential) offset in chemical shift referencing. Note that none of the correlations is significant, most likely due to the large inaccuracies in the predicted values.

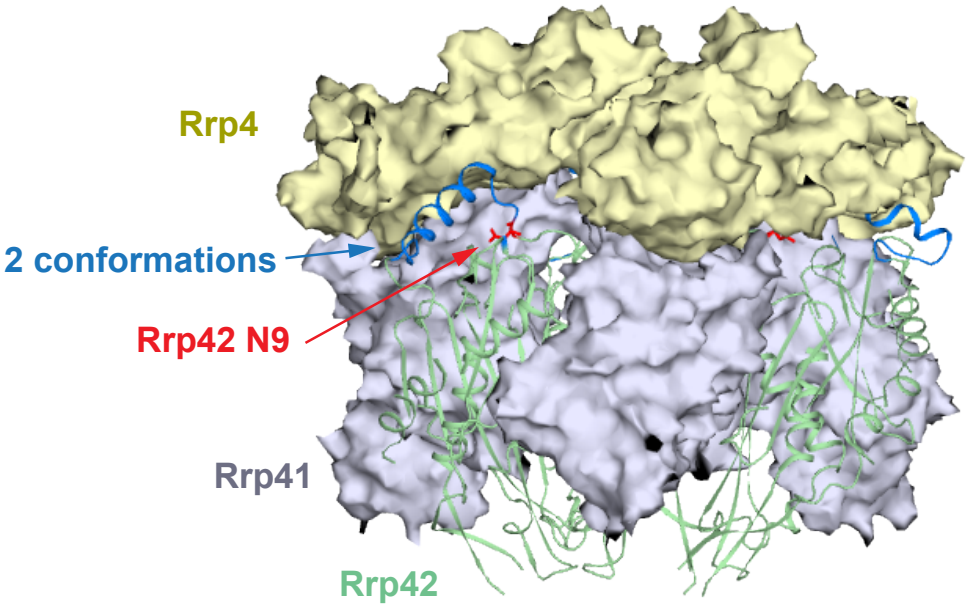

Figure S5. "State A mutant" (N9A) exosome complex. Location of N9 in the exosome complex. N9 is remote from the interaction with the cap structure. Mutations in this residue do thus not change the interaction between the exosome core and the cap directly, but rather indirectly through changes in exosome dynamics.

The identification of the N9A mutant was inspired by the spectra of the assignment mutants (Table S1), where we noticed that the relative intensities of the two sets of peaks varied. This indicated that the equilibrium between the two states could be modified. We then systematically mutated residues that were close, but not directly in the cap-interaction-helix and monitored the state A: state B peak ratio. In this process we identified that the N9A mutation yielded only a single set of resonances.

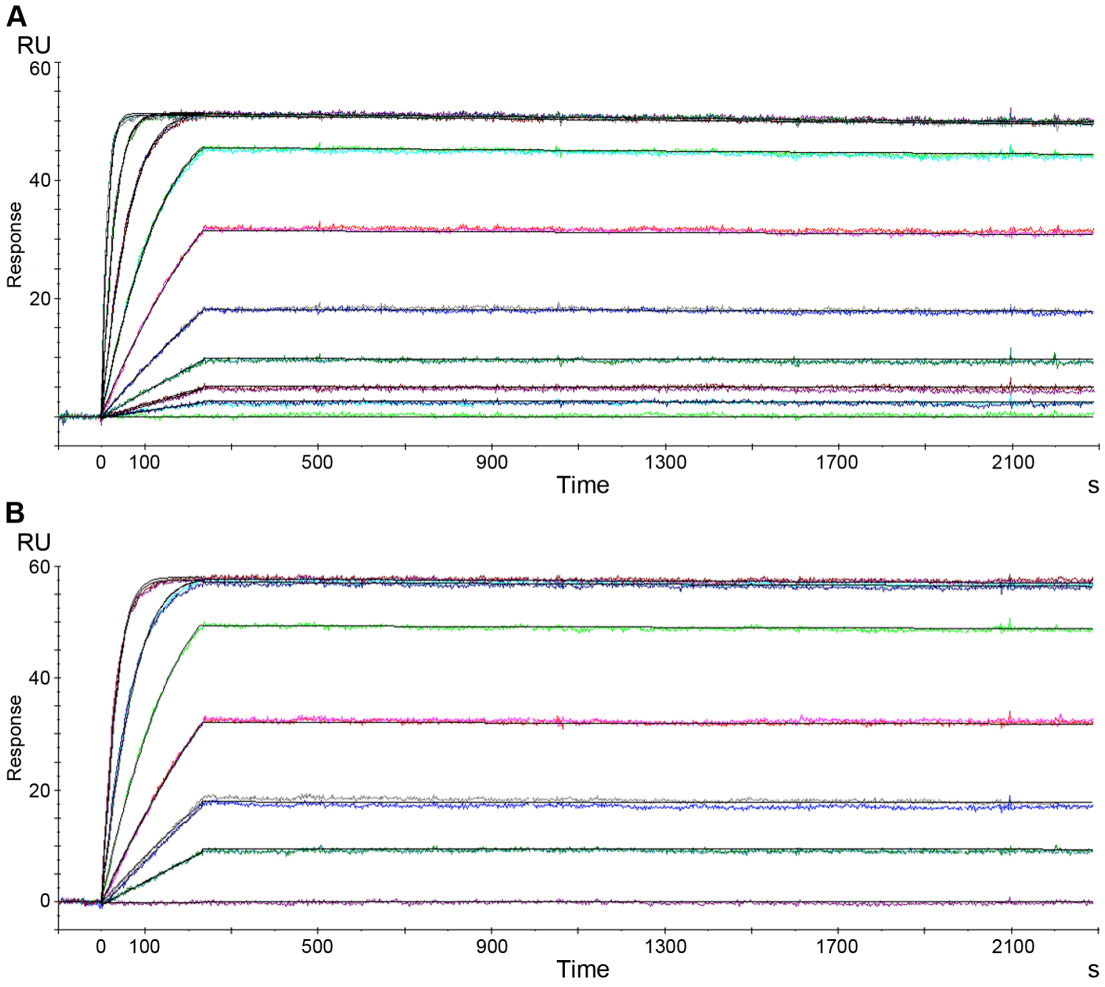

Figure S6 Kinetic SPR analyses of cap protein Rrp4 with His-tagged wildtype (A) or "state A mutant" (B) exosome complex attached to a Ni NTA chip. The double-referenced sensorgrams (indicating that two controls experiment were performed: one without ligand and one without analyte) are overlaid with fits of a "1:1 binding with mass transfer" model.

	wildtype	SE [a]	state A mutant	SE[a]
k_{on} ($M^{-1}s^{-1}$)	$1.7 \cdot 10^{6}$	$2.1 \cdot 10^{3}$	$1.7 \cdot 10^{6}$	$2.8 \cdot 10^{3}$
k_{off} (s^{-1}) [b]	$< 10 \cdot 10^{-5}$		$< 10 \cdot 10^{-5}$	
K_D (M) [c]	**$< 10 \cdot 10^{-11}$**		**$< 10 \cdot 10^{-11}$**	
R_{max} (RU) [d]	51.2	0.01	57.7	0.01
χ^2 (RU^2)	0.17		0.20	

[a] Standard error (obtained from the Biaeval software kit)

[b] The off-rate is at the detection limit of the system

[c] Defined as k_{off} /k_{on}

[d] Theoretical maximum response that is reached when all ligand binding sites are occupied by the analyte; RU refers to response units.

The SPR analyses were performed on a Biacore 2000 system at 15 °C. Two consecutive flow cells (a measurement cell and a reference cell in which no ligand was immobilized) were used. In both the measurement and the reference cell an NTA chip (GE healthcare) was loaded with $NiCl_2$ following the manufacturer's instructions. His-tagged wildtype or "state A mutant" exosome complex was diluted in running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 5 % D20,

50 uM EDTA) and 100-110 RU of ligand were non-covalently bound to the experimental flow cell. Untagged cap protein Rrp4 was serially diluted in running buffer to concentrations ranging from 0.23 to 58 nM and injected for 200 s in both the experimental and the reference cell at a flow rate of 50 μl/min. The dissociation phase was followed for 1800 s. Additionally, we recorded a blank curve, where buffer (without analyte) was injected. Both the curve from the reference cell and the blank injection were subtracted from the SPR signal in the measurement cell. To regenerate the surface regeneration buffer (10 mM HEPES, 150 mM NaCl, 0.005% (v/v) P20, 0.35 M EDTA, pH 8.3) was injected for 3 minutes at a flow rate of 20 μl/min in both the reference and the measurement cell.

In addition to the interaction between the exosome complex and the Rrp4 cap we performed experiments to probe for the interaction between the exosome complex and the reduced Rrp4 cap (that lacks one of the domains; See Figure S7). Unfortunately, this protein interacted unspecifically with the sensorchip surface, which resulted in a strong signal from the reference cell. As a consequence, we were not able to extract any reliable interaction data for the exosome: reduced Rrp4 cap complex.

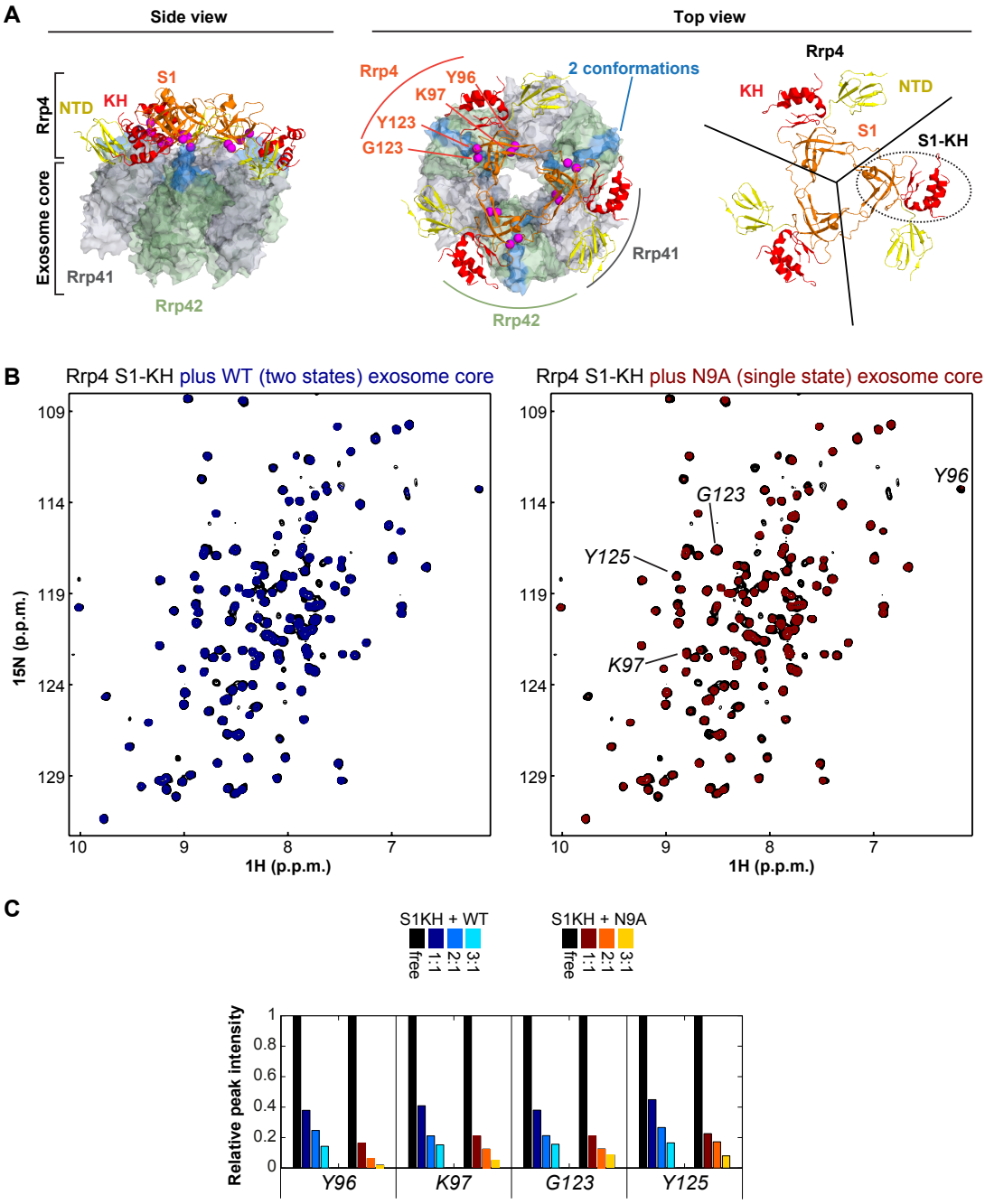

Figure S7 To move the Rrp4 binding affinity into a range where one can discriminate cap binding between WT and "state A mutant" exosome, we deleted one of the three domains from the cap structure (A). This reduced Rrp4 cap contains the domains that interacts with Rrp42 (the S1 and KH domains) but lacks the domain (the NTD; N-terminal domain) that interacts with Rrp41. We then used this reduced Rrp4 cap structure to probe for the interactions with the WT and "state A mutant" exosome. In NMR chemical shift titrations (B), where we added the unlabeled exosome to ^{15}N-labeled reduced cap, we observed a faster decrease in resonance intensity upon addition of the "state A mutant" exosome than upon addition of the WT exosome (C). This implies that the "state A mutant" has a higher affinity for the cap than the WT exosome and establishes that state A plays an important role in the interaction with the Rrp4 cap structure.

(A) Left: Side view of the structure of the exosome-Rrp4 complex. The exosome core is drawn as a surface representation (Rrp41: gray; Rrp42: green), Rrp4 is shown as a ribbon. The three Rrp4 domains are colored separately; NTD (N-terminal domain) in yellow; S1 domain in orange; KH domain in red.

Middle: Top view of the complex, indicating that Rrp42 does not contact the Rrp4 NTD. The region that displays two conformations (blue) contacts both the S1 and KH domains. Selected Rrp4 residues that are in contact with Rrp42 are indicated (see below). Right: Cartoon representation of Rrp4. The S1-KH (reduced Rrp4 protein) region that is used to probe for binding with the exosome is circled.

(B) ^{1}H-^{15}N NMR spectra of the reduced Rrp4 protein in the absence (black) and presence of an equimolar amount of WT exosome (left, blue) or “state A mutant” exosome (red, right). Larger chemical shift perturbations are observed upon addition of the “state A mutant” complex, demonstrating a stronger interaction.

(C) Dependence of reduced Rrp4 peak intensities on the molar excess of exosome (blue-cyan scale) or “state A mutant” exosome (red-yellow scale). Four Rrp4 residues that contact Rrp42 are selected. The signals decrease more rapidly upon addition of the “state A mutant” due to the tighter interaction. Note that the decrease in peak intensity is largely due to fast relaxation in the high molecular weight complex that is formed, preventing accurate extraction of binding constants from the NMR data. After addition of a high excess of (WT or “state A”) exosome the spectra of the reduced cap are no longer visible due to the formation of a large complex.

Residue[1]	Rrp42 monomer		Rrp42 in exosome core		Point mutation[4]
	^{13}C (p.p.m.)	^{1}H (p.p.m.)	^{13}C (p.p.m.)	^{1}H (p.p.m.)	
Ile 10 δ1 (A)[2]	10.349	1.081	10.224	1.050	yes
Ile 10 δ1 (B)	Not applicable[3]		11.412	1.283	yes
Ile 11 δ1	9.836	1.084	9.635	1.026	yes
Ile 13 δ1 (A)	10.394	1.127	10.402	1.153	yes
Ile 13 δ1 (B)	Not applicable		10.004	1.117	yes
Ile 14 δ1	9.972	1.062	10.576	1.079	yes
Ile 19 δ1 (A)	10.265	0.885	10.525	0.879	
Ile 19 δ1 (B)	Not applicable		9.955	0.863	
Val 20 γ1	19.456	1.325	19.86	1.219	
Val 20 γ2	20.208	1.254	19.481	1.296	
Leu 22 δ1	22.933	0.993	22.856	0.98	
Leu 22 δ2	20.245	0.949	20.17	0.942	
Ile 27 δ1 (A)	11.088	1.076	11.066	1.062	yes
Ile 27 δ1 (B)	Not applicable		11.337	1.049	yes
Leu 34 δ1	22.451	1.14	22.385	1.118	
Leu 34 δ2	19.834	0.583	19.668	0.591	
Leu 40 δ1	22.422	1.045	22.356	1.007	
Leu 40 δ2	23.58	1.033	23.504	0.978	
Ile 42 δ1	10.674	0.786	10.659	0.756	
Leu 44 δ1	23.109	1.305	22.995	1.274	
Leu 44 δ2	21.238	1.087	21.066	1.043	
Leu 55 δ1	20.712	0.761	20.731	0.681	
Leu 55 δ2	22.422	0.664	22.084	0.473	
Val 56 γ1	17.994	1.067	17.996	1.042	
Val 56 γ2	19.587	1.04	19.567	1.011	
Leu 58 δ1	---	---	---	---	
Leu 58 δ2	---	---	---	---	
Val 63 δ1	---	---	---	---	yes
Val 63 δ2	---	---	---	---	yes
Leu 64 δ1	21.136	1.165	---	---	
Leu 64 δ2	23.605	1.11	---	---	
Leu 69 δ1	22.232	0.615	22.176	0.603	
Leu 69 δ2	23.035	0.8	23.129	0.772	
Ile 71 δ1	10.394	0.344	10.293	0.317	
Leu 84 δ1	23.372	0.947	---	---	yes
Leu 84 δ2	19.973	0.923	---	---	yes
Ile 85 δ1	10.304	0.23	10.159	0.087	
Val 86 γ1	18.608	1.049	---	---	yes
Val 86 γ2	18.914	1.096	---	---	yes
Val 88 γ1	19.215	1.024	---	---	
Val 88 γ2	18.068	0.996	---	---	
Leu 90 δ1	---	---	---	---	
Leu 90 δ2	---	---	---	---	
Leu 91 δ1	---	---	---	---	

Leu 91 δ2	---	---	---	---	
Leu 93 δ1	---	---	---	---	
Leu 93 δ2	---	---	---	---	
Ile 108 δ1	9.878	1.153	1.123	10.615	yes
Leu 110 δ1	---	---	---	---	yes
Leu 110 δ2	---	---	---	---	yes
Val 113 γ1	19.529	1.046	---	---	yes
Val 113 γ2	20.172	1.215	---	---	yes
Val 114 γ1	19.005	1.036	---	---	yes
Val 114 γ2	19.938	1.023	---	---	yes
Leu 118 δ1	23.912	0.864	24.015	0.856	yes
Leu 118 δ2	19.573	0.81	19.543	0.786	yes
Leu 124 δ1	21.034	0.841	20.822	0.809	
Leu 124 δ2	---	---	---	---	
Leu 126 δ1	23.001	1.205	22.981	1.197	yes
Leu 126 δ2	20.61	0.857	20.554	0.85	yes
Leu 129 δ1	24.014	1.07	23.981	1.045	
Leu 129 δ2	19.804	0.911	19.754	0.908	
Val 130 γ1	19.005	1.061	19.016	1.058	
Val 130 γ2	18.417	1.042	18.379	1.035	
Ile 131 δ1	11.027	-0.082	11.018	-0.067	
Val 137 γ1	19.689	1.018	19.702	1.005	
Val 137 γ2	14.955	0.71	14.945	0.707	
Val 140 γ1	18.184	0.883	18.129	0.871	yes
Val 140 γ2	---	---	---	---	yes
Leu 142 δ1	21.647	0.953	21.474	0.952	
Leu 142 δ2	23.328	0.978	23.246	1.006	
Val 144 γ1	18.637	0.966	---	---	yes
Val 144 γ2	17.527	0.867	---	---	yes
Val 146 γ1	---	---	---	---	yes
Val 146 γ2	---	---	---	---	yes
Leu 147 δ1	23.966	1.049	---	---	yes
Leu 147 δ2	---	---	---	---	yes
Val 153 γ1	---	---	---	---	yes
Val 153 γ2	---	---	---	---	yes
Leu 154 δ1	22.948	1.039	---	---	yes
Leu 154 δ2	---	---	---	---	yes
Leu 159 δ1	23.589	1.042	23.689	1.198	
Leu 159 δ2	21.574	1.133	21.449	1.137	
Val 162 γ1	19.733	1.378	19.821	1.345	
Val 162 γ2	21.222	1.517	21.265	1.502	
Leu 165 δ1	24.657	0.938	24.668	0.937	
Leu 165 δ2	21.267	0.876	21.441	0.86	
Val 170 γ1	18.41	0.657	18.356	0.642	
Val 170 γ2	18.184	0.386	18.138	0.365	
Val 173 γ1	---	---	19.702	1.018	

Val 173 γ2	---	---	19.129	1.152	
Ile 180 δ1	10.997	0.906	11.008	0.899	
Val 182 γ1	19.653	1.274	19.856	1.317	
Val 182 γ2	19.14	1.158	19.106	1.148	
Val 187 γ1	19.683	1.254	19.106	1.158	
Val 187 γ2	---	---	19.585	1.24	
Val 188 γ1	18.785	1.166	18.824	1.162	
Val 188 γ2	16.446	0.945	16.356	0.941	
Leu 191 δ1	21.463	0.773	21.418	0.774	
Leu 191 δ2	24.283	1.01	24.356	1.018	
Leu 193 δ1	21.764	0.16	21.783	0.213	
Leu 193 δ2	21.545	0.673	21.624	0.642	
Val 197 γ1	17.322	1.02	17.904	0.954	yes
Val 197 γ2	18.288	1.055	18.389	1.02	yes
Val 198 γ1	19.719	1.008	19.816	0.97	
Val 198 γ2	16.285	1.082	15.932	1.071	
Ile 200 δ1	9.739	1.061	9.96	1.051	
Val 202 γ1	20.099	1.062	20.237	1.036	
Val 202 γ2	18.552	1.3	18.423	1.253	
Val 205 γ1	17.083	1.125	17.194	1.071	
Val 205 γ2	19.163	1.186	19.147	1.171	
Leu 209 δ1	23.591	1.057	23.697	1.032	
Leu 209 δ2	20.66	0.975	20.254	0.909	
Val 210 γ1	16.637	0.839	16.751	0.8	yes
Val 210 γ2	19.792	0.581	19.637	0.566	yes
Val 211 γ1	---	---	19.504	0.989	yes
Val 211 γ2	---	---	---	---	yes
Leu 215 δ1	---	---	22.561	1.1	yes
Leu 215 δ2	---	---	22.122	1.143	yes
Ile 220 δ1 (A)	13.178	1.06	13.031	1.010	
Ile 220 δ1 (B)	Not applicable		12.753	1.020	
Ile 225 δ1	13.093	0.945	12.911	0.796	
Leu 233 δ1	19.216	0.672	18.774	0.759	
Leu 233 δ2	22.564	0.893	22.668	0.927	
Ile 235 δ1	11.094	1.107	11.543	0.59	yes
Val 236 γ1	17.556	0.975	16.441	0.772	
Val 236 γ2	18.537	1.192	19.566	0.952	
Ile 238 δ1	11.284	1.115	10.106	0.486	yes
Leu 248 δ1	22.292	1.2	---	---	
Leu 248 δ2	20.941	1.14	---	---	
Ile 251 δ1	10.53	1.017	---	---	yes
Val 263 γ1	20.181	1.326	20.106	1.318	
Val 263 γ2	18.537	1.192	18.606	1.184	
Leu 265 δ1	20.508	0.669	19.918	0.625	
Leu 265 δ2	22.422	0.756	22.803	0.733	
Leu 266 δ1	22.82	0.919	22.619	0.881	

Leu 266 δ2	21.659	0.935	22.043	0.93	
Leu 269 δ1	21.151	1.029	20.879	1.029	
Leu 269 δ2	---	---	---	---	
Leu 273 δ1	22.451	0.649	22.502	0.608	
Leu 273 δ2	20.04	0.982	19.957	0.94	
Ile 275 δ1	10.816	0.629	10.867	0.62	

Table S1:
Assigned chemical shifts for Rrp42 as a monomer and in the exosome core.

[1] Leu and Val methyl groups were not stereo-specifically assigned.
[2] A and B refer to the states A and B in the exosome core.
[3] The Rrp42 monomer only displays one state.
[4] Indicated if a point mutation was made to assign (or check the assignment of) the residue.

References

[1] E. Lorentzen, P. Walter, S. Fribourg, E. Evguenieva-Hackenberg, G. Klug, E. Conti, *Nat Struct Mol Biol* **2005**, *12*, 575-581.
[2] E. Lorentzen, A. Dziembowski, D. Lindner, B. Seraphin, E. Conti, *EMBO Rep* **2007**, *8*, 470-476.
[3] S. Hartung, T. Niederberger, M. Hartung, A. Tresch, K. P. Hopfner, *Nucleic Acids Res* **2010**, *38*, 5166-5176.
[4] P. Walter, F. Klein, E. Lorentzen, A. Ilchmann, G. Klug, E. Evguenieva-Hackenberg, *Mol Microbiol* **2006**, *62*, 1076-1089.
[5] V. Tugarinov, V. Kanelis, L. E. Kay, *Nat Protoc* **2006**, *1*, 749-754.
[6] B. Han, Y. Liu, S. W. Ginzinger, D. S. Wishart, *J Biomol NMR* **2011**, *50*, 43-57.

Nucleic Acids Research Advance Access published February 2, 2016

doi: 10.1093/nar/gkw062

The oligomeric architecture of the archaeal exosome is important for processive and efficient RNA degradation

Maxime J. C. Audin, Jan Philip Wurm, Milos A. Cvetkovic and Remco Sprangers*

Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

Received December 14, 2015; Revised January 20, 2016; Accepted January 25, 2016

ABSTRACT

The exosome plays an important role in RNA degradation and processing. In archaea, three Rrp41:Rrp42 heterodimers assemble into a barrel like structure that contains a narrow RNA entrance pore and a lumen that contains three active sites. Here, we demonstrate that this quaternary structure of the exosome is important for efficient RNA degradation. We find that the entrance pore of the barrel is required for nM substrate affinity. This strong interaction is crucial for processive substrate degradation and prevents premature release of the RNA from the enzyme. Using methyl TROSY NMR techniques, we establish that the 3′ end of the substrate remains highly flexible inside the lumen. As a result, the RNA jumps between the three active sites that all equally participate in substrate degradation. The RNA jumping rate is, however, much faster than the cleavage rate, indicating that not all active site:substrate encounters result in catalysis. Enzymatic turnover therefore benefits from the confinement of the active sites and substrate in the lumen, which ensures that the RNA is at all times bound to one of the active sites. The evolution of the exosome into a hexameric complex and the optimization of its catalytic efficiency were thus likely co-occurring events.

INTRODUCTION

The exosome is a large molecular machine that plays a role in the processing and degradation of the 3′ end of a large variety of RNA molecules (1). Complexes that belong to the exosome and exosome-like family share the same three-dimensional architecture and are found in all three domains of life. The simplest form of the complex is the bacterial RNase PH that has a 3′ to 5′ exoribonuclease activity (2) (Supplementary Figure S1A). The biological unit of this complex is a homo-hexamer that comprises three RNase PH dimers that assemble into a ring with six active sites (3). During the degradation reaction, the enzyme uses inorganic phosphate to release nucleotide di-phosphates from the 3′ end of the RNA. The second exosome-like complex is the polynucleotide phosphorylase (PNPase) (Supplementary Figure S1B) that is found in bacteria, chloroplasts and mitochondria. The building block of this enzyme contains two consecutive RNase PH domains, a KH and an S1 domain that are linked in one protein chain (4). Six RNase PH domains from three PNPase monomers assemble into a hexameric ring structure that contains three active sites. The exosome complex itself is found in archaea (5) and eukaryotes (6). In archaea, the core of the complex contains the two RNase PH domain proteins Rrp41 and Rrp42 (7). Three Rrp41:Rrp42 dimers assemble into a hexameric ring structure with three active sites (8,9)(Figure 1, Supplementary Figure S1C). The active sites are located in Rrp41, whereas the Rrp42 protein has lost its catalytic activity. The archaeal exosome core recruits three copies of the cap proteins Rrp4 or Csl4 that contain the RNA binding domains (9,10). The interaction of the archaeal exosome core with these cap proteins enhances the RNA degradation rates and provides substrate specificity (11,12). Besides the similar structures of the PNPase, RNase PH and archaeal exosome, these complexes share a similar phosphorolytic mechanism. The eukaryotic exosome (6) has evolved further into a fully asymmetric complex where all protein chains that form the core and all protein chains that form the cap are different (Supplementary Figure S1D and E). In plants (13), only a single subunit in the core (Exo-9) appears to be catalytically active. In other eukaryotes, all exosome subunits are inactive and form a scaffolding complex (14–16). Catalytic activity is added to the Exo-9 complex by the Rrp44 protein that harbors both exoribonucleolytic and endoribonucleolytic activity (16). Interestingly, the catalytic mechanism of the eukaryotic complex moved from phosphorolytic to hydrolytic. The removal of the phosphorolytic catalytic activities in the eukaryotic exosome barrel might have occurred to prevent polymerase activity that could result in non-specific 3′ elongation of RNA (17).

RNase PH (3), PNPase (4), the archaeal and eukaryotic exosome complexes (8,9,15,18) all assemble into barrel-like

*To whom correspondence should be addressed. Tel: +49 7071 601 1330; Fax: +49 7071 601 1308; Email: remco.sprangers@tuebingen.mpg.de

structures with a (pseudo) 3-fold symmetry (Supplementary Figure S1). The active sites are located inside the lumen of these barrels and these are thus secluded from the cellular environment. As a result access to the active sites can be regulated and erroneous RNA degradation can be prevented. For the exosome and exosome-like complexes, RNA-binding domains that are not part of the catalytic RNase PH ring can be used to provide substrate selectivity (19). Based on previous biochemical data and on published crystal structures, it can be concluded that the RNA substrate is threaded to the catalytic chamber through a central pore (the neck region), which is only large enough to accommodate one single-stranded RNA (10) (Figure 1). The functional advantage of substrate selectivity that results from the formation of the quaternary structure is, however, counterbalanced by catalytic disadvantages, as oligomerization of an enzyme into a multimeric complex reduces the number of substrates that can be degraded at the same time. As an example, three isolated Rrp41:Rrp42 dimers will be able to degrade three substrate RNAs simultaneously, whereas a trimer of Rrp41:Rrp42 dimers, as found in the archaeal exosome, is only able to degrade a single substrate at a time. The oligomerization of enzyme complexes, as is seen in the exosome family of exonucleases, is thus a trade-off between a decrease in the number of available active sites per substrate and an increase in substrate selectivity.

Here, we address whether the quaternary structure of the archaeal exosome complex from *Sulfolobus solfataricus* provides catalytic advantages. We focus on two aspects in the exosome complex that arise due to the oligomerization of the enzyme: the creation of the neck region and the establishment of a high local concentration of active sites in the lumen of the barrel. In brief, we combine methyl TROSY NMR, RNA degradation and binding experiments and find that the neck region is essential for the processivity of the enzyme. In addition, we conclude that the sequestering of active sites inside a small lumen of the complex favors RNA degradation as it ensures that the substrate is always in contact with one of the active sites. The formation of a hexameric complex thus provides significant functional advantages for the exosome and exosome-like complexes.

MATERIALS AND METHODS

Protein expression and purification

The genes for the Rrp41 and Rrp42 proteins from *Sulfolobus solfataricus* were cloned into modified pET vectors that carried an N-terminal TEV cleavable His_6-tag. In addition, a construct for the coexpression of both proteins was constructed in a modified pET vector, where only the Rrp41 protein carried an N-terminal TEV cleavable His_6-tag. Point mutations were introduced using standard site-directed mutagenesis methods.

Escherichia coli BL21 codon plus cells were transformed with the appropriate plasmids (Supplementary Table S1). Cells were grown at 37°C and proteins were over-expressed at 25°C by addition of 1 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) when an OD 600 of 0.8 was reached. Twelve hours later, the cells were pelleted by centrifugation and lysed in buffer A (50 mM $NaPO_4$ pH 7.5, 150 mM NaCl, 1 mM DTT) complemented with 10 mM imidazole, lyzozyme and 0.1% Triton X100. The cell lysate was cleared from insoluble debris by centrifugation and the supernatant was loaded on Ni-NTA resin. The resin was washed with buffer A that was complemented with 10 mM imidazole. The protein bound to the resin was eluted with buffer A complemented with 300 mM imidazole. TEV protease was added to the eluted protein and dialyzed overnight into buffer A. To remove the His_6-tagged TEV protease and the cleaved His_6-tag, the dialysate was applied to Ni-NTA resin. The final purification step was performed using size exclusion chromatography on a Superdex 200 column in buffer B (30 mM Hepes pH 6.9, 100 mM NaCl, 1 mM DTT).

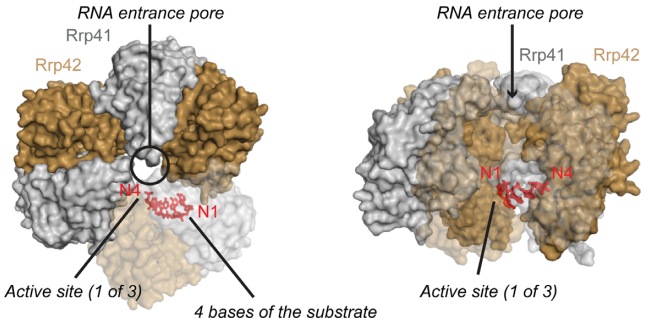

Figure 1. Structure of the *Sulfolobus solfataricus* exosome in complex with a short RNA. (PDB: 2C38) (27). Rrp41 subunits are colored gray, Rrp42 subunits are colored in light brown and the 4 bases of the RNA substrate that are visible in the crystal structure are colored in red. Left: topview of the complex, where the substrate entrance pore is indicated with a circle. Right: sideview of the complex, where the substrate entrance path is indicated with an arrow. One Rrp41:Rrp42 dimer is shown transparent to allow visualization of the inside of the barrel.

The exosome complex was reconstituted from separately purified Rrp42 (after size exclusion chromatography) and Rrp41 (after dialysis and TEV cleavage). Equal amounts of both proteins were mixed and incubated for several hours at room temperature. Uncomplexed Rrp41 or Rrp42 was removed by incubation at 50°C for 2 h. After removal of the precipitated proteins, the sample was further purified using size exclusion chromatography as described above.

Exosome complexes with a different number of active sites (on average) were obtained by mixing catalytically active and catalytically inactive Rrp41 before the addition of Rrp42. The percentage of active Rrp41 was varied between 10 and 100%. After reconstitution, the complexes were purified as described above.

Exosome complexes with exactly one, two or three active sites were obtained with plasmids containing the gene coding for Rrp42 together with three copies of the Rrp41 gene (Supplementary Table S1). These coexpression plasmids were designed as previously described (20). The first copy of Rrp41 contains a His_6-tag, the second one contains a MBP-tag and the third one contains a Strep-tag. For exosomes with a single active site, the His_6- and Strep-tagged versions of Rrp41 were catalytically inactive; for exosomes with two active sites, the His_6-tagged version of Rrp41 was catalytically active; for the exosome with three active sites, all Rrp41 versions were catalytically active (Supplementary Table S1). Cells that coexpressed Rrp42 and the three versions of Rrp41 were grown and induced as described above.

After lysis, the proteins were purified using Ni-affinity chromatography as described above. The protein that eluted from the Ni column contained at least one His_6-tagged version of Rrp41. These complexes were then applied to amylose resin to select for complexes that contained a MBP-tagged version of Rrp41 in addition to a His_6-tagged version of Rrp41. The complexes that were eluted from the amylose resin were subsequently applied to Strep-Tactin resin to select for exosome complexes that contained all three tagged versions of Rrp41. The eluted complex was subsequently treated with TEV protease to remove all affinity tags and dialyzed, prior to performing a size exclusion chromatography as described above. It is worth mentioning that the yield of the exosome complexes purified this way is significantly reduced as only 22% of the exosome complexes that are formed during over-expression contain all three different tags.

Labeling with NMR active nuclei was achieved by over-expression in minimal medium that was based on 100% D_2O. $^{12}C^2H$ glucose was used as the carbon source and methyl labeling was achieved by addition of 100 mg/l U-[^1H,^{13}C] methionine, 60 mg/l 4-methyl $^{13}CH_3$ α-ketobutyric acid (labeled isoleucine precursors) or 100 mg/l methyl $^{13}CH_3$ α-ketoisovaleric acids (labeled valine/leucine precursors) 1 h before induction with 1 mM IPTG.

Coexpression of Rrp41 and Rrp42 was used for the preparation of the complex that contained NMR active methyl groups in both Rrp41 and Rrp42. To that end, both proteins were coexpressed in NMR active growth medium (Supplementary Table S1). Purification was performed as for the single proteins.

RNA *in vitro* transcription and purification

RNA was prepared using *in vitro* transcription with T7 polymerase. The DNA template was obtained from a linearized vector. RNAs used in NMR experiments were transcribed with an HDV-ribozyme that cleaves at the end of the target RNA sequence, resulting in the presence of a 3′ cyclic phosphate. The 3′ cyclic phosphate prevents degradation of the RNA by the exosome and thus allows for long-term NMR measurements. RNAs used in binding and degradation experiments were produced using run-off transcription, where the final RNA contained a 3′ GCT that resulted from the linearization of the template vector with the HindIII restriction enzyme. All RNA constructs contained a hairpin structure (GGCCCCCCCCGAAAGGGGGGGG) followed by 32, 63, 92 or 118 adenines (Supplementary Table S1). The DNA vector containing 63, 92 or 118 adenines were obtained from genesynthesis (GenScript USA Inc.). *In vitro* transcribed RNA was purified natively with weak ion exchange chromatography using a DEAE-sepharose column as described (21). The pooled fractions were concentrated and buffer exchanged into H_2O with a PD10 column, followed by SpeedVac concentration.

Degradation assays

RNA degradation assays were performed in 180 μl reaction buffer (20 mM Hepes pH 6.5, 60 mM KCl, 0.1 mM EDTA, 2 mM DTT, 8 mM $MgCl_2$, 10 mM Na_2HPO_4) that contained 60 nM exosome (hexameric complex) and 25 μM RNA. The 10 μl samples were taken at different time-points and the reaction was quenched by addition of 10 μl 8M Urea, 20 mM EDTA, 2mM Tris pH 8.

HPLC analysis

Ten microliters of the quenched reaction were automatically injected onto an analytical DNAPac PA100 column (Dionex) that was heated to 80°C. Substrate and product were separated using a linear gradient from buffer A (5 M Urea, 20 mM Tris pH 8, 100 mM NaCl) to buffer B (5 M Urea, 20 mM Tris pH 8, 2 M NaCl) and detected using the absorption at 260 nm. To convert peak intensities to absolute concentrations, the detector response was calibrated by injecting known amounts of RNA (Supplementary Figure S6).

Analysis of degradation data

For each time-point the product concentration was divided by the total concentration [product + substrate] to normalize the signal. The progression of the reaction was then fitted from data at several time-points (Supplementary Figures S4 and S5). Based on the known amounts of enzyme and substrate together with the length of the substrate, the progression curves were translated into number of nucleotides cleaved per second per exosome. To estimate the error in the extracted catalytic rates, we used a jackknife approach, where we fitted the data multiple times after randomly removing a subset of the data.

Fluorescence anisotropy

RNA (GCCCCCCCCGAAAGGGGGGGG-A(21)-4-S-U-A(11)-GCU) for Fluorescence anisotropy measurements was obtained from Dharmacon. The attachment of the 6-(Iodoacetamido)-fluorescein (Sigma-Aldrich) to the thio-uridine (4-S-U) was performed according to the Ramos *et al.* (22). Dilution series of the inactive exosome (2000, 1000, 500, 250, 125, 60, 30, 15, 10 and 0 nM) or of the neck mutant exosome (80, 60, 40, 20, 10, 5, 2, 1, 0.5 and 0 μM) were mixed with 10 nM of RNA labeled with 6-(Iodoacetamido)-fluorescein. For the competition assays, 20 nM of either 32, 63, 98, or 118As RNA was added to the exosome:fluorescent RNA mixture. In all measurements, buffer (30 mM KPO4 pH 6.9, 100 mM NaCl, 0.005% Triton X-100) was used as a reference. Fluorescence anisotropy was recorded every 5 min using a plate reader (Tecan, Infinite F200; filter linear polarization XP38: excitation at 485 nm and emission at 535 nm). Affinity constants were obtained from the data using in-house written scripts using standard equations (23).

NMR

All NMR samples were in buffer B, based on 100% D_2O. NMR spectra were recorded on AVIII-600 and AVIII-800 spectrometers with room temperature probe-heads. Methyl TROSY spectra were recorded at 50°C using a carbon

chemical shift evolution time of 40 ms. SQ (single quantum) dispersion experiments were recorded at 600 and 800 MHz using a relaxation delay of 50 ms and Car-Purcell-Meiboom-Gill (CPGM) frequencies ranging from 40 to 1000 Hz. Relaxation dispersion data were fitted numerically using in-house written scripts using published equations (24). For the final analysis, two residues (I71 and I85), two magnetic fields (600 and 800 MHz) and three temperatures (308, 315 and 323 K) were fitted together to one intrinsic R_2 rate per curve, one exchange rate per temperature and one chemical shift difference per residue. Errors in the extracted parameters were obtained using a Monte Carlo analysis, where the measured data-points were randomly varied around the experimental error. Chemical exchange saturation transfer (CEST) experiments were recorded on an 800 MHz spectrometer at 20°C and using a 400 ms, 5 Hz B_1 field, at 42 different carbon offsets that were spaced by 10 Hz. All NMR spectra were processed using the NMRPipe/NMRDraw software suite (25). Figures displaying NMR spectra and molecular structures were produced using NMRview (onemoonscientific.com) and Pymol (pymol.org), respectively.

RESULTS

The exosome interacts tightly with RNA substrate

The first step in an enzymatic cycle is the formation of an enzyme:substrate complex. Here, we used fluorescence anisotropy measurements to determine the affinity between the exosome and an RNA substrate that contains 32 adenines downstream of a stable GC hairpin structure. To visualize the RNA, we introduced a single 4-thiouridine 15 bases downstream of the 3′ end and coupled this base to 6-(Iodoacetamido)-fluorescein. We then added increasing amounts of a catalytically inactive version of the exosome (D182A in Rrp41). Upon substrate:enzyme complex formation, the rotational lifetime of the RNA is changed, from which we extracted an affinity of 11.7 (0.9) nM (Figure 2A) for the interaction between the exosome and the RNA. This indicates that the exosome interacts tightly with substrates and that substrates can thus be recruited to the complex very efficiently.

We then asked if the interaction between the exosome and RNA depends on the length of the RNA substrate. To that end we performed fluorescence anisotropy competition experiments where we added increasing amounts of non-fluorescently labeled RNA to preformed exosome:fluorescently labeled RNA complex (Figure 2B). As competitors, we used RNA species that contain 32, 63, 92 or 118 adenines downstream of the stable GC hairpin. In these experiments, the fluorescently labeled RNA is competed away from the exosome, which results in a decrease in the fluorescence anisotropy. We then used the program DynaFit (26) to extract the kD for the competitor and found that all RNA species we tested interact with an affinities around 2.7 (0.9) nM (Figure 2B). Note that the competition experiments yield somewhat lower affinities, which is potentially due to minor interference of the fluorescence label with the binding. These data show that the affinity of the enzyme for the substrate is independent of the length of

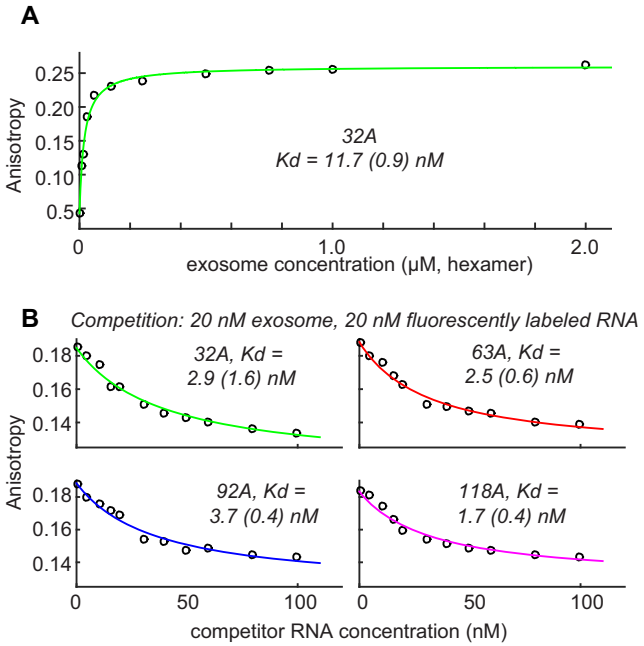

Figure 2. (**A**) Fluorescence anisotropy measurements to determine the affinity between an RNA substrate and the exosome. 32A refers to an RNA that contains a 5′ GC hairpin structure, followed by 32 adenine bases. The extracted error is based on three independent measurements. (**B**) The interaction strength between RNA substrates and the exosome is independent of the length of the RNA. Shown are fluorescence anisotropy measurements, where 20 nM exosome and 20 nM fluorescently labeled 32A RNA (see above), were complemented with increasing amounts of non-fluorescently labeled RNA. The extracted affinities are very similar for RNA species that contain 32, 63, 92 or 118 adenine bases 3′ to a GC hairpin structure. The errors in the extracted parameters result from independent measurements.

the substrate and implies that the 3′ end of the RNA is the prime recognition site for the exosome.

The tight exosome substrate interaction provides processivity

The RNA degradation of the exosome is highly processive, where RNA substrates are not released from the enzyme prior to complete degradation (27). Here, we confirm this processivity and show that an RNA substrate that contains a stable GC hairpin followed by 32 adenines is degraded into an RNA species that contains 10 adenines in addition to the hairpin (Figure 3A). This product results from the fact that the stable hairpin prevents entrance of the substrate into the exosome barrel and from the distance between the entrance pore and the active sites that spans 10 bases (27). During the degradation reaction no intermediate products are observed, indicating that the substrate is not released from the enzyme until degradation has been completed.

To shed light on the interactions that are responsible for this processive degradation we introduced a point mutation in the neck region of the exosome complex (R67G in Rrp41). Interestingly, this mutation that, due to the symmetry of the complex removes three positive charges, caused a reduction in the affinity between the RNA and the exosome from 11.7 nM (Figure 2A) to 27 μM (Figure 3B).

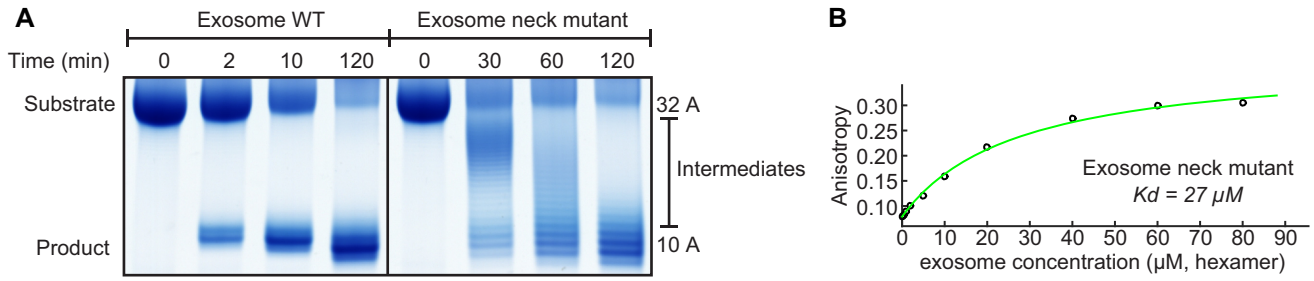

Figure 3. (**A**) Left: Substrate RNA (32A) is processively degraded into product (10A; a GC hairpin with 10 adenine bases) by WT exosome as no intermediate degradation products are detected. Right: a single point mutation in the neck region (Rrp41 R67G) abolishes the processivity as intermediate degradation products appear during the reaction. (**B**) Fluorescence anisotropy binding curve of the neck mutant exosome (Rrp41 R67G) with the fluorescently labeled 32A RNA. The single point mutation results in a 1000-fold reduction of the affinity between the RNA and the enzyme (compare: Figure 2A).

The removal of the positive charge at the entrance pore thus reduces the affinity between the exosome complex and the substrate RNA by three orders of magnitude. At the same time, the processivity of the degradation reaction has been lost as intermediate degradation products appear during the reaction (Figure 3A, right). Over time, these intermediates decrease in length and disappear as they act as substrates in subsequent rounds of degradation by the exosome complex. These results indicate that the neck region of the exosome is responsible for the tight interaction between the enzyme and the substrate and that this tight interaction results in processive RNA degradation.

The RNA is mobile inside the exosome barrel

Based on RNA degradation experiments that use a substrate that contains a stable 5′ hairpin structure, the distance between the neck region of the exosome and the active sites is 10 nucleotides (Figure 3) (27). The exact path of the RNA between the neck region and the active sites has, however, not been revealed. In crystal structures of the *Sulfolobus solfataricus* exosome, the substrate RNA is only visible for the 4 bases (bases N1 to N4; where base N1 is the 3′ terminal base; Figure 1) encompassed in the active site as well as the base N10 in the entrance pore (10,27). Except for these bases, no electron density is observed for the RNA inside the chamber. This suggests that parts of the RNA are not well ordered in the barrel of the exosome.

To obtain additional insights into the path of the substrate RNA within the exosome core, we turned to methyl TROSY NMR spectroscopy (28). This NMR method results in high quality proton-carbon NMR spectra (29) that can be used to identify intermolecular interactions (30). To that end, we recorded methyl TROSY NMR spectra of the exosome in the absence and presence of RNA substrate. Methyl groups that come close to the RNA undergo chemical shift perturbations (CSPs) that can report on the path of the substrate.

First, we prepared an exosome complex that contained NMR invisible Rrp41 and a version of Rrp42 that was NMR active in methyl groups of isoleucine, valine and leucine residues (ILV labeling). We recently obtained resonance assignments for these methyl groups and reported that the RNA interacts with residues around the pore region and with the isoleucine 85 of Rrp42 (31) that is close to the active sites of the enzyme. In case the RNA adopts a static structure with one of the three active sites in the exosome barrel, one would expect that resonances of residues close to the RNA split into two, where 2/3 of the original resonance remains and a novel resonance with intensity 1/3 arises. For residue 85, however, we observe that the resonance intensity is reduced by a factor of ∼8, whereas we failed to identify a novel resonance. This indicates that the RNA substrate is not stably bound to a single active site when it is inside the exosome barrel. To determine if other resonances from Rrp42 also undergo line broadening upon RNA interaction, we divided resonance intensities observed in the free exosome by the resonance intensities of the complex with RNA (Figure 4A). Although most resonance intensities are unaffected by the RNA interaction, we noticed that isoleucine 71 and valine 137 are also significantly weakened in the presence of the RNA. Like isoleucine 85, these residues cluster close to the four nucleotides of the 3′ end of the RNA substrate. The peak broadening of isoleucine 85 is more pronounced because it interacts directly with the RNA, whereas isoleucine 71 and valine 137 are sensing the presence of RNA through conformational changes of the loop underneath the active site (32). Interestingly, the line broadening effect is more prominent at higher temperature (50°C) than at lower temperature (20°C) (Figure 4A). Most likely, this temperature dependence is due to transient interactions between the RNA and the enzyme, where the RNA binding-unbinding causes exchange broadening of the methyl groups in the vicinity. The reduction of the line broadening at lower temperatures can be explained by slower motions of the RNA at 20°C than at 50°C. In summary, our NMR titration data suggest that the RNA is mobile inside the barrel of the exosome.

To identify residues in Rrp41 that are important for the RNA interaction, we prepared an exosome complex that contained NMR active methyl groups for the isoleucine residues of Rrp41 and Rrp42 (Figure 4B). Resonances from Rrp41 can be identified in a straightforward manner based on spectra that were recorded on an exosome that was only labeled in Rrp42 (Supplementary Figure S2). Addition of RNA to the Rrp41:Rrp42 isoleucine labeled sample resulted in a number of CSPs (Figure 4B) where CSPs in Rrp42 confirmed the data that we obtained from the Rrp42 ILV labeled sample. To identify amino acids in Rrp41

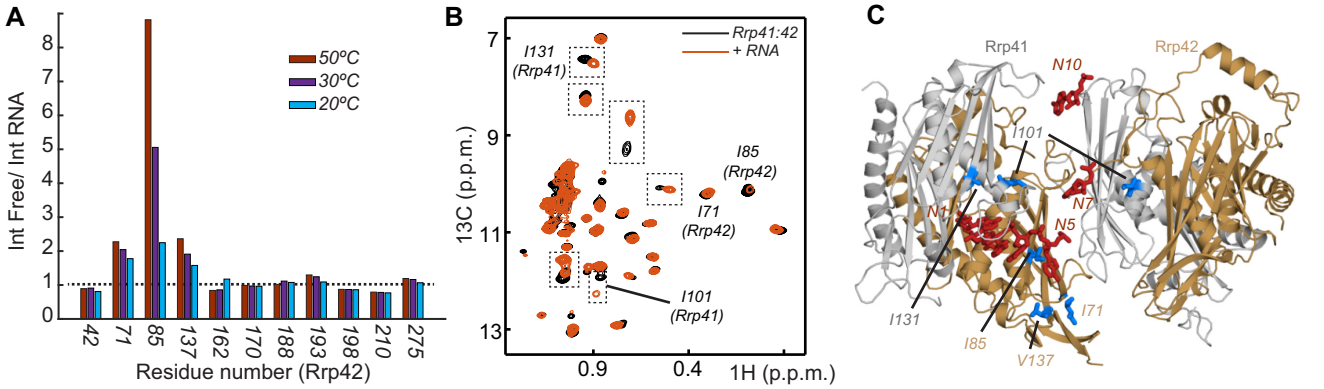

Figure 4. (**A**) The methyl group resonance intensities of Rrp42 resonances in the RNA free exosome divided by the corresponding resonance intensity in the presence of RNA. High bars indicate significant line broadening upon RNA binding, notably observed for residues Ile 71, Ile 85 and Val 137 of Rrp42. The line broadening is shown for three different temperatures, 50, 30 and 20°C. Note that a large number of resonances are not affected by the interaction with RNA (below the dashed bar). (**B**) Methyl TROSY NMR spectrum of the isoleucine region of the exosome that contains NMR active groups in Rrp41 and Rrp42. NMR spectra in the absence (black) and presence (red) of RNA are shown and a number of assignments are indicated. Regions that are highlighted with a dashed box correspond to resonances in Rrp41 that experience large CSPs upon interaction with the RNA substrate. Spectra were recorded at 323 K. (**C**) Structure of the *Sulfolobus solfataricus* exosome (PDB: 2C38) (27) superimposed onto the RNA substrate (in red) visible in the structure of the *Pyrococcus abyssi* exosome (35) (PDB: 2PO1). Assigned residues that show CSPs in the presence of RNA are indicated in blue. Only a single RNA is present per hexameric exosome complex, as the narrow entrance pore does not allow for the recruitment of multiple substrates simultaneously.

that are affected by the RNA, the NMR resonances need to be assigned to the residues in the complex. A full resonance assignment of the Rrp41 methyl groups was, in our hands, not feasible, as Rrp41 in isolation had a high tendency to aggregate. Nevertheless, we assigned a number of Rrp41 methyl group resonances using a mutational approach, where we mutated Rrp41 isoleucine residues into closely related amino acids (33,34). This ideally results in the disappearance of a single isoleucine resonance from the methyl TROSY spectrum. In that way we identified that isoleucines 101 and 131 in Rrp41 interact with the substrate RNA (Figure 4B). Interestingly, these residues shift rather than broaden upon RNA binding. In addition, we noticed that an additional set of unassigned residues in Rrp41 experiences resonance shifts (boxed regions in Figure 4B). Shifting of resonances in NMR titration experiments takes place when the binding–unbinding process is fast on the NMR chemical shift timescale. This suggests that a number of residues in Rrp41, including isoleucine 101 and 131, interact weaker with the RNA than the residues close to the active sites in Rrp42. This observation is in agreement with the lack of electron density for the nucleotides that come close to the Rrp41 protomer.

In comparison with the structure of the *Sulfolobus solfataricus* exosome (27), the structure of the *Pyrococcus abyssi* exosome (35) in complex with RNA shows additional electron density for RNA bases N5 and N7, albeit with low quality. To validate our NMR titration experiments, we superimposed the RNA of the *Pyrococcus abyssi* complex onto the structure of the *Sulfolobus solfataricus* exosome (Figure 4C). This reveals that CSPs that we observe in our NMR titration experiments are in some cases further than 5 Å away from the substrate. This is especially true for isoleucine 101 in Rrp41 that is more than 9 Å away from base N7. We can, however, explain the CSPs of isoleucine 101 in Rrp41 with the mobility of the RNA in the exosome barrel, where structural changes in the RNA result in shorter distances between the substrate and isoleucine 101.

Taken together, our NMR data and the lack of electron density of the RNA substrate (10,27,35) point to a high mobility of the substrate RNA inside the barrel of the exosome.

Quantification of the RNA motions inside the exosome barrel

To directly measure motions of the substrate RNA in the barrel of the exosome, we made use of methyl group relaxation dispersion experiments (36,37). In those experiments, the line broadening that is induced by an exchange process can be quantified and exchange rates can be extracted. We focused our analysis on isoleucine residues of Rrp42 because concentrated samples that are only labeled in Rrp42 can be produced (31). In agreement with the line broadening that we observed in isoleucine 85 and isoleucine 71 (Figure 4A), we detected significant dispersion profiles for these residues in the presence of substrate RNA (Figure 5A). Importantly, these dispersion profiles are solely due to the interaction of the enzyme with the substrate, as they were not observed in the absence of RNA (Supplementary Figure S3). In total, we measured dispersion data at three different temperatures (35, 42 and 50°C) and two magnetic field strengths (600 and 800 MHz; Supplementary Figure S3). To extract the underlying exchange parameters, we fitted all data together and assumed that the chemical shift difference was temperature independent. In addition, we assumed that the excited (RNA bound) state had a population of 1/3, as the substrate RNA can only interact with one of the three active sites at a time. Based on that, we extracted exchange rates of 1021 (163), 1615 (230) and 1744 (249) per second at 35, 42 and 50°C, respectively. The extracted chemical shift differences between the free state and the RNA bound state are 0.25 (0.01) and 0.17 (0.02) p.p.m. for isoleucine 85 and 71, respectively. Isoleucine 71 and isoleucine 85 are located close to the active site of the enzyme and the motions that we detect through those methyl groups thus report on

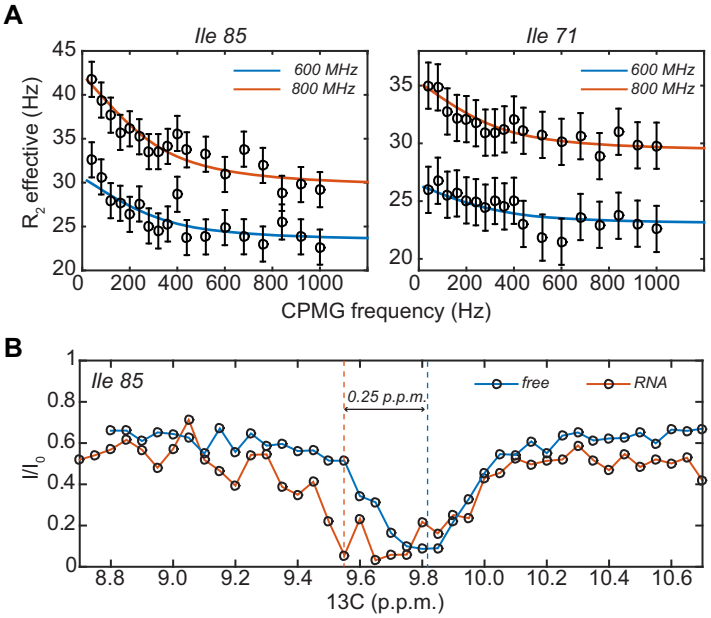

Figure 5. (**A**) Single quantum relaxation dispersion profiles in the presence of RNA. The profiles for isoleucine 71 and isoleucine 85 are shown at two magnetic field strengths and at 50°C. For clarity, the dispersion profiles at 600 MHz are plotted 5 Hz lower than they actually are. The circles indicate the measurement points, where the error bars have been derived from duplicate measurement points. The drawn lines are the best fit to the data, where both residues at three temperatures were fitted simultaneously to one exchange rate and one carbon chemical shift difference per residue. (**B**) CEST profiles for isoleucine 85 in the absence (blue) and presence (red) of the RNA substrate. In the presence of RNA, an invisible state 0.25 downfield of the main resonance of isoleucine 85 appears.

the binding–unbinding of the substrate RNA with the active site region of the exosome. We conclude that the RNA moves from one active site to the next one with a frequency between 1000 and 1700 Hz, depending on the temperature. These data reinforce the notion that the RNA is highly mobile in the lumen of the exosome.

To validate the analysis of the relaxation dispersion experiments, we used CEST experiments (38) to identify the resonance frequency of the invisible RNA-bound state of isoleucine 85. The exchange rates that we determined between 35 and 50°C are too fast for efficient CEST to occur. Therefore, we lowered the experimental temperature to 20°C. Unfortunately, at temperatures lower than 20°C, the signal to noise ratio of the NMR spectra drops significantly, which prevented us from measuring at even lower temperatures. In the CEST experiment, we clearly observe the presence of a second and invisible state for isoleucine 85 (Figure 5B). Interestingly, this invisible state is located 0.25 p.p.m. downfield from the resonance frequency of residue 85. This is in excellent agreement with the chemical shift difference extracted from the relaxation dispersion experiments and confirms that the parameters that we extracted from the relaxation dispersion experiments are accurate and reliable.

The exosome exploits all three active sites during catalysis

Our NMR data show that the 3′ end of the RNA substrate is highly mobile inside the lumen of the exosome. Based on the relaxation dispersion experiments, the substrate interacts with the active sites around 1700 times per second at 50°C. This raises the question whether all three active sites are used or potentially even required during the degradation process. To address this, we used two complementary biochemical experiments where we measured the activity of exosome complexes that contained different numbers of active sites.

First, we reconstituted exosome complexes from separately expressed Rrp42 and Rrp41. To vary the number of active sites in the reconstituted complex, we used different mixtures of catalytically active and catalytically inactive (D182A, that does not interfere with the RNA binding) Rrp41. For example, when a mixture of 40% active and 60% inactive Rrp41 is used in the reconstitution process, statistically the following complexes will form: exosome complexes without any active sites (22%), with one active site (43%), with two active sites (29%) and with three active sites (6%). The relation between the average number of active sites and the activity of the complex depends on the mechanism that is used during the degradation process. In case all active sites equally and independently contribute to the reaction, the activity will linearly increase with the average number of active sites (Figure 6A, red curve). Alternatively, in case only one active site is used in the degradation process (e.g. when the substrate stays on a single active site during the degradation process), the activity will level off with increasing average number of active sites (Figure 6A, blue curve). In case all three active sites are essential for efficient catalysis, the activity of the exosome will only reach high levels when there are on average a high number of active sites (Figure 6A, green curve). Experimentally we can distinguish between these three scenarios by measuring the catalytic activity of exosomes that (on average) contain a different number of active sites. Here, we increased the average number of active sites in a stepwise manner from 10% to 100% in steps of 10%

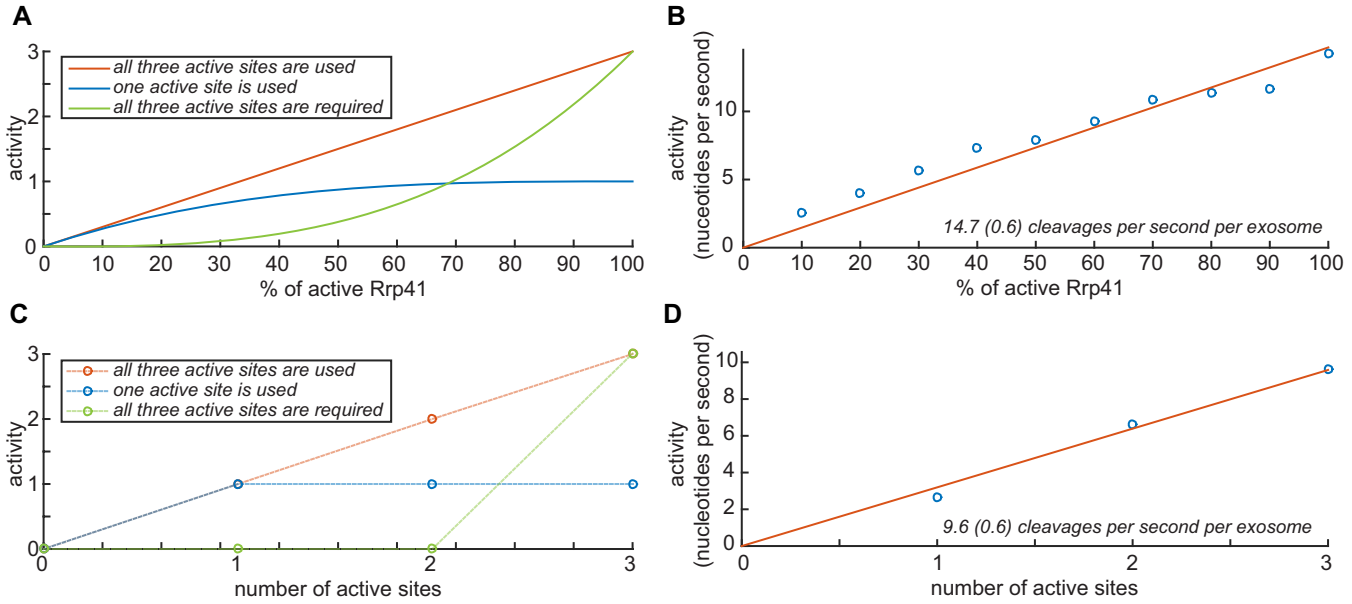

Figure 6. (**A**) Theoretical relationship between the activity and number of active sites inside the exosome complexes, which were reconstituted, with different ratios of active and inactive Rrp41. In case all three active sites in Rrp41 are independently used in the RNA degradation process, the predicted relation between the activity and the percentage of active Rrp41 is shown in red. The blue and green relationships apply to the situation where only one active site is used and where the exosome requires all three sites for full activity, respectively. (**B**) Experimental relationship between the percentage of active Rrp41 in the exosome and the activity. The blue circles indicate the degradation rates (nucleotides cleaved per second per exosome). The red line is the best fit to the data from which the activity of the fully active exosome is determined to be 14.7 cleavages per exosome at 50°C. The data clearly show that all three sites in the exosome are used in the degradation process. (**C**) Theoretical relationship between the activity and the discrete number of active sites inside the exosome. The dashed lines are shown for clarity only. See also (A). (**D**) Experimental relationship between the number of discrete active sites in the exosome complex and the activity. The blue circles refer to the measurements, the red line is the best fit to the data. As in (B), the data show that all three sites in the exosome are used in the degradation process.

(see Methods). Each of these exosome complexes was incubated with substrate RNA and the reaction was quenched at different time-points. In those assays, we choose to use the RNA substrate that contains a 5′ hairpin structure followed by 32 adenine nucleotides. The product of the degradation reaction will then be the hairpin structure with 10 nucleotides, which can be readily detected (see above). Initially, we quantified the levels of substrate and product during the reaction using an Urea PAGE analysis (Supplementary Figure S4), however, we found that quantification using an HPLC approach was more accurate (Supplementary Figure S5) and thus we used this method for all degradation experiments. To obtain degradation rates, we fitted the substrate and product concentrations to a progression curve (Supplementary Figures S4, S5, S6 and methods). In summary, we observe that the number of cleaved nucleotides increases linearly with increasing amounts of active Rrp41 (Figure 6B). These data thus show that all three active sites in the exosome are used during the degradation process. Based on these data, we conclude that one substrate RNA is degraded by three active sites in one exosome. At 50°C, this results in 14.7 (0.6) nucleotide cleavages per second per exosome (Figure 6B).

In a second and independent approach to determine how many active sites in the exosome play a role in the RNA degradation process, we prepared three different samples with exosome complexes that contained exactly one, two or three active sites. These complexes were obtained by overexpression of untagged Rrp42 with three copies of Rrp41, each of them fused to a different affinity tag (His-tag, MBP-tag or Strep-tag). Importantly, we were not able to detect any subunit exchange between different exosome complexes (Supplementary Figure S7), establishing that the exosome complexes with a discrete number of active sites are extremely stable. Subsequently, three consecutive purification steps were used to obtain complexes that contained exactly one of each Rrp41 affinity tags. During expression of the exosome complex, we used one, two or three catalytically active versions of the differently tagged Rrp41 proteins and were thus able to prepare exosome complexes with a discrete number of active sites. As described before, the activity of these complexes can provide information on the mechanism that is used during RNA degradation (Figure 6C, see above). We then experimentally determined the activity of the exosome complexes that harbored a discrete number of active sites and found that all three active sites are equally involved during the degradation process (Figure 6D). This confirms the experiments that we performed using exosomes with mixed number of active sites. The overall activity of the exosome complexes with a discrete number of active sites appears somewhat lower (9.6 cleavages per second per exosome at 50°C) than the activity of the exosome complexes that contain mixture of active and inactive sites. This is most likely due to a loss in activity of the exosome complex during the long purification protocol that is required for the preparation of the exosome complexes with a discrete number of active sites.

The RNA motions are much faster than the enzymatic turnover rates

Our degradation assays show that the RNA substrate uses all three active sites in the exosome barrel. This is in agreement with the NMR data that show that the substrate rapidly exchanges between the active sites. Interestingly, at 50°C the rate of exchange (1700 s^{-1}) is two orders of magnitude faster than the number of cleavages per second per exosome (~10 s^{-1}). It is worth noting that these differences are not a peculiarity of this temperature. Indeed, we compared degradation experiments at 20, 35, 50 and 65°C (Figure 7A) and found that the temperature dependence of the degradation rate follows the Eyring relationship (Figure 7B). The temperature dependence of the RNA hopping frequency in the exosome barrel, as we determined using relaxation dispersion experiments, follows a very similar trend, albeit at much higher frequencies. This indicates that the RNA hopping frequency is significantly larger than the degradation rates for biologically relevant temperatures.

From the temperature dependence of the activity of the exosome core (Figure 7B) the RNA degradation rate at 75°C (the optimal growth temperature of *Sulfolobus solfataricus*) can be predicted to be around 65 nucleotides/second. This degradation rate is comparable to the RNA degradation rate that was previously determined for the PNPase enzyme (120 nucleotides/second) (39). In addition, the rate of RNA degradation by the exosome core is in the same order of magnitude as the elongation rate of the archaeal RNA polymerases that was determined to be around 20 nucleotides per second in *Methanothermobacter thermautotrophicus* (40).

The degradation rate is independent on the RNA length

The catalytic cycle of the exosome includes multiple steps. If substrate binding or product release is very slow, the exosome would require a significant time between finishing the degradation of one substrate and initiating the degradation of the next one. Such a time would reduce the average cleavage frequency that we measured in our biochemical experiments. To probe whether substrate binding or product release are rate limiting in the catalytic cycle of the exosome we performed degradation assays with RNA substrates of increasing length (32, 63, 92 and 118 adenine repeats, that all interact with the exosome with similar affinities, Figure 2B). For short RNA substrates, the exosome would need to reload the substrate significantly more often than for long substrates, which would reduce the overall turnover rate. In our experiments, we designed the substrates such that short (less than ~10 nucleotides) single stranded RNA stretches are not formed (Supplementary Figure S4A), as these have been shown to be degraded at very low rates (41). Interestingly, in our degradation experiments, we find that the activity of the exosome (number of nucleotides cleaved per second) is very similar for the four RNAs used, showing that the activity of the exosome is largely independent of the length of the substrate (Figure 7C). This shows that substrate binding and product release are not significantly limiting catalytic turnover. The nucleotide cleavage frequencies that we measured thus directly report on the activity of the active sites. This validates our conclusion that the RNA jumping frequencies are much faster than the catalytic cleavage rates.

DISCUSSION

Self-compartmentalization is a principle that is exploited not only by the exosome, but also by other enzymes, including the proteasome (42). The sequestering of active sites inside a small space prevents degradation of substrates that are not actively targeted to the enzyme. For the archaeal exosome, substrate selection takes place through RNA interacting proteins that dock around the entrance pore of the barrel (9,10,19). In the eukaryotic exosome, these RNA-binding proteins have evolved further and include helicases that are able to unfold RNA species that contain secondary structure (43).

Despite the functional advantages that are related to substrate selection, self-compartmentalization comes at a cost: multiple protomers in a larger complex can only act on a single substrate whereas the same number of monomeric proteins could act on multiple substrates simultaneously. For the *Archaeoglobus fulgidus* exosome, the activity of the native $(Rrp41:Rrp42)_3$ exosome complex was previously compared to the activity of a version of the complex that only assembles into $(Rrp41:Rrp42)_1$ dimers (41). Interestingly, for long RNA substrates, it was found that the catalytic activity of one $(Rrp41:Rrp42)_3$ hexameric complex is higher than that of three $(Rrp41:Rrp42)_1$ dimers. One $(Rrp41:Rrp42)_3$ exosome that acts on one substrate is thus more efficient than three $(Rrp41:Rrp42)_1$ dimers that act on three substrates simultaneously. For the *Archaeoglobus fulgidus* exosome complex, the advantages of self-compartmentalization thus outweigh the disadvantages. The molecular basis that underlies this gain in activity upon assembly of the Rrp41 and Rrp42 protomers into a barrel like quaternary structure remained undetermined.

Here, we addressed the catalytic advantages of oligomerization of the *Sulfolobus solfataricus* exosome. Two features that appear in the enzyme upon oligomerization of the Rrp41 and Rrp42 proteins into a hexameric barrel are the entrance pore (the neck region) and a lumen that contains a very high concentration of active sites. We find that these two aspects are fundamental to the efficiency of the exosome complex (Figure 8).

Using binding measurements, we show that the neck region, where the RNA enters the exosome lumen, strongly interacts with unstructured RNA. Interestingly, this interaction involves substrate nucleotides that are located 10 bases upstream of the degradation site. This very strong interaction ensures that substrates can be recruited efficiently, and at the same time prevents that substrates are released from the enzyme complex before complete degradation (Figures 2 and 3). Experimentally we have shown this using an exosome complex that contains a single point mutation in the neck region. This mutation results in a 1000-fold decrease of the enzyme substrate interaction strength (Figure 3B). Importantly, also the processivity that is observed in the WT complex is lost upon weakening the neck-substrate interactions (Figure 3A) as substrates can no longer be retained to the enzyme complex during turnover. The importance of the neck region for RNA degradation appears to be

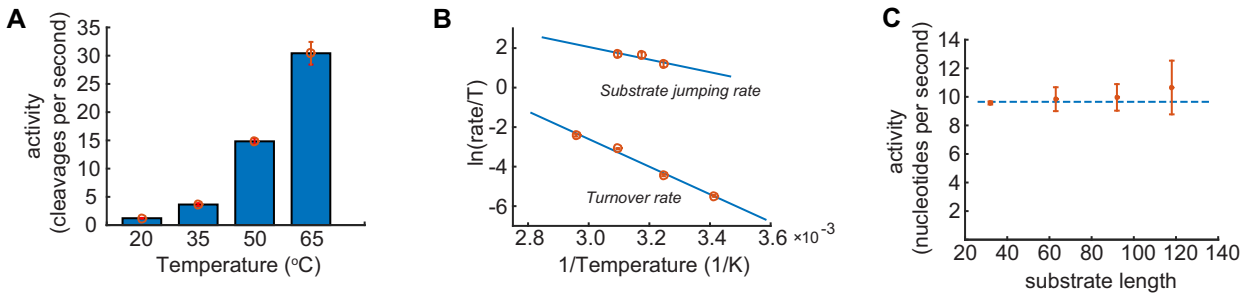

Figure 7. (**A**) Effect of the temperature on the activity of the exosome. (**B**) Eyring analysis of the temperature dependence of the degradation rates (lower line) and NMR jumping rates (upper line). The lines follow the same trend, indicating that the substrate jumping rate is much faster than the degradation rate at all biologically relevant temperatures. (**C**) The number of nucleotides cleaved by the exosome is independent of the length of the RNA. This shows that substrate recruitments or product release are not rate limiting in the reaction. The larger errors for the 118A RNA substrates might result from some inhomogeneity in the exact length of the substrate that resulted from slippage or stalling of the T7 polymerase that was used to prepare this long and highly repetitive sequence.

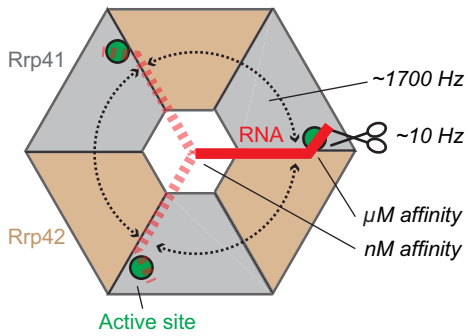

Figure 8. Cartoon that summarizes our findings. The RNA (red, dashed red) enters the exosome barrel (gray, sand) through the neck and can interact with one of three active sites (green circles). The substrate jumps between these three active sites with frequencies of ∼1700 s^{-1} (rounded dashed arrows). The catalytic cleavage rates are in the order of 10 s^{-1} (scissor symbol), two orders of magnitude slower than the RNA jumping rate. Only 1 in 100 RNA:exosome encounters results in a catalytic reaction. The exosome:RNA nM affinity interactions through the neck region keep the 3′ end of the RNA in the exosome barrel. The high concentration of active sites and substrate in the barrel of the exosome ensures that the 3′ end of the RNA is always in contact with one of the active sites.

conserved in exosome and exosome-like complexes (Supplementary Figure S1). First, it has been shown that the hexameric structure of the RNAse PH complex is essential for activity as a dimeric form of the enzyme that does not form a neck region is unable to interact with RNA (44). Secondly, mutations of basic residues in the entrance channel of the PNPase complex result in a loss of activity and processivity (45). Finally, mutations within the Rrp41 neck region of the yeast exosome have been shown to be important for the channeling of substrates through the catalytically inactive Exo-9 complex toward the catalytically active Rrp44 subunit (18).

To address the importance of the high concentration of active sites in the lumen of the exosome, we determined the kinetics of the substrate in the proximity of the active sites. Using methyl TROSY NMR methods we show that the 3′ end of the RNA is highly mobile and that it jumps between the active sites with a frequency of around 1700 Hz at 50°C (Figure 5). This high mobility will allow for a rapid dissociation of the reaction products, thereby facilitating turnover. Based on fluorescence anisotropy experiments that use a mutant where the neck interactions are impaired, the affinity between the active sites and the 3′ end of the RNA is in the μM regime (Figure 3B). This affinity is in line with previous ITC measurements that used a very small RNA substrate (46). Based on this μM affinity and the determined exchange rate of the RNA in the barrel it can be concluded that the on-rate of the binding process is fast (in the order of 10^9 $M^{-1}s^{-1}$). This fast on-rate is probably a direct consequence of the small volume of the lumen of the exosome and the resulting high substrate concentration. Using a series of RNA degradation experiments (Figures 6 and 7) we show that the motions of the RNA allow that all active sites in the exosome lumen equally participate in RNA degradation. Interestingly, the number of nucleotide cleavage events (∼10 s^{-1}) is two orders of magnitude lower than the number of active-site: substrate encounters (∼1700 s^{-1}). This shows that around 100 encounter complex formation events are required for one cleavage event.

When one assumes that the exosome has a spherical lumen with a diameter of 32 Å, the concentration of active sites in the enclosed volume can be estimated to be around 300 mM whereas the local substrate concentration is one-third of that (one RNA can enter the lumen). Based on the μM affinity of the 3′ end of the RNA for the active sites, the active-site occupancy can be predicted to be essentially 100%. As soon as RNA enters the lumen of the exosome it will thus be bound to one of the three active sites. In that light, the oligomerization of the exosome enhances catalytic efficiency by ensuring that the enzyme interacts with the substrate in a highly efficient manner at all times. In the theoretical case, where the archaeal exosome would only contain a single active site, the RNA would still be in contact with this site for more than 98% of the time. The reduction of the number of active sites in exosome(-like) complexes (6 active sites in RNase PH, 3 active sites in the PNPase and archaeal exosome, one in the plant exosome; Supplementary Figure S1) does not result in a reduction of the catalytic activity; the RNA will always be in full contact with an active site. After oligomerization of the exosome-like complexes into a barrel-like quaternary structure, the diversification of the subunits and the removal of active sites thus posed no catalytic disadvantages. It should

be noted that in our degradation experiments (Figure 6) we intentionally removed activity through a single-point mutation that did not alter the interaction with RNA. In that situation, the 3′ end of the RNA partitions between active and inactive sites, which leads to the observed reduction in activity. When the removal of active sites is accompanied by the removal of the substrate interaction, the activity should stay at its maximum, independent of the number of active sites. In that light, it should be noted that the eukaryotic exosome indeed makes no contacts with the substrate RNA as all contact points have been removed (47).

In summary, we show that the oligomerizaiton of the exosome complex into a barrel like structure provides novel catalytic advantages. The basis for this lies in the fact that the enzyme interacts very strongly with the substrate close to base N10, whereas the 3′ end of the RNA close to base N1 remains highly flexible (Figure 8). The evolution of the hexameric exosome complex from single protein chains might have benefited from the increase in catalytic efficiency that is associated with the formation of the quaternary structure.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We acknowledge all members of the laboratory for discussions. We thank Silke Wiesner for helpful suggestions and Lewis E. Kay (University of Toronto) for sharing the code for the CEST NMR experiment with us.

FUNDING

M.A.C. acknowledges funding from the IMPRS ‘From Molecules to Organisms’. This work was supported by the Max Planck Society and the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013), ERC grant agreement no. 616052.
Conflict of interest statement. None declared.

REFERENCES

1. Houseley,J., LaCava,J. and Tollervey,D. (2006) RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.*, **7**, 529–539.
2. Deutscher,M.P., Marshall,G.T. and Cudny,H. (1988) RNase PH: an Escherichia coli phosphate-dependent nuclease distinct from polynucleotide phosphorylase. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 4710–4714.
3. Ishii,R., Nureki,O. and Yokoyama,S. (2003) Crystal structure of the tRNA processing enzyme RNase PH from Aquifex aeolicus. *J. Biol. Chem.*, **278**, 32397–32404.
4. Symmons,M.F., Jones,G.H. and Luisi,B.F. (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure*, **8**, 1215–1226.
5. Koonin,E.V., Wolf,Y.I. and Aravind,L. (2001) Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. *Genome Res.*, **11**, 240–252.
6. Mitchell,P., Petfalski,E., Shevchenko,A., Mann,M. and Tollervey,D. (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′–>5′ exoribonucleases. *Cell*, **91**, 457–466.
7. Evguenieva-Hackenberg,E., Walter,P., Hochleitner,E., Lottspeich,F. and Klug,G. (2003) An exosome-like complex in Sulfolobus solfataricus. *EMBO Rep.*, **4**, 889–893.
8. Lorentzen,E., Walter,P., Fribourg,S., Evguenieva-Hackenberg,E., Klug,G. and Conti,E. (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat. Struct. Mol. Biol.*, **12**, 575–581.
9. Buttner,K., Wenig,K. and Hopfner,K.P. (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol. Cell*, **20**, 461–471.
10. Lorentzen,E., Dziembowski,A., Lindner,D., Seraphin,B. and Conti,E. (2007) RNA channelling by the archaeal exosome. *EMBO Rep.*, **8**, 470–476.
11. Evguenieva-Hackenberg,E., Roppelt,V., Finsterseifer,P. and Klug,G. (2008) Rrp4 and Csl4 are needed for efficient degradation but not for polyadenylation of synthetic and natural RNA by the archaeal exosome. *Biochemistry*, **47**, 13158–13168.
12. Roppelt,V., Klug,G. and Evguenieva-Hackenberg,E. (2010) The evolutionarily conserved subunits Rrp4 and Csl4 confer different substrate specificities to the archaeal exosome. *FEBS Lett.*, **584**, 2931–2936.
13. Chekanova,J.A., Shaw,R.J., Wills,M.A. and Belostotsky,D.A. (2000) 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. *J. Biol. Chem.*, **275**, 33158–33166.
14. Januszyk,K. and Lima,C.D. (2014) The eukaryotic RNA exosome. *Curr. Opin. Struct. Biol.*, **24**, 132–140.
15. Liu,Q., Greimann,J.C. and Lima,C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–1237.
16. Dziembowski,A., Lorentzen,E., Conti,E. and Seraphin,B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.*, **14**, 15–22.
17. Wahle,E. (2007) Wrong PH for RNA degradation. *Nat. Struct. Mol. Biol.*, **14**, 5–7.
18. Bonneau,F., Basquin,J., Ebert,J., Lorentzen,E. and Conti,E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell*, **139**, 547–559.
19. Walter,P., Klein,F., Lorentzen,E., Ilchmann,A., Klug,G. and Evguenieva-Hackenberg,E. (2006) Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon Sulfolobus solfataricus. *Mol. Microbiol.*, **62**, 1076–1089.
20. Mund,M., Overbeck,J.H., Ullmann,J. and Sprangers,R. (2013) LEGO-NMR spectroscopy: a method to visualize individual subunits in large heteromeric complexes. *Angew. Chem. Int. Ed. Engl.* , **52**, 11401–11405.
21. Easton,L.E., Shibata,Y. and Lukavsky,P.J. (2010) Rapid, nondenaturing RNA purification using weak anion-exchange fast performance liquid chromatography. *RNA*, **16**, 647–653.
22. Ramos,A. and Varani,G. (1998) A new method to detect long-range protein-RNA contacts: NMR detection of electron-proton relaxation induced by nitroxide spin-labeled RNA. *J. Am. Chem. Soc.*, **120**, 10992–10993.
23. Johnson,P.E., Tomme,P., Joshi,M.D. and McIntosh,L.P. (1996) Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of Cellulomonas fimi CenC 2. NMR and ultraviolet absorption spectroscopy. *Biochemistry*, **35**, 13895–13906.
24. Korzhnev,D.M., Kloiber,K. and Kay,L.E. (2004) Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: theory and application. *J. Am. Chem. Soc.*, **126**, 7320–7329.
25. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
26. Kuzmic,P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.*, **237**, 260–273.
27. Lorentzen,E. and Conti,E. (2005) Structural basis of 3′ end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol. Cell*, **20**, 473–481.
28. Tugarinov,V., Hwang,P.M., Ollerenshaw,J.E. and Kay,L.E. (2003) Cross-correlated relaxation enhanced 1H[bond]13C NMR

spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.*, **125**, 10420–10428.
29. Sprangers,R. and Kay,L.E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, **445**, 618–622.
30. Wiesner,S. and Sprangers,R. (2015) Methyl groups as NMR probes for biomolecular interactions. *Curr. Opin. Struct. Biol.*, **35**, 60–67.
31. Audin,M.J., Dorn,G., Fromm,S.A., Reiss,K., Schutz,S., Vorlander,M.K. and Sprangers,R. (2013) The archaeal exosome: identification and quantification of site-specific motions that correlate with cap and RNA binding. *Angew. Chem. Int. Ed. Engl.*, **52**, 8312–8316.
32. Lorentzen,E. and Conti,E. (2012) Crystal structure of a 9-subunit archaeal exosome in pre-catalytic states of the phosphorolytic reaction. *Archaea*, 721869.
33. Sprangers,R., Gribun,A., Hwang,P.M., Houry,W.A. and Kay,L.E. (2005) Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 16678–16683.
34. Amero,C., Asuncion Dura,M., Noirclerc-Savoye,M., Perollier,A., Gallet,B., Plevin,M.J., Vernet,T., Franzetti,B. and Boisbouvier,J. (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J. Biomol. NMR*, **50**, 229–236.
35. Navarro,M.V., Oliveira,C.C., Zanchin,N.I. and Guimaraes,B.G. (2008) Insights into the mechanism of progressive RNA degradation by the archaeal exosome. *J. Biol. Chem.*, **283**, 14120–14131.
36. Korzhnev,D.M., Kloiber,K., Kanelis,V., Tugarinov,V. and Kay,L.E. (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J. Am. Chem. Soc.*, **126**, 3964–3973.
37. Lundstrom,P., Vallurupalli,P., Religa,T.L., Dahlquist,F.W. and Kay,L.E. (2007) A single-quantum methyl 13C-relaxation dispersion experiment with improved sensitivity. *J. Biomol. NMR*, **38**, 79–88.
38. Bouvignies,G. and Kay,L.E. (2012) A 2D (1)(3)C-CEST experiment for studying slowly exchanging protein systems using methyl probes: an application to protein folding. *J. Biomol. NMR*, **53**, 303–310.
39. Fazal,F.M., Koslover,D.J., Luisi,B.F. and Block,S.M. (2015) Direct observation of processive exoribonuclease motion using optical tweezers. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 15101–15106.
40. Xie,Y. and Reeve,J.N. (2004) Transcription by an archaeal RNA polymerase is slowed but not blocked by an archaeal nucleosome. *J. Bacteriol.*, **186**, 3492–3498.
41. Hartung,S., Niederberger,T., Hartung,M., Tresch,A. and Hopfner,K.P. (2010) Quantitative analysis of processive RNA degradation by the archaeal RNA exosome. *Nucleic Acids Res.*, **38**, 5166–5176.
42. Makino,D.L., Halbach,F. and Conti,E. (2013) The RNA exosome and proteasome: common principles of degradation control. *Nat. Rev. Mol. Cell Biol.*, **14**, 654–660.
43. Halbach,F., Reichelt,P., Rode,M. and Conti,E. (2013) The yeast ski complex: crystal structure and RNA channeling to the exosome complex. *Cell*, **154**, 814–826.
44. Choi,J.M., Park,E.Y., Kim,J.H., Chang,S.K. and Cho,Y. (2004) Probing the functional importance of the hexameric ring structure of RNase PH. *J. Biol. Chem.*, **279**, 755–764.
45. Shi,Z., Yang,W.Z., Lin-Chao,S., Chak,K.F. and Yuan,H.S. (2008) Crystal structure of Escherichia coli PNPase: central channel residues are involved in processive RNA degradation. *RNA*, **14**, 2361–2371.
46. Oddone,A., Lorentzen,E., Basquin,J., Gasch,A., Rybin,V., Conti,E. and Sattler,M. (2007) Structural and biochemical characterization of the yeast exosome component Rrp40. *EMBO Rep.*, **8**, 63–69.
47. Makino,D.L., Schuch,B., Stegmann,E., Baumgartner,M., Basquin,C. and Conti,E. (2015) RNA degradation paths in a 12-subunit nuclear exosome complex. *Nature*, **524**, 54–58.

The oligomeric architecture of the archaeal exosome is important for processive and efficient RNA degradation

Maxime J. Audin, Jan Philip Wurm, Milos A. Cvetkovic and Remco Sprangers*

Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

* To whom correspondence should be addressed.
Email: remco.sprangers@tuebingen.mpg.de

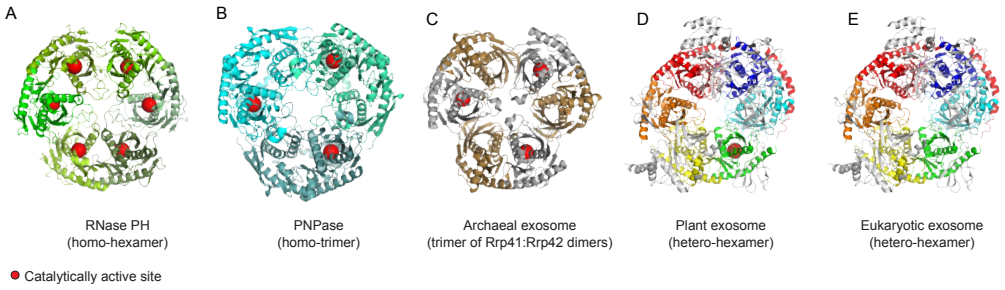

Figure S1. Structures of the exosome and exosome-like complexes.

(A) The homo-hexameric RNase PH (PDB: 1UDN) (1) that harbors six active sites (red spheres). The six identical protein chains are colored in different shades of green.

(B) The home-trimeric PNPase (PDB: 1E3H) (2) that harbors three active sites (red spheres). The three identical protein chains are colored in different shades of cyan.

(C) The archaeal exosome complex (PDB: 2BR2) (3) that contains three Rrp41 (grey) Rrp42 (sand) dimers. Only Rrp41 contains a catalytically active site.

(D) The core of the eukaryotic exosome structure (PDB: 2NN6) (4) that contains six different protein chains. The plant exosome appears to have a single catalytic active site in Rrp41. The three different proteins that form the cap of the complex are colored grey.

(E) As in (D), eukaryotic exosomes from other species have lost all catalytically active sites and act as a scaffolding complex.

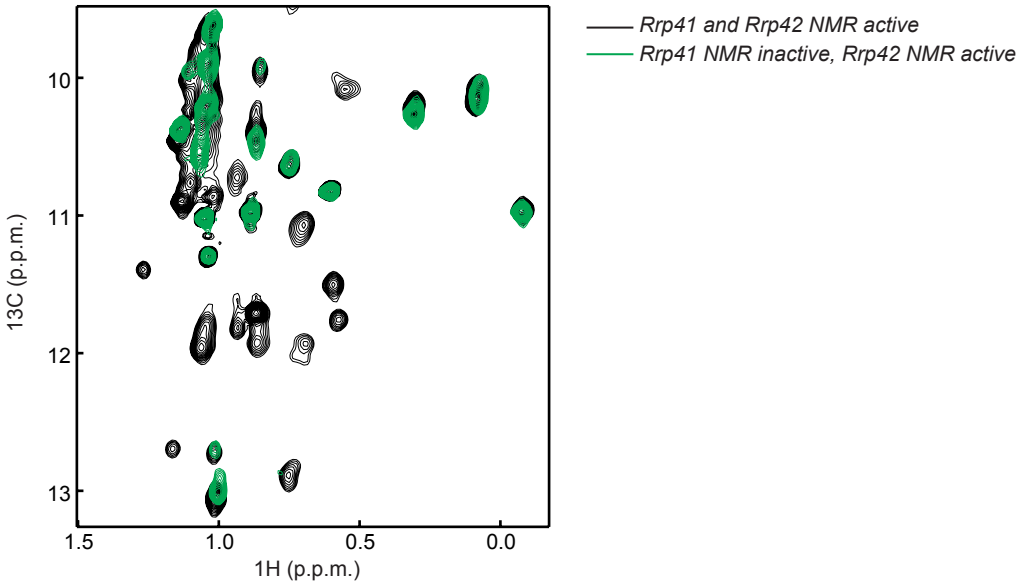

Figure S2.

Overlay of methyl TROSY NMR spectra of the exosome complex that contains NMR active isoleucine residues in both Rrp41 and Rrp42 (black) or only in Rrp42 (green). Resonances that result from residues in Rrp41 can be readlily identified.

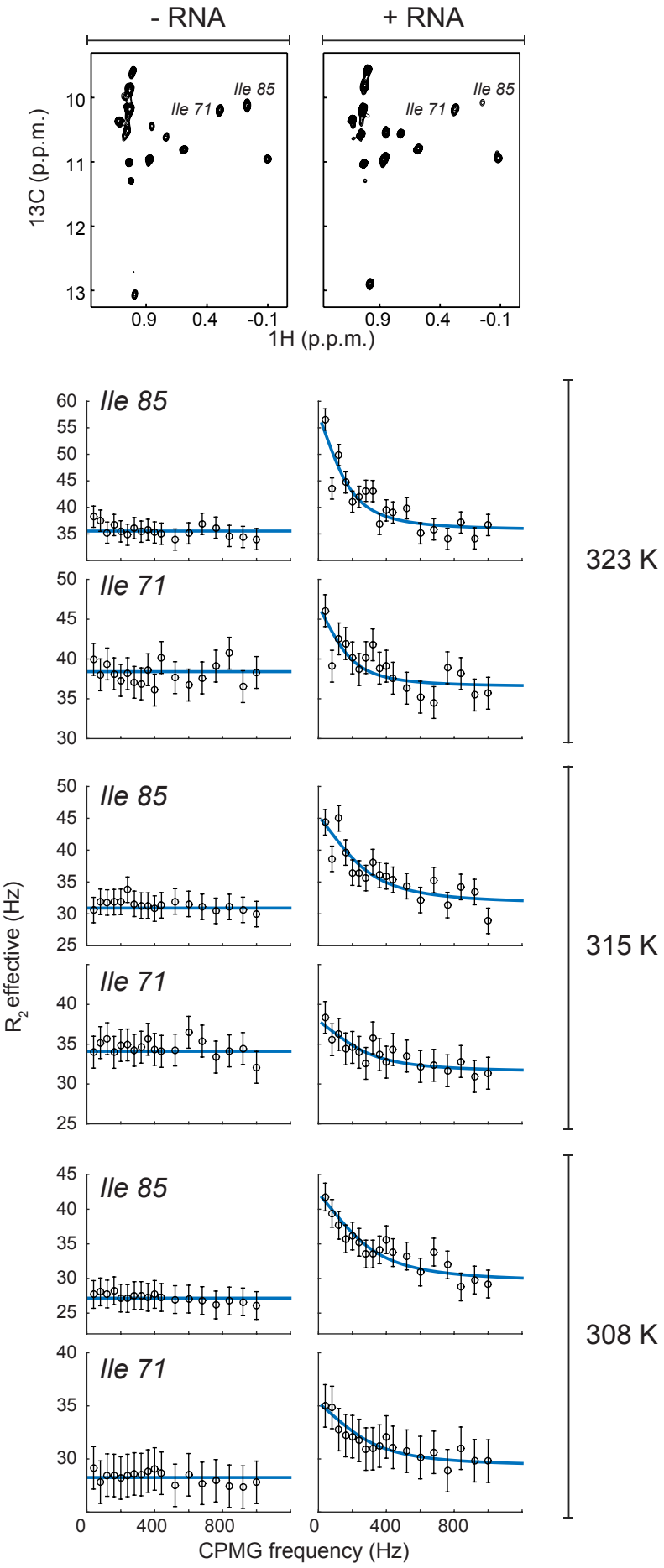
- RNA
+ RNA
Ile 85
Ile 71
13C (p.p.m.)
1H (p.p.m.)
323 K
315 K
308 K
R2 effective (Hz)
CPMG frequency (Hz)

Figure S3.

Top part: methyl TROSY NMR spectra of Rrp42 in the exosome complex in the absence (left) and presence (right) of RNA substrate. Bottom part: Isoleucine 71 and 85 show flat dispersion profiles in the absence of RNA (left column). In the presence of substrate RNA, these residues show motion on the ms timescale. The motions that we observe in the exosome:RNA complex are thus a direct result of the RNA. The shown profiles have been recorded at 800 MHz.

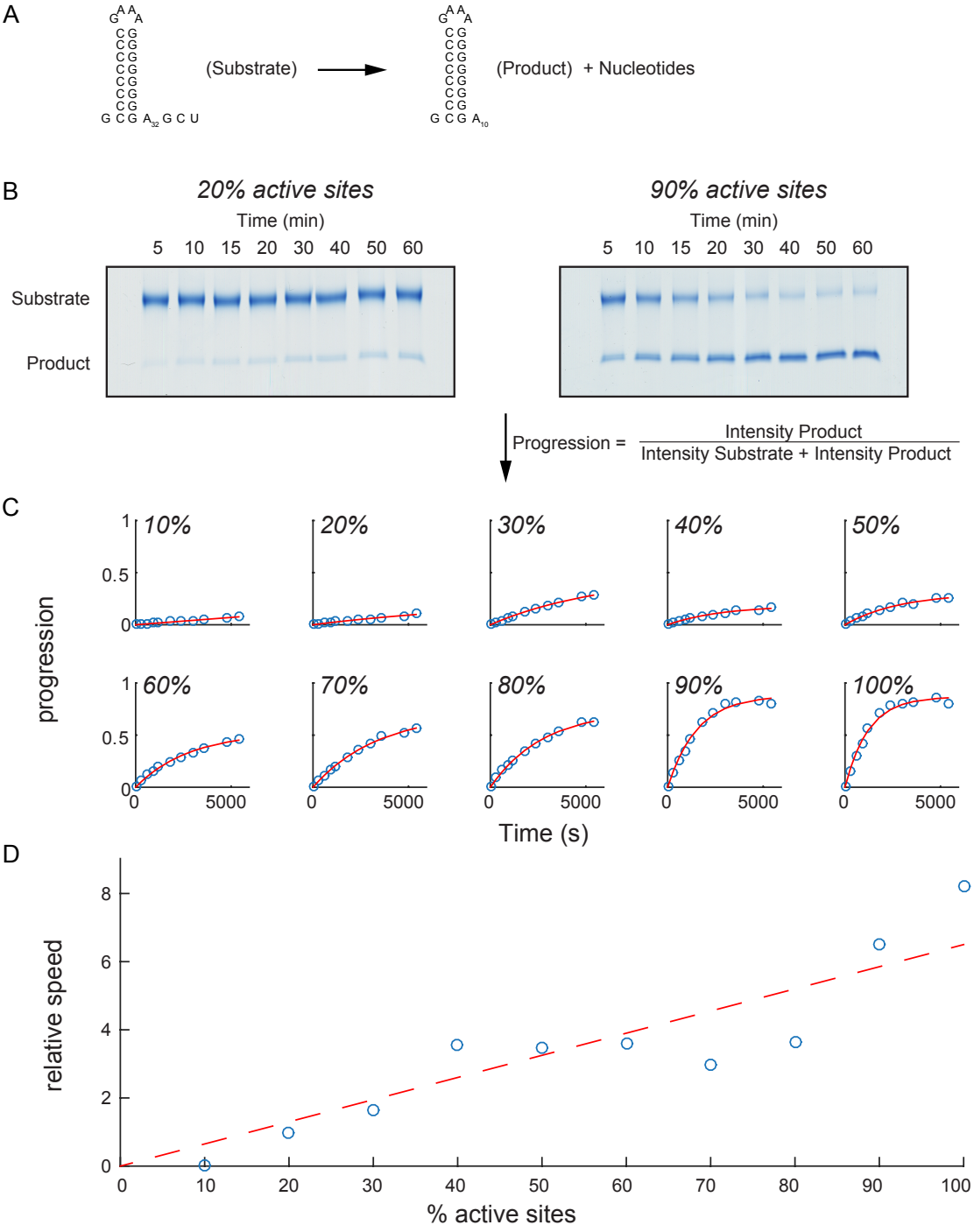

Figure S4.

(A) In our experiments the RNA substrate (a GC hairpin followed by 32 adenines, where the 3' GCU nucleotides result from the linearization of the DNA template used for *in vitro* transcription) is processively degraded into the product RNA (a GC hairpin followed by 10 adenines) and nucleotides.

(B) 30 μM substrate (see A) was incubated with 70 nM exosome (that contained an average number of active sites between 10 and 100%) at 50°C. Aliquots of the reaction were taken at different time points and mixed with 8M urea to quench the reaction. The samples of the different time-points were analyzed using Urea PAGE, after which the RNA was visualized

with methylene blue. The amount of substrate and product was quantified using ImageJ (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2015). The Urea PAGE band intensities were scaled by the length of the RNA to correct for differences in staining efficiency of the substrate and product. Two exemplary gels of the reaction in the presence of the exosome that contains on average 20% (left) or 90% (right) active sites are shown. The time-points at which the reaction was quenched are indicated on top.

(C) The progression of the reaction was determined at each time point by dividing the band corrected band intensity of the product by the sum of the corrected band intensities of the product and the substrate. The rate of the reaction was subsequently determined by fitting the progression of the reaction to the exponential function B*(1-exp(-A*t)), where A is the rate of the degradation reaction and B is a scaling factor.

(D) The relative speeds of the reaction were plotted versus the average number of active sites to determine if the reaction linearly depends on the number of active sites (See Fig 6). The determined catalytic rates show larger spreads, indicative for inaccuracies in the extracted parameters. These inaccuracies can have multiple causes, including inaccuracies in the determination of Urea PAGE band intensities, different efficiencies in the methylene blue staining of the substrate and product, pipetting errors during the loading of the gel and variations in the quality of different urea-page gels.

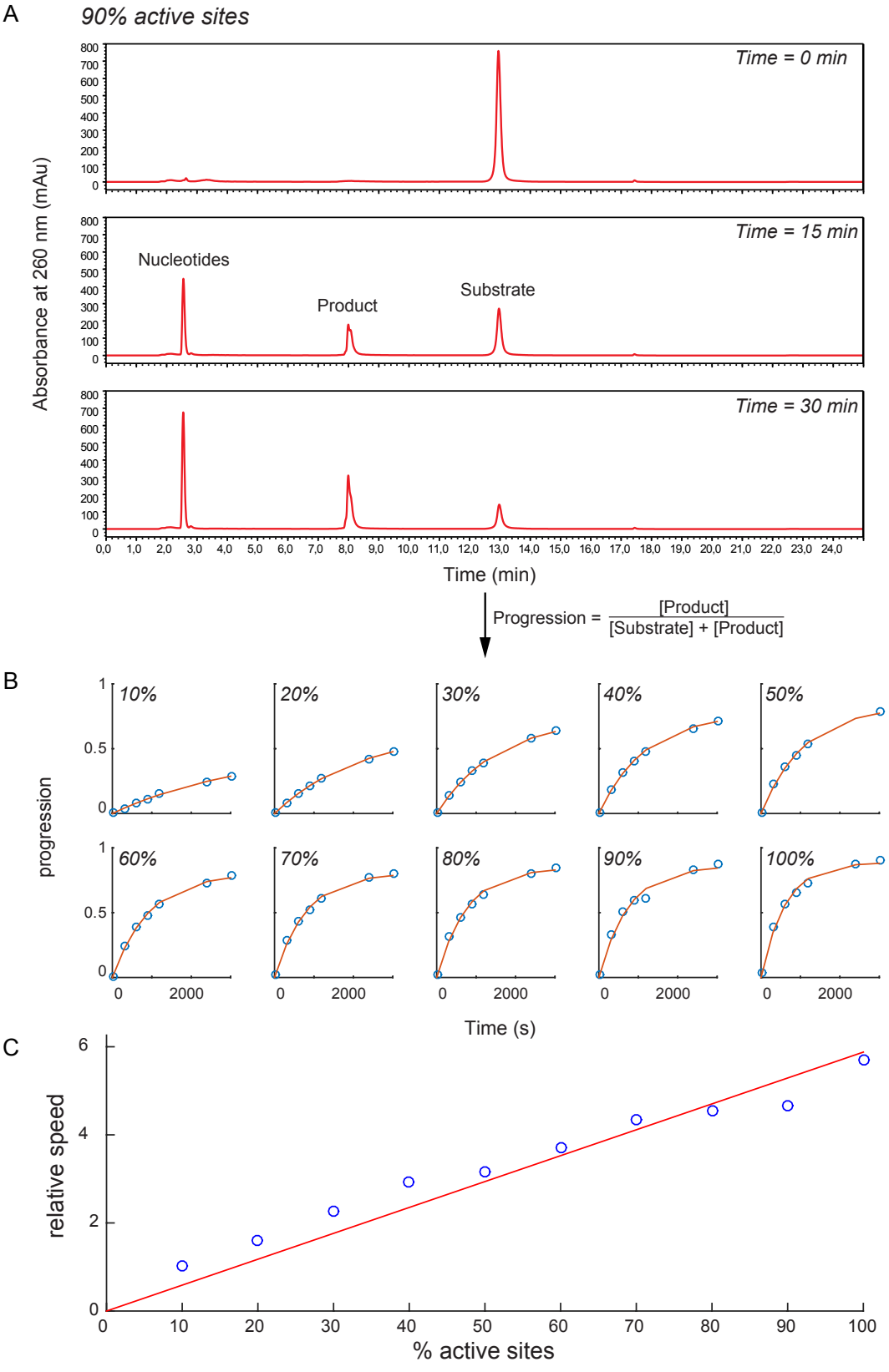

Figure S5.

(A) Exosome and substrate were mixed as described in Figure S4A and S4B, but using 25 μM substrate RNA and 60 nM exosome. The reaction time-points were analyzed on a Dionex PA-100 column that was heated to 80 °C and coupled to an HPLC system that was equipped

with an auto sampler for the precise and reproducible injection of the sample onto the column. This allowed for the accurate quantification of the levels of substrate and product in the reaction mixture, where the peak areas were converted into absolute concentrations based on determined calibration curves (Figure S6). Note that we did not use the concentration of the nucleotides in the analysis as the peak that contained the nucleotides overlapped with small peaks that resulted from the buffer in the reaction mixture (see e.g. top chromatogram taken at time=0 and where no product and nucleotides are present. The progression of the reaction was assessed by dividing the product concentration by the total RNA concentration.

(B) As in Fig S4C, the rate of the reaction was determined at each time point by fitting the progression of the reaction to the exponential function B*(1-exp(-A*t)), see Fig S4. Note that each data point in panel B is extracted from a single HPLC run.

(C) Plot that correlates the activity of the exosome complex versus the average number of active sites in the complex. The spread in the data is significantly reduced compared to the Urea PAGE analysis (See Fig S4). Based on that we decided to analyze all degradation experiments using the HPLC method.

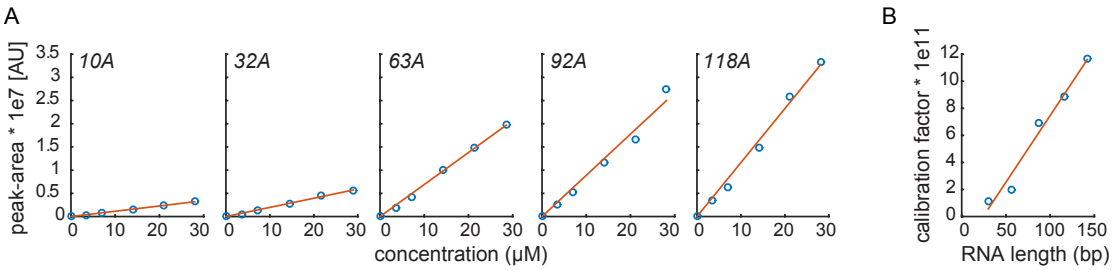

Figure S6.

(A) Calibration of the detector response using known amounts of product (a GC hairpin followed by 10 adenines) and substrates (a GC hairpin followed by 32, 63, 92 or 118 adenines).

(B) Global calibration of the product and substrate amounts that was used to convert all HPLC peak areas into absolute concentrations.

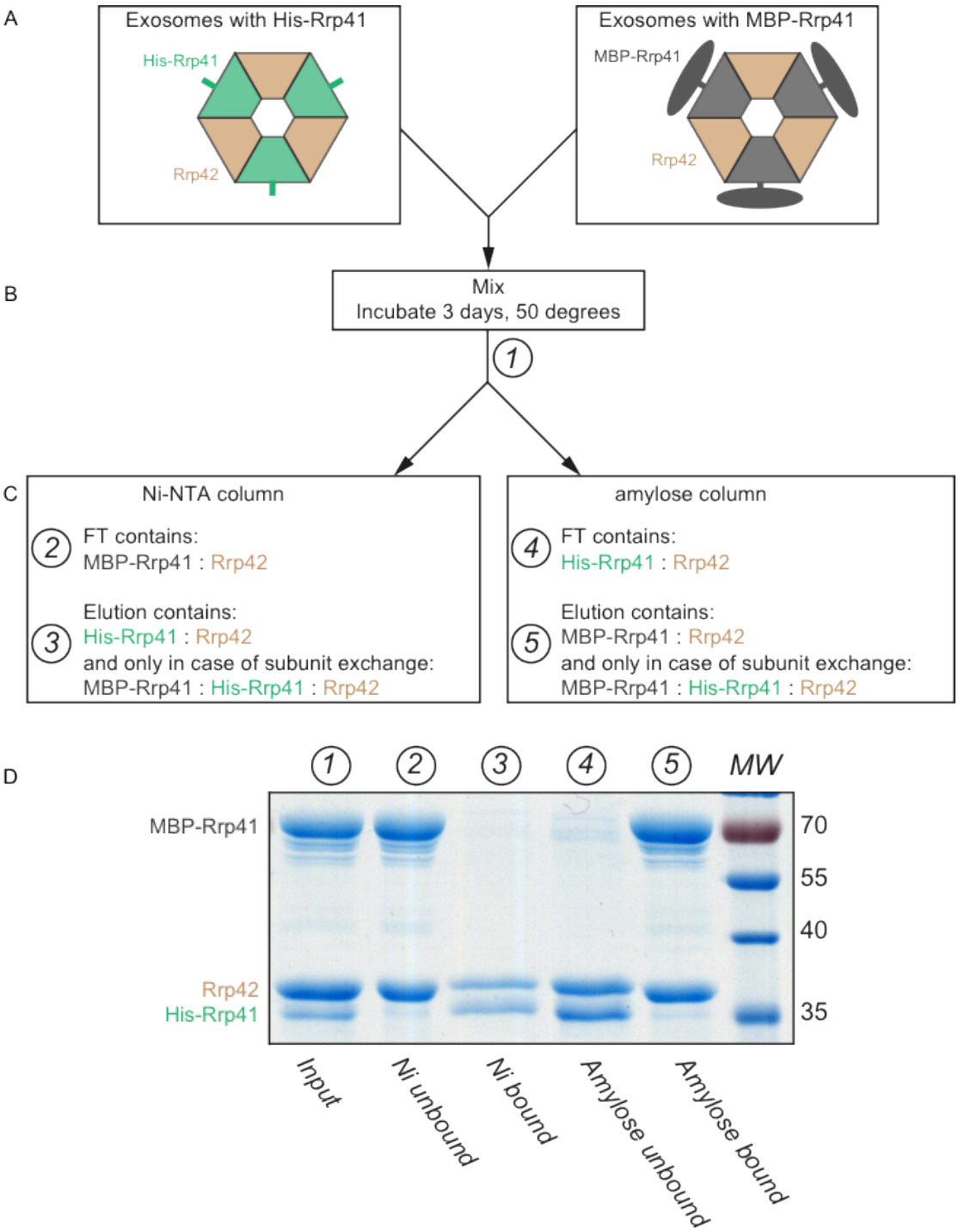

Figure S7.

(A) Exosome complexes that contain only a His-tagged version of Rrp41 (left, green) and exosome complexes that only contain an MBP-tagged version of Rrp41 (right, dark grey) were separately prepared.

(B) The His- and MBP-tagged version of the exosome complex were mixed in a 1:1 ratio and incubated at 50°C for 3 days. In case of subunit exchange between the His and MPB tagged exosome complexes a MBP-Rrp41:His-Rrp41:Rr42 complex would form.

(C) The sample was split into two parts. One part was applied to Ni-NTA resin. Exosome complexes that don't contain any His-Rrp41 subunits (only MBP-Rrp41:Rrp42) will not bind to

the resin. The elution of the Ni-NTA column contains all complexes that contain one or more histidine tags and this fraction thus contains His-Rrp41:Rrp42 complexes. In addition complexes that result from subunit exchange and that thus contain His-Rrp41:MBP-Rrp41:Rrp42 will be found in this fraction. The second half of the incubated sample was applied to amylose resin that specifically interact with the MPB tag. The fraction that does not bind to amylose resin thus only contains His-Rrp41:Rrp42 exosome complexes. The fraction that binds to the amylose resin contains the MBP-Rrp41:Rrp42 complex and potentially the His-Rrp41:MBP-Rrp41:Rrp42 complex that can have formed as a result of subunit exchange during the incubation time.

(D) SDS page analysis. (1): the exosome complexes after incubation at 50°C for 3 days that contains His-Rrp41, MBP-Rrp41 and Rrp42. (2): The FT of the Ni-column, which contains, as expected, only MBP-Rrp41 and Rrp42. (3): The elution of the Ni-column, which only contains His-tagged Rrp41 and Rrp42. MBP-tagged Rrp41 is absent, showing subunit exchange did not take place. (4): The FT of the amylose column, which contains, as expected, only His-Rrp41 and Rrp42. (5): The elution of the amylose column, which contains only MBP-tagged Rrp41 and Rrp42. The His-tagged version of Rrp41 is absent from this fraction, proving that subunit exchange between different exosome complexes does not take place at rates that are significant for the experiments that we performed here.

RNA [1]	Sequence	Note	ID
20A [3]	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAA	3' end after ribozyme cleavage	12
32A	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAGCU	Run-off transcription from linearized plasmid	17
63A	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAGCU	Run-off transcription from linearized plasmid	30
92A	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAGCU	Run-off transcription from linearized plasmid	31
118A	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAGCU	Run-off transcription from linearized plasmid	32
32A FA[2]	GCCCCCCCCGAAAGGGGGGGGAAAAAAAAAAAAAA AAAAAAA-4-S-U-AAAAAAAAAAGCU	Synthesized (Dharmacon)	-

Protein [4]	Sequence	Note	ID
His-Rrp41[5]	MKHHHHHHPMSDYDIPTTENLYFQGAMGREMLQVER PKLILDDGKRTDGRKPDELRSIKIELGVLKNADGSAIFE MGNTKAIAAVYGPKEMHPRHLSLPDRAVLRVRYHMTP FSTDERKNPAPSRREIELSKVIREALESAVLVELFPRTA IDVFTEILQADAGSRLVSLMAASLALADAGIPMRDLIAG VAVGKADGVIILDLNETEDMWGEADMPIAMMPSLNQV TLFQLNGSMTPDEFRQAFDLAVKGINIIYNLEREALKSK YVEFKEEGV	pET based vector	82 422
MBP-Rrp41[5]	MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTV EHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSG LLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEAL SLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNL QEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGA KAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTI NGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLS AGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGA VALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF WYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNN NNNNNPMSENLYFQGAMGREMLQVERPKLILDDGKR TDGRKPDELRSIKIELGVLKNADGSAIFEMGNTKAIAAV YGPKEMHPRHLSLPDRAVLRVRYHMTPFSTDERKNP APSRREIELSKVIREALESAVLVELFPRTAIDVFTEILQA DAGSRLVSLMAASLALADAGIPMRDLIAGVAVGKADG VIILDLNETEDMWGEADMPIAMMPSLNQVTLFQLNGS MTPDEFRQAFDLAVKGINIIYNLEREALKSKYVEFKEEG V	pET based vector	938 993
TwinStrep-Rrp41[5]	MKWSHPQFEKGGGSGGGSGGSSAWSHPQFEKPMS DYDIPTTENLYFQGAMGREMLQVERPKLILDDGKRTD GRKPDELRSIKIELGVLKNADGSAIFEMGNTKAIAAVYG PKEMHPRHLSLPDRAVLRVRYHMTPFSTDERKNPAPS RREIELSKVIREALESAVLVELFPRTAIDVFTEILQADAG SRLVSLMAASLALADAGIPMRDLIAGVAVGKADGVIILD LNETEDMWGEADMPIAMMPSLNQVTLFQLNGSMTPD EFRQAFDLAVKGINIIYNLEREALKSKYVEFKEEGV	pET based vector	989
Rrp42	MGMSSTPSNQNIIPIIKKESIVSLFEKGIRQDGRKLTDY RPLSITLDYAKKADGSALVKLGTTMVLAGTKLEIDKPYE DTPNQGNLIVNVELLPLAYETFEPGPPDENAIELARVV	pET based vector No affinity tags	373

	DRSLRDSKALDLTKLVIEPGKSVWTVWLDVYVLDYGG NVLDACTLASVAALYNTKVYKVEQHSNGISVNKNEVV GKLPLNYPVVTISVAKVDKYLVVDPDLDEESIMDAKISF SYTPDLKIVGIQKSGKGSMSLQDIDQAENTARSTAVKL LEELKKHLGI		
His-Rrp42	MKHHHHHHPMSDYDIPTTENLYFQGAMSSTPSNQNII PIIKKESIVSLFEKGIRQDGRKLTDYRPLSITLDYAKKAD GSALVKLGTTMVLAGTKLEIDKPYEDTPNQGNLIVNVE LLPLAYETFEPGPPDENAIELARVVDRSLRDSKALDLT KLVIEPGKSVWTVWLDVYVLDYGGNVLDACTLASVAA LYNTKVYKVEQHSNGISVNKNEVVGKLPLNYPVVTISV AKVDKYLVVDPDLDEESIMDAKISFSYTPDLKIVGIQKS GKGSMSLQDIDQAENTARSTAVKLLEELKKHLGI	pET based vector	487
His-Rrp41 Rrp42[6]	MKHHHHHHPMSDYDIPTTENLYFQGAMGREMLQVER PKLILDDGKRTDGRKPDELRSIKIELGVLKNADGSAIFE MGNTKAIAAVYGPKEMHPRHLSLPDRAVLRVRYHMTP FSTDERKNPAPSRREIELSKVIREALESAVLVELFPRTA IDVFTEILQADAGSRLVSLMAASLALADAGIPMRDLIAG VAVGKADGVIILDLNETEAMWGEADMPIAMMPSLNQV TLFQLNGSMTPDEFRQAFDLAVKGINIIYNLEREALKSK YVEFKEEGV MGMSSTPSNQNIIPIIKKESIVSLFEKGIRQDGRKLTDY RPLSITLDYAKKADGSALVKLGTTMVLAGTKLEIDKPYE DTPNQGNLIVNVELLPLAYETFEPGPPDENAIELARVV DRSLRDSKALDLTKLVIEPGKSVWTVWLDVYVLDYGG NVLDACTLASVAALYNTKVYKVEQHSNGISVNKNEVV GKLPLNYPVVTISVAKVDKYLVVDPDLDEESIMDAKISF SYTPDLKIVGIQKSGKGSMSLQDIDQAENTARSTAVKL LEELKKHLGI	pET based vector Coexpression of Rrp41 and Rrp42 Catalytically inactive Rrp41 (as indicated with the underlined D182A mutation in the sequence)	1113

Exosome complexes with a discrete number of active sites			
# active sites	Co-expression construct design[6]	Note	ID
3	His-Rrp41 WT TwinStrep-Rrp41 WT MBP-Rrp41 WT Rrp42	pET based vector. See above for the protein sequences	1020
2	His-Rrp41 D182A TwinStrep-Rrp41 WT MBP-Rrp41 WT Rrp42	pET based vector. See above for the protein sequences	1021
1	His-Rrp41 D182A TwinStrep-Rrp41 WT MBP-Rrp41 D182A Rrp42	pET based vector. See above for the protein sequences	1038

Table S1: Used constructs for protein expression and RNA *in vitro* transcription

1 The RNA stem-loop is highlighted in green.
2 RNA used for Fluorescence anisotropy measurements. 4-S-U refers to a thio-uridine that is used to couple the RNA to 6-(Iodoacetamido)-fluorescein.
3 Used for NMR studies as this RNA is not degraded by the exosome due to the presence of a 3' cyclic phosphate that results from ribosomal cleavage.
4 Protein affinity tags are colored red (His-Tag), pink (MBP-Tag) or yellow (STREP-Tag). TEV cleavage sites are highlighted in blue, where the TEV protease cleaves between the Q and G.
5 The inactive version of the Rrp41 construct contains the D182A mutation.
6 Co-expression vectors were constructed as described (5). All four proteins are in one vector and simultaneously induced.

References

1. Ishii, R., Nureki, O. and Yokoyama, S. (2003) Crystal structure of the tRNA processing enzyme RNase PH from Aquifex aeolicus. *The Journal of biological chemistry*, **278**, 32397-32404.
2. Symmons, M.F., Jones, G.H. and Luisi, B.F. (2000) A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure*, **8**, 1215-1226.
3. Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G. and Conti, E. (2005) The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nature structural & molecular biology*, **12**, 575-581.
4. Liu, Q., Greimann, J.C. and Lima, C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223-1237.
5. Mund, M., Overbeck, J.H., Ullmann, J. and Sprangers, R. (2013) LEGO-NMR spectroscopy: a method to visualize individual subunits in large heteromeric complexes. *Angewandte Chemie*, **52**, 11401-11405.