Chaperoning 5S RNA assembly

Clément Madru,[1,5] Simon Lebaron,[1,2,5] Magali Blaud,[1] Lila Delbos,[1] Juliana Pipoli,[3] Eric Pasmant,[4] Stéphane Réty,[1] and Nicolas Leulliot[1]

[1]Laboratoire de Cristallographie et RMN Biologiques, UMR, CNRS 8015, Université Paris Descartes, Faculté de Pharmacie, Sorbonne Paris Cité, 75006 Paris, France; [2]CR2 INSERM; [3]Plateforme Génomique, Institut Cochin, UMR_S1016, INSERM, 75006 Paris, France; [4]Service de Biochimie et Génétique Moléculaire, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, 75006 Paris, France

In eukaryotes, three of the four ribosomal RNAs (rRNAs)—the 5.8S, 18S, and 25S/28S rRNAs—are processed from a single pre-rRNA transcript and assembled into ribosomes. The fourth rRNA, the 5S rRNA, is transcribed by RNA polymerase III and is assembled into the 5S ribonucleoprotein particle (RNP), containing ribosomal proteins Rpl5/uL18 and Rpl11/uL5, prior to its incorporation into preribosomes. In mammals, the 5S RNP is also a central regulator of the homeostasis of the tumor suppressor p53. The nucleolar localization of the 5S RNP and its assembly into preribosomes are performed by a specialized complex composed of Rpf2 and Rrs1 in yeast or Bxdc1 and hRrs1 in humans. Here we report the structural and functional characterization of the Rpf2–Rrs1 complex alone, in complex with the 5S RNA, and within pre-60S ribosomes. We show that the Rpf2–Rrs1 complex contains a specialized 5S RNA E-loop-binding module, contacts the Rpl5 protein, and also contacts the ribosome assembly factor Rsa4 and the 25S RNA. We propose that the Rpf2–Rrs1 complex establishes a network of interactions that guide the incorporation of the 5S RNP in preribosomes in the initial conformation prior to its rotation to form the central protuberance found in the mature large ribosomal subunit.

[*Keywords*: 5S RNP; ribosome assembly; p53; Brix domain]

Supplemental material is available for this article.

Received February 13, 2015; revised version accepted June 2, 2015.

The production of ribosomes is one of the most energy-consuming processes in eukaryotic cells. Ribosome biogenesis is a highly organized and regulated pathway that requires the sequential action of >200 nonribosomal factors (Gamalinda et al. 2014). Ribosome biogenesis starts by the transcription of a pre-RNA precursor in the nucleolus, a specialized non-membrane-bound nuclear compartment. Three of the four ribosomal RNAs (rRNAs)—the 5.8S, 18S, and 25S rRNAs—are cotranscribed as a single 35S precursor by polymerase I. This precursor is cotranscriptionally modified, folded, cleaved, and assembled with both ribosomal proteins and nonribosomal factors to generate the mature ribosomes. During ribosome production, preribosomal particles are exported first to the nucleus and then to the cytoplasm. The final maturation step occurs in a translation-like cycle involving the pre-40S complex and the mature 60S subunit (Lebaron et al. 2012; Strunk et al. 2012). The production of ribosomes has been suggested to be a sequential series of quality control steps that block and recycle nonoptimally assembled preribosomes. Defects in the ribosome assembly pathway have been identified in a number of inherited hematopoietic disorders, collectively called ribosomopathies, which have been linked to progression into cancer (Armistead et al. 2009; Narla and Ebert 2011; Sondalle and Baserga 2014).

In contrast to the 5.8S, 18S, and 25/28S rRNAs, the maturation of 5S RNA follows a totally different pathway before incorporation into preribosomes. The 5S rRNA is transcribed by a different polymerase (RNA polymerase III) at a locus not associated with nucleoli in metazoans or located in proximity to the nucleolus in yeast and has to be imported into the nucleolus for integration in preribosomal particles. A significant fraction of the 5S rRNA in the cell is not associated with ribosomes (Knight and Darnell 1967; Sloan et al. 2013) and has been shown to interact with several proteins (for review, see Ciganda and Williams 2011). The La protein binds immature 5S RNA containing additional uridines at the 3′ terminus. The transcription factor TFIIIA, which binds the 5S rDNA and activates its transcription, also binds 5S RNA to form a 7S ribonucleoprotein particle (RNP) in several organisms (for review, see Layat et al. 2013). In *Xenopus*

[5]These authors contributed equally to this work.
Corresponding authors: nicolas.leulliot@parisdescartes.fr, simon.lebaron@parisdescartes.fr
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.260349.115.

oocytes, the 7S RNP is found in the cytoplasm; TFIIIA protects the RNA from degradation and "stores" the 5S RNA for later assembly into ribosomes.

The 5S rRNA associates with the Rpl5 and Rpl11 proteins (also referred to as uL18 and uL5), the main 5S rRNA-interacting proteins in mature ribosomes, prior to its incorporation into ribosomes (Steitz et al. 1988; Zhang et al. 2007; Sloan et al. 2013). Rpl5 and Rpl11 are imported and probably assembled on the 5S RNA by the specialized importin Syo1 (Calviño et al. 2015). The 5S RNP has recently been found to be a major actor in the p53–MDM2 regulation pathway in humans. MDM2 is an E3 ubiquitin ligase that ubiquitinates p53 and targets it for proteasomal degradation. The 5S RNP binds to and inhibits MDM2, resulting in the stabilization and activation of p53. This pathway regulates energy metabolism, senses nucleolar stress such as impairment of ribosome biogenesis, and is an essential control for oncogene overexpression (Donati et al. 2013; Sloan et al. 2013; Liu et al. 2014). The regulation of the 5S RNP/p53 pathway plays a major role in ribosomopathies, such as Diamond Blackfan anemia, 5q syndrome, and Treacher Collins syndrome, which are believed to be caused by the misregulation of p53. As a master regulator of the tumor suppressor p53 and ribosome biogenesis, the 5S RNP pathway is therefore a promising therapeutic target for the treatment of both cancer and ribosomopathies (Miliani de Marval and Zhang 2011; Sasaki et al. 2011).

Assembly of the 5S RNP into preribosomes requires two nonribosomal proteins: Rpf2 (Bxdc1 in humans) and Rrs1 (Zhang et al. 2007). These factors form a binary complex and interact with the 5S RNP to form the complex that is incorporated into preribosomes. Rpf2/Bxdc1 contains a Brix domain that is found in a family of proteins involved in ribosome biogenesis that includes Rpf1, Imp4, Brix1, and Ssf1/Ssf2 (Wehner and Baserga 2002). The Brix domain is predicted to be an RNA-binding motif, although no specific RNA target has been determined so far for any of the family members. The structure and function of Rrs1 are unknown. Depletion of Rpf2/Bxdc1 or Rrs1/hRrs1 leads to a processing defect of the large ribosomal subunit rRNAs, mislocalization of Rpl5/Rpl11, impairment of 5S RNP recruitment into preribosomes, and impairment of nuclear export of preribosomes (Zhang et al. 2007; Donati et al. 2013; Sloan et al. 2013). The cryo-electron microscopy (cryo-EM) structure of late pre-60S particles has shown that the central protuberance (CP), which contains the 5S RNP, is structurally very different from that found in the mature ribosome (Leidig et al. 2014). In the preribosome, the 5S RNP is docked to helix 84 in the 25S rRNA but is rotated by ~180° relative to its position in the mature 60S complex. The conformational state of the 5S RNP in the preribosome is locked by the ribosome biogenesis factor Rsa4, which interacts with Rpl5/uL18 and the 25S rRNA. Remodeling of the CP to generate the conformation present in the mature 60S subunit and removal of Rsa4 is performed by the dynein-related AAA ATPase Rea1 (Baßler et al. 2014) and is coupled to the action of the Nug2 GTPase to generate an export-competent particle (Matsuo et al. 2014).

In this study, we set out to elucidate the function of the Rpf2–Rrs1 complex in the incorporation of the 5S RNP into preribosomes. We solved the structure of the Rpf2–Rrs1 complex alone, bound to 5S RNA, and within pre-60S particles using X-ray crystallography and small angle X-ray scattering in solution (SAXS) and fitting these structures to pre-60S cryo-EM envelopes. These structural studies, complemented by in vivo and in vitro study of RNA and protein interactions, enable us to propose a model for the function of Rpf2–Rrs1 in ribosome biogenesis.

Results

Structure solution of the Rpf2–Rrs1 complex

The full-length yeast Rpf2–Rrs1 proteins were expressed in a bacterial system using a polycistronic construct. A two-step purification procedure using nickel affinity purification (Rpf2 was expressed with an N-terminal histidine tag) followed by size exclusion chromatography yielded a sample purified to homogeneity that was suitable for structural studies (Supplemental Fig. S1). However, no crystals were obtained using the full-length Rpf2–Rrs1 proteins. This was likely due to the presence of unstructured regions in the N-terminal and C-terminal regions of both proteins. Limited trypsin proteolysis identified a stable structural core comprising both Rpf2 and Rrs1 (Supplemental Fig. S1). This sample crystallized under several of the tested conditions without any further purification. The structure was phased at 3.4 Å resolution by single isomorphous replacement with anomalous scattering using a platinum derivative and was further refined to 2.4 Å resolution in a nonisomorphous native data set. Data collection and refinement statistics are reported in Table 1.

The Rpf2 and Rrs1 proteins are present in a 1:1 complex (Fig. 1) in agreement with the elution volume from size exclusion chromatography and molecular weight derived from SAXS data (see later). The final model includes the conserved domains identified by sequence analysis for both Rpf2 and Rrs1, and the residues observed are consistent with the fragments identified by mass spectrometry analysis of the proteolyzed sample. The crystallized Rpf2 fragment is a single polypeptide chain spanning residues 23–252, with missing electron density for the β_9–β_{10} loop (residues 212–216), englobing the predicted Brix domain (32–237) (Supplemental Fig. S1). The Rrs1 fragment contained residues 9–106, with missing density for the β_4–β_5 loop (residues 74–84) (Supplemental Fig. S1).

The Rpf2 structure shows the same overall α–β–α–β sandwich fold as the archaeal Brix domain protein Mth680 (9% sequence identity) (Ng et al. 2005) and superposes with 3 Å root mean square deviation over 180 residues. As in the archaeal protein, the structure is composed of an extended U-shaped β sheet forming a taco shell-shaped structure (Fig. 1). This sheet forms two layers: The first β layer (L1) (light blue in Fig. 1) is formed by the β_2–β_1–β_3–β_4–β_5–β_{10b} strands, and the second β layer (L2) (dark blue in Fig. 1) is formed by the β_6–β_8–β_9–β_{10a} strands. β_{10} forms an extended strand that spans and

Table 1. *X-ray diffraction statistics*

	Native	PtCl6	PtCl4
Wavelength	0.934 Å	1.0716 Å	1.0716 Å
Resolution	45.51–2.40	45.51–3.40	48.72–4.01
Space group	P 43 21 2	P 43 21 2	P 43 21 2
Unit cell	72.28 Å, 72.28 Å, 175.71 Å	72.28 Å, 72.28 Å, 175.71 Å	72.28 Å, 72.28 Å, 175.71 Å
Total reflections	217,634 (20,401)	49,898 (4494)	32,462 (2861)
Unique reflections	18,921 (1811)	6270 (585)	4141 (363)
Multiplicity	11.5 (11.3)	8.0 (7.7)	7.8 (7.9)
Completeness	99.72% (97.52%)	92.83% (90.56%)	98.85% (92.84%)
Mean I/σ(I)	17.82 (1.92)	19.19 (4.01)	21.51 (4.69)
Solvent content	52.80%		
R-merge	0.098(1.09)	0.084 (0.57)	0.097 (0.60)
CC 1/2	0.999 (0.680)	0.998 (0.776)	0.999 (0.936)
CC*	1 (0.9)	0.999 (0.935)	1 (0.983)
R-work	0.19 (0.28)		
R-free	0.23 (0.35)		
RMS (bonds)	0.009		
RMS (angles)	1.17°		
Ramachandran favored	97%		
Ramachandran allowed	3%		

connects the two layers. The central α-helical layer filling the taco shell is composed of the α_1 and α_3 helices inserted between the first and the second strand in each β layer (β_1–β_2 and β_6–β_8). The α_2 helix inserted between the β_2–β_3 strands packs on the outward face of β layer L1. No structural homologs could be identified for Rrs1.

Rpf2 and Rrs1 form an intimate complex

The Rpf2–Rrs1 complex forms an unexpected and surprising structure where the structural elements of Rrs1 extend the Rpf2 β sheets. Rrs1 is composed of a long α helix ($\alpha_{1'}$) at the N terminus and four β strands ($\beta_{1'}$ to $\beta_{4'}$). The $\beta_{1'}$ and $\beta_{2'}$ strands extend the Rpf2 β layer L2 after the Rpf2 β_6 strand (Fig. 1C). The Rrs1 $\alpha_{1'}$ helix packs on the exterior of this extended β layer L2 (Fig. 1B). The Rrs1 $\beta_{3'}$ and $\beta_{4'}$ strands form an additional β sheet with the Rpf2 β_7 strand and pack on the C terminus of the $\alpha_{1'}$ helix.

The structure of the archaeal protein Mth680 revealed that Brix domains harbor an internal duplication of an α–β layer related by a twofold symmetry, indicating that the Brix family proteins could have arisen from a gene duplication and fusion event (Ng et al. 2005). This duplicated unit, hereafter called the Brix internal duplicated (BID) domain, corresponds to the two halves of the taco shell (colored light and dark blue in Fig. 1). The BID domain is structurally homologous to the anti-codon loop-binding domain of class IIa aminoacyl-tRNA synthetases (aaRSs) (Ng et al. 2005). A remarkable feature of the Rpf2–Rrs1 complex is that Rrs1 completes the structure of the Rpf2 BID domain. Indeed, superposition of the two β layers reveals that the Rrs1 $\beta_{2'}$ strand in β layer L2 mirrors the position of Rpf2 β_2 in β layer L1 (i.e., the β_2–β_1–β_3–β_4–β_5 layer superposes with the $\beta_{2'}$–β_6–β_8–β_9–β_{10a} strands) (Fig. 1D). In contrast, in the Mth680 protein, the β_7 strand completes the L2 β layer before β_6, while the Rpf2 β_7 is looped out and forms the extra β sheet with the Rrs1 $\beta_{3'}$ and $\beta_{4'}$ strands (Fig. 1D). The structural change in the topology of the β layer L2 could be induced by Rrs1 binding.

In addition, Rrs1 extends the BID domain by providing an extra α-helical layer. Superposition of the two β layers in Rpf2 and Mth680 shows that the α_1 and α_3 helices composing the central α layer in the taco shell (Fig. 1D) are structural equivalents: They superpose in the same region and are inserted between equivalent β strands. However, in the CTDs of both Mth680 and Rpf2, there is no equivalent α helix to α_2, which packs on the exterior surface of β layer L1. In the Rpf2–Rrs1 complex, the Rrs1 $\alpha_{1'}$ helix packs on the exterior face of β layer L2 in the same position and orientation as α_2 (Fig. 1D). Moreover, $\alpha_{1'}$ follows the $\beta_{2'}$ strand in L2, equivalent to β_2 in L1. This structural and topological analogy defines $\alpha_{1'}$ as a structural equivalent to the α_2 helix and extends the duplicated BID domain to the entire α–β–α layer. Recently, the crystal structure of the Rpf2–Rrs1 complex from *Aspergillus nidulans* has been reported (Asano et al. 2015). Despite being from two different organisms, the structures obtained are very similar. Interestingly, the same portion of the proteins was stabilized after proteolysis treatment, and the overall structural elements are conserved.

Rpf2 is a 5S rRNA-specific RNA-binding protein

The Brix domain proteins have been suggested to bind RNA, but their cognate RNA-binding sites have not been determined. The electrostatic potential mapped to the surface of the Rpf2–Rrs1 complex shows an extended basic surface along one side of the complex, principally centered within the Rpf2 protein that encompasses the α_1 and α_2 helices; the β_1–α_1, β_3–β_4, and β_9–β_{10} loops; and the β_5–β_6 linker between the two domains (Supplemental Fig. S2). This basic patch is likely the RNA-binding surface of Rpf2 and corresponds, on the first BID, to the

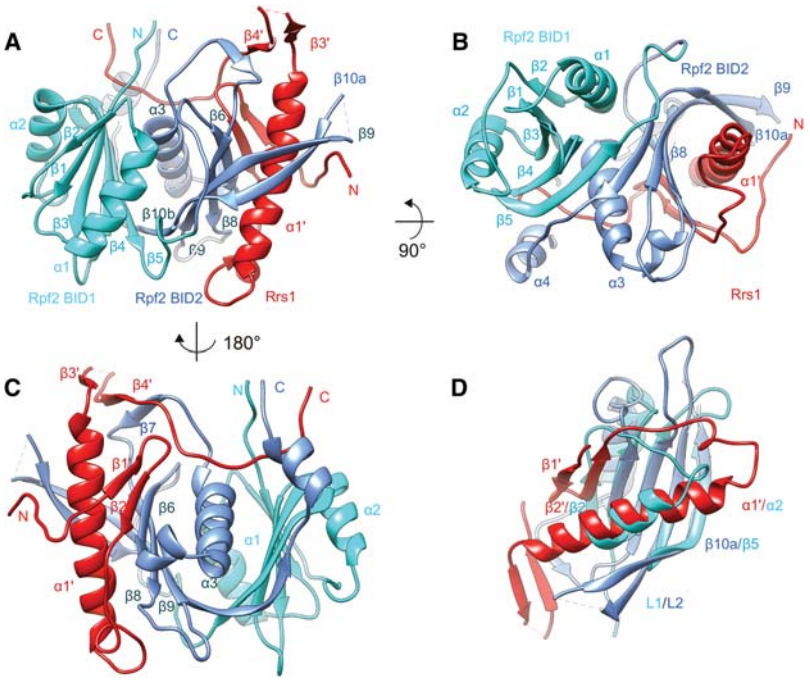

Figure 1. Structure of the Rpf2–Rrs1 complex. (*A–C*) Ribbon representation of three orthogonal views of the Rpf2–Rrs1 complex. The Brix internal duplicated (BID) domains are represented in light and dark blue, and Rrs1 is shown in red. (*D*) Superposition of the two Rpf2 BID domains shows that Rrs1 completes the structural elements in the second BID domain.

RNA binding surface of the homologous aaRS domain (Ng et al. 2005).

To identify Rpf2 and Rrs1 substrates in the cell, we performed genome-wide in vivo identification of the RNA-binding sites of Rpf2 and Rrs1 using the cross-linking and cDNA (CRAC) analysis methodology. Briefly, HTP-tagged Rpf2 or Rrs1 was cross-linked to RNA using UV irradiation, the protein-bound RNA was purified using denaturating conditions using the protein affinity tag and trimmed to the region protected by the protein using a cocktail of RNAses, cDNAs were generated, and the sequences were revealed by high-throughput sequencing (Granneman et al. 2009). During the process, the cross-linked RNAs were labeled with ^{32}P and analyzed by autoradiography. No enrichment for RNAs associated with Rrs1 was seen compared with the control. The majority of RNA fragments bound to Rpf2 correspond to the 5S RNA, as was expected, since Rpf2 is involved in 5S RNP recruitment into the ribosome (Fig. 2A,B). In addition to the 5S rRNA fragments, sequences corresponding to two neighboring regions of 25S rRNA were found bound to Rpf2 (Fig. 2C,D). Interestingly, these two regions flank helix 84—the binding site of both Rpl5 and Rpl11 and the main docking site for the 5S RNP in the pre-60S subunit.

We next set out to confirm the interaction of the recombinant Rpf2–Rrs1 complex with the 5S RNA in vitro. The 5S RNA is composed of a three-way junction containing five helical regions (I–V) and five loops (A–E) with coaxial packing of helix I onto helix II (Fig. 2C). As these types of structures can be hard to fold in vitro, we purified the 5S RNA from yeast to obtain RNA samples containing both post-transcriptional modifications and native-like structures (as evidenced form SAXS) (data not shown). The interaction of the Rpf2–Rrs1 complex with the 5S rRNA was performed by both electrophoretic mobility shift assays (EMSAs; using radioactive-labeled RNA to determine the strength of the interaction) and circular dichroism (CD) spectroscopy to assess the conformational changes in RNA and protein components upon complex formation. We found that the Rpf2–Rrs1 complex efficiently bound to the 5S rRNA with a K_d of ~57 nM (Fig. 3A,B), and CD spectroscopy showed that the interaction involves a structural rearrangement of the RNA (Fig. 3C). In agreement with the in vivo data, no change in CD spectra was observed when Rrs1 alone was used (Supplemental Fig. S3C), indicating that Rrs1 does not contact this RNA directly. Rpf2 alone is very unstable but was still able to induce a small change in the RNA CD spectra (Supplemental Fig. S3B). We conclude that the contacts to the 5S rRNA principally involve Rpf2 and that, at least in vitro, Rpf2 is stabilized by the presence of Rrs1.

Assembly of the Rpf2–Rrs1 complex onto the 5S rRNA

Since the Rpf2–Rrs1 complex can interact with the 5S rRNA in vitro, we performed structural studies using SAXS to study the assembly of the complex in solution (data collection and modeling statistics are shown in Table 2). We first recorded SAXS data for both the full-length and the proteolyzed Rpf2–Rrs1 complex. The data clearly show that the proteolyzed complex is globular, and modeling confirms that the solution structure is the same structure as in the crystal (Fig. 4A). Analysis of the full-length complex revealed that the C-terminal regions of Rpf2 and Rrs1 form extended and unstructured extensions that protrude away from the structural core (data not shown). Because of the difficulty in modeling a protein–RNA complex with these unstructured regions, we performed EMSA

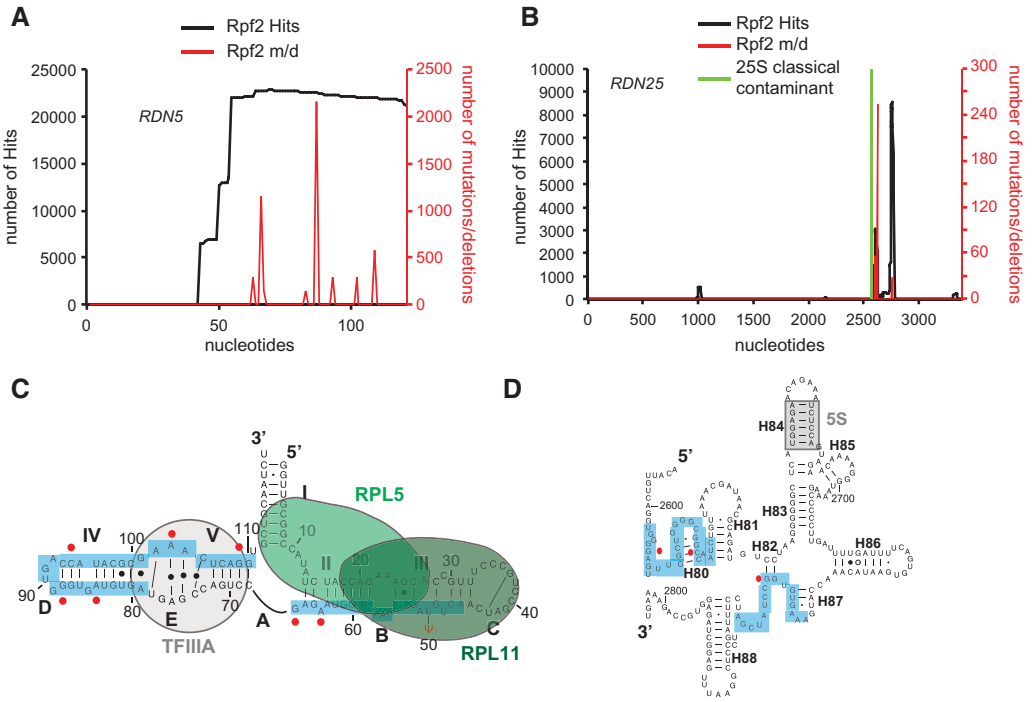

Figure 2. In vivo RNA-binding sites of Rpf2. (*A*,*B*) RNAs in Rpf2-HTP cells were UV cross-linked in cells growing in culture medium, trimmed and ligated to linkers, amplified by RT–PCR, and sequenced with an ionTorrent. Sequences were aligned with main hits to the 5S rRNA (nucleotides 1–120) (*A*) and the 25S rRNA (nucleotides 1–3396) (*B*). The frequency of recovery (hits per 100,000 mapped reads) is plotted for each individual nucleotide (shown in black). The locations of mutations/deletions that are likely due to RNA cross-linking to the residue are shown in red. The location of a classical contaminant sequence found in the 3′ end of the 25S rRNA (recovered with the "no tag" control experiment) is represented by a green bar. (*C*) Secondary structure of 5S RNA in yeast. The binding sites recovered for Rpf2-HTP are indicated on the sequence (blue). Mutated nucleotides that indicate a direct cross-link are indicated by red dots alongside the sequence. Known binding sites for 5S-binding protein TFIIIA, Rpl5, and Rpl11 are indicated as gray, light-green, and dark-green circles, respectively. (*D*) Secondary structure of the 25S RNA in yeast. The binding sites recovered for Rpf2-HTP are indicated on the sequence (blue). Mutated nucleotides are indicated by red dots alongside the sequence. Transient interaction between 5S and 25S in pre-60S particle is indicated as a gray square (Leidig et al. 2014).

using the trypsin proteolyzed samples that we used for the crystallization of the Rpf2–Rrs1 complex to determine whether it could be used for structural studies using RNA. Surprisingly, we found that while the proteolyzed sample still binds 5S RNA, it does so with eightfold less affinity compared with the complex containing the full-length proteins (Fig. 3A,B). The conformational changes in the RNA observed by CD were also less pronounced with the proteolyzed sample (Fig. 3D). Since Rrs1 does not bind RNA, we attribute the higher affinity of the full-length complex to the C terminus of Rpf2. These residues are not part of the consensus core Brix domain but are conserved in Rpf2 homologs. We conclude that in our in vitro assays, the nonconserved extensions of Rpf2 and Rrs1 likely refold when bound to the RNA and contribute to the binding affinity.

We further characterized the 5S rRNA–Rpf2–Rrs1 complex interaction using the proteolyzed Rpf2–Rrs1 complex and yeast 5S rRNA. Interestingly, the SAXS-derived molecular envelope of the free 5S RNA clearly shows an unstacking of the RNA helices compared with the ribosome-bound structure (Fig. 4C). We interpreted and modeled the extended conformation of the RNA as the unbending and unstacking of helices I, II, and III in the absence of the Rpl11 and Rpl5 proteins. In the 5S rRNA–Rpf2–Rrs1 complex, inspection of the molecular envelopes for the RNA and protein phases revealed that the RNA adopts an extended conformation and that the protein complex binds within the first half of the RNA length around 5S RNA helix I (Fig. 4C). The modeling of the complex was performed using the ribosome-bound Rps2–Rrs1 structure (see the next section).

Structure of the preribosome-bound Rpf2–Rrs1 complex

Since the Rpf2–Rrs1 complex recruits the 5S RNP to the early preribosome complexes and is still present in late pre-60S particles, we reasoned that it should be present in the cryo-EM structures of Tap-Alb1-purified pre-60S particles (Leidig et al. 2014). Indeed, the presence of Rpf2 in these preribosomes was confirmed by mass spectrometry, and Rrs1 was probably not identified because of overlap with ribosomal protein bands. These particles contain ~20 ribosomal assembly factors and the CP/5S RNP rotated 180° in respect to its final position in mature ribosomes (Supplemental Fig. S5; Leidig et al. 2014). We identified an

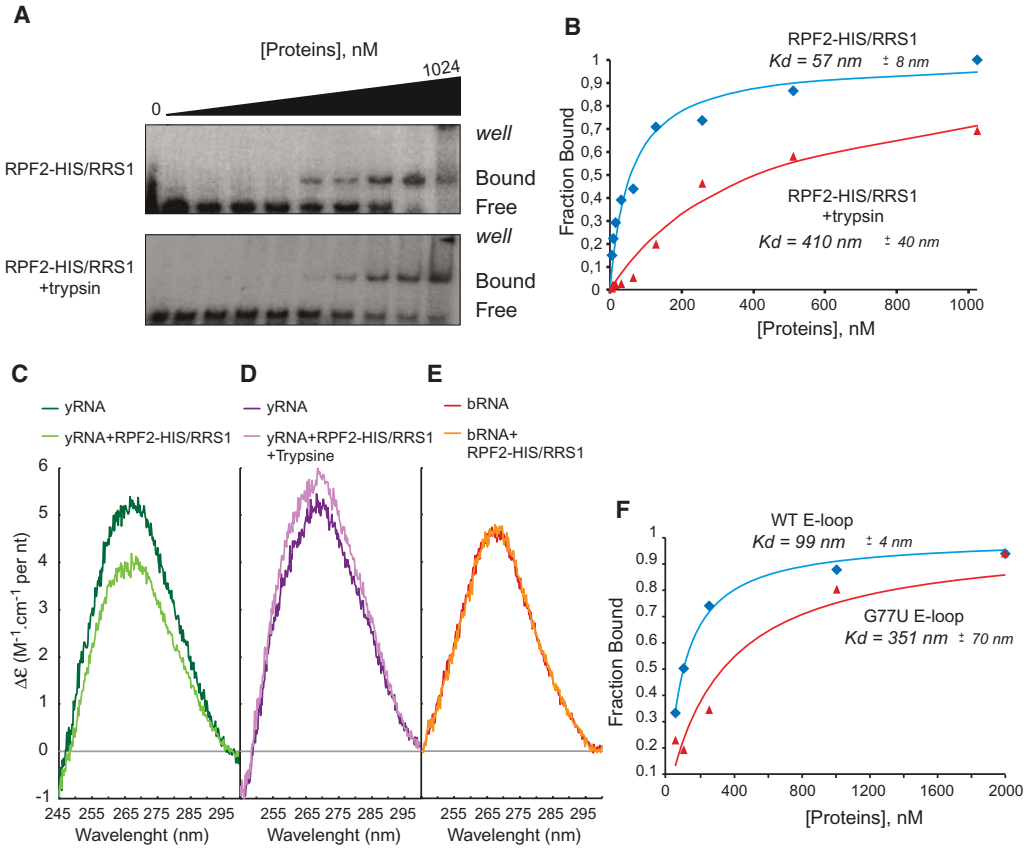

Figure 3. In vitro interaction of the Rpf2–Rrs1 complex with 5S rRNA. (*A*) EMSA of the 5S rRNA. 5′-^{32}P-labeled 5S rRNA was bound with 0, 4, 8, 16, 32, 64, 128, 256, 512, and 1024 nM indicated proteins. (*B*) Fractions of bound 5S rRNA for each protein concentration were quantified using PhosphoImager and plotted for the Rpf2/Rrs1 complex before (blue diamonds) or after (red triangles) 1 h of trypsin proteolysis. The theoretical curves are represented for the Rpf2/Rrs1 complex before (blue) and after (red) trypsin treatment. (*C–E*) CD analysis of the RNA conformational changes upon Rpf2/Rrs1 binding. (*C*) CD spectra of the yeast 5S rRNA (yRNA) obtained when the protein complex Rpf2/Rrs1 and the yRNA are present in two separate cuvette compartments (dark green) or after mixing the two compartments (light green). (*D*) The same experiment as in *C* but with the protein complex Rpf2/Rrs1 after trypsin treatment. Spectrum were recorded before (dark purple) and after (light purple) mixing proteins and yRNA. (*E*) The same experiment as in *C* with bacterial 5S rRNA (bRNA) and the complex Rpf2/Rrs1 before (red) or after (orange) mixing the two compartments. (*F*) Filter-binding assay of the 5S rRNA E loop to the RPF2/RRS1 complex. The fluorescently labeled E loop of wild type or the G77U mutant was bound with 0, 25, 50, 100, 250, 1000, and 2000 nM protein complex. Fractions of the bound E loop of wild type or the G77U mutant of the 5S rRNA for each protein concentration were quantified using odyssey (Li-COR) and are plotted for the wild-type E loop (blue diamonds) and G77U mutant (red triangles). The theoretical curves are represented for the wild type(blue) and G77U mutant (red).

unattributed density in the pre-60S structure contacting the 5S RNA helices I and V (Fig. 5A) in a region compatible with both the CRAC and SAXS data. We were able to fit the Rpf2–Rrs1 complex into this density (Fig. 5B). We then refined the structures of the neighboring molecules against the cryo-EM density, which provided us with a molecular model of Rpf2–Rrs1 in complex with the 5S rRNA, ribosomal proteins, and preribosomal assembly factors within the pre-60S ribosome (Fig. 5B,C). The high resolution obtained in these cryo-EM studies (8.7 Å) gives us a model of the Rpf2–Rrs1–preribosome structure in which the overall binding interfaces are unambiguously defined, but the detailed interaction at the residue level will require higher-resolution structures.

In this model, the Rpf2–Rrs1 complex contacts the 5S rRNA at the junction of the helices I and IV (Fig. 5B), a region distinct from the Rpl5- and Rpl11-binding regions (Figs. 2C, 5B). The structure shows that the four proteins Rpf2, Rrs1, Rpl5, and Rpl11 can coexist in the same complex with 5S RNA. Furthermore, in this complex, Rpl5 and Rpl11 are in the same conformation (relative to the 5S rRNA) as found in mature ribosomes. A model of the ribosome-bound 5S–Rpf2–Rrs1 complex was used in refinement against SAXS data, and we found a good fit to the scattering data and in the molecular envelope with a conformation in which the Rpf2–Rrs1 complex interaction is conserved, and helices II and III form an extended helix (Fig. 4D).

In summary, the preribosome-bound and solution structure of the 5S–Rpf2–Rrs1 complex shows that Rpf2–Rrs1 can bind to the 5S rRNA in the absence of Rpl5 and/or Rpl11 and in the same overall conformation as the

Table 2. *SAXS data collection and modelling statistics*

	5S RNA	Rpf2–Rrs1 trypsinized	5S–Rpf2–Rrs1 trypsinized
Data collection parameters			
Instrument	SWING	BM29	BM29
Beam geometry	0.4 mm × 0.1 mm	0.5 mm × 0.5 mm	0.5 mm × 0.5 mm
Wavelength	1.03 Å	1.00 Å	1.0 Å
q range	0.007–0.5 $Å^{-1}$	0.0025–0.5 $Å^{-1}$	0.0025–0.5 $Å^{-1}$
Exposure time/nb frames	1 sec/100	1 sec/10	1 sec/10
Concentration range	10 mg/mL	5–25 mg/mL	5–25 mg/mL
Temperature	288	288	288
Structural parameters			
I(0) [from P(r)]	0.02 cm^{-1}	0.06 cm^{-1}	0.17 cm^{-1}
Rg [from P(r)]	37.8 Å	22.49 Å	38.06 Å
I(0) (from Guinier)	0.02 cm^{-1}	0.06 cm^{-1}	0.17 cm^{-1}
Rg (from Guinier)	35.2 Å ± 2.3 Å	22.89 Å ± 1.23 Å	36.84 Å ± 0.83 Å
D_{max}	122.3 Å	71.5 Å	128.9 Å
Porod estimate	62,902 $Å^3$	63,190 $Å^3$	104,848 $Å^3$
Molecular mass determination			
Partial specific volume	0.54 cm^3/g	0.745 cm^3/g	0.639 cm^3/g
Contrast ($\Delta\rho \times 10^{10}$ cm^{-2})	6.283	2.736	4.510
Molecular mass M_r [from I(0)]	39,314	39,494	75,530
Calculated monomeric M_r from sequence	39,155	37,558	76,713
Data processing			
Primary data reduction	FOXTROT	EDNA–SAS pipeline	EDNA–SAS pipeline
Data processing	PRIMUS	PRIMUS	PRIMUS
Ab initio analysis	DAMMIF	DAMMIF	MONSA
Number of models	50	50	20
Model χ^2	2.196 ± 0.020	0.855 ± 0.050	2.38 ± 0.068
Validation and averaging	DAMAVER	DAMAVER	DAMAVER
Normalized spatial discrepancies	0.786 ± 0.145	0.978 ± 0.169	0.629 ± 0.112
Rigid body modeling	MC-SYM	DADIMODO	SASREF
Computation of model intensities	CRYSOL	CRYSOL	CRYSOL
Model χ^2	1.866	1.289	1.677

ribosome-bound conformation for helices I and V in the 5S rRNA but that helices II and III require the interaction with Rpl5 and Rpl11 in order to attain their mature conformation. Our data are consistent with the model in which the 5S RNP–Rpf2–Rrs1 complex is incorporated as an independent structural module in preribosomes.

Rpf2 recognizes a eukaryote-specific E loop in the 5S rRNA

Close analysis of the preribosome-bound Rpf2–Rrs1 complex reveals that the main contact to 5S RNA is with the E loop between helices IV and V (Fig. 2C). While all 5S rRNA species contain an E loop, the sequence is significantly different between eukaryotes and prokaryotes. The E loop forms a G-bulged cross-strand stack structure consisting of three stacked non-Watson-Crick base pairs, a bulged-out guanine (G77) that forms a base triple with A78 and U102, and two cross-strand purine stacking interactions. Loop E is therefore a well-folded and rigid structure (Lee et al. 2006) that provides structural features that could constitute a specific protein-binding site: an unusual backbone geometry, a widened major groove, unique hydrogen-bonding possibilities to the non-Watson-Crick base pairs, and a bulged-out guanine. To confirm that the E loop is important for the specific recognition of the 5S RNA by Rpf2, we performed in vitro binding assays with bacterial-purified 5S rRNA (Fig. 3E), which contains the same overall structure as the eukaryote rRNA but a different E loop. No interaction between Rpf2–Rrs1 and bacterial 5S rRNA was observed by gel shift assays (data not shown) or CD spectra (Fig. 3E), indicating that the eukaryotic E loop is the specific RNA-binding element for Rpf2. To confirm this result, we analyzed the interaction between the Rpf2–Rrs1 complex and a labeled RNA probe containing only the E loop and its flanking helices (Supplemental Fig. S3E) using a filter-binding assay. The Rpf2/Rrs1 complex was able to bind this construct with a K_d of 99 nM, and the substitution of the looped-out guanine residue (G77) (Fig. 6A) for a uridine leads to a threefold decrease of the binding affinity (Fig. 3F; Supplemental Fig. S3D).

Rpf2 contacts the 5S rRNA along the previously observed conserved basic patch (Supplemental Fig. S2). The contacts to loop E in the major groove involve the β_1–α_1 and β_3–β_4 loops. The β_2–α_2 loop also contacts the loop E 5′ strand backbone and the minor groove of helix IV. The β_3–β_4 loop contains conserved and charged amino acids that bind the RNA backbone on both sides of the loop E and certainly contribute to the specific recognition of the loop E structure. It also places residues of the β_3–β_4 loop within hydrogen-bonding distance of the bulged guanine (Fig. 6A). All of the protein interactions with the

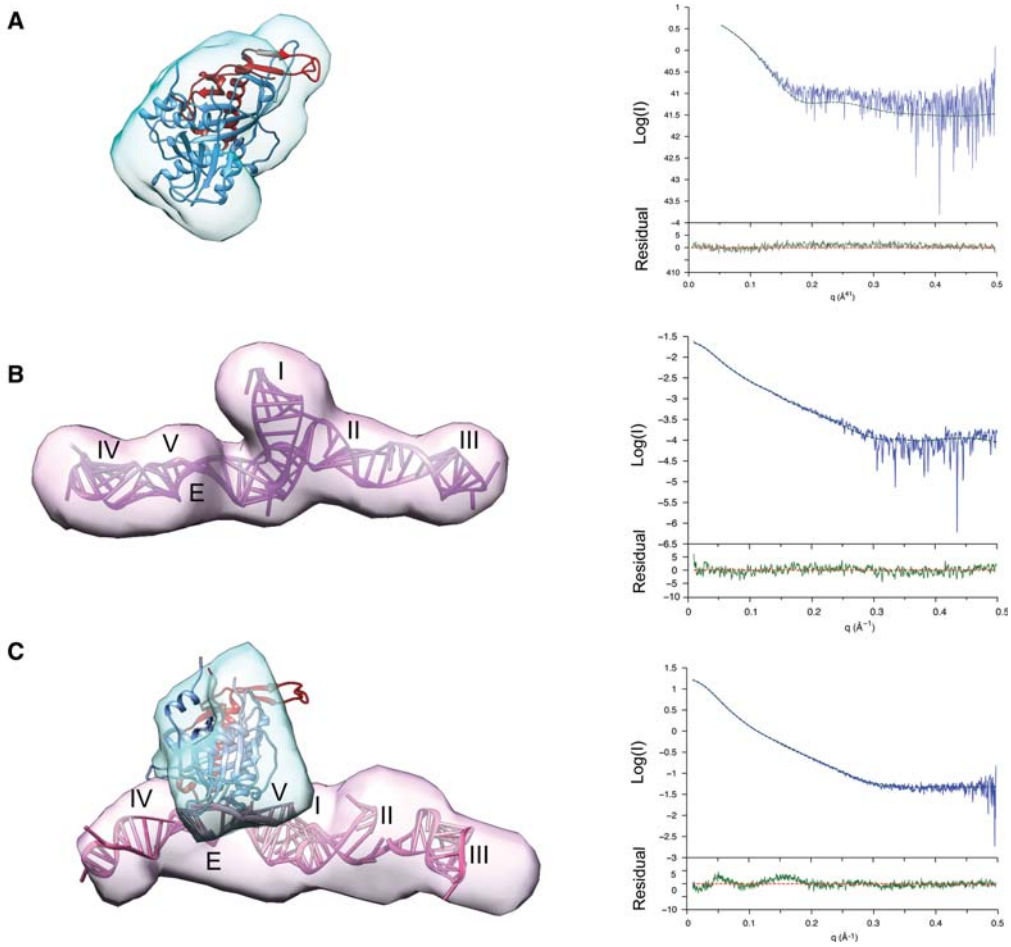

Figure 4. SAXS-derived solution structures of the Rpf2 subcomplexes. Solution structure model of the Rpf2–Rrs1 complex in the proteolyzed state (χ^2 = 1.289) (*A*), the full-length Rpf2–Rrs1 complex (χ^2 = 1.701) (*B*), the free 5S RNA (χ^2 = 1.866) (*C*), and the Rpf2–Rrs1–5S complex with proteolyzed proteins (χ^2 = 1.677). The structures are superposed to a representative envelope calculated by DAMMIN (*A*,*B*) or MONSA (*C*). The proteins and RNA are represented in the same orientation. The corresponding calculated X-ray scattering curves (dashed green) superposed to the experimental scattering curves (blue) are shown in the *right* panel. The locations of interesting 5S structural elements are indicated by their names to refer to Figure 2C.

E loop lie within the first BID domain of Rpf2. Superposition of the Rfp2–5S RNA and aaRS–synthetase–tRNA complex shows that the proteins recognize very different RNA substrates with different regions of the protein domain: Binding of Rpf2 is primarily to the E-loop major groove, while aaRS binds the anti-codon loop of a stem–loop structure (Supplemental Fig. S4B). The only common RNA-binding element maps to the β_2–α_2 loop contacting the 3′ strand of the loop E in the Rpf2–5S complex.

Loop E-like structures are known to mediate specific protein–RNA interactions, such as the sarcin–ricin loop (SRL) of the 25S rRNA with ricin or EF-G and, more interestingly, TFIIIA to the 5S rRNA E loop (Lu et al. 2003). Superposition of the structures of Rpf2 and TFIIIA bound to the 5S rRNA shows that although the two proteins use different strategies to bind the 5S loop E, the binding surfaces of the two proteins overlap on the RNA. This indicates that binding of TFIIIA and Rpf2 to the 5S rRNA is mutually exclusive (Supplemental Fig. S4A). It has also been reported that Rpl5 displaces TFIIIA on the 5S rRNA due to overlap of their binding sites. These mutually exclusive interactions might ensure the proper sequential interaction of proteins with the 5S rRNA along the maturation pathway. Indeed, TFIIIA could prevent Rpf2 from binding the naked 5S rRNA early in the pathway before the TFIIIA/Rpl5 exchange.

The Rpf2 complex docks the 5S RNP into the pre-60S particles using a network of RNA and protein interactions

The structure of the Rpf2–Rrs1 complex bound to the pre-60S defines its interaction with the 5S rRNA but also suggests additional contacts to the 25S rRNA and neighboring proteins. Analysis of the CRAC data revealed that, in addition to the 5S rRNA, Rpf2 contacts the 25S rRNA around helices 80 and 87. No contacts between the conserved Brix domain structural core and the 25S RNA are

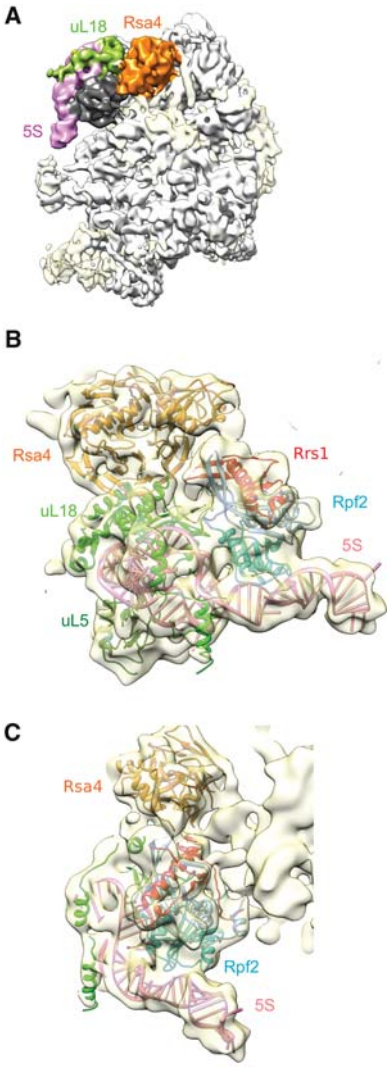

Figure 5. Electron microscopy structure of pre-60S bound Rpf2–Rrs1. (*A*) Cryo-EM structure of the Alb1-TAP purified pre-60S particles (emd 2528) (Leidig et al. 2014). The unidentified density in contact with 5S (pink) and Rsa4 (orange) is shown in dark gray. (*B*, *C*) Orthogonal views of the Rpf2–Rrs1 crystal structure fit in the cryo-EM density. No optimization of the Rpf2–Rrs1 structure was performed.

observable in the pre-60S complex in this region. However, the position of the RNA regions found is compatible with an interaction with the "unstructured" C-terminal region of Rpf2, which was not visible in the crystal structure. Unidentified electron density links the last observed residues of Rpf2 with the 25S rRNA, suggesting that the conserved region following the canonical Brix domain binds this region of the rRNA (Fig. 6B). In the in vitro binding assays, we showed that this region has RNA-binding capabilities because it improves binding of Rpf2 to the 5S rRNA (Fig. 3A–D). We propose that the Rpf2 C-terminal extension contributes to 5S RNP assembly in preribosomes by binding an exposed region of the 25S RNA in proximity to the 5S-binding site.

The importance of the C-terminal domain (CTD) of Rpf2 has been tested by expressing either PTH-tagged versions of full-length (FL) Rpf2, the N-terminal domain (NTD) 1–252 of Rpf2, or the CTD 253–344 of Rpf2 in the presence or absence of the endogenous protein in yeast (Supplemental Fig. S6). As expected, only the full-length protein was able to complement the absence of the endogenous copy (Supplemental Fig. S6A). Interestingly, expression of the NTD mutant exerts a clear dominant-negative effect on growth (Supplemental Fig. S6A). This result supports a model in which the NTD competes with the endogenous copy for 5S RNP binding but fails to progress further in the maturation pathway. This could be due to the impossibility of recruiting the 5S RNP to the ribosome or the inhibition of a subsequent step mediated by the Rpf2 C terminus. In order to distinguish between these two possibilities, we affinity-purified Rpf2-FL, Rpf2-NTD, and Rpf2-CTD and analyzed the bound RNA. We found a 1.6-fold decrease in the ratio of bound 5.8S versus 5S RNA between Rpf2-FL and Rpf2-NTD, indicating that the recruitment of Rpf2 in the ribosome is affected in the absence of Rpf2-CTD (Supplemental Fig. S6C,D). No significant amount of RNA was retrieved with Rpf2-CTD alone, indicating that this region is not

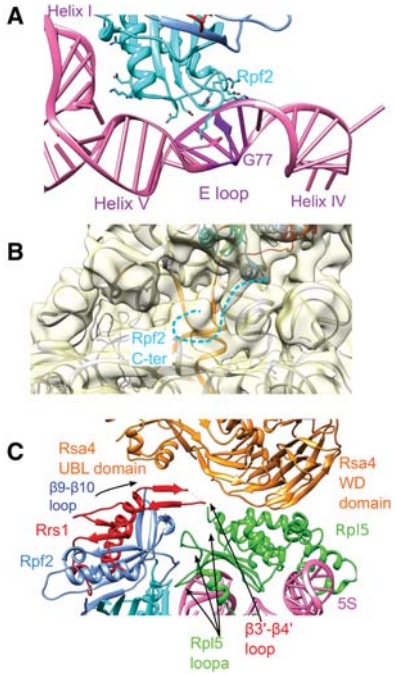

Figure 6. RNA and protein contacts in the preribosomes. (*A*) Rpf2 binds the 5S RNA E loop. The 5S RNA and the Rpf2–Rrs1 complex were fitted simultaneously in the electron density with flexible molecular dynamics using MdFF. The E loop is colored in purple, and the looped-out guanine is depicted in a solid cartoon. (*B*) Rpf2 contacts to the 25S rRNA. The predicted C-terminal extension of Rpf2, not visible in the crystal structure, is shown by a dashed line. The rRNA sequence identified by CRAC is colored in orange. (*C*) Protein–protein interaction with the Rpf2–Rrs1 complex in the preribosome.

sufficient for binding preribosomal particles but, in synergy with the Rpf2-NTD, helps the recruitment of the 5S RNP.

In the pre-60S-bound complex, Rpf2 and Rrs1 are in close proximity to the ribosomal protein Rpl5 and the nonribosomal protein Rsa4. Rpl5/uL18 has already been shown to interact with Rsa4 (Leidig et al. 2014), and interactions have been observed between Rpl5, Rpl11, Rpf2, and Rrs1 (Zhang et al. 2007). The structure of the ribosome-bound Rpf2 complex shows that the Rpf2/Rpl5 interaction involves the Rpl5 eukaryote-specific loops that are also involved in Rsa4 binding (Leidig et al. 2014). This suggests that Rpf2 specifically binds to the 5S rRNA/Rpl5 complex for integration in the preribosomes. In addition, Rrs1 and Rpf2 interact with not only Rpl5 but also the Rsa4 assembly factor. In the pre-60S structure, the Rrs1 $\beta_{3'}$–$\beta_{4'}$ and Rpf2 β_9–β_{10} loops, which are not visible in the crystallographic electron density of the Rpf2–Rrs1 complex, are in proximity to the WD repeat domain and the ubiquitin-like (UBL) domain of Rsa4 (Fig. 6C). These interactions could participate in the network of protein–protein and protein–RNA interactions that anchor the Rpf2–Rrs1–5S RNP complex in the preribosomes containing the Rsa4 assembly factor and the correct local structure of the 25S rRNA.

Discussion

Brix domain proteins serve as structural hubs that bind both RNA and protein

The structure of the Rpf2–Rrs1 complex provides us with a model for the function of Brix domain proteins. The first striking observation is the unexpected nature of the interaction between the two proteins. Rrs1 forms a tight complex with Rpf2 and completes the Rpf2 structure. The absence of a free pool of Bxdc1 and hRrs1 in vivo (Sloan et al. 2013), the poor solubility and stability of Rpf2 in vitro (this study), and the instability of Rpf2 in vivo upon depletion of Rrs1 (Zhang et al. 2007) suggest that the proteins always exist in a complex. It is tempting to speculate that other members of the Brix domain family use the same interaction surface to bind Rrs1-like partner proteins. The Rpf2 residues forming salt bridges and/or hydrogen bonds with Rrs1 are not strictly conserved in the Brix domain family, suggesting that Rrs1 does not bind other Brix domain proteins. However, conserved residues specific to each family lie along the Rrs1-interacting surface, suggesting that they bind other proteins, which could define a Brix domain-associated protein (BAP) family. Members of the Brix family have been shown to form specific complexes with other ribosome assembly factors: Imp4 with Imp3/Mpp10 and Brx1 with Ebp2. It remains to be determined whether these Brix domains use the same binding strategy as Rpf2.

The identification of the Rpf2–Rrs1 complex in the pre-60S structure revealed how this complex binds RNA. All of the protein–RNA contacts are made by Rpf2 in both the Brix domain and the C-terminal extension. On the other hand, Rrs1 does not contact RNA; it faces the exterior of the 5S RNP and is available to interact with other proteins. The data presented here on Rpf2 provide the first evidence that Brix domain proteins can bind specific RNA elements with high affinity. The binding site for Rpf2 in the 5S rRNA, the three-way junction, and loop E structure suggest that Brix domain proteins bind both double-helical and loop/bulge RNA structures and can therefore recognize complex RNA three-dimensional structures. The C-terminal extension also provides long-range RNA-binding capabilities to either recruit the complex to a specific loci or monitor the proper folding of distal RNA structures. In agreement with this hypothesis, the Brix domain of Imp4 is sufficient to bind the U3 snoRNA, but the full-length Imp4 is required to form the duplex between the U3 snoRNA and the pre-rRNA (Gérczei and Correll 2004).

From our data, we conclude that the Brix domain is a docking platform that mediates both RNA and protein contacts. Domain duplication in the Brix domain was followed by function specialization: The first domain provides RNA specificity (Fig. 6A), and the second serves as a protein-binding module (Fig. 6C). The function of Rpf2 is to bring together different regions of the preribosomes by acting as a structural hub for preribosomal proteins, rRNA, and ribosomal proteins. In the case of the Rpf2–Rrs1 complex, the question arises of why the Rpf2 Brix domain binds the Rrs1 BAP, which in turn provides binding specificity to Rsa4. An intriguing possibility is that the ancestral form of the Brix domain was able to bind multiple BAPs that acted as adapters for an RNA substrate structure resembling the E loop and protein interaction sites. In the course of evolution, these complexes specialized to form different Brix domain proteins having different RNA substrate specificities. This is in agreement with the observation that archaeal genomes contain only one Brix domain protein, yeasts contain half a dozen, and metazoan genomes code for >10 members of this family of protein. The differences in the number of Brix domain proteins between archaea, yeast, and metazoans correlate with the increasing complexity of ribosome biogenesis in these organisms. A speculative hypothesis is that, in eukaryotes, a single Brix domain protein could be addressed to different RNAs through protein–protein interactions using different BAPs. This would be compatible with the ribosome-independent roles of several ribosome assembly factors, including Rrs1 (Gambe et al. 2009), Brix protein Ssf1 homolog PeterPan (Bugner et al. 2011), and Imp4 (Furtado et al. 2007).

The Rpf2–Rrs1–5S RNP complex acts as a structural probe for correct ribosome assembly

The ribosome-bound structure of the Rpf2–Rrs1 complex not only defines it as a 5S rRNA binding module but shows that Rpf2 and Rrs1 anchor the 5S rRNA and Rpl5 in a network of interactions with the Rsa4 assembly factor and 25S RNA. This work has enabled us to assign possible function to Rpf2 and Rrs1 during 5S RNP incorporation into the preribosome (Schematic model in Fig. 7). Rpf2

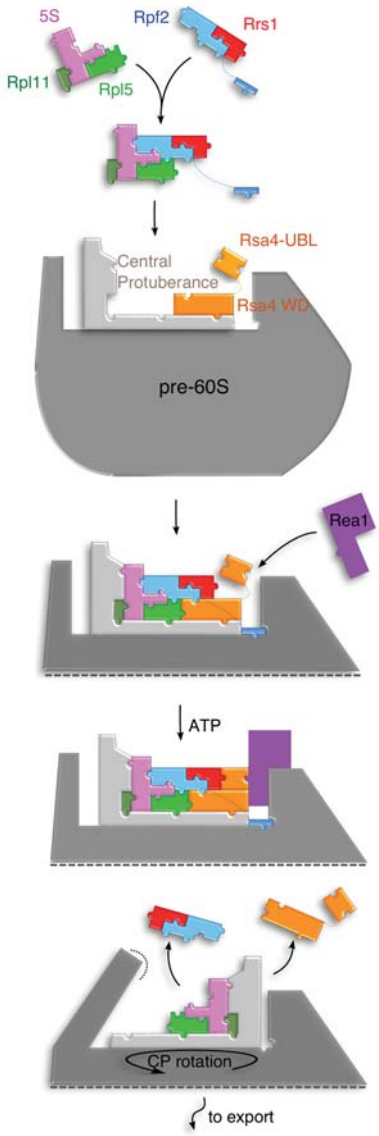

Figure 7. Jigsaw puzzle model for Rpf2/Rrs1 complex function in 5S RNP integration in preribosomes. Strong affinity of the Rpf2/Rrs1 complex for the 5S rRNA suggests that it recruits the free pool of 5S RNP. In this complex, the 5S RNP can only assemble the preribosomes containing the correct conformation. The preassembled Rpf2/Rrs1 targets the 5S RNP onto pre-60S ribosomes using interactions between Rpf2 and the 25S rRNA and between Rpf2/Rrs1/Rpl5 and Rsa4. The Rsa4 UBL domain is positioned by Rrs1 to allow the interaction with Rea1, which recycles Rsa4 and Rpf2/Rrs1 from the pre-60S particle and allows the 5S RNP to adopt its final configuration state through a 180° rotation of the CP.

specifically binds to the 5S rRNA/Rpl5 RNP and helps dock this complex onto the 25S RNA. The 5S assembly pathway prior to this step has been omitted in Figure 7, and it is not known whether Rpf2/Rrs1 can have a role in the assembly of Rpl5/Rpl11 on the 5S RNA. It is also possible that the import factor Syo1 is involved in chaperoning the assembly of Rpl11 on the H84 helix (Calviño et al. 2015). The Rpf2-binding surface on the 25S rRNA identified by CRAC is exposed only in the rotated state of the pre-60S CP and not in the mature 60S particles, which confirms that this state is indeed an on-pathway intermediate. Additional contacts to the Rsa4 factor by Rpl5, Rpf2, and Rrs1 also ensure that 5S RNP is addressed to ribosomes with the rotated CP. This interaction could provide selectivity for recruiting the 5S RNP only to the particles that have the proper conformation (Fig. 7). Indeed, because the 5S rRNA–Rpl5–Rpl11 complex has the same structure as mature ribosome-bound 5S RNP, it could very well associate with malformed preribosomal particles that have bypassed quality control mechanisms, for example, by the premature action of Rea1 to rotate the CP before the 5S RNP and other factors have been assembled. However, since the 5S RNP–Rpf2–Rrs1 complex is not compatible with the position of the 5S in the mature 60S (Supplemental Fig. S5), the association of Rpf2–Rrs1 to 5S RNP in the nucleolus would function in preventing the association with malformed preribosomes. The high affinity of Rpf2–Rrs1 to the 5S RNP and the requirement of Bxdc1/hRrs1 for nucleolar localization of the 5S RNP (Sloan et al. 2013) suggest that it binds the 5S RNP nucleolar pool to control and direct its incorporation in the correct preribosomes. We suggest that the interaction of the Rpf2–Rrs1 complex with the 5S RNP provides a quality control checkpoint by defining a building block that can interact with only specific preribosomal particles (Fig. 7).

It has been hypothesized that Rsa4 relays structural information at various regions of the pre-60S. It is thought that removal of this central "pin" by the Rea1 AAA ATPase rips out the interacting proteins and leads to a major remodeling of the pre-60S particles by rotation of the CP (Baßler et al. 2014). It is not yet known what triggers the action of the Rea1 ATPase. We suggest that the Rpf2–Rrs1 complex, positioned at the interface between Rsa4 and 5S rRNA, contributes to this major rearrangement. Indeed, The Rpf2–Rrs1 complex binds the UBL domain of Rsa4, which is involved in recruiting the Rea1 ATPase (Ulbrich et al. 2009). In the Rsa4 crystal structure, this domain is found in different orientations with respect to the WD repeat domain (Leidig et al. 2014). Interaction with the Rfp2–Rrs1 complex could stabilize an orientation of the UBL domain, which offers an extended interaction surface for Rea1 recruitment and activation. In this model, recruitment of the 5S RNP by the Rpf2–Rrs1 complex would act to both select the correct preribosomes and enable the ribosome biogenesis to proceed to the next step by activation of Rea1.

Recent advance in the understanding of the Mdm2–p53–ribosomal protein revealed that regulation of p53 is mediated by the 5S RNP. The 5S RNP appears to be an important target for anti-cancer and anti-ribosomopathy drugs. In addition, it was shown that depletion of either hRrs1 or Bxdc1 induces the p53 response. We therefore propose that targeting the Rpf2–Rrs1/5S rRNA interaction would both impair ribosome biogenesis and activate p53 by the Mdm2–p53–5S RNP pathway and could constitute a novel therapeutic target for cancer.

Materials and methods

Cloning, expression, and purification

The ORFs of the *RRS1* and *RPF2* genes from *Saccharomyces cerevisiae* were synthesized commercially by Genscript Corp. and inserted into pET21(a^+) (Novagen) as a pET21-Rpf2–Rrs1 polycistronic construct with an N-terminal 6xHis-tagged Rpf2 fusion protein. The expression and purification were essentially as described (Loc'h et al. 2014). Briefly, the proteins were expressed in the Rosetta 2DE3 strain from *Escherichia coli* (Invitrogen) at 37°C in LB medium (Sigma) supplemented with 100 µg/mL ampicillin and 25 µg/mL chloramphenicol until OD_{600} between 0.6 and 0.8. Recombinant protein expression was induced by adding 1 mM isopropyl-β-D-1-thiogalactopyranoside, incubating overnight at 20°C, harvesting by centrifugation, and resuspending in buffer A (50 mM Tris-HCl at pH 8, 500 mM NaCl, 20 mM imidazole) supplemented with complete EDTA-free protease inhibitors (Roche). Cells were lysed by sonication, and lysate was centrifuged at 20,000 rpm for 30 min.

The clear lysate containing the Rrs1–Rpf2 6xHis-tagged complex was loaded onto a 5-mL HisTrap (GE Healthcare) connected to an ÄKTA pure (GE Healthcare). The Rrs1–Rpf2 6xHis-tagged complex was eluted with a linear gradient of imidazole (buffer B, 50 mM Tris-HCl at pH 8, 500 mM NaCl, 500 mM imidazole). Gel filtration was then performed on the eluted fractions containing protein complex using buffer C (50 mM Tris-HCl at pH 8, 150 mM NaCl) on a Superdex 200 26/60 (GE Healthcare).

Crystallization, data collection, and processing

Crystallization trials were performed at 18°C using the hanging drop vapor diffusion technique in 1-µL drops (with a 1:1 protein: precipitant ratio) equilibrated against 500 µL of reservoir solution. The Rrs1–Rpf2 His-tagged complex was first digested for 1 h at 30°C using a bovine trypsin solution at 1/1000 ratio (w/w) and was used in crystallization trials without further purification. Crystals were obtained in 0.2 M LiSO4, 30% (w/v) polyethylene glycol 4000, and 0.1 M TrIs-HCl (pH 8.5) with a complex solution at 15 mg/mL containing trypsin. Crystals were cryoprotected using successive soaking steps in increasing concentrations of ethylene glycol.

X-ray data were tested and optimized at the European Synchrotron Radiation Facility (ESRF) and collected at the Soleil synchrotron on beamlines Proxima1 and Proxima2. For phasing, crystals were soaked in a reservoir solution containing 10 µM potassium tetrachloroplatinate (PtCl4) or hexachloroplatinate (PtCl6), and data were collected at the absorption threshold of platinum (1.0716 Å). Native and derivative data sets were indexed using XDS (Kabsch 2010) and Mosflm (Leslie and Powell 2007). The structure was solved by single isomorphous replacement with anomalous scattering using the PtCl4 derivative. Since native data sets were not isomorphous, the PtCl6 data set, which did not contain anomalous signal, was used as a native data set and phased with SHARP (Bricogne et al. 2003). Experimental phasing and molecular replacement were carried out with the Autosol and Phaser programs from Phenix (Adams et al. 2010). The initial rebuilding was carried out with Buccaneer from the CCP4 program suite (Collaborative Computational Project 1994) and subsequent rebuilding and refinement were carried out with COOT (Emsley and Cowtan 2004) and the Refine module from Phenix. Structures were fitted in the cryo-EM electron density (emdb 2528) using Chimera (Pettersen et al. 2004) and were refined using MdFF (Trabuco et al. 2008).

Yeast strains

The strains used are listed in Supplemental Table S1. For the CRAC approach, the genomic copies of Rpf2 and Rrs1 were C-terminally tagged (Longtine et al. 1998) for expression as a Rpf2-His6-TEV-2ProteinA (HTP) and a Rrs1-HTP fusion in the BY4741 (MATa; *his3Δ1*; *leu2Δ0*; *met15Δ0*; *ura3Δ0*) background. The HTP tags did not detectably affect cell growth. To obtain plasmids expressing PTH-tagged mutants of Rpf2, cDNAs corresponding to each mutant were amplified by PCR and cloned into pRS415-PTH via XmaI restrictions sites.

Yeast RNA isolation

The wild-type BY4741 strain from *S. cerevisiae* was grown at 25° C in 1 L of YPD medium to an OD_{600} of 8. The cell culture was then divided into six pots and harvested by centrifugation. Extracts were prepared from pellets resuspended in 1.5 mL of guanidinium thyocianate–phenol mix (1:1 [v/v]) with 1.5 mL of zirconia beads and vortexed for 5 min. Twelve microliters of guanidinium thyocianate–phenol mix (1:1 [v/v]) was then added, and extracts were incubated for 10 min at 65°C and left for 10 min on ice. Four microliters of 100 mM sodium acetate and 6 mL of chloroform were added, and extracts were centrifuged at 4600 rpm for 30 min. Aqueous phases were removed and subjected to ethanol precipitation. Precipitated RNA were resuspended in buffer A (10 mM Tris-HCl at pH 8, 400 mM NaCl) and loaded onto a 13-mL 15Q chromatography column connected to an ÄKTA pure system (GE Healthcare). RNAs were then eluted with a linear gradient of NaCl (buffer B, 10 mM Tris-HCl at pH 8, 700 mM NaCl). Gel filtration was then performed on the eluted fractions containing the 5S RNA using buffer C (10 mM Tris-HCl at pH 8, 150 mM NaCl) on a Superdex 200 16/60 (GE Healthcare).

Immunoprecipitations

Immunoprecipitations using the ProtA tag on IgG sepharose beads were performed as described before (Lebaron et al. 2005).

EMSAs

5S RNA was purified from yeast, dephosphorylated, labeled with ^{32}P at the 5′ end, and purified on G50 column. Labeled RNA (0.1 nM) was incubated with protein in a 10-µL reaction containing 10 mM Tris-HCl (pH 7.6), 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 1 mM EDTA, and 5% glycerol for 10 min at room temperature. The reactions were resolved in 5% native polyacrylamide gels running in 0.5× TBE (pH 7.8) buffer at 4°C. The gels were dried and autoradiographed using a BAS-5000 PhosphorImager (Fugi).

CD

CD spectra were recorded using a Jobin-Yvon Mark VI circular dichrograph at a scan speed of 0.2 nm/sec. Quartz spare split-compartment cuvettes with a 0.437-cm path length per compartment were used. The relevant protein solution was placed in one compartment of the cuvette, and the RNA target solution was placed in the other one. The CD spectra were recorded before and after mixing the cuvette contents. Blanks were run before each spectrum and subtracted from the raw data. Three spectra were averaged to increase the signal to noise ratio. The final protein and/or RNA concentration was 1 µM in 10 mM Tris-HCl buffer (pH 8) containing 150 mM NaCl, and the assays were carried out at 20°C. The protein and RNA spectra alone were recorded by replacing one or the other compartment by buffer. The results are presented as normalized Δε values on the basis of the nucleotide mean residue mass of 330 Da. Taking into account a sensitivity of $\delta(\Delta A) = 10^{-6}$ for the apparatus, the nucleotide concentration, and the optical path length of the cuvette, measurements were obtained at a precision of $\delta(\Delta\varepsilon) = \pm 0.01\ M^{-1}\ cm^{-1}$ per nucleotide.

Filter-binding assay

The fluorescently 5′ IRD800-labeled RNA was ordered to IDT. After reception, the RNA was heated for 2 min at 65°C, immediately placed for 10 min on ice, and diluted in binding buffer containing 1× PBS (pH 7.6), 2 mM $MgCl_2$, 6% glycerol, 0.5 mM DTT, 0.1 mM EDTA, 5 µg of *E. coli* tRNA per milliliter, and 50 µg of BSA per milliliter. Binding reactions consisted of 10 µL of RNA at 0.5 nM and 5 µL of proteins (final concentration from 25 to 2000 nM). Binding reactions were incubated for 15 min at 20°C and then applied directly to filters containing the two membranes under gentle vacuum. Before and after application of the binding reactions, 200 µL of binding buffer was used to equilibrate and rinse the system. Binding was quantified using an Odyssey apparatus (Li-Cor) and Image Lite program (Li-Cor). The intensity was corrected for background and fit for K_d using Gnuplot (http://www.gnuplot.info) using the following equation: $f(x) = x/(x + K_d)$, where x is the concentration of protein.

Cross-linking and analysis of Torrent sequence data

Cells expressing Rpf2-HTP, Rrs1-HTP, or wild type were grown to an OD_{600} of 0.5 in synthetic dextrose (SD) medium lacking Trp and Ura with 2% glucose. Cells were directly cross-linked inside culture media using Megatron (Granneman et al. 2011). Cells were processed as previously described (Granneman et al. 2009). Torrent sequencing data were aligned to yeast genome using TMAP (http://www.iontorrent.com). Downstream analyses, including the pileups presented here, were performed using the pyCRAC tool suite (Webb et al. 2014).

SAXS data collection

SAXS data on *S. cerevisiae* free 5S RNA were collected on beamline SWING (Soleil Synchrotron) at an energy of 13 keV. SAXS data on the Rpf2–Rrs1 protein complex and the 5S–Rpf2–Rrs1 RNA–protein complex, either full-length or proteolyzed, were collected on beamline BM29 (ESRF) at an energy of 12.5 keV. Scattering data were collected at 20°C at sample concentrations between 1 and 25 mg/mL. For measurement on the 5S RNA at SWING, the sample was injected on a gel filtration column (bio-SEC 3, Agilent), and data were recorded on the in-line elution profile (David and Perez 2009). For proteins and protein–RNA complexes, samples were injected directly on the BM29 flow cell (Pernot et al. 2013). Buffer background scattering was collected on the gel filtration buffers used for the RNA, proteins, and RNA–protein complexes. Background subtraction, averaging, and scaling were carried out using Foxtrot on the SWING beamline or the EDNA pipeline available on the BM29 beamline.

SAXS data processing

Further processing and data analysis were done using the programs of the ATSAS suite (Petoukhov et al. 2012). Guinier analysis was carried out in PRIMUS on data in the scattering range consistent with $Q_{max} \times Rg < 1.3$. Distance distribution functions and D_{max} were determined using the program Gnom. Ab initio free atom modeling was performed using the program DAMMIF protein complex. For RNA–protein complexes, the multiphase approach implemented in MONSA was used for shape modeling using the information from free RNA, Rpf2–Rrs1 alone, and RNA–protein complexes. Multiple MONSA calculations were run on the ATSAS online server. Several models (20–50) were computed for each macromolecule. Superposition, averaging, filtering, and computation of normalized spatial discrepancies (NSDs) were done using the program DAMAVER for each model ensemble.

Molecular modelling

The missing parts of the proteins were modeled by an ab initio approach using Rosetta (Rohl et al. 2004). Structures of each missing domain were calculated with Rosetta and clustered with a root mean square deviation (RMSD) radius of 5 Å. The top cluster as each domain was included as a template for homology modeling with Modeller for full atom model generation (Sali and Blundell 1993). The structures calculated with Modeller were ranked by the DOPE score (Shen and Sali 2006). SAXS data were then used for further modeling with DADIMODO (Evrard et al. 2011). The free RNA structure was modeled by generating 2000 models based only on secondary structure restraints using MC-SYM (Parisien and Major 2008). This data set explored a wide range of 5S RNA structures in both the structure of each helix and the relative orientation of the helices of the three-way junction. The RNA–protein complexes were modeled by superposing the structure of the protein complex in the same orientation as the cryo-EM bound model on all of the RNA structures generated. For each model, the goodness of fit was estimated by χ^2 calculation with Crysol.

Accession numbers

The Gene Expression Omnibus accession number for all sequence data reported in this study is GSE68431. The structure of the Rpf2–Rrs1 complex was deposited at the Protein Data Bank under code 5a53.

Acknowledgments

We thank Professor Nicolas James Watkins from Newcastle University for his careful reading of the manuscript and fruitful suggestions. We thank Dr. Frank Letourneur for his technical support on deep-sequencing analysis. Work at Université Paris Descartes was supported by Centre National de la Recherche Scientifique, University Paris Descartes, the RNPGenesis grant from the Agence Nationale de la Recherche (ANR JC RNPGenesis), and the Institut Universitaire de France. We acknowledge SOLEIL and ESRF for provision of synchrotron radiation facilities and thank the beamline scientists on the PROXIMA I and II, SWING, and BM29 beamlines. C.M. and S.L. expressed, purified, and crystallized the Rpf2 complex. C.M., N.L., and S.R. solved the crystal structure. S.L., M.B., L.D., and C.M. performed RNA- and protein-binding assays. S.R., M.B., N.L., S.L., and C.M. performed the SAXS experiments. S.L., J.P., and E.P. performed the CRAC experiments. N.L. designed the studies, interpreted data, and wrote the paper with contributions from S.L., C.M., M.B., and S.R.

References

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. Phenix: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66:** 213–221.

Armistead J, Khatkar S, Meyer B, Mark BL, Patel N, Coghlan G, Lamont RE, Liu S, Wiechert J, Cattini PA, et al. 2009. Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. *Am J Hum Genet* **84:** 728–739.

Asano N, Kato K, Nakamura A, Komoda K, Tanaka I, Yao M. 2015. Structural and functional analysis of the Rpf2-Rrs1 complex in ribosome biogenesis. *Nucleic Acids Res* **43:** 4746–4757.

Baßler J, Paternoga H, Holdermann I, Thoms M, Granneman S, Barrio-Garcia C, Nyarko A, Stier G, Clark SA, Schraivogel D, et al. 2014. A network of assembly factors is involved in remodeling rRNA elements during preribosome maturation. *J Cell Biol* **207:** 481–498.

Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W. 2003. Generation, representation and flow of phase information in structure determination: recent developments in and around Sharp 2.0. *Acta Crystallogr D Biol Crystallogr* **59:** 2023–2030.

Bugner V, Tecza A, Gessert S, Kühl M. 2011. Peter Pan functions independently of its role in ribosome biogenesis during early eye and craniofacial cartilage development in *Xenopus laevis*. *Development* **138:** 2369–2378.

Calviño FR, Kharde S, Ori A, Hendricks A, Wild K, Kressler D, Bange G, Hurt E, Beck M, Sinning I. 2015. Symportin 1 chaperones 5S RNP assembly during ribosome biogenesis by occupying an essential rRNA-binding site. *Nat Commun* **6:** 6510.

Ciganda M, Williams N. 2011. Eukaryotic 5S rRNA biogenesis. *Wiley Interdiscip Rev RNA* **2:** 523–533.

David G, Perez J. 2009. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J Appl Crystallogr* **42:** 892–900.

Donati G, Peddigari S, Mercer CA, Thomas G. 2013. 5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2–p53 checkpoint. *Cell Rep* **4:** 87–98.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126–2132.

Evrard G, Mareuil F, Bontems F, Sizun C, Perez J. 2011. DADIMODO: a program for refining the structure of multidomain proteins and complexes against small-angle scattering data and NMR-derived restraints. *J Appl Crystallogr* **44:** 1264–1271.

Furtado C, Regis-da-Silva CG, Passos-Silva DG, Franco GR, Macedo AM, Junho Pena SD, Machado CR. 2007. *Schistosoma mansoni*: the IMP4 gene is involved in DNA repair/tolerance after treatment with alkylating agent methyl methane sulfonate. *Exp Parasitol* **116:** 25–34.

Gamalinda M, Ohmayer U, Jakovljevic J, Kumcuoglu B, Woolford J, Mbom B, Lin L, Woolford JL. 2014. A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. *Genes Dev* **28:** 198–210.

Gambe AE, Matsunaga S, Takata H, Ono-Maniwa R, Baba A, Uchiyama S, Fukui K. 2009. A nucleolar protein RRS1 contributes to chromosome congression. *FEBS Lett* **583:** 1951–1956.

Gérczei T, Correll CC. 2004. Imp3p and Imp4p mediate formation of essential U3–precursor rRNA (pre-rRNA) duplexes, possibly to recruit the small subunit processome to the pre-rRNA. *Proc Natl Acad Sci* **101:** 15301–15306.

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* **106:** 9613–9618.

Granneman S, Petfalski E, Tollervey D. 2011. A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J* **30:** 4006–4019.

Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132.

Knight E, Darnell JE. 1967. Distribution of 5 s RNA in HeLa cells. *J Mol Biol* **28:** 491–502.

Layat E, Probst AV, Tourmente S. 2013. Structure, function and regulation of transcription factor IIIA: from *Xenopus* to *Arabidopsis*. *Biochim Biophys Acta* **1829:** 274–282.

Lebaron S, Froment C, Fromont-Racine M, Rain J-C, Monsarrat B, Caizergues-Ferrer M, Henry Y. 2005. The splicing ATPase prp43p is a component of multiple preribosomal particles. *Mol Cell Biol* **25:** 9269–9282.

Lebaron S, Schneider C, van Nues RW, Swiatkowska A, Walsh D, Böttcher B, Granneman S, Watkins NJ, Tollervey D. 2012. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat Struct Mol Biol* **19:** 744–753.

Lee BM, Xu J, Clarkson BK, Martinez-Yamout MA, Dyson HJ, Case DA, Gottesfeld JM, Wright PE. 2006. Induced fit and 'lock and key' recognition of 5S RNA by zinc fingers of transcription factor IIIA. *J Mol Biol* **357:** 275–291.

Leidig C, Thoms M, Holdermann I, Bradatsch B, Berninghausen O, Bange G, Sinning I, Hurt E, Beckmann R. 2014. 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. *Nat Commun* **5:** 3491.

Leslie AG, Powell HR. 2007. Processing diffraction data with Mosflm. In *Evolving methods for macromolecular crystallography* (ed. Read R, Sussman JL), pp. 41–51. Springer, Dordrecht, The Netherlands.

Liu Y, He Y, Jin A, Tikunov AP, Zhou L, Tollini LA, Leslie P, Kim T-H, Li LO, Coleman RA, et al. 2014. Ribosomal protein–Mdm2–p53 pathway coordinates nutrient stress with lipid metabolism by regulating MCD and promoting fatty acid oxidation. *Proc Natl Acad Sci* **111:** E2414–E2422.

Loc'h J, Blaud M, Réty S, Lebaron S, Deschamps P, Bareille J, Jombart J, Robert-Paganin J, Delbos L, Chardon F, et al. 2014. RNA mimicry by the Fap7 adenylate kinase in ribosome biogenesis. *PLoS Biol* **12:** e1001860.

Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14:** 953–961.

Lu D, Alexandra Searles M, Klug A. 2003. Crystal structure of a zinc-finger–RNA complex reveals two modes of molecular recognition. *Nature* **426:** 96–100.

Matsuo Y, Granneman S, Thoms M, Manikas R-G, Tollervey D, Hurt E. 2014. Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. *Nature* **505:** 112–116.

Miliani de Marval PL, Zhang Y. 2011. The RP–Mdm2–p53 pathway and tumorigenesis. *Oncotarget* **2:** 234–238.

Narla A, Ebert BL. 2011. Translational medicine: ribosomopathies. *Blood* **118:** 4300–4301.

Ng CL, Waterman D, Koonin EV, Antson AA, Ortiz-Lombardía M. 2005. Crystal structure of Mil (Mth680): internal duplication and similarity between the Imp4/Brix domain and the anticodon-binding domain of class IIa aminoacyl-tRNA synthetases. *EMBO Rep* **6:** 140–146.

Parisien M, Major F. 2008. The MC-sold and MC-sym pipeline infers RNA structure from sequence data. *Nature* **452:** 51–55.

Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, et al. 2013. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Radiat* **20:** 660–664.

Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HD, Konarev PV, Svergun DI. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **45:** 342–350.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612.

Rohl CA, Strauss CEM, Misura KMS, Baker D. 2004. Protein structure prediction using Rosetta. *Methods Enzymol* **383:** 66–93.

Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234:** 779–815.

Sasaki M, Kawahara K, Nishio M, Mimori K, Kogo R, Hamada K, Itoh B, Wang J, Komatsu Y, Yang YR, et al. 2011. Regulation of the MDM2–P53 pathway and tumor growth by PICT1 via nucleolar RPL11. *Nat Med* **17:** 944–951.

Shen M-Y, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Sci* **15:** 2507–2524.

Sloan KE, Bohnsack MT, Watkins NJ. 2013. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell Rep* **5:** 237–247.

Sondalle SB, Baserga SJ. 2014. Human diseases of the SSU processome. *Biochim Biophys Acta* **1842:** 758–764.

Steitz JA, Berg C, Hendrick JP, La Branche-Chabot H, Metspalu A, Rinke J, Yario T. 1988. A 5S rRNA/L5 complex is a precursor to ribosome assembly in mammalian cells. *J Cell Biol* **106:** 545–556.

Strunk BS, Novak MN, Young CL, Karbstein K. 2012. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* **150:** 111–121.

Trabuco LG, Villa E, Mitra K, Frank J, Schulten K. 2008. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16:** 673–683.

Ulbrich C, Diepholz M, Baßler J, Kressler D, Pertschy B, Galani K, Böttcher B, Hurt E. 2009. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell* **138:** 911–922.

Webb S, Hector RD, Kudla G, Granneman S. 2014. PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol* **15:** R8.

Wehner KA, Baserga SJ. 2002. The σ70-like motif: a eukaryotic RNA binding domain unique to a superfamily of proteins required for ribosome biogenesis. *Mol Cell* **9:** 329–339.

Zhang J, Harnpicharnchai P, Jakovljevic J, Tang L, Guo Y, Oeffinger M, Rout MP, Hiley SL, Hughes T, Woolford JL. 2007. Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rpL5 and rpL11 into nascent ribosomes. *Genes Dev* **21:** 2580–2592.