# A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease

**The EMBO Journal (2011) 30,** 4006–4019 | © 2011 European Molecular Biology Organization | Some Rights Reserved 0261-4189/11
www.embojournal.org

THE EMBO JOURNAL

EMBO open

# A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease

**Sander Granneman*, Elisabeth Petfalski and David Tollervey***

Wellcome Trust Centre for Cell Biology and Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh, Scotland

**The 5′-exonuclease Rat1 degrades pre-rRNA spacer fragments and processes the 5′-ends of the 5.8S and 25S rRNAs. UV crosslinking revealed multiple Rat1-binding sites across the pre-rRNA, consistent with its known functions. The major 5.8S 5′-end is generated by Rat1 digestion of the internal transcribed spacer 1 (ITS1) spacer from cleavage site $A_3$. Processing from $A_3$ requires the '$A_3$-cluster' proteins, including Cic1, Erb1, Nop7, Nop12 and Nop15, which show interdependent pre-rRNA binding. Surprisingly, $A_3$-cluster factors were not crosslinked close to site $A_3$, but bound sites around the 5.8S 3′- and 25S 5′-regions, which are base paired in mature ribosomes, and in the ITS2 spacer that separates these rRNAs. In contrast, Nop4, a protein required for endonucleolytic cleavage in ITS1, binds the pre-rRNA near the 5′-end of 5.8S. ITS2 was reported to undergo structural remodelling. *In vivo* chemical probing indicates that $A_3$-cluster binding is required for this reorganization, potentially regulating the timing of processing. We predict that Nop4 and the $A_3$ cluster establish long-range interactions between the 5.8S and 25S rRNAs, which are subsequently maintained by ribosomal protein binding.**

*The EMBO Journal* (2011) **30,** 4006–4019. doi:10.1038/emboj.2011.256; Published online 2 August 2011
*Subject Categories*: RNA; proteins
*Keywords*: CRAC; ribosome synthesis; rRNA processing; RNP structure; yeast

## Introduction

During ribosome subunit biogenesis in eukaryotes, the nascent pre-rRNA can either undergo cotranscriptional cleavage by the small subunit processome or be transcribed into the 35S pre-rRNA, which is cleaved within the 90S pre-ribosome complex. In either pathway, the early cleavages separate the pre-40S and pre-60S complexes, which are then matured independently (Figure 1A). During 60S maturation, internal transcribed spacers 1 and 2 (ITS1 and ITS2) are removed from the pre-rRNA by endonuclease cleavage followed by exonuclease digestion to yield the mature 5.8S and 25S rRNA. The 60S ribosome maturation pathway involves several intermediate particles and a many different proteins. Several late steps in 60S maturation have recently been characterized (see Lo *et al*, 2010 and Panse and Johnson, 2010), but much less is known about early pre-60S particles or the actual function of most of the protein components. Understanding of yeast ribosome assembly will be greatly facilitated by the recent determination of the crystal structure of the mature 80S ribosomes (Ben-Shem *et al*, 2010).

Early in the eukaryotic lineage, the 5.8S rRNA was derived from the 5′-end of an ancestral 23S rRNA-like molecule by insertion of the ITS2 region. Reflecting this, the 5.8S rRNA associates with the 25S rRNA via two regions of extended base pairing between the 5′-region of 5.8S (nts 4–13) and 25S rRNA (nts 404–413), and between the 3′-region of 5.8S (nts 139–155) and 25S rRNA (nts 4–19). 5.8S–25S base pairing is established before or at $C_2$ cleavage, since the 7S pre-rRNA remains associated with 25S rRNA following deproteinization (DT, unpublished). However, secondary structure models of the ITS1 region of the 27SA pre-rRNAs strongly predict that the 5′-region of 5.8S is base paired to the 3′-end of ITS1 in a stem-loop that is incompatible with base pairing to 25S rRNA, implying the presence of a conformational switch (Yeh *et al*, 1990). It is currently unclear whether the transfer of the 5′-region of 5.8S from ITS1 base pairing to 25S base pairing is provoked by exonuclease digestion of ITS1 from $A_3$ to $B_{1S}$, or is required for this processing to take place.

The 5′-end of the yeast 5.8S rRNA is heterogeneous due to the use of two alternative processing pathways (Henry *et al*, 1994). The major, short form of the rRNA ($5.8S_S$) has a 5′-end at site $B_{1S}$ (Figure 1A). This is generated by exonuclease digestion from an upstream cleavage site ($A_3$), which is the target for cleavage by RNase MRP. Processing from $A_3$ to $B_{1S}$ involves two exonucleases: the Rat1–Rai1 heterodimer (Schmitt and Clayton, 1993; Henry *et al*, 1994; Lygerou *et al*, 1996; Xue *et al*, 2000) and Rrp17 (Oeffinger *et al*, 2009). In addition, the cytoplasmic exonuclease Xrn1 can degrade the $A_3$-$B_{1S}$ region when Rat1 is inactive (Henry *et al*, 1994), but probably does not contribute to normal processing (El Hage *et al*, 2008). Formation of the 5′-end of the less abundant $5.8S_L$ rRNA does not require RNase MRP or the 5′-exonucleases and is believed to involve an unidentified endonuclease (Faber *et al*, 2006). Similarly, 5′ heterogeneity is observed for 5.8S rRNA in metazoans, plants and other eukaryotes (Henry *et al*, 1994), suggesting the use of conserved processing pathways.

In addition to the exonucleases, $A_3$ processing requires components of the pre-60S particles (Dunbar *et al*, 2000;

*Corresponding authors. S Granneman, Centre for Systems Biology at Edinburgh (CSBE), University of Edinburgh, CH Waddington Building, Mayfield Road, Kings Buildings, Edinburgh EH9 3JD, Scotland.
Tel.: +44 131 651 9082; Fax: + 44 131 651 9068;
E-mail: sgrannem@staffmail.ed.ac.uk or D Tollervey, Wellcome Trust Centre for Cell Biology, University of Edinburgh, Michael Swann Building, Mayfield Road, Kings Buildings, Edinburgh EH9 3JR, Scotland.
Tel.: +44 131 650 7092; Fax: +44 131 650 7040;
E-mail: d.tollervey@ed.ac.uk

**Received: 23 December 2010; accepted: 29 June 2011; published online: 2 August 2011**

Pestov *et al*, 2001; Gadal *et al*, 2002; Oeffinger *et al*, 2002; Fatica *et al*, 2003; Oeffinger and Tollervey, 2003; Miles *et al*, 2005). Loss of these factors (including Rlp7, Cic1/Nsa3, Erb1, Nop7 and Nop15) leads to accumulation of the $27SA_3$ pre-rRNA and reduced synthesis of $27SB_S$ and mature $5.8S_S$ without concomitant loss of $27SB_L$ or $5.8S_L$. A further factor, Nop12, is non-essential but required for efficient 25S rRNA synthesis (Wu *et al*, 2001) and, together with Brx1 and Ebp2 is predicted to be physically close to Rlp7, Cic1/Nsa3, Erb1, Nop7 and Nop15 within the pre-60S particle (Tarassov *et al*, 2008; Figure 1B). The role of Nop12 in 5.8S rRNA maturation was not assessed (Wu *et al*, 2001). A further protein, Nop4 (Nop77) was predicted to bind in this region (Tarassov *et al*, 2008), but its depletion confers a different phenotype, with specific inhibition of cleavage at both $A_3$ and $B_{1L}$ (Bergès *et al*, 1994; Sun and Woolford, 1994, 1997).

To better understand the roles of the non-enzymatic ribosome synthesis factors implicated in 5.8S 5′-maturation, we identified their sites of interaction with the pre-rRNA using the crosslinking and analyses of cDNA (CRAC) crosslinking technique (Supplementary Figure S1; Granneman *et al*, 2009). Unexpectedly, binding sites for five factors required for exonucleolytic removal of the $A_3$-$B_{1S}$ fragment clustered near the 3′-end of the 5.8S rRNA (E-site), not the 5′-end. In contrast, Nop4, required for endonucleolytic cleavages in ITS1, crosslinked near the 5′-end of 5.8S. This suggests that a ribonucleoprotein complex, hereafter referred to as '$A_3$ cluster', assembled near the 3′-end of 5.8S rRNA is required for the activity of the 5′-exonucleases acting at the 5′-end of 5.8S, whereas endonucleolytic cleavage in ITS1 requires Nop4 binding to the 5′-end of 5.8S. High-throughput CRAC on the Rat1 5′-exonuclease indicated significant crosslinking at the 3′-end of 5.8S and in ITS2 and we furthermore demonstrate that Rat1 can bind the 5′-end of 25S independently of Nop4 and $A_3$-cluster protein Nop15. We propose that $A_3$-cluster proteins act to coordinate events at the 3′- and 5′-ends of 5.8S rRNA, ensuring that 5′-end maturation of 5.8S precedes ITS2 cleavage.

## Results

### *Rat1 is associated with multiple sites in the pre-rRNA*

Rat1 has several distinct targets in the pre-rRNA; it accurately processes the 5′-ends of the mature 5.8S ($A_3$-$B_{1S}$) and 25S rRNAs ($C_2$-$C_1$) and completely degrades the excised $A_0$-$A_1$ and $A_2$-$A_3$ pre-rRNA fragments (Figures 1 and 2). In addition, Rat1 functions in the 5′-maturation of intronic snoRNAs, surveillance of mRNAs with defects in splicing and 5′-capping, and transcription termination on both Pol I and Pol II genes. Given this large number of substrates, we anticipated that Rat1 binding would be very transient. To maximize the recovery of RNA binding sites, we therefore developed a novel UV crosslinking system that allows UV irradiation of large volumes of cells actively growing in culture medium. This crosslinking device (the 'Megatron') consists of a 1.2-m metal tube with a central 205 W, 254 nm UV lamp, and can irradiate 2.5 l of culture in ~1 min (Supplementary Figure S1A). Cells expressing HTP-tagged Rat1 and a non-tagged negative control strain (BY4741) were UV irradiated in the Megatron. $A_3$-$B_{1S}$ processing also requires the 5′-exonuclease Rrp17 and the Rat1 cofactor Rai1 (Fang *et al*, 2005; Oeffinger *et al*, 2009). Both were tested as HTP-tagged constructs but neither gave usable crosslinking efficiencies. Illumina-Solexa deep sequencing of two independent Rat1 CRAC cDNA libraries revealed that around a quarter of the reads mapped to rDNA repeats. A multiple sequence alignment of rDNA-mapped reads is shown in Supplementary Table S2. As expected, many other putative Rat1 substrates were identified, which will be discussed elsewhere (manuscript in preparation).

Prominent peaks of Rat1 crosslinking were observed over the 5′-region of the 5′-ETS (Figure 2B), which is degraded by pathways including endonuclease cleavage (Lebreton *et al*, 2008; Schaeffer *et al*, 2009; Schneider *et al*, 2009). These data indicate that, like its human counterpart Xrn2 (Wang and Pestov, 2010), Rat1 is involved in degradation of the cleaved 5′-ETS fragments. A very prominent peak was present at 3′-end of the 5′-ETS, directly upstream of the $A_1$ cleavage site, in agreement with the reported role of Rat1 in degradation of the excised $A_0$-$A_1$ fragments (Petfalski *et al*, 1998).

Inspection of the regions surrounding the known Rat1 pre-rRNA processing substrates, $A_3$-$B_1$ in ITS1 and $C_2$-$C_1$ in ITS2 (Figure 2) revealed apparent similarities. High levels of crosslinking were seen immediately 5′ to the $A_3$ and $C_2$ cleavage sites; over the $A_2$-$A_3$ region in ITS1 and over the 3′-region of 5.8S and between the 3′-end of 5.8S (site E) and cleavage site $C_2$ in ITS2. Reads including the 3′-region of 5.8S predominately extended through site E, at least 2 nt into ITS2 (Supplementary Figure S2A), indicating that binding occurred on the pre-rRNA, rather than on mature 5.8S rRNA. Substantial Rat1 crosslinking was also observed 3′ to the Rat1 targets, over the 5′-regions of the mature 5.8S (helices H3 and H4) and 25S rRNAs (H11) (Figure 2). In contrast, crosslinking was much lower over the Rat1 processing substrates $A_3$-$B_1$ and $C_2$-$C_1$. The intermediates in $A_3$-$B_1$ and $C_2$-$C_1$ processing are almost undetectable in wild-type yeast indicating high processivity during Rat1 processing. Finally, hits located at H66, H79 and H99 were also frequently found in negative control experiments (marked with asterisks in Figures 2A, B and 3C) and were therefore considered background.

Rat1 also participates in degradation of the excised $A_2$-$A_3$ spacer fragment (Petfalski *et al*, 1998). This may contribute to Rat1 crosslinking over this region, but mutational analyses (below) indicate that this is not the major source.

The location of site $C_2$ was originally inferred from fingerprinting of *in vivo* labelled RNA and predicted to lie within the G133–G136 region (Veldman *et al*, 1980). However, Rat1-associated sequences frequently extended ~8 nt further 3′ to terminate at U140 and A141 in ITS2 (Supplementary Figure S2A and C). This end was previously detected by primer extension in a *rat1* mutant (Geerlings *et al*, 2000) and the same end is observed in the wild-type (see below). 7S pre-rRNA was immunoprecipitated using HTP-tagged Nop7 and 3′-ends were mapped by RNA cloning (Supplementary Figure S2B). The results confirmed that the 3′-end of 7S predominately falls at U140 and A141. We conclude that the major endonuclease cleavage site lies between A141 and G142, as proposed previously (Geerlings *et al*, 2000), which we refer to herein as site $C_2$.

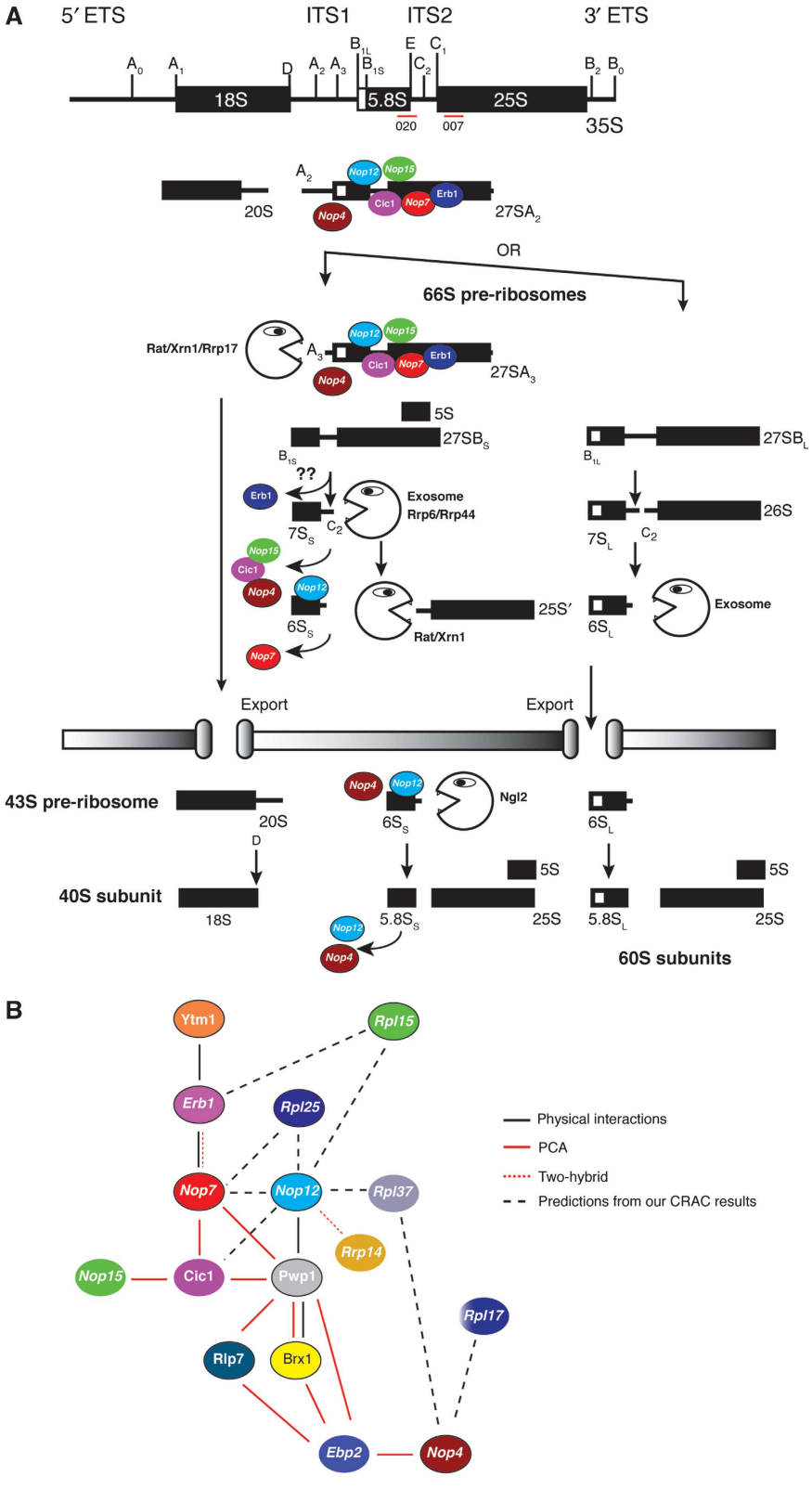
We conclude that Rat1 is associated with sites flanking its target regions for pre-rRNA processing.

### *Binding sites for Nop7, Nop12, Erb1, Cic1 and Nop15 cluster around the 3′-end of the 5.8S rRNA*

Based on published data (Miles *et al*, 2005; Tarassov *et al*, 2008), a group of synthesis factors were predicted to physi-

cally interact in early pre-60S particles: Brx1, Cic1/Nsa3, Erb1, Ebp2, Nop4, Nop7, Nop12, Nop15, Pwp1 and Ytm1 (see Figure 1B for interaction map).

Except for Rlp7 and Ebp2, all of these factors were tested in CRAC analyses, and six (Cic1, Erb1, Nop4, Nop7, Nop12 and Nop15) efficiently crosslinked to RNA. The locations of crosslinking sites along the rDNA are indicated in Figure 3A and C, and are displayed on the predicted secondary structures of 25S, 5.8S and ITS2 in Figure 3D. CRAC experiments were performed 2–5 times and to be considered a *bona fide* RNA binding site, a nucleotide sequence had to be significantly enriched in every experiment. As negative con-



trols, four independent CRAC experiments were performed with the non-tagged parental strain (Figure 3B). Sanger sequencing of Nop4 CRAC cDNA libraries had indicated multiple pre-rRNA binding sites and therefore high-throughput Solexa sequencing was performed together with a negative control to obtain sufficient sequence depth for each binding site (Figure 3C). Recovered sequences that overlapped with control peaks in H66, H79 and H99 at the 3′-end of the 25S were considered background (asterisks in Figure 3). Crosslinking sites were precisely identified by the presence of multiple point deletions or substitutions at a specific position in sequence reads (see Granneman *et al*, 2009), or a minimal RNA binding site was determined from overlapping sequences.

In CRAC analyses, Erb1 was crosslinked to the 5′-region of the 25S rRNA over H21 and H22 (Figure 3A and D). In the yeast 60S structure, this region forms a compact structure, probably associated with Rpl42, Rpl8, Rpl28 and Rpl15 (Figure 4A and B; Ben-Shem *et al*, 2010). The average length of RNAs crosslinked to Erb1 was relatively long (59 nt), possibly because this structure confers protection during the RNase digestion used to generate the cloned fragments.

Nop7 predominately crosslinked to H54 in 25S rRNA (Figure 3A and D). This site is within Domain III, in close

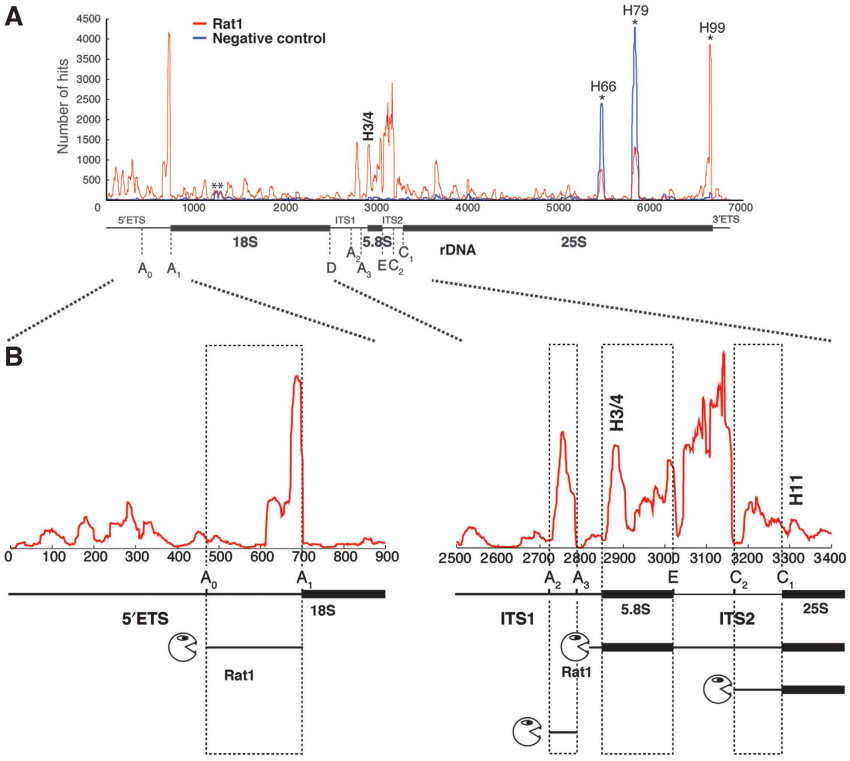
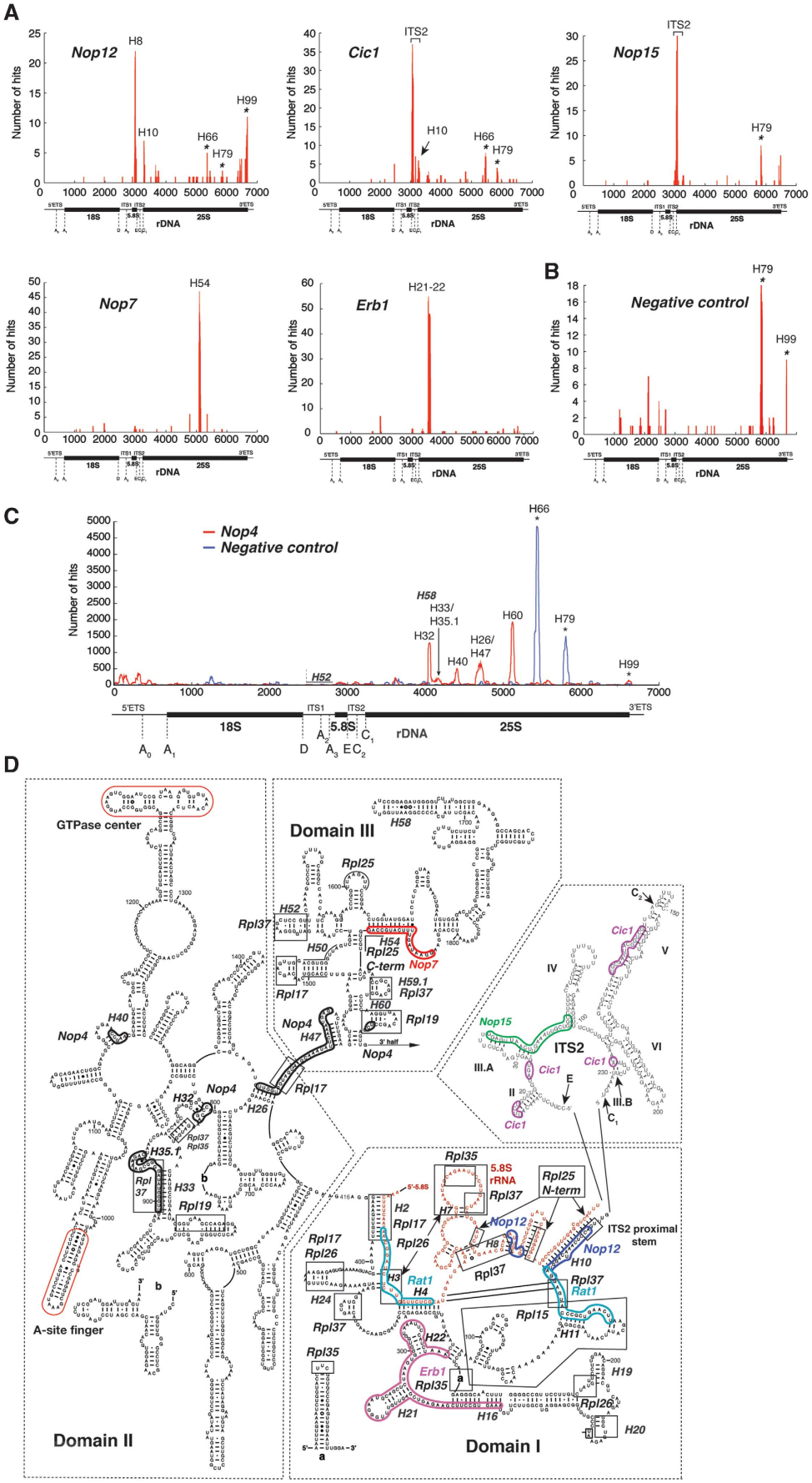



**Figure 2** Rat1 crosslinking sites over the pre-rRNAs. (**A**) Rat1 crosslinks primarily to spacer regions in the pre-rRNA. A histogram that displays the distribution of rDNA-mapped reads along the entire rDNA sequence is shown. The red line indicates 100 000 averaged hits from two independent Rat1 CRAC experiments. The blue line indicates the distribution of 10 000 rDNA hits from the negative control experiment. The asterisks indicate frequent contaminants. The rDNA is schematically represented below the *x* axis, with processing sites included. The *y* axis displays the total number of times a nucleotide within an RNA fragment was mapped to the rDNA sequence. (**B**) The dashed lines point to expanded views of hits over the 5′-ETS and ITS1-25S region with schematics showing Rat1 substrates. Positions of potential crosslinking sites in spacer regions are shown in Supplementary Figure S9.

**Figure 1** Pre-rRNA processing and protein interactions. (**A**) Schematic representation of pre-rRNA processing in yeast. The locations of processing sites on the 35S pre-rRNA are indicated. The positions of oligonucleotides used for northern hybridization (020) and primer extension (007) are shown. In the nucleolus, the 35S pre-rRNA, part of a 90S-sized complex called the SSU processome or 90S pre-ribosome, is processed at sites $A_0$, $A_1$ and $A_2$, leading to the formation of 43S and 66S pre-ribosomes. Proteins studied in the CRAC analysis and their association with pre-ribosomal complexes are indicated. The 43S pre-ribosome is exported to the cytoplasm where Nob1 cleaves at site D, yielding the mature 18S rRNA and 40S subunit. 66S pre-ribosomes containing $27SA_2$ pre-rRNA are either processed at $A_3$ or at $B_{1L}$, which requires the presence of Nop4. Pre-ribosomes containing $27SA_3$ pre-rRNA are exonucleolytically trimmed by the Rat1, Rrp17 and Xrn1 5′–3′ exonucleases, yielding the $27SB_S$ pre-rRNA. This maturation step requires the presence of Nop12, Nop7, Nop15 and Erb1 (depicted as coloured circles). After this step, the 27SB is cleaved at $C_2$ by an unknown endonuclease, followed by exonucleolytic and endonucleolytic trimming of 7S by Rrp6 and Rrp44, two components of the exosome complex. Pre-ribosomes containing the 6S pre-rRNA are exported to the cytoplasm and matured by Ngl2 after which Nop12 and Nop4 dissociate. (**B**) Overview of known and predicted protein–protein interactions in 66S pre-ribosomes. The interaction map depicts interactions between the various assembly factors and r-proteins in 66S pre-ribosomes. Black lines; physical interactions among proteins shown to be part of subcomplexes or interacting as recombinant proteins (Krogan *et al*, 2004; Miles *et al*, 2005). Dashed red lines; yeast two-hybrid interactions (Ito *et al*, 2001; Miles *et al*, 2005). Red lines; interactions from protein-fragment complementation assays (PCAs), which detect proteins located within ~80 Å (Tarassov *et al*, 2008). Dashed black lines; protein–protein interactions predicted from our CRAC data. Note that these may be mediated by RNA.



proximity to the binding sites for Rpl25 and Rpl35 (Figures 3D, 4A and B). Nop7 is genetically linked to Rpl25 (Oeffinger *et al*, 2002) and cells lacking Rpl35A accumulate $27SA_3$ pre-rRNA and are defective in ITS2 processing (Babiano and de la Cruz, 2010). In the yeast 60S structure, the N-terminal domain of Rpl25 makes extensive contacts with 5.8S, whereas the C-terminal region mainly contacts Domain III of 25S (Figure 3D). Both CRAC and genetic data, therefore, support direct interactions between Nop7, Rpl25 and Rpl35. Notably, Rpl35, together with Rpl17 and Rpl26, binds in the vicinity of the peptide exit tunnel in mature 60S subunits (Figure 4E and F).

Cic1 and Nop15 crosslinked preferentially to the 5′-end of the ITS2 spacer region (Figure 3A and D). Although many of the RNAs crosslinked to Nop15 and Cic1 originated from the same region, analysis of shorter fragments and specific mutations in the RNAs allowed us to distinguish the respective RNA binding sites (Supplementary Figure S3). Cic1 also consistently crosslinked to H10 in 25S (Figure 3A), albeit less frequently. High-throughput data suggested that Cic1 and Nop15 interact (Tarassov *et al*, 2008), and the locations of their RNA binding sites strongly support their direct association in pre-60S ribosomes.

The major Nop12 binding sites were located at H8 (Figure 3A), near the 3′-end of the 5.8S rRNA, and in H10, which is formed by base pairing between the 3′-end of 5.8S and the 5′-end of the 25S rRNA (Figure 3D, ITS2 proximal stem). Thus, Nop12 could have a role in formation or stabilization of the interaction between the 5.8S and 25S rRNAs. The proximity of Nop12 binding sites to $A_3$-cluster protein crosslinking sites prompted us to re-analyze the role of Nop12 in 5.8S processing. Because cells lacking Nop12 are cold-sensitive (Wu *et al*, 2001) pre-rRNA processing was analysed in *nop12*Δ cells grown at 18 °C for 24 h. Similar to $A_3$-cluster-depleted cells, *nop12*Δ cells accumulated 35S and $27SA_3$ pre-rRNAs (Supplementary Figure S4B and C, lanes 11 and 12), albeit modestly. We conclude that Nop12 is a novel $A_3$-cluster protein.

Nop4 crosslinking sites were dispersed in the primary rRNA sequence and mapped to 25S rRNA domains II and III (Figure 3C and D). In the yeast 60S crystal structure, these sites are ∼30–80 Å apart and cluster near the 5′-end of 5.8S (Figure 4C and D). Nop4 is an ∼80-kDa protein and could conceivably contact many of these sites simultaneously, bringing domains II and III in proximity to each other. The Nop4 binding sites in H26, H32, H33 and H47 are close to 5.8S rRNA and to binding sites for Rpl17, Rpl35 and Rpl37. We predict Nop4 directly contacts these ribosomal proteins. The binding site in H60 is adjacent to Rpl19 binding sites (Figures 3D, 4C and D). Consistent with previous data (Bergès *et al*, 1994), Nop4 depletion leads to delayed cleavage at site $A_2$, with more severe inhibition at $A_3$, $B_{1L}$ and concomitant decrease in 27SB levels (Supplementary Figure S4B and C, lanes 7 and 8). The $27SB_{1L}/27SB_{1S}$ ratio was not significantly altered in these cells (Supplementary Figure S4D, compare lane 2 with lane 1), suggesting Rat1 can process residual $27SA_3$ generated in the absence of Nop4.

We conclude that cleavage at $A_3$ and $B_{1L}$ requires Nop4 binding near the 5′-end of 5.8S, whereas Rat1-dependent processing from $A_3$ requires $A_3$-cluster proteins binding near the 3′-end of 5.8S.

### *Validation of CRAC data*

To validate the CRAC data, we performed immunoprecipitation experiments with HTP-tagged proteins and the non-tagged parental strain and analysed co-precipitated RNAs by northern blot analysis, primer extension analysis and EtBr staining (Figure 5). The data were quantified by calculating fold enrichment over background levels observed in experiments with untagged strains (Supplementary Figure S5). Consistent with previous reports, tagged Nop7, Cic1 and Nop15 co-precipitated 27SA/B and 7S pre-rRNAs (Nissan *et al*, 2002; Fatica *et al*, 2003; Oeffinger & Tollervey, 2003; Figure 5B and C). HTP-tagged Nop7 also associated with 6S pre-rRNA, a 5.8S rRNA precursor that is 3′-extended by only 8–9 nt (Figure 5B, lane 12). These observations are consistent with the CRAC data and indicate that ITS2 may be a major determinant for pre-60S binding by Cic1 and Nop15, but not by Nop7.

The association of Nop4, Nop12 and Erb1 with pre-rRNA species has not previously been analysed in detail. Both Nop4 and Nop12 co-precipitated 27SA/B and 7S pre-rRNAs, albeit in modest amounts compared with other proteins tested (Figure 5B and C, lanes 11 and 13). Nop12 and Nop4 also reproducibly showed modest enrichment of mature 5.8S and 25S (Figure 5A and B, lanes 11 and 13; Supplementary Figure S5A), suggesting that these proteins remain associated with pre-ribosomes after removal of ITS2. HTP-tagged Erb1 efficiently precipitated 27S pre-rRNA, but precipitated low amounts of 7S and 5′-extended 25S pre-rRNAs (25S' and 26S) compared with other proteins tested (Figure 5B and C, lane 10; Supplementary Figure S5B and C), implying that Erb1 dissociates from pre-ribosomes following cleavage of 27SB pre-rRNA at site $C_2$.

In the 25S rRNA secondary structure, the Erb1 and Nop7 binding sites appear distant from the 3′-end of 5.8S and ITS2 (Figure 3D). However, in the yeast 60S structure (Ben-Shem *et al*, 2010), these sites are in the vicinity of the Nop12 binding sites and ∼30–50 Å (equivalent to 10–15 nt) from the 3′-end of 5.8S rRNA (Figure 4A and B). Nop7 and Erb1 directly interact in pre-ribosomes (Tang *et al*, 2008), suggest-

**Figure 3** Overview of CRAC results and locations of protein–RNA interaction sites in the 25S and 5.8S rRNA secondary structures. (**A**) Results from 2 to 5 independent CRAC experiments. (**B**) Results from untagged strain. (**C**) Illumina-Solexa results from Nop4 (red line) and negative control (untagged strain; blue line). Sequences were aligned to the rDNA reference sequence using blast and plotted using gnuplot. Locations of mature rRNA sequences, spacers and cleavage site are indicated below the *x* axis. The *y* axis displays the total number of times each nucleotide within an RNA fragment was mapped to the reference sequence. The location of the peaks in the secondary structure of the rRNA is indicated with helix (H) numbers (Klein *et al*, 2004). The asterisks indicate frequent contaminants. (**D**) Locations of minimal binding sites for the ribosome synthesis factors are displayed on the 5.8S/25S rRNA secondary structures (http://www.rna.ccbb.utexas.edu/) and the 'ring model' for yeast ITS2 structure (Joseph *et al*, 1999; Cote *et al*, 2002). Large 25S rRNA domains are indicated with dashed boxes. The 5.8S rRNA sequence is coloured red. Locations of r-protein binding sites are boxed, based on their locations in the yeast 60S crystal structure (Ben-Shem *et al*, 2010) and previous genetic studies (van Beekvelt *et al*, 2000). Two Rat1 binding sites in helices 3/5 and 11 are shown in light blue. Crosslinking sites in the spacer regions are shown in Supplementary Figure S9.

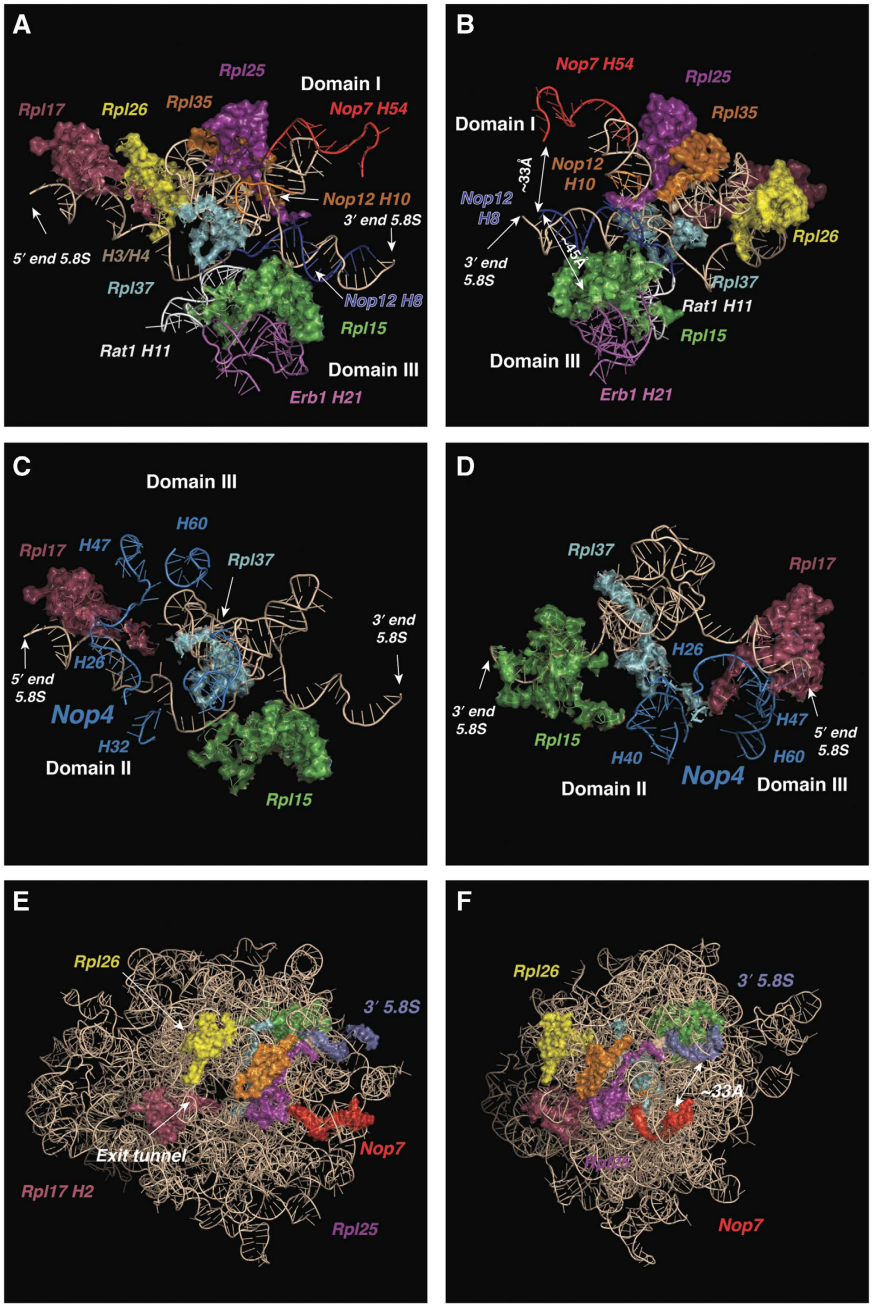



**Figure 4** Location of Nop4, Nop7, Erb1 and Nop12 and RNA binding sites in 60S crystal structure. (**A**, **B**) Images, rotated by 90°, showing the protein neighbourhood around the 3′-end of the yeast 5.8S rRNA. The 5.8S rRNA is shown in wheat color, with binding sites for Nop12 (orange and blue), Nop7 (red) and Erb1 (purple) indicated as coloured nucleotide strands. Images were generated using pymol. Ribosomal proteins (Rpl15, Rpl17, Rpl25, Rpl26, Rpl35 and Rpl37) binding to this region are indicated as surface representations. Helix numbers are indicated with 'H'. The double arrows indicate the distance (in Å) between the 3′-end region of 5.8S and the Nop7 and Erb1 binding sites. (**C**, **D**) Images showing the location of the Nop4 (blue) binding sites that surround the 5′-end of 5.8S. Ribosomal proteins Rpl15, Rpl17 and Rpl37 that bind in this region are indicated as surface representations. Helix numbers are indicated with 'H'. (**E**, **F**) Images showing the location of the Nop7 (red) binding site in the structure of the 25S rRNA (wheat color) and 5.8S 3′-end (light blue). Superimposed are structures for r-proteins Rpl35 (orange), Rpl25 (dark purple) and Rpl15 (green), Rpl26 (yellow) and Rpl17 (light purple). The polypeptide exit tunnel is indicated. The double arrow indicates the distance (in Å) between the 3′-end of 5.8S and the base of the helix containing the Nop7 binding site.

ing their participation in long-range interactions bringing domains III and I of 25S close to the 3′-end of 5.8S (Figure 4B).

Collectively, UV crosslinking, protein–protein interaction and pre-rRNA processing data strongly indicate that $A_3$-cluster proteins physically and functionally interact with the pre-60S region surrounding the 3′-end of 5.8S and with ITS2.

### *Altered Rat1 binding in strains depleted of Nop4 or Nop15*

Sites of Rat1 crosslinking may reflect targets undergoing active degradation, pre-rRNA docking sites directed by protein association or both. To distinguish between these possibilities, Rat1 crosslinking was analysed in strains depleted of Nop4 or the $A_3$ cluster protein Nop15.

Rat1-HTP strains carrying $P_{GAL}$*::3xHA-NOP4* or $P_{GAL}$*::3xHA-NOP15* were compared with the parental Rat1-HTP strain. Cells were harvested at an $OD_{600}$ of ~0.5 after 12 h of depletion in glucose-containing minimal medium. At this depletion time point, only modest defects in growth and pre-rRNA processing were observed (Supplementary Figure S6A and B). Crosslinked RNAs were cloned using barcoded 5′ linkers (Supplementary Table S1) and pooled cDNA libraries were Illumina deep sequenced. Depletion of Nop4 or Nop15 caused a ~1.6-fold reduction in the total number of reads that mapped to the rDNA (Figure 6A). To normalize the data sets, 2000 reads mapping to the rDNA from two independent experiments were randomly selected and the averaged number of hits for each nucleotide in the rDNA was plotted. Only binding sites present in all four Rat1-HTP CRAC experiments were considered.

The results for the 5′ ETS and a region comprising ITS1, 5.8S and ITS2 are presented in Figure 6C and D. Average hit densities (sum of the hits for nucleotides in a region) for the 5′ ETS and 5.8S-ITS2 boundary region (nts 2970–3170, shaded region in Figure 6D) from the normalized data sets are plotted in Figure 6B. Numerous changes were observed, but these were strikingly similar following depletion of either Nop4 or Nop15.

Three classes of changes were apparent. Loss of Nop4 or Nop15 did not dramatically alter the distribution of Rat1 crosslinking to the 5′-ETS region including $A_0$-$A_1$ (Figure 6C), whereas crosslinking was increased 2–3-fold over H3/4 at the 5′-end of 5.8S and H11 near the 5′-end of 25S, and also over the 5′-region of ITS1. Binding of Rat1 at all of these sites, therefore, appears to be independent of Nop4 and the $A_3$ cluster. In marked contrast, crosslinking was dramatically (~70%) reduced, across the $A_2$-$A_3$ and 5.8S 3′ to $C_2$ (Figure 6B and D). Notably, association of 5′ with $A_3$ was equally reduced in strains depleted of Nop4 (which inhibits processing before $A_3$ cleavage) and Nop15 (which inhibits processing following $A_3$ cleavage), indicating that loss of $A_2$-$A_3$ fragment synthesis is not solely responsible for the effect. Rat1 association over the 5.8S 3′- to ITS2 5′-region and the number of reads that terminated at $C_2$ (the 3′-end of 7S pre-rRNA) and the level of 7S were similarly reduced (Supplementary Figure S6C). We could not detectably co-precipitate 7S with HTP-tagged Rat1 (unpublished observation), but binding to ITS2 may be transient *in vivo* and unstable during the extended incubations needed for precipitation.

We propose that Rat1 binding 5′ to $A_3$ at least in part represents binding to $27SA_2$ and precedes cleavage at $A_3$, while binding 5′ to $C_2$ represents binding to $27SA_2$ and/or 27SB and precedes cleavage at $C_2$. Pre-binding of Rat1 at the $A_3$ and $C_2$ cleavage sites may explain the strikingly low steady-state abundance of the 3′ products of these cleavages ($27SA_3$ and 26S pre-rRNA, respectively) in wild-type cells. Rat1 association at these sites is apparently dependent on both Nop4 and the $A_3$ cluster.

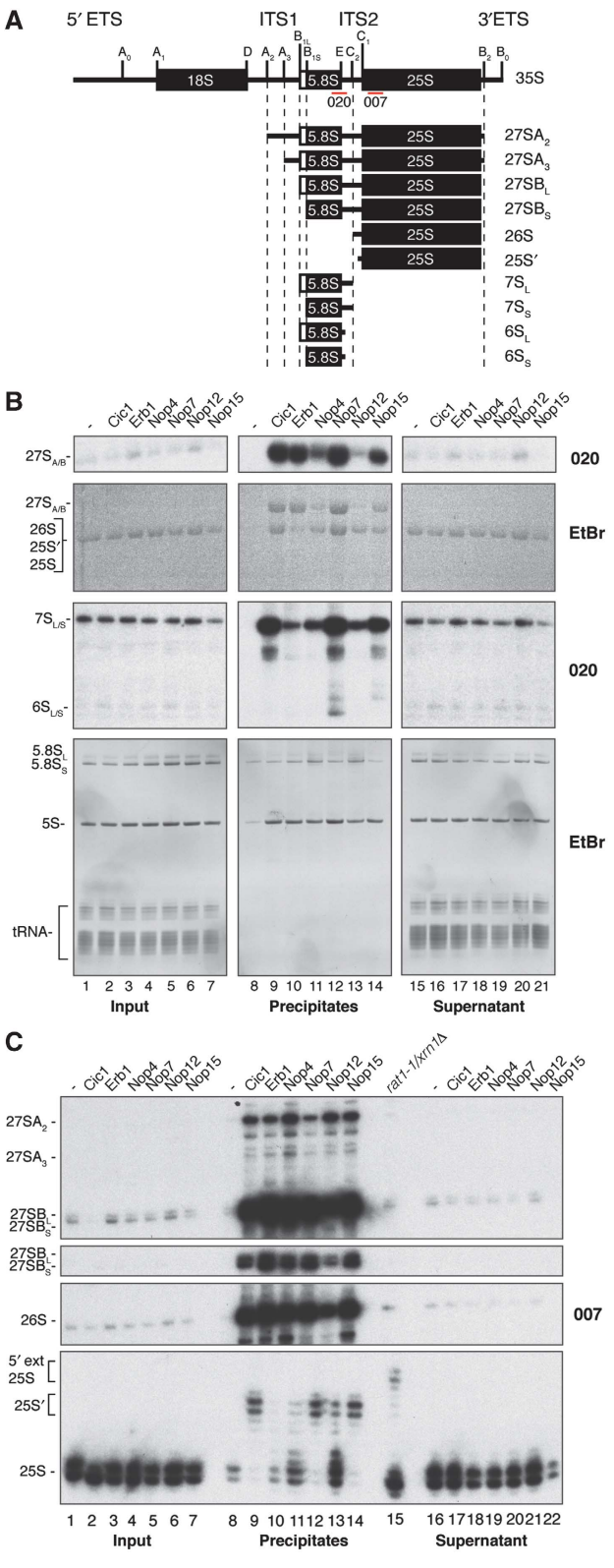



**Figure 5** Nop4, Nop7, Nop12, Erb1 and Nop15 associate with 66S pre-rRNAs. Immunoprecipitations were performed with HTP-tagged proteins and the non-tagged parental strain (−lanes). (**A**) Schematic representation of pre-rRNA species that are detected by oligonucleotides 020 and 007. Red lines indicate the positions of these oligonucleotides on the rDNA. RNAs were resolved on 1.2% agarose and 8% polyacrylamide/7 M urea gels and detected by northern hybridization (**B**) or primer extension (**C**) using oligonucleotides 020 and 007, respectively (Supplementary Table S1). Input and supernatant indicate 0.1% of RNA extracted from cell lysates and supernatants after immunoprecipitation. Note that in the northern blots shown in (**B**) longer exposures are shown for inputs and supernatants. *Rat1-1/xrn1*Δ primer extension products were loaded (**C**, lane 15) as markers for 26S and 5′-extended (5′-ext) 25S.

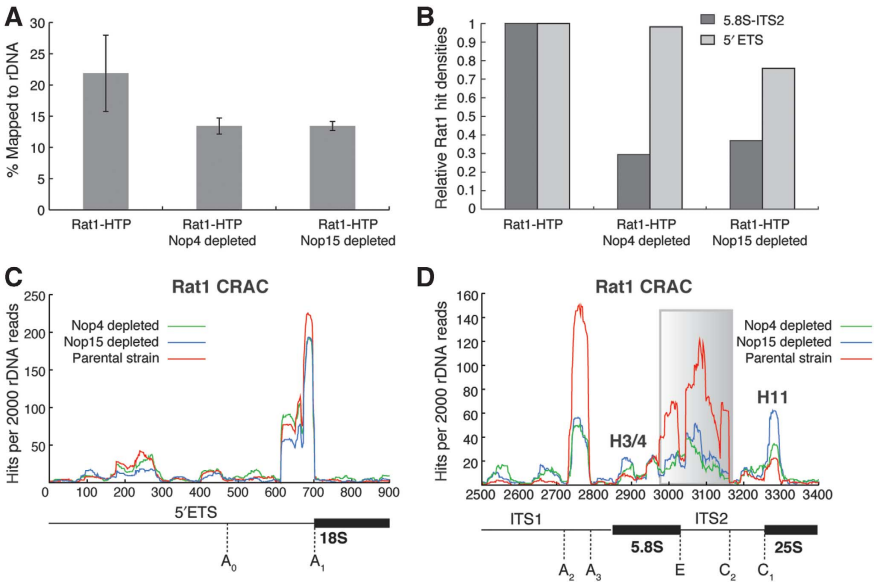



**Figure 6** Rat1 can bind helices 3, 4 and 11 in pre-rRNA independently of Nop4 and Nop15. (**A**) Depletion of Nop4 and $A_3$-cluster protein Nop15 reduces binding to the 3′ 5.8S–5′ ITS2 region but does not affect crosslinking to the 5′-ETS and helices 3, 4 and 11. Average hits for each nucleotide in the rDNA were calculated using 2000 rDNA reads from two independent experiments. Hit densities are the sum of hits for each nucleotide in a nucleotide sequence and were plotted as percentages relative to the results of the Rat1-HTP CRAC experiments. (**B**, **C**) Distribution of reads mapped to the 5′-ETS (**B**) and the 5.8S-ITS2 regions (**C**, **D**). Average hits per nucleotide from 2000 rDNA-mapped reads from two independent experiments are shown. Reads mapped to the 5′ ETS are shown in (**C**), whereas reads mapped to ITS1-5.8S-ITS2 region are shown in (**D**). The shaded area in (**D**) covers the region from which hit densities were calculated.

### *Cic1 and Nop15 are required to establish the conformation of ITS2*

Based on phylogenetics and structure probing, two distinct structures were proposed for ITS2, termed the 'ring' (Figure 7C) and 'hairpin' (Figure 7D) structures (Yeh and Lee, 1990; Joseph *et al*, 1999). These share the same distal stems but differ in the 25S and 5.8S proximal regions. Notably, the 5′-end of 25S and 3′-end of 5.8S are base paired in the mature ribosome, and are brought into proximity by the hairpin but not by the ring structure. Subsequent genetic data indicated that ITS2 undergoes a conformational change from the ring to hairpin structure during maturation (Cote *et al*, 2002). Structurally conserved regions of ITS2 are important for pre-rRNA processing (shaded in blue in Figure 7C and D; van der Sande *et al*, 1992; Van Nues *et al*, 1995; Peculis and Greer, 1998; Joseph *et al*, 1999; Cote *et al*, 2002), including domains II, III.A and III.B. These were proposed to be binding sites for ribosome synthesis factors (Van Nues *et al*, 1995; Cote *et al*, 2002) and they are indeed the crosslinking sites for Cic1 and Nop15 (see Figure 7C and D). Deletions and mutations in these regions abrogated 25S rRNA synthesis and caused 27SA accumulation, with a strong block in conversion of 27SA to 27SB (Van Nues *et al*, 1995; Cote *et al*, 2002); processing defects similar to those observed $A_3$-cluster protein-depleted cells (Supplementary Figure S4B and C; Van Nues *et al*, 1995). These findings strongly implicated Cic1 and Nop15 in the structural rearrangement of ITS2.

To assess the RNA secondary structure, we performed *in vivo* and *in vitro* chemical modification experiments using dimethyl-sulphate (DMS) (Figure 7), which preferentially methylates adenines and cytosines in single stranded or flexible RNA not bound by proteins. *In vitro* chemical probing reactions were performed on refolded total RNA, which helps to identify RNA regions potentially protected by protein binding *in vivo*. We analysed DMS modifications in conditional ($P_{GAL}$) Nop15 and Cic1 strains, grown for 12 h in glucose medium to deplete the proteins. As controls, we analysed the parental strain (BY4741) and the strain depleted of Nop4, which did not significantly crosslink to ITS2 (Figure 3C). DMS modifications were mapped by primer extension analysis and autoradiography and quantified by normalizing signal intensities for each band to remove differences between lanes (Figure 7E). Note that primer extension was performed on total RNA, revealing the 'average' secondary structure of all 27S pre-rRNA species.

The 3′ most region of ITS2 could not be analysed but the distal region (stems VI and V) adopts virtually the same secondary structure *in vitro* or *in vivo* and in the depleted and parental strains (Figure 7B). In contrast, ITS2 stem II has a more accessible structure *in vivo* than in refolded RNA, as judged by the intensity of the primer extension stops at nucleotides C2 and C5 (indicated by open circles in Figure 7A, compare lane 9 with lane 1, and see Supplementary Figure S8). This is consistent with C2 and C5 base pairing with G22 and G25 *in vitro* (dotted arrows in Figure 7), forming a longer stem II very similar to that proposed in the hairpin model (Figure 7D).

In Cic1- and Nop15-depleted cells, C2 and C5 in stem II were almost completely protected from DMS modification *in vivo* (Figure 7A, compare lanes 3, 4 with lane 1). We conclude that binding of Cic1 (and perhaps also Nop15) to this region blocks formation of these base-pairing interactions, maintaining the shorter stem II structure of the ring structure. Notably, these nucleotides are located in a highly conserved region required for $A_3$-$B_{1S}$ processing (Van Nues *et al*, 1995).

Stem III (stem III.A in the ring model) was more protected from DMS modification *in vivo* compared with the refolded RNA (Figure 7A, compare lane 9 with lane 1), but this protection was lost in Nop15-depleted cells (Figure 7A, compare lane 4 with lane 1). The CRAC data indicate that

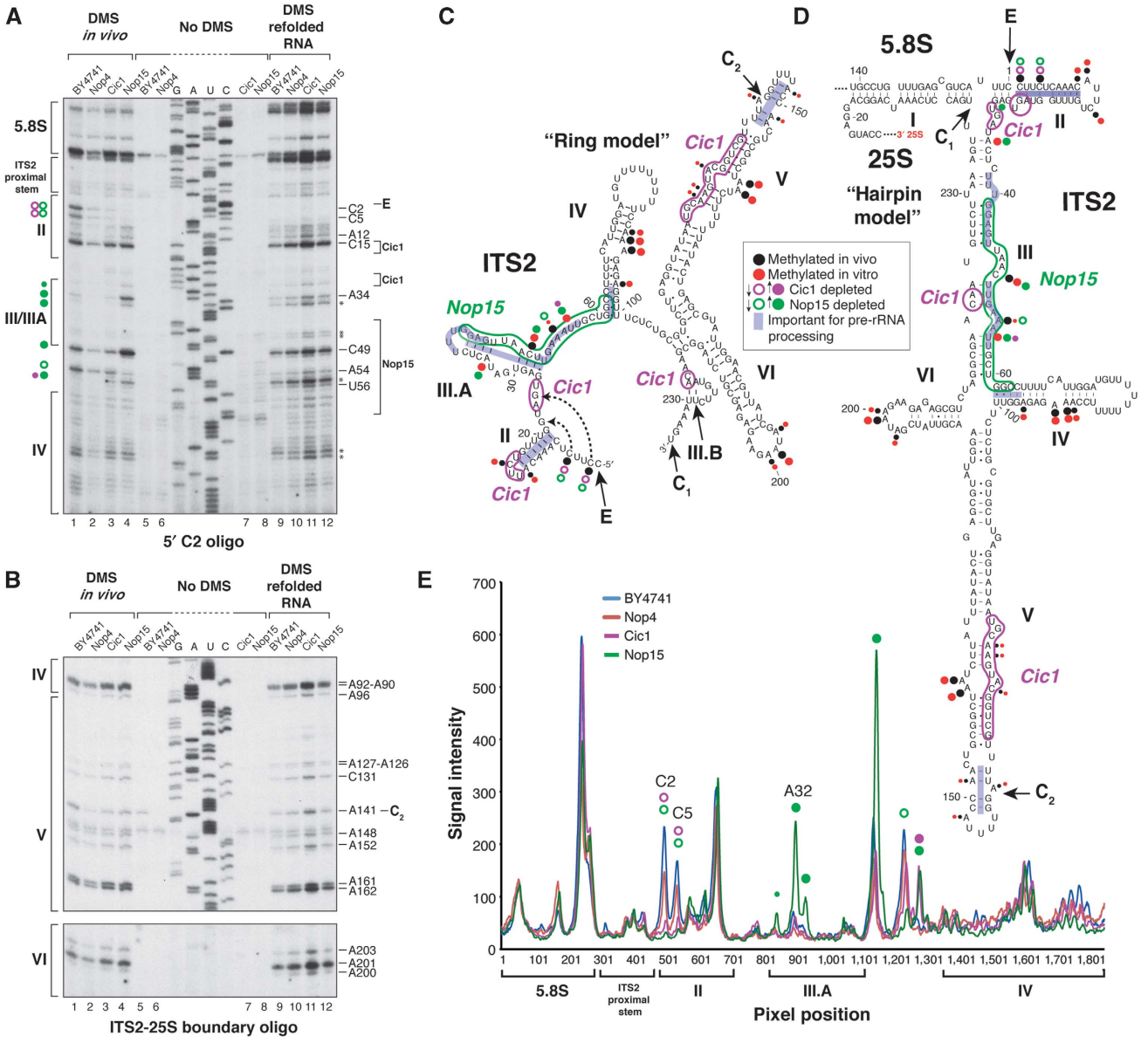



**Figure 7** Cic1 and Nop15 are required to maintain a flexible RNA conformation in the 5′-end of ITS2. (**A**, **B**) The 5′-end of ITS2 adopts a different conformation in Nop15- and Cic1-depleted cells. Parental strain (BY4741), and Nop4, Nop15 and Cic1 GAL depletion strains were grown in YPG/R to exponential phase and subsequently grown for 12 h in glucose. Dimethyl-sulphate (DMS) probing was performed *in vivo* in cell culture (lanes 1–4) or on *in vitro* refolded total RNA extracted from the same cultures (lanes 9–12). To control for natural primer extension stops, primer extension was performed on unmodified total RNA (lanes 5–8). DMS modifications were mapped by primer extension and detected by autoradiography following gel electrophoresis. The positions of the modified nucleotides are indicated on the right side of each panel. Results for the ITS1 and 5.8S are shown in Supplementary Figure S7. Oligonucleotides used for primer extensions (Supplementary Table S1) are indicated below each panel. Position of helices and RNA domains is indicated on the left of each panel. Coloured circles indicate changes in DMS reactivity in Nop15 (green circles) and Cic1-depleted cells (purple circles). Small circles depict weak DMS modifications; closed circles indicate an increased DMS reactivity; open circles indicate a decreased DMS reactivity. (**C**, **D**) Overview of the chemical probing results on the 'ring' (**C**) and 'hairpin' secondary structure models for ITS2. Positions of processing sites ($C_1$, $C_2$ and E) are indicated by arrows. Changes in DMS modifications are highlighted as described above. Red circles indicate nucleotides that were methylated *in vitro*, whereas black circles indicate *in vivo* modified nucleotides. (**E**) Quantification of primer extension results shown in (**A**). The data were normalized by removing differences in signal intensity between lanes. *Y* axis indicates the signal intensities calculated by the AIDA software. *X* axis indicates pixel positions in the lanes. Positions of 5.8S sequences and ITS2 domains II, III and IV are indicated below the *x* axis.

Nop15 binds to this region, which would largely explain the observed differences in DMS modification in stem III.

Depletion of Nop4 did not dramatically affect the *in vivo* DMS modification pattern in ITS2 compared with the other strains tested (Figure 7A and B). Notably, we did observe a reduction in DMS reactivity in depleted cells at G319 in ITS1, between the $A_3$ and $B_{1L}$ cleavage sites (Supplementary Figure S7, indicated with arrows); however, primer extension stops at this position were also found in unmodified RNA, indicating that it represents a natural primer extension stop (Supplementary Figure S7, lanes 5–8). The significance of this observation remains unclear.

We conclude that Cic1 and Nop15 binding is necessary to maintain or establish a more flexible and open structure in the 5′-end of ITS2, potentially preventing premature formation of the hairpin structure.

## Discussion

The eukaryotic pre-rRNA processing and ribosome assembly pathways are strikingly complex making coordination and timing of key importance. Following cleavage of site $A_3$ by the endonuclease RNase MRP, the 5′-exonuclease Rat1 digests the pre-rRNA back to the major 5′-end of the 5.8S rRNA at site $B_{1S}$ (Schmitt and Clayton, 1993; Henry *et al*, 1994; Lygerou *et al*, 1996; Figure 1A). We initially assumed that $A_3$ cleavage to generate a free 5′-end would be necessary and sufficient for recruitment of Rat1, which shows robust, processive activity *in vitro* on non-specific substrates (Kenna *et al*, 1993). However, subsequent data revealed that $A_3$ processing requires a group of eight proteins (Nop12, Nop7, Erb1, Ytm1, Rlp7 Nop15, Nsa3/Cic1 and Rrp1), termed the $A_3$ cluster (Tang *et al*, 2008) (Sahasranaman *et al*, submitted) (this work). Similarly, Nop4 and other factors are required *in vivo* for $A_3$ cleavage (Bergès *et al*, 1994; Sun and Woolford, 1994), even though this site can be cleaved by RNase MRP alone *in vitro* (Lygerou *et al*, 1996).

Rat1 showed strong pre-rRNA binding over several different regions, reflecting its diverse targets for processing and degradation. Strong signals were seen over the $A_0$-$A_1$ region of the 5′-ETS, a known Rat1 target, and at sites further 5′, presumably reflecting a role in spacer degradation following endonuclease cleavage by the PIN domain of the Rrp44 component of the exosome complex (Lebreton *et al*, 2008; Schaeffer *et al*, 2009; Schneider *et al*, 2009; Wang and Pestov, 2010). In contrast, fewer hits were recovered over the $A_3$-$B_{1S}$ and $C_2$-$C_1$ regions that are targets for Rat1 processing during pre-rRNA maturation. The steady-state levels of $A_3$-cleaved and $C_2$-cleaved pre-rRNAs ($27SA_3$ and 26S, respectively, in Figure 1A) are very low, indicating that exonuclease digestion very rapidly follows endonuclease cleavage, and processing intermediates are essentially undetectable. These observations indicate that Rat1 is actively recruited to, and highly processive on, the $27SA_3$ and 26S pre-rRNAs.

Rat1 was strongly crosslinked to regions immediately 5′ to the $A_3$ and $C_2$ cleavage sites, suggesting that it is bound to these locations before cleavage. Rat1 binding 5′ to $A_3$ presumably also partly reflects its role is $A_2$-$A_3$ fragment degradation (Petfalski *et al*, 1998). However, the similar effects of depletion of Nop4 and Nop15, which differ in their effects on $A_3$ processing, strongly indicates that loss of $A_2$-$A_3$ is not solely responsible for the reduction. In contrast, crosslinking of Rat1 to the 5′-region of 5.8S (H3/4) and H11 in the 5′-region of 25S rRNA was independent of the $A_3$-cluster protein Nop15 and of Nop4. This suggests that these helices provide distinct access points for Rat1 in pre-ribosomes, potentially representing initial binding sites.

### *Model for links between pre-ribosome structure and processing from site $A_3$*

CRAC analyses placed Nop7, Erb1, Cic1, Nop12 and Nop15 on the surface of the pre-60S ribosomes in the vicinity of the 5.8S 3′–25S 5′ stem. Using the CRAC data, we can extend the interaction map for the $A_3$-cluster proteins (see Figure 1B). Some of these interactions are superimposed onto a schematic of the pre-rRNA in Figure 8, allowing potential mechanisms to be proposed.

In the mature 60S structure, both the 5′- and 3′-regions of 5.8S rRNA are base paired to the 25S rRNA. However, based on *in vitro* chemical modification experiments, it was proposed that in the 35S pre-rRNA the 5.8S 5′-region is initially base paired to ITS1 (Yeh *et al*, 1990), forming a stable stem-loop structure that will be strongly favoured over the long-range interaction with 25S during folding of the nascent transcript. Our *in vivo* chemical probing experiments are largely consistent with this model (Supplementary Figure S7), as are analyses of the intermediates detected in *rat1* mutants *in vivo* (Henry *et al*, 1994) and phylogenetic analyses (Van Nues *et al*, 1994). Rat1 processing of ITS1 will free the 5′-end of 5.8S, which might be a prerequisite for the establishment of base pairing with 25S (Figure 8).

The 3′-end of 5.8S and 5′-end of 25S are separated by ITS2, which was previously reported to be refolded during pre-rRNA maturation from a more open 'ring' conformation into a closed 'hairpin' structure (Yeh and Lee, 1990; Peculis and Greer, 1998; Joseph *et al*, 1999). In the hairpin structure, a proximal stem brings the 5.8S and 25S rRNAs into close proximity, presumably promoting base pairing between them. The ring-to-hairpin transition was shown to be an active process involving unidentified ribosome synthesis factors (Peculis and Greer, 1998). The locations of the binding sites for Nop15 and Cic1 are in strikingly good agreement with the binding sites predicted for these factors. Moreover, the phenotypes of ITS2 mutations expected to block these interactions confer processing defects similar to those seen in $A_3$-cluster mutants. *In vivo* structure probing indicated that depletion of Nop15 and Cic1 indeed shifted the conformation of ITS2 from the ring to the hairpin structure. The hairpin structure is predicted to be more thermodynamically stable and is very likely to fold spontaneously. Binding of Nop15 and Cic1 blocks folding of the hairpin, potentially establishing the timing of base pairing between 5.8S 3′- and 25S 5′-regions (Figure 8A).

$A_3$-cluster proteins show interdependent pre-rRNA binding (i.e., depletion of a single protein causes loss of many or all $A_3$-cluster proteins; Tang *et al*, 2008) (Sahasranaman *et al*, submitted). This implies that they interact—so their binding sites, which are dispersed in the pre-rRNA sequence, must all be brought into proximity. Notably, several 60S r-proteins make contact with multiple sites in the rRNA, which must also be brought together during subunit assembly. We speculate that this is a key function of the ribosome synthesis factors.

Nop7 directly binds Erb1, forming an inter-domain bridge in the 27S pre-rRNA. Cells lacking these $A_3$-cluster proteins fail to assemble r-proteins Rpl35, Rpl37, Rpl26 and Rpl17 in pre-60S complexes (Sahasranaman *et al*, submitted). All of these r-proteins have binding sites in close proximity to Nop12, Erb1 and Nop7 crosslinking sites (and presumably also Cic1 and Nop15; Figure 3D) and contact both the 25S rRNA and the 5.8S rRNA (Ben-Shem *et al*, 2010). Nop4 crosslinking sites were located near the binding sites for Rpl17, Rpl35 and Rpl37 in each of two different domains. In addition, Nop7 was genetically linked to Rpl25 and crosslinked adjacent to an Rpl25 binding site in 25S domain III. The Nop4, Nop7 and Erb1 interactions with pre-rRNA bring together sites that are distant in the primary sequence. We speculate that this establishes an assembly platform for Rpl17, Rpl25, Rpl26, Rpl35 and Rpl37, allowing these proteins to simultaneously interact with multiple regions of

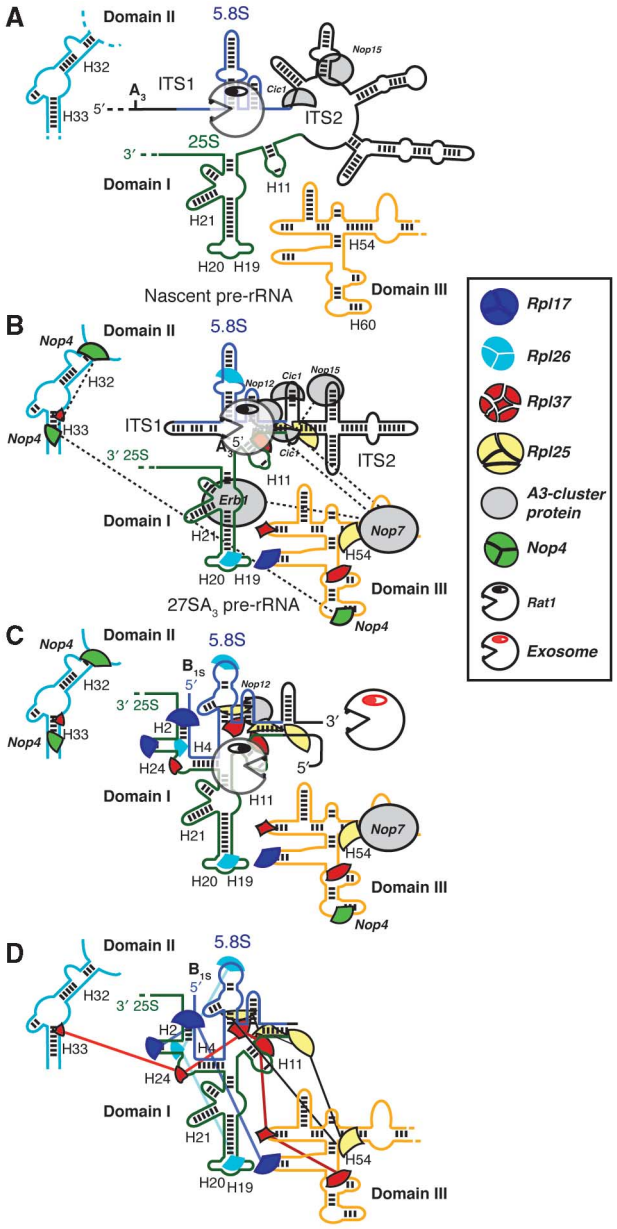



**Figure 8** Speculative model for the reorganization of the ITS1-5.8S–ITS2-25S region during Rat1-dependent removal of ITS1. (**A**) Schematic representation of the secondary structure of the nascent pre-rRNA. Domain I of the 25S rRNA sequence is indicated in green, domain II is indicated in light blue, domain III is indicated in orange and 5.8S rRNA is indicated in dark blue. ITS2 is drawn in the 'ring' conformation predicted for the nascent transcript (Joseph *et al*, 1999). Rat1 binds before $A_3$ cleavage and is depicted on 5.8S H3/4, whereas Cic1 and Nop15 are depicted in grey bound to the ITS2 'ring' conformation. (**B**) Binding sites for the $A_3$-cluster proteins are indicated in grey. Nop4 binding sites are indicated in green. Interactions between these factors (dashed black lines) are predicted to bring all of these positions into close proximity. Coloured segments represent binding sites for r-proteins Rpl25, Rpl26, Rpl37 and Rpl17. These proteins each contact multiple rRNA sequences that are dispersed in the 1° and 2° structures but associate in the ribosomal particles. Stable pre-rRNA binding by these r-proteins requires the $A_3$-cluster proteins (Sahasranaman *et al*, submitted), which may bring together the dispersed components of the r-protein binding sites. Major Rat1 binding sites are located in the 5′-ends of 5.8S and 25S, which are independent of the $A_3$-cluster protein Nop15 and Nop4. In the absence of the $A_3$ cluster, we speculate that the 5′-end of ITS2 rapidly adopts a highly base-paired conformation preventing access to $A_3$ and $C_2$ for Rat1, bound to 5.8S H3/4 and/or 25S H11, which are in close proximity. Interactions between Rat1 and sequences 5′ to $A_3$ and $C_2$ likely precede cleavage, but have been omitted for simplicity. (**C**) Degradation of ITS1 by Rat1 makes the 5′-end of the 5.8S rRNA (blue) accessible for base pairing with the 25S rRNA sequences (green), resulting in the formation of 5.8S–25S interactions H3 and H4. We speculate that formation of these stem structures triggers stable binding of r-protein Rpl17, Rpl26 and Rpl37. After removal of ITS1, ITS2 is cleaved at $C_2$. The ITS2 3′-region is very rapidly processed back to the 25S 5′-end, and we predict that this involves Rat1 bound at 25S H11. Degradation of ITS2 is likely accompanied by removal of Nop15 and Cic1. (**D**) The remaining $A_3$-cluster proteins dissociate from the mature 25S rRNA and 6S pre-rRNA. Long-range interactions that were initially established by the $A_3$-cluster proteins are now maintained by the r-proteins via their multipartite binding elements (indicated by coloured lines).

the pre-rRNA. The *H. marismortui* homologue of Rpl26 (HmRpl24) is an initiator protein for 50S assembly (Nowotny and Nierhaus, 1982) and appears critical for the tertiary structure of 23S Domain I (Klein *et al*, 2004). HmRpl24 binds in the centre of 5.8S rRNA and to the region corresponding to H24 near the 5′-end of 5.8S (Klein *et al*, 2004). Notably, the binding sites for Rpl26 include at least four rRNA regions that are dispersed in the primary and secondary structures, including contacts at H19, in proximity to the Erb1 binding site, and the 5′-end of 5.8S (Figures 3D and 8). We propose that interactions between the $A_3$-cluster proteins bring these regions together to promote Rpl26 binding. In turn, Rpl26 may be required for correct 5.8S rRNA folding, correct positioning of the 5′-region of 5.8S and binding of further r-proteins, including Rpl17, Rpl35 and Rpl37.

Erb1 appears to dissociate from pre-ribosomes after $C_2$ is cleaved (Figure 5), suggesting that this bridging interaction is required only at early stages of 60S assembly. This is consistent with bridging by the $A_3$-cluster proteins establishing long-range interactions that are subsequently maintained by the r-proteins. Future structural analyses and assays for r-protein association in strains depleted of individual ribosome synthesis factors will be required to test these models.

## Conclusions

Previous analyses indicated that recruitment of the 18S rRNA 3′-endonuclease Nob1 to the pre-ribosomes is distinct from its activation for pre-rRNA cleavage (Lamanna and Karbstein, 2009; Pertschy *et al*, 2009; Granneman *et al*, 2010). This also appears to be the case for Rat1, suggesting that this might be a more general property of eukaryotic pre-rRNA processing enzymes. Pre-rRNA cleavage or digestion is an irreversible step and separation of binding and activation of the processing enzymes presumably permits better regulation and quality control during pre-ribosome assembly.

Rat1 can terminate Pol I transcription by degradation of the nascent transcript (El Hage *et al*, 2008). Rat1 also degrades the excised $A_2$-$A_3$ fragment (Petfalski *et al*, 1998), and is therefore capable of initiating degradation at site $A_2$. This makes it important that exonuclease degradation from $A_2$ and $A_3$ be delayed until the transcript is released from the polymerase by cleavage at site $B_0$, almost 4 kb away (Figure 1A). We propose that the structural reorganization of ITS2, induced by $A_3$-cluster binding, is important for

regulating the timing of 5′-processing of the 5.8S rRNA. ITS2 was introduced into an ancestral 23S-like rRNA early in eukaryotic evolution and this mechanism offers a potential rationale for the advantage of having, and retaining, an ITS2 region.

## Materials and methods

### *Yeast strains and media*

*Saccharomyces cerevisiae* strain BY4741 (MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0) was used as the parental strain (Brachmann *et al*, 1998). The HTP carboxyl-tagged strains were generated by PCR as described (Rigaut *et al*, 1999). Strains were grown in YPD (1% yeast extract, 2% peptone, 2% dextrose), YPG/R (1% yeast extract, 2% peptone, 2% galactose, 2% raffinose) or yeast minimal medium (Formedium) lacking uracil and tryptophan, at 30 °C.

### *CRAC and sequence analyses*

*In vivo* CRAC, western and northern blot analyses were performed essentially as described previously (Granneman *et al*, 2009) with the following modifications: For the Rat1 CRAC experiments shown in Figure 6, cells were pre-grown in filter sterilized synthetic minimal medium containing galactose and raffinose, but lacking uracil and tryptophan (SG/R-URA-TRP), to an $OD_{600}$ of 1.0, subsequently shifted to glucose containing medium (SD-URA-TRP) for 12 h to an $OD_{600}$ of ~0.5. Cultures were UV irradiated in the Megatron (Supplementary Figure S1A) at room temperature for 100 s (equivalent to an average dose of ~1.6 mJ/cm$^2$) and cells were harvested by centrifugation, washed with phosphate-buffered saline (PBS), and stored at −80 °C. Megatron parts were purchased from UVO3 (http://www.uvo3.co.uk; contact Peter Wadsworth http://Peter@uvo3.co.uk). Sanger sequencing of cDNAs was performed as described (Granneman *et al*, 2009). Negative control cDNAs were generated from RNA extracted from excised membrane fragments from mock experiments run in parallel. For the high-throughput sequencing analysis, crosslinked RNAs were sequentially ligated to L3 and barcoded L5 adaptors and amplified by RT/PCR (see Supplementary Table S1 for oligonucleotide sequences). Oligonucleotides were purchased from Integrated DNA technologies (IDT). Illumina sequencing (single end 50-bp reads) was performed according to the manufacturer's procedure and reads were aligned to the yeast genome using novoalign (http://www.novocraft.com). Data analyses were performed using pyCRAC, a set of Python tools for high throughput sequence analysis (Webb, Tollervey and Granneman, manuscript in preparation).

### *Chemical footprinting experiments*

Parental (BY4741) and GAL depletion strains (GAL::3HA-*nop4*, GAL::3HA-*cic1* and GAL::3HA-*nop15*) were grown in YPG/R to logarithmic phase, shifted to YPD and grown for 12 h to an $OD_{600}$ of ~0.5. For the *in vivo* modification experiment, 200 μl of 33% DMS solution (diluted in ethanol) was added to 10 ml of cells and shaken for 3 min at room temperature. DMS was quenched using 10 ml of 0.7 M β-mercaptoethanol and 5 ml of water-saturated isopropanol. Cell pellets were subsequently washed once with 10 ml 0.7 M β-mercaptoethanol and once with ice-cold phosphate saline (PBS). RNA extractions were performed as previously described (Tollervey, 1987). For the *in vitro* DMS experiments, RNA extracted from 5 ODs of cells was dissolved in 400 μl of refolding buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaAc, 1.5 mM $MgCl_2$). RNA was refolded by slowly cooling heat denatured RNA (65 °C 10 min) to room temperature. Refolded RNA was incubated with 20 μl 33% DMS solution (diluted in ethanol) for 3 min at room temperature and reactions were quenched by adding 160 μl of 0.7 M β-mercaptoethanol. One microgram of total RNA was used for primer extension reactions. Primer extension products were resolved on 6% polyacrylamide/7 M urea gels and visualized by autoradiography. Chemical probing data were quantified as described in the main text using a Fuji FLA5000 phosphoimager and the Aida software package.

### *Supplementary data and methods*

Supplementary data and Materials and Methods are available at The *EMBO Journal* Online (http://www.embojournal.org).

## Acknowledgements

We thank Ernst Granneman, Peter Davis and the Darwin workshop for help with designing and modifying the Megatron; John Woolford and Aarti Sahasranaman for communicating unpublished results; Adam Ben-Shem, Lasse Jenner, Gulnara Yusupova and Marat Yusupov for making the yeast 80S ribosome crystal structure available; Tollervey laboratory members for helpful suggestions; the Edinburgh Gene Pool Sanger and Solexa Sequencing Facilities; and the Swann Building kitchen staff. This work was supported by the Welcome Trust, the BBSRC (BB/D019621/1), EMBO long-term fellowship (SG) and a Marie Curie EIF fellowship (SG).

*Author contributions*: SG and DT designed the research. SG, EP performed the research and analysed the data. SG and DT wrote the paper.

## Conflict of interest

The authors declare that they have no conflict of interest.

## References

Babiano R, de la Cruz J (2010) Ribosomal protein L35 is required for 27SB pre-rRNA processing in Saccharomyces cerevisiae. *Nucleic Acids Res* **38:** 5177–5192

Ben-Shem A, Jenner L, Yusupova G, Yusupov M (2010) Crystal structure of the eukaryotic ribosome. *Science* **330:** 1203–1209

Bergès T, Petfalski E, Tollervey D, Hurt EC (1994) Synthetic lethality with fibrillarin identifies NOP77p, a nucleolar protein required for pre-rRNA processing and modification. *EMBO J* **13:** 3136–3148

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14:** 115–132

Cote CA, Greer CL, Peculis BA (2002) Dynamic conformational model for the role of ITS2 in pre-rRNA processing in yeast. *RNA* **8:** 786–797

Dunbar DA, Dragon F, Lee SJ, Baserga SJ (2000) A nucleolar protein related to ribosomal protein L7 is required for an early step in large ribosomal subunit biogenesis. *Proc Natl Acad Sci USA* **97:** 13027–13032

El Hage A, Koper M, Kufel J, Tollervey D (2008) Efficient termination of transcription by RNA polymerase I requires the 5′ exonuclease Rat1 in yeast. *Genes Dev* **22:** 1069–1081

Faber AW, Vos HR, Vos JC, Raue HA (2006) 5′-end formation of yeast 5.8SL rRNA is an endonucleolytic event. *Biochem Biophys Res Commun* **345:** 796–802

Fang F, Phillips S, Butler JS (2005) Rat1p and Rai1p function with the nuclear exosome in the processing and degradation of rRNA precursors. *RNA* **11:** 1571–1578

Fatica A, Oeffinger M, Tollervey D, Bozzoni I (2003) Cic1p/Nsa3p is required for synthesis and nuclear export of 60S ribosomal subunits. *RNA* **9:** 1431–1436

Gadal O, Strauss D, Petfalski E, Gleizes PE, Gas N, Tollervey D, Hurt E (2002) Rlp7p is associated with 60S preribosomes, restricted to the granular component of the nucleolus, and required for pre-rRNA processing. *J Cell Biol* **157:** 941–951

Geerlings TH, Vos JC, Raue HA (2000) The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5′->3′ exonucleases. *RNA* **6:** 1698–1703

Granneman S, Kudla G, Petfalski E, Tollervey D (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high throughput analysis of cDNAs. *Proc Natl Acad Sci USA* **106:** 9613–9818

Granneman S, Petfalski E, Swiatkowska A, Tollervey D (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNA-protein cross-linking. *EMBO J* **29:** 2026–2036

Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S, Tollervey D (1994) The 5′ end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J* **13:** 2452–2463

Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* **98:** 4569–4574

Joseph N, Krauskopf E, Vera MI, Michot B (1999) Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. *Nucleic Acids Res* **27:** 4533–4540

Kenna M, Stevens A, McCammon M, Douglas MG (1993) An essential yeast gene with homology to the exonuclease-encoding *XRN1/KEM1* gene also encodes a protein with exoribonuclease activity. *Mol Cell Biol* **13:** 341–350

Klein DJ, Moore PB, Steitz TA (2004) The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J Mol Biol* **340:** 141–177

Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP, Beattie BK, Lalev A, Zhang W, Davierwala AP, Mnaimneh S, Starostine A, Tikuisis AP, Grigull J, Datta N, Bray JE *et al*. (2004) High-definition macromolecular composition of yeast RNA-processing complexes. *Mol Cell* **13:** 225–239

Lamanna AC, Karbstein K (2009) Nob1 binds the single-stranded cleavage site D at the 3′-end of 18S rRNA with its PIN domain. *Proc Natl Acad Sci USA* **106:** 14259–14264

Lebreton A, Tomecki R, Dziembowski A, Seraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* **456:** 993–996

Lo K-Y, Li Z, Bussiere C, Bresson S, Marcotte EM, Johnson AW (2010) Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. *Mol Cell* **39:** 196–208

Lygerou Z, Allmang C, Tollervey D, Séraphin B (1996) Accurate processing of a eukaryotic pre-rRNA by RNase MRP *in vitro*. *Science* **272:** 268–270

Miles TD, Jakovljevic J, Horsey EW, Harnpicharnchai P, Tang L, Woolford Jr JL (2005) Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes. *Mol Cell Biol* **25:** 10419–10432

Nissan TA, Baßler J, Petfalski E, Tollervey D, Hurt E (2002) 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J* **21:** 5539–5547

Nowotny V, Nierhaus KH (1982) Initiator proteins for the assembly of the 50S subunit from Escherichia coli ribosomes. *Proc Natl Acad Sci USA* **79:** 7238–7242

Oeffinger M, Leung A, Lamond A, Tollervey D (2002) Yeast Pescadillo is required for multiple activities during 60S ribosomal subunit synthesis. *RNA* **8:** 626–636

Oeffinger M, Tollervey D (2003) Yeast Nop15p is an RNA-binding protein required for pre-rRNA processing and cytokinesis. *EMBO J* **22:** 6573–6583

Oeffinger M, Zenklusen D, Ferguson A, Wei KE, El Hage A, Tollervey D, Chait BT, Singer RH, Rout MP (2009) Rrp17p is a eukaryotic exonuclease required for 5′ end processing of Pre-60S ribosomal RNA. *Mol Cell* **36:** 768–781

Panse VG, Johnson AW (2010) Maturation of eukaryotic ribosomes: acquisition of functionality. *Trends Biochem Sci* **35:** 260–266

Peculis BA, Greer CL (1998) The structure of the ITS2-proximal stem is required for pre-rRNA processing in yeast. *RNA* **4:** 1610–1622

Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J Biol Chem* **284:** 35079–35091

Pestov DG, Stockelman MG, Strezoska Z, Lau LF (2001) ERB1, the yeast homolog of mammalian Bop1, is an essential gene required for maturation of the 25S and 5.8S ribosomal RNAs. *Nucleic Acids Res* **29:** 3621–3630

Petfalski E, Dandekar T, Henry Y, Tollervey D (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol Cell Biol* **18:** 1181–1189

Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17:** 1030–1032

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat Struct Mol Biol* **16:** 56–62

Schmitt ME, Clayton DA (1993) Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13:** 7935–7941

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

Sun C, Woolford Jr JL (1997) The yeast nucleolar protein Nop4p contains four RNA recognition motifs necessary for ribosome biogenesis. *J Biol Chem* **272:** 25345–25352

Sun C, Woolford JLJ (1994) The yeast *NOP4* gene product is an essential nucleolar protein required for pre-rRNA processing and accumulation of 60S ribosomal subunits. *EMBO J* **13:** 3127–3135

Tang L, Sahasranaman A, Jakovljevic J, Schleifman E, Woolford Jr JL (2008) Interactions among Ytm1, Erb1, and Nop7 required for assembly of the Nop7-subcomplex in yeast preribosomes. *Mol Biol Cell* **19:** 2844–2856

Tarassov K, Messier V, Landry CR, Radinovic S, Serna Molina MM, Shames I, Malitskaya Y, Vogel J, Bussey H, Michnick SW (2008) An *in vivo* map of the yeast protein interactome. *Science* **320:** 1465–1470

Tollervey D (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J* **6:** 4169–4175

van Beekvelt CA, Kooi EA, de Graaff-Vincent M, Riet J, Venema J, Raue HA (2000) Domain III of Saccharomyces cerevisiae 25 S ribosomal RNA: its role in binding of ribosomal protein L25 and 60 S subunit formation. *J Mol Biol* **296:** 7–17

van der Sande CAFM, Kwa M, van Nues R, van Heerikhuizen H, Raué HA, Planta RJ (1992) Functional analysis of internal transcribed spacer 2 of Saccharomyces cerevisiae ribosomal DNA. *J Mol Biol* **223:** 899–910

Van Nues RW, Rientjes JMJ, Morré SA, Mollee E, Planta RJ, Venema J, Raué HA (1995) Evolutionarily conserved structural elements are critical for processing of internal transcribed spacer 2 from *Saccharomyces cerevisiae* precursor ribosomal RNA. *J Mol Biol* **250:** 24–36

Van Nues RW, Rientjes JMJ, Van der Sande CAFM, Zerp SF, Sluiter C, Venema J, Planta RJ, Raué HA (1994) Separate structural elements within internal transcribed spacer 1 of Saccharomyces cerevisiae precursor ribosomal RNA direct the formation of 17S and 26S rRNA. *Nucleic Acids Res* **22:** 912–929

Veldman GM, Brand RC, Klootwijk J, Planta RJ (1980) Some characteristics of processing sites in ribosomal precursor RNA of yeast. *Nucleic Acids Res* **8:** 2907–2920

Wang M, Pestov DG (2010) 5′-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. *Nucleic Acids Res* **39:** 1811–1822

Wu K, Wu P, Aris JP (2001) Nucleolar protein Nop12p participates in synthesis of 25S rRNA in Saccharomyces cerevisiae. *Nucleic Acids Res* **29:** 2938–2949

Xue Y, Bai X, Lee I, Kallstrom G, Ho J, Brown J, Stevens A, Johnson AW (2000) Saccharomyces cerevisiae RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol Cell Biol* **20:** 4006–4015

Yeh L-CC, Thweatt R, Lee JC (1990) Internal transcribed spacer 1 of the yeast precursor ribosomal RNA Higher order structure and common structural motifs. *Biochemistry* **29:** 5911–5918

Yeh LC, Lee JC (1990) Structural analysis of the internal transcribed spacer 2 of the precursor ribosomal RNA from Saccharomyces cerevisiae. *J Mol Biol* **211:** 699–712