THE C-TERMINAL REGION OF THE EXOSOME-ASSOCIATED PROTEIN RRP47 IS SPECIFICALLY REQUIRED FOR BOX C/D SNORNA 3' MATURATION

Joe L. Costello1, Jonathan A. Stead2, Monika Feigenbutz, Rebecca M. Jones3 and Phil Mitchell

From the Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK

Running head: Rrp47 interacts with snoRNP proteins

1Present address: The University of Manchester, Faculty of Life Sciences, Manchester M13 9PT, UK
2Present address: The Institute of Cancer Studies, The University of Sheffield, Sheffield S10 5GB, UK
3Present address: The School of Cancer Studies, The University of Birmingham, Birmingham B15 2TT, UK

Address correspondence to: Phil Mitchell, Department of Molecular Biology and Biotechnology, The University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK; E-mail: p.j.mitchell@shef.ac.uk

Cells lacking the exosome-associated protein Rrp47 show similar defects in stable RNA processing to those observed in the absence of the catalytic subunit Rrp6, but the precise mechanism(s) by which Rrp47 functions together with Rrp6 remains unclear. Deletion complementation analyses defined an N-terminal region of Rrp47, largely coincident with the bioinformatically defined Sas10/C1D domain, which was sufficient for protein function in vivo. In vitro protein interaction studies demonstrated that this domain of Rrp47 binds the PMC2NT domain of Rrp6. Expression of the N-terminal domain of Rrp47 in yeast complemented most RNA processing defects associated with the rrp47Δ mutant but failed to complement the defect observed in 3' end maturation of box C/D snoRNAs. Consistent with these results, protein capture assays revealed an interaction between the C-terminal region of Rrp47 and the snoRNP proteins Nop56 and Nop58. Filter binding assays demonstrated that deletion of the lysine-rich sequence at the C-terminus of Rrp47 blocked RNA binding in vitro. Furthermore, a protein mutated both at the C-terminus and within the N-terminal domain showed a synergistic defect in RNA binding without impacting on its ability to interact with Rrp6. These studies provide evidence for a role of Rrp47 in registering snoRNP particle assembly, functionally characterise the Sas10/C1D domain of Rrp47 and show that both the C-terminus of Rrp47 and the N-terminal domain contribute to its RNA binding activity.

The exosome ribonuclease complex was initially characterised through its function in the processing of 3' extended precursors to stable RNAs such as 5.8S rRNA, small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs), and in the degradation of excised RNA fragments that are released during such RNA processing reactions (1,2). It has since become clear that the exosome also plays a crucial role in nuclear RNA surveillance mechanisms that recognise and degrade transcripts that are incorrectly processed, do not receive the appropriate posttranscriptional modification or that are not correctly assembled into ribonucleoprotein particles (3-5). More recently, studies have shown that the exosome is responsible for the degradation of short-lived transcripts that are the result of pervasive transcription throughout the genome (6-9). In the cytoplasm, the exosome also functions in normal mRNA turnover and in translation-coupled RNA surveillance mechanisms. The role of the exosome in RNA processing and surveillance pathways has been the subject of a number of recent reviews (3,4,10-14).

Catalytic activity of the eukaryotic exosome is attributed to two subunits, Rrp44/Dis3 and Rrp6 (15,16). Rrp44 is a member of the RNase II family of 3'→5' exoribonucleases but also has an endonucleolytic activity associated with its N-terminal PIN domain (17-19). Rrp6 is related to another bacterial 3'→5' exoribonuclease, RNase D (20-22). These enzymes have differing substrate preferences in vitro and yeast strains mutant for Rrp44 or Rrp6 show distinct stable RNA processing phenotypes, rrp6 mutants accumulating characteristic precursor species including a 3' extended form of 5.8S rRNA and adenylated forms of snoRNAs and U6 snRNA (1,2,6,23,24). RNA analyses of yeast mutants expressing a C-
terminally truncated Rrp6 mutant that cannot bind stably to the exosome suggest that this enzyme might carry out some of its functions independently of the exosome complex (25).

Yeast strains lacking the small nuclear protein Rrp47 (also known as Lrp1) show comparable RNA processing phenotypes to those observed in the absence of Rrp6, and Rrp47 is found associated with exosome complexes that contain Rrp6 (26-29). We have previously shown that Rrp47 interacts directly with the N-terminal region of Rrp6 (30) comprising the PMC2NT domain (31). Like other exosome factors, Rrp47 is conserved throughout eukaryotes and the homologous human protein, C1D, functions in stable RNA processing pathways and interacts with the Rrp6-homologous protein PM/Scl-100 (32). Both Rrp47 and C1D have RNA- and DNA-binding activity, with an apparent specificity for double-stranded DNA and structured RNA substrates (30,32,33). Rrp47 has been shown to bind both Rrp6 and nucleic acid concomitantly, suggesting that it might promote Rrp6 activity by enabling the enzyme to bind structural elements within RNA (30).

Rrp47 lacks previously characterised domains that are associated with nucleic acid or protein binding activity. In the absence of available structural data for Rrp47, we initiated a mutagenesis study to map the regions of the protein that are required for its characterised RNA processing functions in vivo and for its Rrp6-binding and RNA binding activities in vitro. Mutually supportive data from biochemical and genetic experiments define a domain within the N-terminal portion of Rrp47 that can interact with Rrp6 and which is sufficient for Rrp47 function in yeast. In contrast, stable binding of Rrp47 to RNA in vitro required the full-length protein. Finally, RNA analyses of rrp47 mutants and protein interaction studies in vitro suggest an unexpected role for Rrp47 in sensing the assembly of box C/D small nucleolar ribonucleoprotein (snoRNP) particles.

**Experimental procedures**

**Bioinformatics** The secondary structure of Rrp47 was predicted using the web server Phyre (34). Rrp47 homologous sequences were identified by WU-BLAST2, using the EBI server (35). Multiple sequence alignments were created using ClustalW and viewed using Jalview (36).

**Construction and growth of strains** Details of plasmid constructions, primer sequences and strain descriptions are provided in the Supplementary Data. Plasmid constructs were generated by standard restriction digestion cloning techniques or site-directed mutagenesis, using the Quik-change mutagenesis kit (Stratagene). Yeast strains carrying plasmids were grown under standard selective conditions, while strains expressing genomically encoded TAP-tagged proteins (Open Biosystems) were grown in YPD medium.

**Expression and Purification of Recombinant Proteins** GST, GST-22, GST-C, GST-PMC2NT and His(6)-Rrp47 proteins were induced in BL21(DE3)lysS cells and cell lysates were made in 20 mM HEPES pH 7.6, 300 mM NaCl, 10 mM imidazole, 1 mM PMSF, as described previously (30). Full-length His(6)-Rrp47 and the N* sextuple point mutant (E79A, R82G, K84I, Y86S, K89M, K91I) were purified by Ni-NTA affinity chromatography and ion exchange chromatography over SP sepharose, as described (30). For the ΔC, S100X, I162X, V176X, G181X and N* G181X mutant Rrp47 proteins, the Ni-NTA eluates were dialysed against 20 mM HEPES pH 7.6, 150 mM NaCl, passed over SP sepharose beads and the nonbound fraction retained. Proteins were analysed by SDS-PAGE and western blotting using the penta-His monoclonal antibody (Qiagen) or anti-GST antiserum (Sigma).

Expression of the heteromeric His(9)-Nrd1/MBP-Nab3 complex was performed as described (37). This expression construct does not encode the C-terminal 26 residues of Nrd1, 191 residues from the N-terminus of Nab3 and 238 residues from the C-terminus of Nab3. The Nrd1/Nab3 complex was purified by Ni-NTA affinity chromatography and detected in pull-down assays by western blot analyses, using monoclonal antibodies against the His epitope (Qiagen) or MBP (Sigma).

**Protein and RNA analyses** The SL-AU RNA (ggguuuccucaaacauucuguuuggauguauuuaa aagauau) was generated by in vitro transcription of EcoRI-linearised plasmid p299 using T7 RNA polymerase. Transcription products were dephosphorylated by incubation with alkaline phosphatase, gel-purified and 5' 32P-labelled with γ[32P]-ATP using T4 polynucleotide kinase. Yeast
tRNA\textsuperscript{Phe} (Sigma) was dephosphorylated and 5' labelled, as above. RNA binding experiments were performed using a modified double filter binding assay (38). Purified recombinant proteins were serially diluted in RNA binding buffer (10 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM MgCl\textsubscript{2}; 0.1 mM EDTA, 10 % glycerol) and 20 µl aliquots were mixed with 5 µl \textsuperscript{32}P-labelled RNA substrate (final concentration, 10 nM). After incubation on ice for 10 minutes, the reaction mixtures were filtered through a double membrane consisting of Hybond N' overlaid with Hybond C (GE Healthcare) using a minifold II slot blot apparatus (Schleicher and Schuell) and washed three times with 0.5 ml binding buffer. RNA retained on each membrane was visualised using a Storm PhosphorImager (GE Healthcare) and the data quantified using ImageQuant software (GE Healthcare). Triplicate assays were performed in parallel.

Pull-down experiments using extracts from Escherichia coli cells expressing recombinant GST-PMC2NT and Rrp47 proteins were performed essentially as described (30). Yeast cell extracts were made in lysis buffer comprising 10 mM Tris-HCl pH 7.6, 150 mM NaCl, 5 mM MgCl\textsubscript{2}. Protein capture assays were performed by adding 0.5 ml of cell extract from yeast strains expressing epitope-tagged proteins to 20 µl Ni-NTA beads or glutathione sepharose beads pre-charged with recombinant protein. Binding reactions were carried out for 2 hours at 4 °C. After washing 5 times with 1 ml cell lysis buffer containing 0.1 % NP-40 for 10 minutes, the beads were routinely incubated with 1 µg RNase A for 30 minutes at 20 °C and then rewarshed. Retained protein was eluted in 50 µl elution buffer (20 mM HEPES pH 7.6, 300 mM NaCl containing 250 mM imidazole or 25 mM reduced glutathione). In the experiment shown in Fig. S3, proteins from yeast cell extracts that were retained on Rrp47-charged beads were treated with DNase I, micrococcal nuclease or RNase A before rewashing and elution with imidazole-containing buffer. Samples were resolved by SDS-PAGE and analysed by western blotting using the peroxidase/anti-peroxidase (PAP) antibody (Sigma).

Total cellular RNA was resolved through 8 % polyacrylamide gels containing 50 % urea, transferred to Hybond N' membrane and hybridised with 5' \textsuperscript{32}P-labelled oligonucleotide probes, as described (25). The probes used were: 5.8S rRNA (o236), gcttggtctagctatg; 5'ETS (o274), gcgtgcacacagg; snR38 (o272) gaggtgaatctttgacgctagc; snR50 (o494) ctgcgtcaattgctatctac; snR52 (o318) gataaggatcttggtagatagtc; U6 (o517) atctgctatgctacaaatcag; SCR1 (o242) aaggacccaagacttg.

RESULTS

The N-terminal Sas10/C1D domain of Rrp47 is sufficient for normal growth. Yeast strains lacking Rrp47 exhibit a moderate slow growth phenotype in selective medium (30), while strains lacking both RRP47 and REX1 genes are nonviable (27). To determine the region of Rrp47 required for function in vivo, plasmids encoding N- or C-terminal deletion mutants were assayed for their ability to complement the synthetic lethality of rrp47\textsuperscript{Δ} rex1\textsuperscript{Δ} double mutants in a plasmid shuffle assay, using a host strain that also allowed visualisation of plasmid loss through a red/white colony sectoring phenotype (39), and for their ability to rescue the slow growth phenotype of rrp47\textsuperscript{Δ} single mutants. The deletion mutants generated in this study are depicted in Figure 1.

Protein structure prediction analyses suggested that the N-terminal two thirds of Rrp47 are largely α-helical in nature (Fig. S1). We therefore initially generated and tested a mutant encoding the N-terminal 120 residues, denoted ΔC. This mutant fully complemented the synthetic lethality of the rrp47\textsuperscript{Δ} rex1\textsuperscript{Δ} double mutant in the colony sectoring assay and plasmid shuffle assay (Fig. 2A, and data not shown), demonstrating that the C-terminal region of Rrp47 is not required for the essential function of Rrp47 in vivo. Further consideration of the predicted secondary structure suggested possible looped regions within the Rrp47 structure around residues 95-100 and 68-74. Truncation at position S100 generated a construct that complemented the synthetic lethality of the rrp47\textsuperscript{Δ} rex1\textsuperscript{Δ} double mutant in the colony sectoring assay and plasmid shuffle assay (Fig. 2A, and data not shown), demonstrating that the C-terminal region of Rrp47 is not required for the essential function of Rrp47 in vivo as in the absence of Rex1. Further consideration of the predicted secondary structure suggested possible looped regions within the Rrp47 structure around residues 95-100 and 68-74. Truncation at position S100 generated a construct that complemented the synthetic lethality of the double mutant, the colonies obtained on medium containing 5'-FOA for the rrp47\textsuperscript{Δ} rex1\textsuperscript{Δ} double mutant that expressed the S100X rrp47 allele being of similar size to the wild type control (Fig. 2B). Truncation at residue V70 also complemented the rrp47\textsuperscript{Δ} rex1\textsuperscript{Δ} double mutant, but the colonies obtained were significantly smaller. A slightly larger C-terminal truncation (L60X) failed to support growth (Fig.
2B). Consistent with this data, \textit{rrp47}Δ transformants expressing the S100X mutant grew comparably to cells expressing the wild-type allele, while V70X and L80X mutants only partially restored the slow growth phenotype of the \textit{rrp47}Δ strain and the L60X mutant grew at a rate comparable to the vector-transformed control (Fig. 2C). A short deletion (Δ2-9) at the N-terminus allowed complementation in the plasmid shuffle assay and rescued the slow growth phenotype, whereas the larger Δ2-19 deletion did not (Fig. 2B,C). The region of Rrp47 that is required for function in vivo, as defined by these genetic complementation assays, closely maps to residues 10-89 that constitute the bioinformatically predicted Sas10/C1D domain (Pfam accession number PF04000) (40,41) (see Fig. 1). We conclude from these data that residues 10-100 of Rrp47, encompassing the Sas10/C1D domain, are critical for protein function in vivo.

**The Sas10/C1D domain of Rrp47 binds Rrp6.**

We had previously shown that recombinant Rrp47 binds to the N-terminal PMC2NT domain of Rrp6 in \textit{E. coli} cell lysates (30). To test whether the Sas10/C1D domain of Rrp47 is sufficient for this interaction with Rrp6, purified His-tagged full-length protein and the ΔC mutant were added to beads charged with GST or GST-PMC2NT fusion protein. Both purified full length Rrp47 and the ΔC construct bound to the GST-tagged PMC2NT domain, whereas neither protein was retained on beads charged with GST (Fig. 3A). As observed previously (30), a number of polypeptides shorter than the ~50 kDa GST fusion of the Rrp6 N-terminal domain are detected upon Western analysis that presumably arise through proteolytic degradation. The S100X truncated Rrp47 protein was stably expressed and also bound to the GST-PMC2NT fusion protein. In control reactions, the unrelated Nrd1/Nab3 protein complex failed to purify with the tagged Rrp6 N-terminal domain (Fig 3B). Similar results were obtained when pull-downs were performed on nonfractionated cell lysates containing the recombinant proteins (data not shown). Proteins with more extensive C-terminal truncations (L80X, V70X) showed a specific but progressively decreasing level of binding to the GST-PMC2NT fusion protein, but were not expressed at sufficient levels to allow a direct comparative analysis using the binding assay (data not shown).

**Both N- and C-terminal regions of Rrp47 contribute to RNA binding.** We then asked whether the N-terminal region of Rrp47 has RNA binding activity. Filter binding assays were performed on full length Rrp47 or C-terminally truncated mutants, using either yeast tRNA\textsubscript{Phe} or a model structured RNA (denoted SL-AU) comprising the stem loop at the 5' end of the yeast ITS2 sequence (42) linked to a 3' A/U-rich sequence.

Consistent with previous gel shift assays (30), full length Rrp47 bound to tRNA\textsubscript{Phe} and the SL-AU RNA substrates with an apparent dissociation constant of approximately 1 µM (as calculated using the concentration of the monomeric protein) (Fig. 4A,B) and showed no binding activity with a control substrate transcribed from the polylinker of a cloning vector (data not shown). This selective binding probably reflects the degree of structure in the RNAs assayed, since Rrp47 has previously been shown to bind structured RNAs but not poly(A) (30). In contrast to the wild-type protein, the ΔC mutant showed no significant binding to either tRNA\textsubscript{Phe} or the SL-AU RNA substrates (Fig. 4A,B). The C-terminal region of Rrp47 and its homologous proteins is typically very basic, 12 of the final 22 amino acids in the protein from \textit{Saccharomyces cerevisiae} being either lysine or arginine residues (shown schematically in Fig. 1; for a sequence alignment, see Fig. S2). Mutant proteins lacking the C-terminal 22 residues (denoted I162X) or 9 residues (denoted V176X) showed no detectable RNA binding activity, while a mutant lacking only the last 4 residues (denoted G181X) exhibited a substantial decrease in RNA binding activity (Fig. 4A,B). A GST fusion of the C-terminal 22 residues showed weak, nonspecific binding activity to all RNA substrates tested, including the polylinker transcript (data not shown). Thus, the basic C-terminus of Rrp47 is required for stable interaction with RNA but is, in itself, not sufficient for the stability and specificity of binding.

There are several well conserved, charged or aromatic amino acid residues within the N-terminal region of Rrp47 that lie clustered within the primary sequence, namely E79, R82, K84, Y86, K89 and K91 (Fig. S2). We mutated these residues, singularly and in combination, and tested the resultant mutant proteins for their ability to interact with RNA. The single point mutants E79A, R82A, K84A, Y86A, K89A and K91A showed no defect...
in RNA binding activity in vitro (data not shown),
neither did a sextuple mutant wherein each of the
above residues was mutated (denoted N*, Fig. 4B).
However, combining the sextuple mutant with the
compromised G181X mutant (denoted N* G181X)
caused an essentially complete loss of RNA
binding activity in the resultant protein (Fig. 4C).
This effect does not reflect a gross structural
alteration of the N-terminal region in the N*
G181X mutant protein, since the N*, G181X and
N* G181X mutants bound to the PMC2NT region
of Rrp6 in a comparable manner to the wild-type
protein in pull-down assays (Fig. 4C). These
results demonstrate that the N-terminal region and
the extreme C-terminus both contribute to the RNA
binding activity of Rrp47.

The C-terminal region of Rrp47 has a specific
function in snoRNA maturation. Yeast rrp47Δ
strains accumulate 3’ extended forms of 5.8S
rRNA and box C/D snoRNAs, and are defective in
the degradation of the excised 5’ ETS pre-rRNA
fragment (27,28). To determine which regions of
the Rrp47 protein are required for its function in
these processes, northern blot hybridisation
analyses were performed on total cellular RNA
from a wild-type strain, an rrp47 Δ strain and a
series of N- and C-terminal rrp47 truncation
mutants.

As reported previously (27,28), the rrp47Δ
mutant accumulated the 3’ extended 5.8S+30
species to high levels, with a corresponding
decrease in the level of mature 5.8S rRNA, and
accumulated readily detectable amounts of the 5’
ETS fragment (Fig. 5). Consistent with previous
data (28,30), this mutant also showed significantly
increased levels of a 3’ unprocessed form of snR38
and expressed an snR38 species slightly longer
than the mature form of wild-type strains (labelled
snR38-3’ and snR38+3, respectively, in Fig. 5C,D), as well as accumulating both discrete and
diffuse snR38 species of intermediate length. The
rrp47Δ mutant also accumulated extended U6
snRNA species. The diffuse, extended forms
observed for U6 and snR38 are consistent with
polyadenylated forms of these RNAs previously
observed in rrp6Δ mutants (1,2,6,24).

Mutant rrp47 alleles encoding proteins that do
not encompass the complete Sas10/C1D domain
(Δ2-19, L40X, I50X, L60X, V70X, L80X)
exhibited RNA processing phenotypes that are
characteristic of rrp47Δ mutants (Fig. 5, lanes 5-
10). These mutants either failed to complement the
rrp47 Δ allele genetically or exhibited slow growth
(Fig. 2 and data not shown) and so their expression
levels may well be compromised. In contrast, the
S100X mutant showed no defect in 5.8S rRNA
processing or 5’ ETS degradation, and did not
accumulate the extended forms of snR38 or U6
(Fig. 5, compare lanes 10 and 11). Similarly, the
Δ2-9 mutant showed no detectable defect in 5.8S
rRNA processing or in the maturation of box C/D
snoRNAs, and did not accumulate extended U6
snRNA species. These results correlate closely
with the genetic complementation analyses (Fig. 2)
and suggest that N- or C-terminal deletions that
encroach into the N-terminal Sas10/C1D domain
either impact on Rrp47 expression levels or inhibit
Rrp47 function in stable RNA processing and
degradation pathways in vivo.

Intriguingly, strains expressing the S100X, ΔC,
I130X and T140X rrp47 mutants accumulated the
snR38+3 species but not the longer, extended
forms of snR38 (Fig. 5C,D, lanes 11-14), showed
no defect in 5.8S rRNA processing or 5’ ETS
degradation, and did not accumulate 3’ extended
forms of U6 snRNA. To test whether this effect
was specific to intron-encoded snoRNAs, the blots
were also hybridised with probes complementary
to sequences within the independently transcribed
snoRNAs snR50 and snR52. Slightly extended
forms of snR50 and snR52 also accumulated in the
S100X, ΔC, I130X and T140X mutants, albeit to a
lesser extent in the case of snR52 (Fig. 5). Comparable phenotypes were observed for other
intron-encoded (snR39) and independently
transcribed (snR54, snR13) snoRNAs (data not
shown).

In contrast, strains expressing the longer
S150X, I162X and G181X mutant proteins showed
little or no accumulation of aberrant forms of any
of the RNAs tested (Fig. 5, lanes 15-17).
Inspection of a multiple alignment of Rrp47
sequences (Fig. S2) suggested that a short
sequence showing a very high degree of
conservation across yeast species (amino acids
128-144 in S. cerevisiae, consensus yeast sequence
PAISKSNFQGKHTKFED) maps to the region
around position 140, with residues I130, S131,
F135, H139 and F142 being highly conserved. The
two phenylalanine residues F135 and F142 were
individually mutated to alanine residues and RNA
was analysed from strains expressing the mutant
proteins. The F142A mutant showed a specific RNA processing defect in snoRNA 3' end maturation, with a marked accumulation of the +3 species for snR38, snR50 and snR52, while the F135A mutant had a considerably weaker effect (Fig. 5D-F, lanes 18 and 19). These data support a specific role for residues within the C-terminal region of Rrp47 in the final maturation of box C/D snoRNAs and demonstrate that the basic C-terminus, although required for stable RNA binding in vitro, is dispensable for RNA processing functions in vivo.

The C-terminal region of Rrp47 interacts with snoRNP proteins. To test whether the C-terminal region of Rrp47 mediates an interaction between the exosome and snoRNP particles, we assayed immobilised recombinant Rrp47 for its ability to bind epitope-tagged forms of the snoRNP proteins Nop56 and Nop58 in yeast cell lysates. Both Nop56 and Nop58 were readily detected in the eluate fractions from beads charged with full length Rrp47 but not in eluates from beads charged with the ∆C mutant (Fig. 6A). In control reactions, the full length and truncated Rrp47 proteins bound full length Rrp6 and both failed to retain a truncated Rrp6 mutant lacking the N-terminal PMC2NT domain. Both the C-terminally tagged Nop56 and Nop58 proteins consistently migrated in SDS-PAGE gels as one major band with additional, slower migrating, diffuse bands. The reason for this is not clear but presumably reflects a common structural feature in these proteins, such as the C-terminal KKD/E repeats (43). The beads were routinely incubated with RNase A prior to elution of the retained material, and incubation with micrococcal nuclease or DNase I also failed to release Nop56 from the Rrp47-charged beads (Fig. S3). To address whether the C-terminal region of Rrp47 is sufficient for this interaction, residues 121-184 were expressed as a GST fusion (denoted GST-CT) and the resultant protein was assayed for Nop56 and Nop58 binding, as above. The GST-CT fusion protein bound to both Nop56 and Nop58 present in cell extracts, while no binding was observed for the GST control to either snoRNP protein (Fig. 6B). These experiments show that Rrp47 can bind to box C/D snoRNP proteins in an RNase- and DNase-insensitive manner, and that the C-terminal region of Rrp47 is necessary and sufficient for this interaction.

DISCUSSION

Genetic complementation assays and northern blot analyses of cellular RNAs from rrp47 mutants provide independent, corroborative data sets demonstrating that the N-terminal region of Rrp47 encompassing residues 10-100 comprises a domain that is critical for its function. Biochemical protein binding studies demonstrated that the N-terminal domain of Rrp47 interacts directly with the PMC2NT domain of Rrp6. These data are consistent with a key biochemical activity of Rrp47 being its ability to interact with Rrp6. Indeed, previous studies from our laboratory showed that the interaction with Rrp6 is required for normal expression levels of Rrp47 (30). This empirically determined N-terminal domain correlates with the bioinformatically defined Sas10/C1D domain (residues 10-89) (40) that is also found in the 18S rRNA synthesis factors Sas10/Utp3 and Lcp5 (44,45) and in the metazoan protein neuroguidin, a factor involved in translational regulation during neuronal development (41,46). We have found no evidence for an interaction between Sas10 or Lcp5 and Rrp6 (M. Turner and P. M., unpublished data), but it remains feasible that Sas10/C1D family members interact with PMC2NT-related domains within other proteins.

An alternative, but not mutually exclusive, hypothesis is that the Sas10/C1D family may represent a group of RNA-binding proteins. Both Sas10/Utp3 and Lcp5 are found associated with U3 snoRNA, while neuroguidin is associated with the mRNA cap binding protein eIF4E and the mRNA polyadenylation element binding protein CPEB, and is therefore presumably bound to mRNP particles. Our studies reported here show that the RNA binding activity of Rrp47 is dependent upon both basic residues at the extreme C-terminus of the protein and residues within the N-terminal domain. Notably, Sas10, Lcp5 and neuroguidin also have clusters of basic residues close to their C-termini (45-47), and basic residues are conserved across the Sas10/C1D domain family at positions equivalent to K84 and K89 of Rrp47 (339 and 290 residues of all 382 available sequences being either lysine or arginine at these positions, respectively) (http://pfam.sanger.ac.uk//family/PF04000) (41).

Given the RNA-binding activity of Rrp47 and the requirement for this protein in stable RNA processing events, an unexpected observation was
that short truncations from the C-terminus of Rrp47 that prevented stable RNA binding in vitro (e.g. the I162X mutant) had no discernible effect on Rrp47-dependent RNA processing events in vivo. One interpretation of these data would be that the RNA binding activity of Rrp47 is not strictly required for its function in cellular stable RNA processing pathways, perhaps due to redundancy with other RNA binding proteins. Alternatively, interactions with other protein factors may enable the RNA binding activity of the Sas10/C1D domain of Rrp47 to be sufficient for protein function in vivo.

Our data reveal specific interactions between the C-terminal region of Rrp47 and proteins of box C/D snoRNP complexes, and demonstrate that rrp47 mutants lacking the C-terminal region have specific defects in final 3' maturation of box C/D snoRNAs. Previous studies have revealed similar snoRNA processing phenotypes for conditional mutants of the snoRNP proteins Nop58 and Nop1 (48,49). The 3'-extended snoRNAs have been shown to be generated by posttranscriptional oligoadenylation (24). Together, these data suggest a model (Fig. 7) wherein during normal snoRNP assembly, Rrp47 interacts with the RNP particle during 3' maturation and allows release of the mature complex. Conversely, if assembly of the snoRNP particle is delayed, the lack of interaction between Rrp47 and the snoRNP particle promotes snoRNA oligoadenylation and/or impedes the removal of the oligoadenylate tail by Rrp6. Oligoadenylated snoRNAs can be detected at low levels in wild-type cells, and transcriptional pulse chase analyses have shown that oligoadenylated products are processed to mature snoRNAs in wild-type cells (24). Thus, oligoadenylation represents a step in a productive processing pathway for at least a fraction of the snoRNA transcripts under normal physiological conditions. Our results point to an important novel function of Rrp47 as a sensor for RNP assembly during 3' end maturation of snoRNAs.

Delays in RNA processing or RNP assembly pathways can facilitate targeting of RNA substrates to quality control systems (50-52). In many cases, adenylation of RNA substrates is an initial step in such RNA degradation mechanisms, some of which involve Rrp6. Our data is consistent with the model (24) that snoRNA adenylation stalls the fate of the RNA and provides a kinetic opportunity for the delayed assembly of RNP particles to be completed (Fig. 7). Indeed, adenylation of stable RNAs during productive processing appears quite widespread (2,24,53-55). Similar interactions to that described here between Rrp47 and snoRNP proteins might also underlie the coupling of RNP assembly and processing of other RNAs.

To date, Rrp47 has been envisaged to function as an RNA-binding protein that allows Rrp6 to degrade structured RNA (13,30). The data presented here reveal an unsuspected role for Rrp47 in RNP substrate recognition during final 3' end maturation of box C/D snoRNAs. It is tempting to speculate that polyadenylated species that accumulate in rrp47Δ and rrp6Δ mutants, such as primary snoRNA transcripts (24), cryptic unstable transcripts (CUTs) or extended forms of U6 snRNA (6) (see Fig. 5), arise due to the absence of similar interactions between the Sas10/C1D domain of Rrp47 and other factors involved in RNA transcription termination or processing.

REFERENCES


**FOOTNOTES**

We thank Jeff Corden (John Hopkins School of Medicine, Baltimore) for providing the Nrd1/Nab3 expression construct. J.A.S. was supported by a research grant from the BBSRC (BB/D001161). M.F. was supported by Yorkshire Cancer Research and by a research grant from the Wellcome Trust (088306/Z/09/Z). J.L.C. and R.M.J. were supported by PhD studentships from the BBSRC and the University of Sheffield, respectively.

The abbreviations used are: GST, glutathione S-transferase; TAP, tandem-affinity purification; 5'-FOA, 5'-fluoro-orotic acid; ITS2, internal transcribed sequence 2; 5' ETS, 5' external transcribed sequence.

**FIGURE LEGENDS**
Fig. 1. Mutations generated in Rrp47. The position of the Sas10/C1D domain within the Rrp47 protein and the sequence of the lysine-rich C-terminal peptide are indicated. The scale indicates the position along the length of the protein in amino acid residues. The coding regions present in each of the C- or N-terminal truncation mutants is indicated; truncations that encroach into the Sas10/C1D domain or the lysine-rich C-terminal peptide are shown by serrated lines.

Fig. 2. The N-terminal Sas10/C1D domain of Rrp47 is required for protein function. A Colony sectoring assay. An rrp47 Δ rex1Δ double mutant harbouring a plasmid-borne copy of the RRP47 gene was transformed with plasmids encoding the wild-type RRP47 gene, the ΔC mutant and a vector control. Loss of the original plasmid through growth on YPD was followed using the red/white colony sectoring assay. B Plasmid shuffle assay. The rrp47 Δ rex1Δ double mutant was transformed with plasmids encoding S100X, V70X, L60X, Δ2-9 and Δ2-19 rrp47 truncations and then grown on solid medium containing 5'-FOA to purge the cells of the original RRP47 plasmid. C Growth rate analyses. An rrp47Δ strain was transformed with plasmids encoding rrp47 truncation mutants, the wild-type RRP47 gene or the vector and the growth rates of the resultant transformants were measured in selective medium. Growth assays were performed in duplicate, with the average values plotted. The r² values for all individual data sets lay between 0.995 and 0.999.

Fig. 3. The N-terminal domain of Rrp47 binds to Rrp6. A Pull-down assays of His-tagged full-length (WT), S100X and ΔC Rrp47 proteins with a GST fusion of the PMC2NT domain of Rrp6 (lanes 1,3,5 and 7) and the GST control (lanes 2,4 and 6). Glutathione sepharose beads charged with ~100 µg GST or GST fusion protein were washed and then incubated with ~50 µg purified recombinant Rrp47. Eluates were resolved by SDS-PAGE and analysed by western blotting using antibodies against the His(6) epitope (upper panel) or GST (lower panel). The electrophoretic mobility of bait proteins and molecular weight markers (in kDa) are indicated on the right and left of the panels, respectively. B Pull-down assays of Rrp47 and the Nrd1/Nab3 complex control with the GST-PMC2NT fusion protein. Equivalent amounts of input, supernatant (sup.) and eluate fractions were analysed by western blot analyses using epitope tag-specific antibodies. Proteins detected in the western analyses are indicated. Multiple bands are detected for the GST-PMC2NT fusion protein in both panels due to proteolytic degradation.

Fig. 4. RNA binding of Rrp47. A,B Filter binding saturation curves obtained for wild-type or mutant Rrp47 proteins with the SL-AU and tRNA Phe substrates. All assays were performed in triplicate, the range of values being indicated by the vertical bars. The sextuple mutant denoted N* carries the point mutations E79A, R82G, K84I, Y86S, K89M and K91I. C Upper panel, parallel filter binding assays of the G181X truncation mutant and the N* G181X mutant. The amount of radiolabelled tRNA retained is plotted against the concentration of Rrp47 protein. Assays were performed in triplicate; the vertical bars indicate the ranges of values. Centre panel, pull-down assays of the purified proteins used in the RNA binding assays. Purified Rrp47 proteins were incubated with beads charged with GST control (lanes 1,3,5 and 7) or the GST-PMC2NT fusion protein (lanes 2,4,6 and 8) and the eluates were analysed by SDS-PAGE and western blotting using the anti-His antibody. Lower panel, Commassie-stained SDS-PAGE gel of the purified recombinant Rrp47 proteins used in these assays.

Fig. 5. RNA Analyses of rrp47 mutants. RNA was resolved through an 8 % denaturing polyacrylamide gel and analysed by northern blot hybridisation. Hybridisation patterns were visualised by autoradiography. Lane 1, RNA from a wild-type strain. Lanes 2-19, RNA from an rrp47Δ strain transformed with plasmids encoding rrp47 alleles, as indicated. Probes used were complementary to the following RNAs: A 5.8S rRNA; B 5’ ETS; C,D snR38; E snR50; F snR52; G U6; H SCR1 RNA. The major species detected are indicated on the right. Panel D shows a shorter exposure than C that reveals the relative levels of snR38 and the 3’ extended snR38+3 species. SCR1 serves as a loading control.
Fig. 6. Rrp47 interacts with Nop56 and Nop58. Cell extracts from yeast strains (~5 mg total protein) expressing TAP-tagged proteins were incubated with beads precharged with ~500 µg recombinant proteins. After binding, the beads were washed and then incubated with RNase A prior to rewashing and elution of the retained protein. Equivalent fractions of cell extract and supernatant, and a 100-fold equivalent of the eluate fractions, were then resolved by SDS-PAGE and analysed by western blotting. The epitope-tagged protein expressed in yeast is indicated to the right. A Pull-downs with full-length Rrp47, the ΔC mutant and a vector control. Lysates from strains expressing epitope-tagged full-length Rrp6 or an Rrp6 mutant protein lacking the PMC2NT domain were used as positive and negative controls, respectively. B Pull-downs with a GST fusion of the C-terminal region of Rrp47 (denoted GST-CT) or GST alone.

Fig. 7. Model for Rrp47 function in snoRNP assembly. SnoRNP particles assemble cotranscriptionally through binding of core proteins to the snoRNA. Correct 3’ end formation of snoRNAs requires the 3’ exonuclease Rrp6. Correct assembly is sensed through interactions between Rrp47 and proteins of the snoRNP particle, allowing release of the mature snoRNP complex. In the absence of the Rrp47/snoRNP interaction, the snoRNA is adenylated by TRAMP and maintained in an oligoadenylated form by the competing action of polymerase and exonuclease activities. This provides an extended period for assembly to be completed. Other exonucleases can substitute for Rrp6 in trimming of the pre-snoRNA but not complete removal of the oligoadenylate tail. For simplicity, Rrp6, rather than the complete exosome complex, is shown.
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Figure 7
The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D snoRNA 3’ maturation
Joe L. Costello, Jonathan A. Stead, Monika Feigenbutz, Rebecca M. Jones and Phil Mitchell

J. Biol. Chem. published online December 6, 2010

Access the most updated version of this article at doi: 10.1074/jbc.M110.162826

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2010/12/06/M110.162826.DC1