



Published in final edited form as:

Structure. 2017 February 07; 25(2): 329–340. doi:10.1016/j.str.2016.12.011.

Structural Heterogeneity in Pre-40S Ribosomes

Matthew C. Johnson^{*,1,2}, Homa Ghalei^{*,3}, Katelyn A. Doxtader^{3,4}, Katrin Karbstein³, and M. Elizabeth Stroupe^{**},¹

¹Department of Biological Science and Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306

³Department of Cancer Biology, The Scripps Research Institute, Jupiter, FL 33458

SUMMARY

Late stage 40S ribosome assembly is a highly regulated, dynamic process that occurs in the cytoplasm, alongside the full translation machinery. Seven assembly factors (AFs) regulate and facilitate maturation, but the mechanisms through which they work remain undetermined. Here, we present a series of structures of the immature small subunit (pre-40S) determined by three-dimensional (3D) cryogenic electron microscopy with 3D sorting to assess the molecule's heterogeneity. These structures demonstrate extensive structural heterogeneity of interface AFs that likely regulates subunit joining during 40S maturation. We also present structural models for the beak and the platform, two regions where the low resolution of previous studies did not allow for localization of AFs, and the rRNA, respectively. These models are supported by biochemical analyses using point variants and suggest that maturation of the 18S 3'-end is regulated by dissociation of the AF Dim1 from the subunit interface, consistent with previous biochemical analyses.

Graphical abstract

^{**}Lead contact to whom correspondences should be addressed, CONTACT: mestroupe@bio.fsu.edu phone: 850-644-1751, address: 91 Chieftain Way, Tallahassee, FL 32306 USA.

²Current address: University of Washington, Seattle, WA 98195

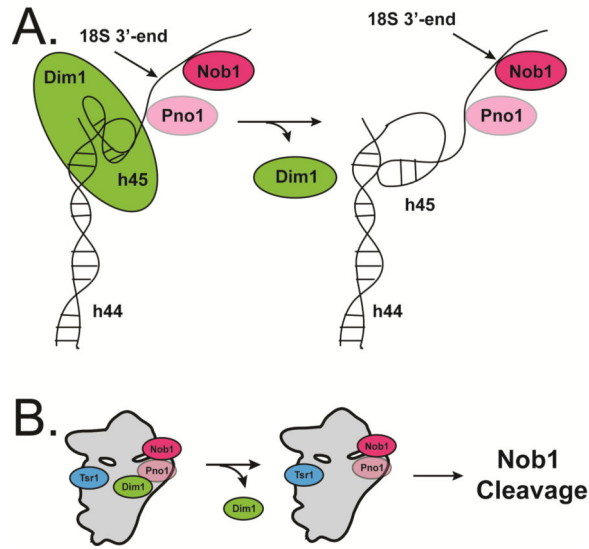
⁴Current address: UT Southwestern Medical Center, Dallas, TX 75390

*Co-first author

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

AUTHOR CONTRIBUTIONS

MCJ performed microscopy/image analysis. HG and KAD performed biochemical/*in vivo* assays. KK and MES supervised experimentation. MCJ, HG, KK and MES contributed to manuscript preparation.



INTRODUCTION

Assembly of the eukaryotic ribosome is a complex process requiring over 200 conserved AFs, the majority of which are essential (Henras et al., 2015; Strunk and Karbstein, 2009; Woolford and Baserga, 2013; Wu et al., 2016). AFs transiently bind to nascent ribosomes and comprise an assembly pathway promoting modification, cleavage, and folding of rRNA, as well as the co-transcriptional and post-transcriptional binding of ribosomal proteins (de la Cruz et al., 2015; Gamalinda and Woolford, 2015; Henras et al., 2015; Nerurkar et al., 2015; Talkish et al., 2016). That is, AFs serve both structural and catalytic functions during ribosomal development.

The assembly of both ribosomal subunits occurs in two phases, first in the nucleolus and then in the cytoplasm. Thus, late-stage assembly occurs in the presence of the full translation machinery (Karbstein, 2011). As a result, late-stage AFs promote both ribosome maturation and inhibit premature translation initiation (Strunk et al., 2011; Strunk et al., 2012). Preventing premature translation initiation is an especially important role of late-stage pre-40S AFs because translation is initiated by the binding of mRNA, the 60S subunit, and translation initiation factors to the small subunit.

Seven AFs promote and regulate late-stage 40S assembly, and their approximate binding sites were previously observed by low-resolution cryogenic electron microscopy (cryo-EM) coupled with analysis of partially depleted assembly intermediates (Strunk et al., 2011): the kinase Rio2, the methylase Dim1, and the GTPase-like Tsr1 bind to the subunit interface; the endonuclease Nob1 and its regulator Pno1 at the platform; and the export adaptor Ltv1 and Enp1 bind to the beak and somehow reposition the small subunit ribosomal protein (Rps) 3. In addition to promoting the final stages of the assembly cascade and inhibiting premature translation initiation, late-stage 40S maturation involves a translation-like quality control cycle in which the translation initiation factor eIF5B promotes the binding of mature 60S subunits to pre-40S, forming an 80S-like complex that contains some of the AFs

(Lebaron et al., 2012; Strunk et al., 2012). Quality control during ribosome assembly is essential because if misassembled ribosomes are released into the translation pool they can cause either mistakes during translation or stalling of ribosomes on mRNAs, which leads to degradation of both the ribosome and the mRNA (Cole et al., 2009). The detrimental effects from translational mistakes are illustrated by the lethality induced from antibiotics that increase the error rate during translation (Drummond and Wilke, 2009), as well as from diseases like Diamond Blackfan Anemia, 5q- syndrome, and congenital asplenia that increase the occurrence of misassembled ribosomes due to haploinsufficiency of ribosomal proteins (Armistead and Triggs-Raine, 2014; Freed et al., 2010; Narla and Ebert, 2010).

One essential step in understanding the role of AFs in pre-40S maturation is to know where they are located relative to the immature 18S RNA and other ribosomal proteins. Low-resolution cryo-EM identified these approximate binding positions (Strunk et al., 2011), but limited resolution of the structure precluded making specific predictions about interactions between the AFs and the growing 40S subunit or the path of the rRNA from one face of the pre-40S to the other. Additionally, the dynamic nature of the pre-40S complex with all seven AFs suggested that resolution-limiting conformational heterogeneity of those AFs might hold clues as to structural changes that need to occur during progression from pre-40S to the 80S-like quality control intermediate that contains the 60S subunit and multiple pre-40S AFs.

To address questions about the extent of heterogeneity in pre-40S, we used cryo-EM and 3D computational analysis to determine multiple sub-states of AFs from a single population of pre-40S molecules. At its core, the single composite map is at about 9 Å resolution. After focused 3D classification at the AF positions, each AF is at about 8–10 Å resolution, sufficient resolution to place the AFs and unambiguously predict regions that modulate pre-40S binding. Biochemical and *in vivo* analyses of variants in each AF confirm these predictions. Local classification revealed significant structural heterogeneity at the small subunit interface, which includes multiple conformations of the AFs Tsr1 and Rio2. Additionally, we determined the architecture of the beak and platform, localizing the AFs known to regulate assembly of the mRNA entry channel and identifying the path of the rRNA through the head of the pre-40S. These structural models have lead us to propose that maturation of the 18S 3'-end is regulated by dissociation of the AF Dim1 from the subunit interface.

RESULTS

AFs associated with pre-40S are flexible

Pre-40S ribosomes were prepared via tandem affinity purification (TAP) with the tag positioned either on Rio2 or Ltv1 as previously described (Pertschy et al., 2009; Schafer et al., 2006; Strunk et al., 2011). The final dataset consisted of 131,842 particles after normalizing angular sampling to account for anisotropy from the particle's preferred orientation that initially resulted in an unreliably high resolution measurement (Fig. S1). The consensus structure reached a resolution of 9.4–11.1 Å based on the 0.143 and 0.5 Fourier Shell Correlation (FSC) criteria, with a local resolution range of 8–16 Å, assessed by ResMap (Kucukelbir et al., 2014) (Fig. 1 and S1). The highest resolution regions were at the

core whereas peripheral AFs were the least well resolved, suggesting conformational heterogeneity and/or partial occupancy in AF binding as observed in other 40S ribosome structures (Anger et al., 2013; Aylett et al., 2015; Erzberger et al., 2014; Khatter et al., 2015; Llacer et al., 2015). We used Molecular Dynamics Flexible Fitting (MDFF) (Phillips et al., 2005) to relax the mature yeast 40S structure (Ben-Shem et al., 2011) into the consensus map, and to provide a framework for interpreting the core in relation to the AFs. We also modeled the AFs based on existing structures (Table S1).

To discriminate between the possible reasons for degraded peripheral resolution (heterogeneity in binding pose or composition), we used hierarchical 3D classification, first on a global scale and then focusing on individual pre-40S features, to independently analyze the interface, platform, and beak. The interface showed inter-related, discrete positioning of its three AFs, Tsr1, Rio2, and Dim1. In contrast, the beak and platform did not reveal discretely different positions of beak proteins Rps3, Enp1, and Ltv1 or platform proteins Pno1 and Nob1. The implications of these observations are described below.

Bi-modal positioning of interface AFs

Tsr1—The subunit interface of pre-40S is an important barrier to 80S assembly because Tsr1, Rio2, and Dim1 physically bind where mRNA, translation factors, and 60S will dock. Despite the steric clash, at least Tsr1 and Dim1 remain bound to 80S-like ribosomes (Karbstein, 2013; Strunk et al., 2012). Thus, the preformed interface must rearrange to accommodate the 60S subunit. Tsr1 is a four-domain homolog of the translation factor GTPase SelB but without the amino acids responsible for GTPase activity (Gelperin et al., 2001). Flexible linkers join the four domains, potentially allowing it to adopt different conformations. In fact, partial deletion of the domain II insertion (amino acids 410 to 476) was required to rigidify the molecule for crystallization, locking it in the characteristic shape shared with SelB (McCaughan et al., 2016).

In the pre-40S consensus map, the Tsr1 density was not as well resolved as the neighboring rRNA decoding helix 44 (h44) (Fig. 1). Therefore, we used focused 3D classification to assess structural heterogeneity at the Tsr1 position (Amunts et al., 2014). This classification produced two dominant subclasses with Tsr1 densities rotated relative to one another by about 28° (Fig. 2A–B), which together accounted for 56% of the data (Fig. S2). Subclass structures were reconstructed with the Euler angles determined from the consensus map, with local resolution variation of 8–14 Å (Fig. S2). The remaining 44% of particles were split between six low-resolution, minority classes (Fig. S2).

The resulting contours of the Tsr1 density improved to reveal unambiguous features that define the Tsr1 position. Specifically, one of the two major classes, T1, fit the *S. cerevisiae* Tsr1 X-ray crystal structure (PDB 5IW7, McCaughan et al., 2016) without any domain-domain rearrangements (Fig. 2A). The GTPase-like domain I reached around the side of pre-40S, anchoring Tsr1 to the pre-40S body. The β -barrel of domain II, nestled into domains I and III, filled a density that is distinctly separated from h44 by about 20 Å. Domain IV reached up from the base formed by domains I, II, and III to interact with the side of the beak.

In contrast, rigid-body fitting of the same structure into the Tsr1 density of the other well-defined subclass, T2, did not result in a good match (Fig. 2B). The overall features were similar, but clearly the relative orientations of the domains needed to change to accommodate the density (Fig. 2B). Specifically, domain I did not fill the density near the body whereas density adjacent to domain II now linked Tsr1 to h44. Furthermore, additional density at the backside of the molecule visible in this substructure might account for the insertion domain missing in the X-ray crystal structure (PDB 5IW7, McCaughan et al., 2016). Finally, the density would better accommodate domain IV if it rotated towards the beak. Thus, to move from T1 to T2, domain II appeared to roll towards h44 by about 28°, whereas domain IV rotated by about 13° toward the beak and domains I and III rotated by about 19° away from domain II (Fig. 2B and Movie 1).

To test our prediction about interactions between Tsr1 domain II and pre-40S, we mutated residues K201, K203, R245 and R248 to glutamates (Tsr1_KKRR) and probed this variant for its effect on yeast growth and Tsr1 binding to pre-40S. Although Tsr1-KKRR was expressed as well as wild type Tsr1 (Fig. S6) it did not complement the absence of Tsr1 (Fig. 2C), and gradient centrifugation showed that Tsr1_KKRR no longer co-sedimented with pre-40S (Fig. 2D), demonstrating that it bound more weakly than wild type Tsr1. Thus, the effect from the Tsr1_KKRR mutation supported the placement and orientation of Tsr1 in the EM density.

Rio2—Like Tsr1, low-resolution density corresponding to Rio2 and its connection to the head suggested that 3D local classification would illuminate multiple conformations (Fig. 1). Indeed, after testing various sorting strategies using 2–12 classes, local 3D classification into eight classes resulted in two dominant, structurally distinct classes. When combined, these classes represent 55% of the data (Fig. S2). Subclass structures were reconstructed with the Euler angles determined from the consensus map, with local resolution variation of 8–14 Å, as in Tsr1 (Fig. S2). The remaining 45% of particles were spread among six low-resolution minority classes (Fig. S2).

In subclass R1 the density corresponding to Rio2 was elongated, bridging between the body and the head and projecting into the cavity formed by Tsr1 domain IV, the beak, and h44 (Fig. 3A). In subclass R2 the density corresponding to Rio2 was U-shaped, reaching from domain IV of Tsr1 across h44 to bind between the body and the head (Fig. 3B). The density that attaches Rio2 to the head also improved in local resolution in each class after separating the subclasses (Fig. S2).

Rio2 is a multi-domain protein with an N-terminal winged helix-turn-helix (wHTH) domain fused to a bilobal kinase domain, where the ATP binding pocket sits deep within a cavity formed between the two middle domains (Ferreira-Cerca et al., 2012; LaRonde-LeBlanc et al., 2005). Specifically, the model of *Chaetomium thermophilum* ADP-bound Rio2 (PDB 4GYI, Ferreira-Cerca et al., 2012) docked well into the elongated Rio2 (R1), positioning the N-terminal wHTH domain (purple in Fig. 3A) near domain IV of Tsr1. Further, mutagenesis experiments by Ferreira-Cerca et al., 2012 showed that K105 is important for Rio2 binding to the pre-40S (Ferreira-Cerca et al., 2012). Importantly, this arrangement positioned K105 in proximity to h29 and h30 in the head of the pre-18S rRNA (Fig. 3A).

The same Rio2 model fit less well to the U-shaped density (R2) (Fig. 3B). The common density from the Rio2-N lobe could remain in place in transition from one form to the other, with the wHTH and Rio2-C lobe each rotating away from the central ATP binding cleft to fill the U-shaped contours (Fig. 3B, arrows, and Movie 2). The unknown structure of the C-terminal extension in the yeast homolog likely explains additional densities in both fittings not accommodated by the existing model.

In addition to differences in the Rio2 density, we also observed differences in the density that connects the C-terminal extension to the head (Fig. 3). This density was near the mature position of Rps15, which we have previously shown to bind Rio2 directly (Campbell and Karbstein, 2011). Thus, changes in Rio2 structure could be communicated to Rps15.

Dim1—Like Rio2 and Tsr1, Dim1 also showed signs of heterogeneity manifested by resolution loss in the consensus map (Fig. 1). Thus, we performed focused 3D classification on this region. Rather than revealing structural heterogeneity, however, the resulting major subclasses showed a dominant Dim1 position and Dim1-free pre-40S (Fig. S2). The elongated Dim1 molecule, shown as the human Dim1 X-ray crystal structure (PDB 1ZQ9, (Dong et al., 2005)), stretched from Rio2 across the interface to the body, where it bound at the platform (Fig. 4A). The absence of Dim1 in the second subclass did not appear to degrade the overall integrity of the sample (Fig. 4B).

We used electrostatic calculations (Fig. S3A) and the orientation of the Dim1 active site (Fig. S3B) to suggest the surface of Dim1 used for 40S binding, which oriented the N-terminus towards Rio2 and the C-terminus towards the platform. We then used mutagenesis to confirm the molecular placement of Dim1. The docked human Dim1 structure (PDB 1ZQ9, (Dong et al., 2005) predicted that R233, K234, N235, and K236 interact with 18S rRNA (Fig. 4A). Mutagenesis of these residues to aspartate (Dim1_RKNK) did not allow for complementation of Dim1 depletion (Fig. 4C), although it is expressed as well as wild type Dim1 (Fig. S6). Further, gradient sedimentation of yeast cells expressing wild type or variant Dim1 confirmed that Dim1_RKNK bound nascent 40S more weakly than wild-type Dim1 (Fig. 4D). These data validated our positioning of the Dim1 crystal structure in the EM density.

Tsr1 repositioning can ameliorate clashes with eIF5B

We previously showed that Tsr1 depletion allows for premature subunit joining (Strunk et al., 2011). Furthermore, Tsr1 blocks binding of eIF5B, which promotes subunit joining (Strunk et al., 2012; Lebaron et al., 2012; and McCaughan et al., 2016). Similarly, in 80S-like ribosomes, Dim1 is predicted to clash with H69 of the 60S subunit (Boehringer et al., 2012; Karbstein, 2013; Strunk et al., 2011). Nevertheless, both Tsr1 and Dim1 are found in 80S-like ribosomes (Strunk et al., 2012), indicating that their positions must be plastic during 40S maturation. To address how the three interface AFs can accommodate eIF5B and the 60S subunit, we docked the eIF5B-bound 80S structure (PDB 4V8Z, Fernandez, Bai et al. 2013) onto our model (Fig. S4). Specifically, we superimposed Rps21 from both structures using the “matchmaker” feature of Chimera (Pettersen et al., 2004). Importantly, nearly identical docked structures were obtained if Rps4, Rps7, Rps8, or Rps13 was used as

the reference, demonstrating the robustness of this procedure. We then looked to see how the eIF5B position in the 80S structure compared to the two Tsr1 subclasses. In both subclasses, Tsr1 fit in a cavity formed between the small and large subunits (Fig. S4); however, as previously noted, eIF5B's domain IV was oriented at about a 30° angle towards the interface relative to where Tsr1 binds, with steric conflict between eIF5B domains I and II and Tsr1 domain I (Karbstein, 2013; McCaughan et al., 2016; Strunk et al., 2012). Interestingly, these conflicts were worse in T2 than in T1, suggesting that these binding modes might represent stages of a larger rearrangement, where Tsr1 rotates away from h44 to allow for formation of the 80S-like complex that accommodates both Tsr1 and eIF5B.

Pno1 positions Nob1 at the 3' end of the small subunit rRNA

We next turned to the platform and attempted to localize the endonuclease Nob1 and its accessory factor Pno1, which together direct the final cleavage to produce mature 18S rRNA (Lamanna and Karbstein, 2009, 2011; Pertschy et al., 2009; Woolls et al., 2011). As with the interface, lower resolution density for these AFs led us to perform focused 3D classification, revealing two major maps that both had density corresponding to Pno1 but only one of which had density corresponding to Nob1 (Fig. S5).

Pno1 is composed of three RNA-binding K Homology (KH)-like domains (Vanrobays et al., 2004; Zhou et al., 2004). In our structure resulting from focused classification of the platform region, two sequential lobes were found on the platform (Fig. 5A), into which the two-domain *Pyrococcus horikoshii* RNA-bound Pno1 fit well (PDB 3AEV, (Jia et al., 2010)). The N-terminal, eukaryotic-specific KH domain is highly degenerate and dispensable (Fig. 5B), suggesting it bound peripherally and was therefore unresolved in our maps. Based on the known interaction between the C-terminal-most KH domain and rRNA (Jia et al., 2010), we assigned the density closest to the 3' end of the rRNA as the third domain. To test this hypothesis, we generated a Pno1 point variant in which K208, K211, and K213 were altered to aspartate and F214 was altered to alanine (Pno1_KKKF). This variant is expressed as well or more than wild type Pno1 (Fig. S6) did not fully complement the absence of Pno1 (Fig. 5B), and gradient sedimentation confirmed that this was due to weaker binding to pre-40S (Fig. 5C).

The remaining, rectangular density on the platform accommodated the structure of the PIN domain of the *P. horikoshii* Nob1 homolog, which positioned Nob1 adjacent to the second Pno1 KH domain (Fig. 5A), consistent with our previous biochemical data (Woolls et al., 2011). Density was also visible for the C-terminal zinc ribbon domain of Nob1 at the tip of the platform (Fig. 5A). These two domains, the only domains present in the NMR structure of *P. horikoshii* Nob1, represent the endonuclease's evolutionarily conserved core. The yeast homolog contains a large insertion that joins the conserved domains, but was unresolved in this structure, likely due to inherent flexibility.

Rps3 bridges the head and beak, held in an immature position by Ltv1/Enp1

Deletion mutagenesis previously identified the Rps3/Enp1/Ltv1 complex as part of the beak in pre-40S. Nevertheless, the beak architecture and complex structure was obscured by poor resolution of the initial cryo-EM structure (Strunk et al., 2011), and a lack of structural

information on the subcomplex or its members. By use of an Ltv1-TAP tagged pre-40S preparation, which allows for enrichment of pre-40S complexes containing Ltv1, coupled with 3D classification, we identified the position of each AF and confirm our model by *in vivo* and *in vitro* analysis of AF variants.

Of the three proteins in the beak complex, Rps3 is the best structurally characterized because it is part of mature 40S ribosomes, the X-ray crystal structure of which has been determined (Ben-Shem et al., 2011; Rabl et al., 2011). A structure of Rps3 in complex with its chaperone Yar1 is also known (Holzer et al., 2013). Rps3 has a distinct bow-tie shape and density that fit each domain was clearly visible, peripheral to its mature position (Fig. 6A and B). Neither Enp1 nor Ltv1 have experimentally determined structures. The only region of either protein that can be reliably modeled is a series of TPR motifs in Enp1, which we did by use of the Phyre algorithm (Kelley et al., 2015). The resulting TPR motifs fit well into an extended density that stretches from the tip of the beak to the body (Fig. 6A and B).

To validate this placement of the Enp1-TPR domain, we altered three conserved residues (Enp1_KKY: K378E, K379E and Y380I) that we predicted to contact the ribosome. As expected if the altered residues were important for Enp1 and pre-40S binding, growth of Enp1-deficient yeast cells expressing Enp1_KKY was severely hampered compared to those expressing wild type protein (Fig. 6C), even though the variant protein was expressed as well as wild type protein (Fig. S6). Further, gradient fractionation of extracts from yeast cells expressing wild-type Enp1 or Enp1_KKY, combined with Western blot analysis, showed less Enp1_KKY bound to 40S ribosomes (Fig. 6D). These data confirmed that Enp1_KKY bound more weakly to ribosomes than wild-type Enp1, supporting placement of the Enp1-TPR domain in the EM density as the anchor to pre-40S.

To further substantiate this proposed topological arrangement of the Rps3/Enp1/Ltv1 complex, we performed biochemical experiments to define sites of interaction with the nascent 40S subunit and delineate regions of interactions between the proteins. First, to test if the TPR repeat domain of Enp1 was sufficient for binding to the nascent 40S subunit, we generated this fragment of Enp1 (residues 154–483, Enp1-TPR, Fig. 6E), identified as a stable core after limited proteolysis of the full-length protein (data not shown). Enp1 or Enp1-TPR were mixed with Ltv1 pre-40S subunits, fractionated over a gradient followed by Western blot analysis (Fig. 6F). Enp1-TPR co-migrated with pre-40S, demonstrating that Enp1-TPR was sufficient for binding to pre-40S, as expected from the placement of this domain in the cryo-EM structure at the interface between the Rps3/Enp1/Ltv1 complex and the pre-40S scaffold.

Second, to test if Enp1-TPR also bound Ltv1, we used *in vitro* pulldowns with purified, recombinant proteins as previously described (Campbell and Karbstein, 2011). Enp1-TPR bound Ltv1, but was not retained on resin alone (Fig. 6G), demonstrating an interaction with Ltv1. Thus, the TPR domain of Enp1 is sufficient to make Enp1's known functional interactions with 40S subunits and Ltv1. The presence of just the Enp1 TPR domain in the EM map is, therefore, explained by it forming a well-defined structural unit.

Third, to map the position of Ltv1 relative to Enp1-TPR and Rps3, we generated two Ltv1 fragments that were identified by limited proteolysis (data not shown). One encompassed the conserved N-terminal 180 residues (Ltv1-N) and the other spanned the middle residues 185–394 (Ltv1-M) (Fig. 6E), including the Ltv1 phosphorylation sites important for release from pre-40S (Ghalei et al., 2015). *In vitro* binding experiments using recombinant Ltv1-M and Enp1 revealed that Ltv1-M was sufficient for interaction with both Enp1 and Enp1-TPR (Fig. 6G). In contrast, Ltv1-N did not bind Enp1 (Fig. S6A).

Fourth, to probe the direct interaction between Ltv1 and Rps3 (Ghalei et al., 2015; Strunk et al., 2011), we performed binding experiments with Ltv1-M or Ltv1-N and Rps3. Both Ltv1 fragments interacted with Rps3 (Fig. S6B and C), consistent with yeast two-hybrid data (Fassio et al., 2010; Mitterer et al., 2016). Further, our data showed that Ltv1 and Rps3 interacted through Rps3's N-terminal domain (Fig. S6D), also consistent with yeast two-hybrid data (Mitterer et al., 2016).

Finally, we used gradient sedimentation to test for binding of Ltv1-N or Ltv1-M to pre-40S (Fig. 6H). Ltv1-M co-sedimentation with pre-40S was indistinguishable from full length Ltv1. In contrast, Ltv1-N did not co-sediment with pre-40S, demonstrating its weaker binding. Thus, we defined Ltv1-M as the structural core that retained its ability to bind Enp1, Rps3 and pre-40S, filling the unassigned electron density near Enp1 (Fig. 6A) and allowing it to bind Enp1, Rps20 (Mitterer et al., 2016), and pre-40S, marked with a red oval.

DISCUSSION

Despite recent publication of near-atomic resolution 80S ribosome structures by cryo-EM, the small ribosomal subunit presents unique technical limitations that affect the absolute resolution of the final reconstruction. First, 40S subunits bind preferentially to the carbon support in a preferred orientation and are known to adopt a preferred orientation in ice (Passmore et al., 2007), creating a bias in sampling and resulting resolution limitation. We addressed this by using a tilt-scheme in data acquisition (van Heel et al., 2000), which helps but does not completely eliminate the angular bias because there is a limit to the degree to which one can tilt a specimen in a TEM. Including the tilted data did, however, improve the anisotropic resolution in the initial reconstruction (Fig. S1). Second, 40S is a biochemically more flexible molecule than its 60S partner because this flexibility is part of its function during protein synthesis (Culver, 2003). Third, the pre-40S specimen we are analyzing is particularly fragile because of the biologically transient nature of AF binding. Consequently, small subunit structures have historically been of lower resolution than large subunit or assembled ribosome structures that do not suffer from these physical limitations, and require binding of translation factors to stabilize specific conformations (Anger et al., 2013; Aylett et al., 2015; Erzberger et al., 2014; Khatter et al., 2015; Larburu et al., 2016; Llacer et al., 2015). Despite these limitations, our structural analysis presented here delineated the underlying heterogeneity in AF binding that contributes to the dynamic function of pre-40S.

Tsr1, Dim1 and Rio2 sterically inhibit binding of translation initiation factors, initiator tRNA and mRNA and deform the decoding helix h44 (Strunk et al., 2011). Furthermore, Tsr1 and Dim1 physically block binding of eIF5B and 60S subunits, thereby blocking the

formation of 80S complexes during 40S maturation. Despite these physical obstacles, Tsr1 and Dim1 are present within 80S-like complexes whereas Rio2 appears to be released around the time of 60S joining (Karbstein, 2013; Strunk et al., 2012). This suggests the possibility that regulated conformational changes of these AFs allow for eIF5B and 60S subunit binding to form 80S-like complexes. Here we demonstrate that Tsr1 and Rio2 bind in at least two distinct conformations. The two Tsr1 conformations display differing degrees of overlap with eIF5B, indicating they might provide snapshots into rearrangements required for eIF5B accommodation and therefore, ultimately, subunit joining. Similarly, Rio2 connects differently with Rps15 in the two structures. This rearranged connection might explain how Ltv1 release is linked to subunit joining, as previously observed (Ghalei et al., 2015), because Rps15 binds both Rio2 and Ltv1 (Campbell and Karbstein, 2011).

Compared to the distinct conformational states of the interface AFs, the AFs bound at the platform (the endonuclease Nob1 and its regulator Pno1) are less well resolved and appear to assume a continuum of similar conformations, some of which lack Nob1. The multiplicity of states indicates flexibility in this region. Interestingly, in those classes that contain Nob1 and Pno1, an unassigned density (blue in Fig. 5A) is visible at the base of the Pno1-KH₃ domain, near the mature 3'-end of 18S rRNA (red in Fig. 5A and B). Previous work has indicated that Pno1-KH₃ binds the GGAUCA sequence three nucleotides 5' to the cleavage site (Jia et al., 2010) and the density fits a single stranded RNA well. Importantly, this density also connects to h45 in the front of the molecule. Thus, we suggest that this density represents the rRNA strand near the 18S 3'-end. Further, we also note that this unassigned density, and its connection to h45, is positioned differently in the two Dim1 subclasses (Fig. S3B–D).

Together, these observations suggest a model for a Dim1-dependent regulation of Nob1 cleavage (Fig. 7): in mature 40S subunits, the loop of h45, which contains the adenosines that are modified by Dim1, is docked into h44. Our structures show that there is no electron density where the adenosines are expected to be located, suggesting that Dim1 binding disrupts this tertiary contact to put the adenosines in the Dim1 active site, thereby repositioning h45 (Fig. S4B–C). Thus, Dim1 uses its binding energy to position the rRNA substrate away from the Nob1 active site. Dissociation of Dim1 from pre-40S within the 80S-like complex then allows h45 to relax to its mature conformation (Fig. S4D), threading the 3'-end of 18S rRNA to the Nob1 active site. This model is consistent with our previous footprinting data, which indicate that the main Nob1 footprint is centered about 3 nucleotides 3' of the 18S end in the pre-40S visualized here (Lamanna and Karbstein, 2009). Additional support comes from biochemical data in bacteria that links Dim1 activity to S21 binding adjacent to the 3'-end of 16S rRNA via manipulation of the rRNA structure (Thammana and Held, 1974; Van Knippenberg et al., 1984).

Previous structural and biochemical data suggest that during maturation Rps3 repositions (Mitterer et al., 2016; Schafer et al., 2006; Strunk et al., 2011). More specifically, it was suggested that the position of the C-terminal domain is fixed, but the N-terminal domain rotates, a hypothesis based on the observation that interaction between the N-terminal domain of Rps3 and Rps20 is important for 40S maturation (Mitterer et al., 2016). Whereas our data strongly support repositioning of Rps3, they are not consistent with the C-terminus of Rps3 being fixed because our structure shows neither domain at its mature conformation

in this precursor form. However, slight overlap between the immature and mature position of the N-terminal domain indicates that it might 'slide' into position after Ltv1/Enp1 are released. Nevertheless, the importance of Rps3-Rps20 interactions are consistent with our structural observations because these interactions might instead be required for a) recruitment or stabilization of Rps20, b) the modest repositioning of Rps3-N, or c) Hrr25-dependent recognition and phosphorylation of Ltv1.

Conclusions

A variety of structural alterations in pre-40S ribosomes must occur for maturation, including Ltv1 release, Rps3 repositioning, and conformational changes in Dim1 and Tsr1. The biochemical triggers for these changes to occur are unclear, however our structures reveal snapshots into the range of structural changes that ultimately lead to the dynamic assembly of the small subunit. This new understanding of the conformational flexibility of the pre-40S ribosomal subunit will form the basis for future biochemical experiments to explore the implications of these motions.

EXPERIMENTAL PROCEDURES

Yeast strains and cloning

Yeast strains (Table S1) were made via PCR-based recombination (Longtine et al., 1998) and confirmed by PCR and Western blot analysis with available antibodies. Vectors (Table S2) were generated using standard cloning protocols and confirmed by sequencing.

Protein purification

Ltv1, Enp1 and truncations were expressed and purified as described (Ghalei et al., 2015). SUMO-Rps3 was expressed from a pET28a vector as an N-terminal His₆/SUMO tag fusion in Rosetta DE3 cells (EMD Millipore). Cells were grown at 37°C to OD₆₀₀=0.5 in 2×YT media before adding 1 mM isopropyl β-D-thiogalactoside (IPTG) at 30°C. Cells were harvested after 4 h, lysed by sonication and purified on Ni²⁺-NTA resin (Qiagen) per the manufacturer's protocol. Eluted protein was dialyzed against 50 mM Tris, 6.8, 100 mM NaCl and 2 mM DTT and purified by ion exchange chromatography (Mono S; GE Healthcare). Protein was eluted with a linear 200–600 mM NaCl gradient before size-exclusion chromatography (Superdex 75; GE Healthcare) in 50 mM Tris, 6.8, 200 mM NaCl and 2 mM DTT.

Ltv1-Flag was co-expressed with His₆-Rps3 or His₆-Rps3^N (residues 1–95) from pETDuet-1 vectors (Novagen) in Rosetta DE3 cells. Cells were grown in LB media at 37°C to an OD₆₀₀=0.4 before induction with 0.3 mM IPTG. Cultures were harvested after 20 h of growth at 16 °C. Cells were lysed by sonication in Ni²⁺-NTA buffer and purified per the manufacturer's protocol. Ni eluates were diluted 10X with TBS buffer (50 mM Tris, 7.5, 150 mM NaCl and 1 mM DTT), purified on anti-Flag agarose (Sigma) per the manufacturer's protocol. Bound proteins were eluted with TBS buffer containing 100 µg/mL Flag-peptide (Sigma).

Density gradient analysis

Sucrose gradient fractionation of cell lysates and subsequent Western blot analyses were carried out as described (Strunk et al., 2012). Briefly, cells were grown to mid log phase in YPD, harvested, washed, and lysed in gradient buffer (20 mM Hepes/7.4, 5 mM MgCl₂, 100 mM KCl, and 2 mM DTT) with 0.1 mg/mL cycloheximide and Complete protease inhibitor cocktail (Roche). Cleared lysate was applied to 10–50% sucrose gradients and spun for 2 h at 40,000 RPM in an SW41Ti rotor before fractionation.

Binding assays

3 μM MBP or MBP-tagged protein was mixed with 5 μM untagged protein in binding buffer (50 mM Tris, 7.5, 150 mM NaCl, and 2 mM DTT) and incubated for 15 m at 4°C. The mixture was applied to 25 μl of amylose resin (New England Biolabs) pre-equilibrated in binding buffer and incubated on a rotating platform at 4 °C for 30 m. Resin was washed four times with 200 μl of binding buffer, and bound proteins were eluted with 25 μl of binding buffer plus 20 mM maltose.

Purification of pre-40S

Yeast strains expressing TAP-tagged proteins were grown in a reactor at 30°C in YPD to an OD₆₀₀=1.0. Affinity purification of TAP-tagged ribosomes was performed as described (Ghalei et al., 2015). Following elution from calmodulin resin, 20 pmoles of purified pre-ribosomes were layered onto a fixation gradient (5–20% glycerol; 0.05 – 0.25% glutaraldehyde, 50 mM HEPES/7.6, 100 mM NaCl, 10 mM MgCl₂, 10 mM β-ME, 0.075% NP-40) (Kastner et al., 2008) and centrifuged in a SW41 rotor at 288,000 x *g* for 4 h at 4°C. Peak fractions were buffer exchanged (PD Mini Trap, GE Healthcare) into cryo buffer (50 mM HEPES/7.6, 100 mM NaCl, 10 mM MgCl₂, 10 mM β-ME).

Grid Preparation and Data Collection

3 μL sample was applied to a glow-discharged cryo-well grid (C-Flat EM grids (Protochips) with a layer of continuous carbon) and plunged into liquid ethane with an FEI Vitrobot. Attempts to use other substrates like the UltraAuFoil (FEI) failed to concentrate the specimen sufficiently. Image acquisition was performed using an FEI Titan equipped with a Direct Electron DE20 detector, at a nominal magnification of 22,500x (1.62 Å/pixel) and a total electron dose of 60 e⁻/Å² over 30 frames. For Rio2-TAP purified ribosomes, 3092 images were recorded at 0° tilt, 818 at 15°, and 727 at 45°. For Ltv1-TAP purified ribosomes, 2281 were recorded at 0° tilt and 781 images at 45°.

Single Particle Analysis

Pre-processing was performed in Appion (Lander et al., 2009; Shrum et al., 2012), including full-frame alignment and dose compensation using the DE_process_frames-2.5.1 software (Spear et al., 2015). Contrast transfer function and defocus estimation were performed with CTFFind3 or CTFTilt (Mindell and Grigorieff, 2003). Particles were picked from frame-aligned micrographs with FindEM (Roseman, 2004) using a 100 Å low-pass filtered pre-40S model (Strunk et al., 2011). Individual particles from tilted images were further subjected to frame alignment and dose compensation (Spear et al., 2015).

Image processing continued in RELION (Scheres, 2012). Rio2-TAP or LTV-TAP particles were sorted in 2D in 3 rounds of selection, resulting in 144,902 and 177,379 particles for the Rio2-TAP and Ltv1-TAP datasets, respectively. These particles were aligned independently in 3D using the pre-40S model (Strunk et al., 2011) low pass filtered to 50 Å as a template. After 25 iterations each, the Rio2-TAP and Ltv1-TAP datasets reached nominal resolutions of ~16 Å and ~14 Å, respectively (Fig. S1).

Angle Normalization

To further alleviate resolution anisotropy due to a preferred orientation, particles from overrepresented Euler angles were culled (Fig. S1). A range of particles-per-angle cutoffs was tested, reducing the overall number of particles by 10–90%. 47,637 Rio2-TAP particles and 97,033 LTV1-TAP particles were removed based on the quality of the resulting structures and 2D projections.

Refinement and classification of the Rio2-TAP dataset

Serial focused classification was employed to classify the Rio2-TAP structure. First, classification with a mask that excluded the rRNA and core ribosomal protein densities found 3 classes to discarded and 7 similar classes to save. Similar classes (76,310 particles) were pooled, refined further (~11 Å), and sorted again by classification with a mask excluding AF densities but including the ribosomal core. 2 of the resulting classes (24,814 particles) exhibited disruptions to h44 so were discarded. 4 remaining classes (51,496 particles) were pooled. Then, frequency-limited refinement of the Rio2-TAP dataset continued with FREALIGN (Grigorieff, 2007; Lyumkis et al., 2013), bringing the resolution to 11.5 Å. 2 rounds of final Rio2-TAP classifications were performed in parallel, using custom masks surrounding each interface AF (Fig. S2).

Refinement and classification of the Ltv1-TAP dataset

Unlike Rio2-TAP, classification of the 80k particle Ltv1-TAP dataset did not identify a large number of damaged particles. Therefore, the angle-normalized particles were directly subjected to frequency-limited refinement in FREALIGN, reaching 14.8 Å (Fig. S1). 2 serial rounds of locally masked classification of the platform followed.

Refinement of pooled datasets and final reconstruction

To improve resolution, the Rio2-TAP and Ltv1-TAP stacks were combined (131,842 particles) and subjected to further frequency-limited refinement in FREALIGN to a final resolution of 9.4–11.1 Å (Fig. S1). Angles and shifts from the combined dataset were used to reconstruct the separately determined Rio2-TAP and Ltv1-TAP focused classes. The reconstruction resolution cutoff was chosen from the in-mask resolution calculated with `relion_postprocess` (Scheres, 2012). Local resolution variation was calculated with ResMAP (Kucukelbir et al., 2014). MDFF (Phillips et al., 2005) was used to relax the yeast mature 40S structure (Ben-Shem et al., 2011) into the Rio2-TAP/Ltv1-TAP consensus map. Rps3, Rps10, Rps17, Rps20, Rps26, Rps29, Rps30, Rack1, and Stm1 were removed manually and MDFF was run for 500,000 steps using a G-scale value of 0.3.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants GM117093 and GM086451 (KK) and NSF award MCB1149763 (MES). HG was partially supported by a PGA National Cancer Awareness Fellowship. KAD was supported by a SURF fellowship. We thank Michael Zawrotny for helpful discussions and acknowledge Florida State University support of the Biological Science Imaging Resource. The Titan Krios was purchased partially by funds from NIH grant S10 RR025080 and the DE-20 direct electron detector was purchased by funds from NIH grant S10 OD018142. The composite Rio2-TAP and Ltv1-TAP map and individual maps from focused classification of the interface AFs, platform, and beak were deposited in the EMDB with IDs 8346, 8347, 8348, 8349, 8350, 8351, and 8353.

REFERENCES

- Amunts A, Brown A, Bai XC, Llacer JL, Hussain T, Emsley P, Long F, Murshudov G, Scheres SHW, Ramakrishnan V. Structure of the Yeast Mitochondrial Large Ribosomal Subunit. *Science*. 2014; 343:1485–1489. [PubMed: 24675956]
- Anger AM, Armache JP, Berninghausen O, Habeck M, Subklewe M, Wilson DN, Beckmann R. Structures of the human and Drosophila 80S ribosome. *Nature*. 2013; 497:80. [PubMed: 23636399]
- Armistead J, Triggs-Raine B. Diverse diseases from a ubiquitous process: The ribosomopathy paradox. *FEBS Letters*. 2014; 588:1491–1500. [PubMed: 24657617]
- Aylett CHS, Boehringer D, Erzberger JP, Schaefer T, Ban N. Structure of a Yeast 40S–eIF1–eIF1A–eIF3–eIF3j initiation complex. *Nat. Struct. Mol. Biol.* 2015; 22:269–271. [PubMed: 25664723]
- Ben-Shem A, de Loubresse NG, Melnikov S, Jenner L, Yusupova G, Yusupov M. The Structure of the Eukaryotic Ribosome at 3.0 angstrom Resolution. *Science*. 2011; 334:1524–1529. [PubMed: 22096102]
- Boehringer D, O’Farrell HC, Rife JP, Ban NN. Structural Insights into Methyltransferase KsgA Function in 30S Ribosomal Subunit Biogenesis. *Journal of Biological Chemistry*. 2012; 287:10453–10459. [PubMed: 22308031]
- Campbell MG, Karbstein K. Protein-Protein Interactions within Late Pre-40S Ribosomes. *Plos One*. 2011; 6:10.
- Cole SE, LaRiviere FJ, Merrih CN, Moore MJ. A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. *Mol Cell*. 2009; 34:440–450. [PubMed: 19481524]
- Culver GM. Assembly of the 30S ribosomal subunit. *Biopolymers*. 2003; 68:234–249. [PubMed: 12548626]
- de la Cruz J., Karbstein K., Woolford, JL. Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo. In: Kornberg, RD., editor. *Annual Review of Biochemistry*. Vol. 84. Palo Alto: Annual Reviews; 2015. p. 93-129.
- Dong A, Wu HHZ, Loppnau P, Sundstrom M, Arrowsmith C, Edwards A, Bochkarev AA, Plotnikov A. 2005 unpublished data.
- Drummond DA, Wilke CO. The evolutionary consequences of erroneous protein synthesis. *Nat Rev Genet*. 2009; 10:715–724. [PubMed: 19763154]
- Erzberger JP, Stengel F, Pellarin R, Zhang SY, Schaefer T, Aylett CHS, Cimermancic P, Boehringer D, Sali A, Aebersold R, et al. Molecular Architecture of the 40S center dot eIF1 center dot eIF3 Translation Initiation Complex. *Cell*. 2014; 158:1123–1135. [PubMed: 25171412]
- Fassio CA, Schofield BJ, Seiser RM, Johnson AW, Lycan DE. Dominant Mutations in the Late 40S Biogenesis Factor Ltv1 Affect Cytoplasmic Maturation of the Small Ribosomal Subunit in *Saccharomyces cerevisiae*. *Genetics*. 2010; 185:199–209. [PubMed: 20215468]
- Ferreira-Cerca S, Sagar V, Schafer T, Diop M, Wesseling AM, Lu HY, Chai E, Hurt E, LaRonde-LeBlanc N. ATPase-dependent role of the atypical kinase Rio2 on the evolving pre-40S ribosomal subunit. *Nat. Struct. Mol. Biol.* 2012; 19:1316. [PubMed: 23104056]

- Freed EF, Bleichert F, Dutca LM, Baserga SJ. When ribosomes go bad: diseases of ribosome biogenesis. *Mol Biosyst.* 2010; 6:481–493. [PubMed: 20174677]
- Gamalinda M, Woolford JL. Paradigms of ribosome synthesis: Lessons learned from ribosomal proteins. *Translation.* 2015; 3:11.
- Gelperin D, Horton L, Beckman J, Hensold J, Lemmon SK. Bms1p, a novel GTP-binding protein, and the related Tsr1p are required for distinct steps of 40S ribosome biogenesis in yeast. *Rna-a Publication of the Rna Society.* 2001; 7:1268–1283.
- Ghalei H, Schaub FX, Doherty JR, Noguchi Y, Roush WR, Cleveland JL, Stroupe ME, Karbstein K. Hrr25/CK1 delta-directed release of Ltv1 from pre-40S ribosomes is necessary for ribosome assembly and cell growth. *J. Cell Biol.* 2015; 208:745–759. [PubMed: 25778921]
- Grigorieff N. FREALIGN: High-resolution refinement of single particle structures. *J. Struct. Biol.* 2007; 157:117–125. [PubMed: 16828314]
- Henras AK, Plisson-Chastang C, O’Donohue MF, Chakraborty A, Gleizes PE. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev.-RNA.* 2015; 6:225–242. [PubMed: 25346433]
- Holzer S, Ban N, Klinge S. Crystal structure of the yeast ribosomal protein rpS3 in complex with its chaperone Yar1. *J Mol Biol.* 2013; 425:4154–4160. [PubMed: 24021814]
- Jia MZ, Horita S, Nagata K, Tanokura M. An Archaeal Dim2-Like Protein, aDim2p, Forms a Ternary Complex with a/eIF2 alpha and the 3’ End Fragment of 16S rRNA. *Journal of Molecular Biology.* 2010; 398:774–785. [PubMed: 20363226]
- Karbstein K. Inside the 40S ribosome assembly machinery. *Current Opinion in Chemical Biology.* 2011; 15:657–663. [PubMed: 21862385]
- Karbstein K. Quality control mechanisms during ribosome maturation. *Trends in cell biology.* 2013; 23:242–250. [PubMed: 23375955]
- Kastner B, Fischer N, Golas MM, Sander B, Dube P, Boehringer D, Hartmuth K, Deckert J, Hauer F, Wolf E, et al. GraFix: sample preparation for single-particle electron cryomicroscopy. *Nature Methods.* 2008; 5:53–55. [PubMed: 18157137]
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015; 10:845–858. [PubMed: 25950237]
- Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP. Structure of the human 80S ribosome. *Nature.* 2015; 520:640–U338. [PubMed: 25901680]
- Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EMEM density maps. *Nature Methods.* 2014; 11:63. [PubMed: 24213166]
- Lamanna AC, Karbstein K. Nob1 binds the single-stranded cleavage site D at the 3’-end of 18S rRNA with its PIN domain. *Proceedings of the National Academy of Sciences of the United States of America.* 2009; 106:14259–14264. [PubMed: 19706509]
- Lamanna AC, Karbstein K. An RNA Conformational Switch Regulates Pre-18S rRNA Cleavage. *Journal of Molecular Biology.* 2011; 405:3–17. [PubMed: 20934433]
- Lander GC, Stagg SM, Voss NR, Cheng A, Fellmann D, Pulokas J, Yoshioka C, Irving C, Mulder A, Lau PW, et al. Appion: An integrated, database-driven pipeline to facilitate EM image processing. *J. Struct. Biol.* 2009; 166:95–102. [PubMed: 19263523]
- Larburu N, Montellese C, O’Donohue MF, Kutay U, Gleizes PE, Plisson-Chastang C. Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. *Nucleic Acids Research.* 2016
- LaRonde-LeBlanc N, Guszczynski T, Copeland T, Wlodawer A. Autophosphorylation of *Archaeoglobus fulgidus* Rio2 and crystal structures of its nucleotide-metal ion complexes. *Febs Journal.* 2005; 272:2800–2810. [PubMed: 15943813]
- Lebaron S, Schneider C, van Nues RW, Swiatkowska A, Walsh D, Bottcher B, Granneman S, Watkins NJ, Tollervey D. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.* 2012; 19:744–753. [PubMed: 22751017]
- Llacer JL, Hussain T, Marler L, Aitken CE, Thakur A, Lorsch JR, Hinnebusch AG, Ramakrishnan V. Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Molecular Cell.* 2015; 59:399–412. [PubMed: 26212456]

- Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*. 1998; 14:953–961. [PubMed: 9717241]
- Lyumkis D, Brilot AF, Theobald DL, Grigorieff N. Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* 2013; 183:377–388. [PubMed: 23872434]
- McCaughan UM, Jayachandran U, Shchepachev V, Chen ZA, Rappsilber J, Tollervey D, Cook AG. Pre-40S ribosome biogenesis factor Tsr1 is an inactive structural mimic of translational GTPases. *Nat. Commun.* 2016; 7:8.
- Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* 2003; 142:334–347. [PubMed: 12781660]
- Mitterer V, Murat G, Rety S, Blaud M, Delbos L, Stanborough T, Bergler H, Leulliot N, Kressler D, Pertschy B. Sequential domain assembly of ribosomal protein S3 drives 40S subunit maturation. *Nat. Commun.* 2016; 7:15.
- Narla A, Ebert BL. Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. 2010; 115:3196–3205. [PubMed: 20194897]
- Nerurkar, P., Altvater, M., Gerhardy, S., Schutz, S., Fischer, U., Weirich, C., Panse, VG. Eukaryotic Ribosome Assembly and Nuclear Export. In: Jeon, KW., editor. *International Review of Cell and Molecular Biology*. Vol. 319. San Diego: Elsevier Academic Press Inc; 2015. p. 107-140.
- Passmore LA, Schmeing TM, Maag D, Applefield DJ, Acker MG, Algire MA, Lorsch JR, Ramakrishnan V. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell*. 2007; 26:41–50. [PubMed: 17434125]
- Pertschy B, Schneider C, Gnadig M, Schafer T, Tollervey D, Hurt E. RNA Helicase Prp43 and Its Co-factor Pfa1 Promote 20 to 18 S rRNA Processing Catalyzed by the Endonuclease Nob1. *Journal of Biological Chemistry*. 2009; 284:35079–35091. [PubMed: 19801658]
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. Ucsf chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004; 25:1605–1612. [PubMed: 15264254]
- Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 2005; 26:1781–1802. [PubMed: 16222654]
- Rabl J, Leibundgut M, Ataide SF, Ban N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science*. 2011; 331:730–736. [PubMed: 21205638]
- Roseman AM. FindEM - a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.* 2004; 145:91–99. [PubMed: 15065677]
- Schafer T, Maco B, Petfalski E, Tollervey D, Bottcher B, Aebi U, Hurt E. Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature*. 2006; 441:651–655. [PubMed: 16738661]
- Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 2012; 180:519–530. [PubMed: 23000701]
- Shrum DC, Woodruff BW, Stagg SM. Creating an infrastructure for high-throughput high-resolution cryogenic electron microscopy. *J. Struct. Biol.* 2012; 180:254–258. [PubMed: 22842049]
- Spear JM, Noble AJ, Xie Q, Sousa DR, Chapman MS, Stagg SM. The influence of frame alignment with dose compensation on the quality of single particle reconstructions. *J. Struct. Biol.* 2015; 192:196–203. [PubMed: 26391007]
- Strunk BS, Karbstein K. Powering through ribosome assembly. *Rna*-a Publication of the Rna Society. 2009; 15:2083–2104.
- Strunk BS, Loucks CR, Su M, Vashisth H, Cheng SS, Schilling J, Brooks CL, Karbstein K, Skiniotis G. Ribosome Assembly Factors Prevent Premature Translation Initiation by 40S Assembly Intermediates. *Science*. 2011; 333:1449–1453. [PubMed: 21835981]
- Strunk BS, Novak MN, Young CL, Karbstein K. A TranslationLike Cycle Is a Quality Control Checkpoint for Maturing 40S Ribosome Subunits. *Cell*. 2012; 150:111–121. [PubMed: 22770215]
- Talkish J, Biedka S, Jakovljevic J, Zhang JY, Tang L, Strahler JR, Andrews PC, Maddock JR, Woolford JL. Disruption of ribosome assembly in yeast blocks cotranscriptional pre-rRNA

- processing and affects the global hierarchy of ribosome biogenesis. *Rna*. 2016; 22:852–866. [PubMed: 27036125]
- Thammana P, Held WA. Methylation of 16S RNA during ribosome assembly in vitro. *Nature*. 1974; 251:682–686. [PubMed: 4610401]
- van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, et al. Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys*. 2000; 33:307–369. [PubMed: 11233408]
- Van Knippenberg PH, Van Kimmenade JMA, Heus HA. Phylogeny of the conserved 3' terminal structures of the RNA of small ribosomal subunits. *Nucleic Acids Research*. 1984; 12:2595–2604. [PubMed: 6709501]
- Vanrobays E, Gelugne JP, Caizergues-Ferrer M, Lafontaine DLJ. Dim2p, a KH-domain protein required for small ribosomal subunit synthesis. *Rna-a Publication of the Rna Society*. 2004; 10:645–656.
- Veith T, Martin R, Wurm JP, Weis BL, Duchardt-Ferner E, Saffenthal C, Hennig R, Mirus O, Bohnsack MT, Wohnert J, et al. Structural and functional analysis of the archaeal endonuclease Nob1. *Nucleic Acids Research*. 2012; 40:3259. [PubMed: 22156373]
- Woolford JL, Baserga SJ. Ribosome Biogenesis in the Yeast *Saccharomyces cerevisiae*. *Genetics*. 2013; 195:643–681. [PubMed: 24190922]
- Wools HA, Lamanna AC, Karbstein K. Roles of Dim2 in Ribosome Assembly. *Journal of Biological Chemistry*. 2011; 286:2578–2586. [PubMed: 21075849]
- Wu S, Tutuncuoglu B, Yan KG, Brown H, Zhang YX, Tan D, Gamalinda M, Yuan Y, Li ZF, Jakovljevic J, et al. Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature*. 2016; 534:133. [PubMed: 27251291]
- Zhou GJ, Zhang Y, Wang J, Guo JH, Ni J, Zhong ZM, Wang LQ, Dang YJ, Dai JF, Yu L. Cloning and characterization of a novel human RNA binding protein gene PNO1. *DNA Seq*. 2004; 15:219–224. [PubMed: 15497447]

Highlights

A 9-Å structure of pre-40S ribosomes shows the position of seven assembly factors.

Interface conformational heterogeneity suggests motions during 40S assembly.

The beak architecture shows how Rps3 binds in precursor molecules.

18S 3'-end formation may be regulated by Dim1 leaving the subunit interface.

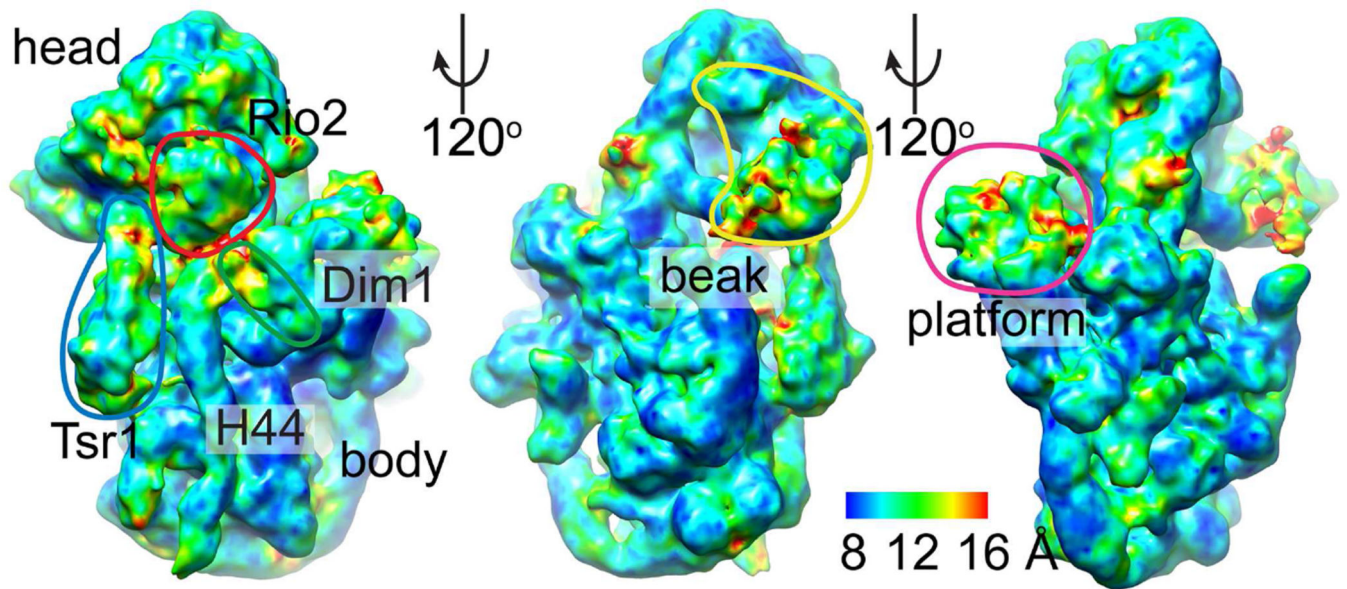


Fig. 1. Anatomy of pre-40S ribosomes

Late intermediate pre-40S ribosomes are not uniformly well ordered. Local resolution ranges from 8–16 Å. Resmap analysis (Kucukelbir et al., 2014) shows the rRNA core is resolved best, whereas peripheral AFs are less-well organized. See also Figure S1.

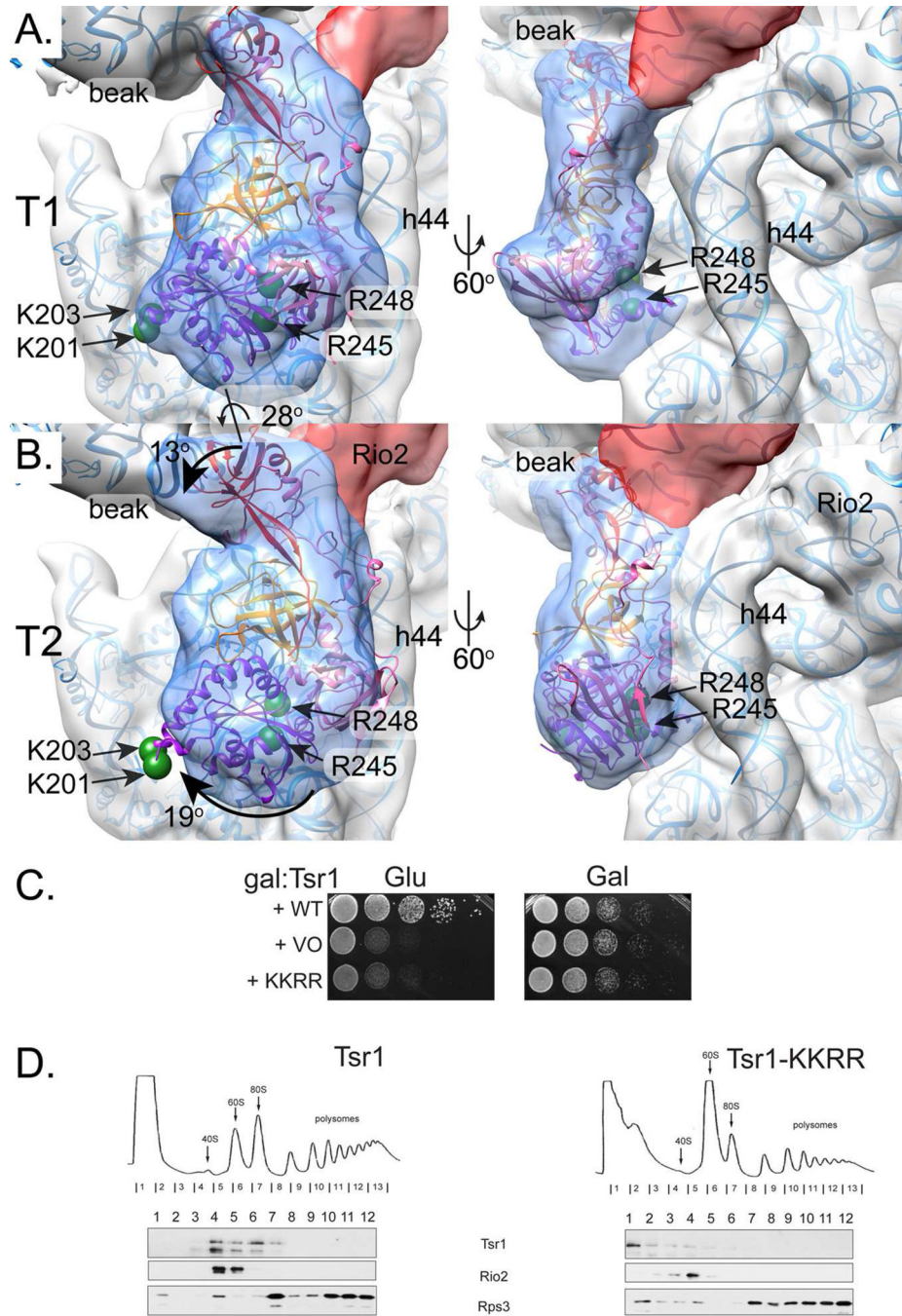


Fig. 2. Tsr1 adopts two discrete conformations

A. Local 3D classification revealed density shaped like Tsr1 (transparent blue) rotated away from h44 (subclass T1). Tsr1's X-ray crystal structure (PDB 5IW7, (McCaughan et al., 2016)) fit as a rigid body, positioning the four amino acids that are essential for its pre-40S binding and *in vivo* function on the ribosome-binding face of Tsr1 (two green spheres represent Ca atom position of R245 and R248) and at the edge of the molecule, far from h44 (two green spheres represent Ca atom position of K201 and K203). Tsr1 domains are colored per McCaughan et al., 2016: Tsr1^I is purple, Tsr1^{II} is pink, Tsr1^{III} is orange, and Tsr1^{IV} is red.

MDFF-relaxed rRNA and 40S ribosomal proteins are shown as blue ribbons to clarify the interpretation of the pre-40S structural features. See also Figure S2.

B. A second subclass of Tsr1 density, T2, shows Tsr1 rotated toward h44 by about 28° but the shape of the density did not correspond as well to the crystal structure. Rotation of Tsr1^I, Tsr1^{II}, and Tsr1^{IV} relative to Tsr1^{III}, colored as in A, would be needed to adjust Tsr1 from its T1-form. Arrows demark these suggested movements and the relevant amino acids are represented as in A. Angles were estimated by independently positioning each domain into the density corresponding to that feature and then calculating the angle relative to the rigid body fit with the “measure rotation” function in Chimera (Pettersen et al., 2004).

C. Growth of galactose-inducible/glucose-repressible Tsr1 (GAL::Tsr1) yeast cells carrying an empty vector (VO), or vectors coding wild-type (WT) or Tsr1_KKRR, is compared on glucose and galactose-containing plates.

D. Sucrose gradients of cell extracts from GAL::Tsr1 cells transformed with vectors carrying WT-Tsr1 or Tsr1_KKRR, grown in glucose for 16 h. Shown are absorbance profiles at 254 nm and Western blots for Tsr1, Rio2 and Rps3.

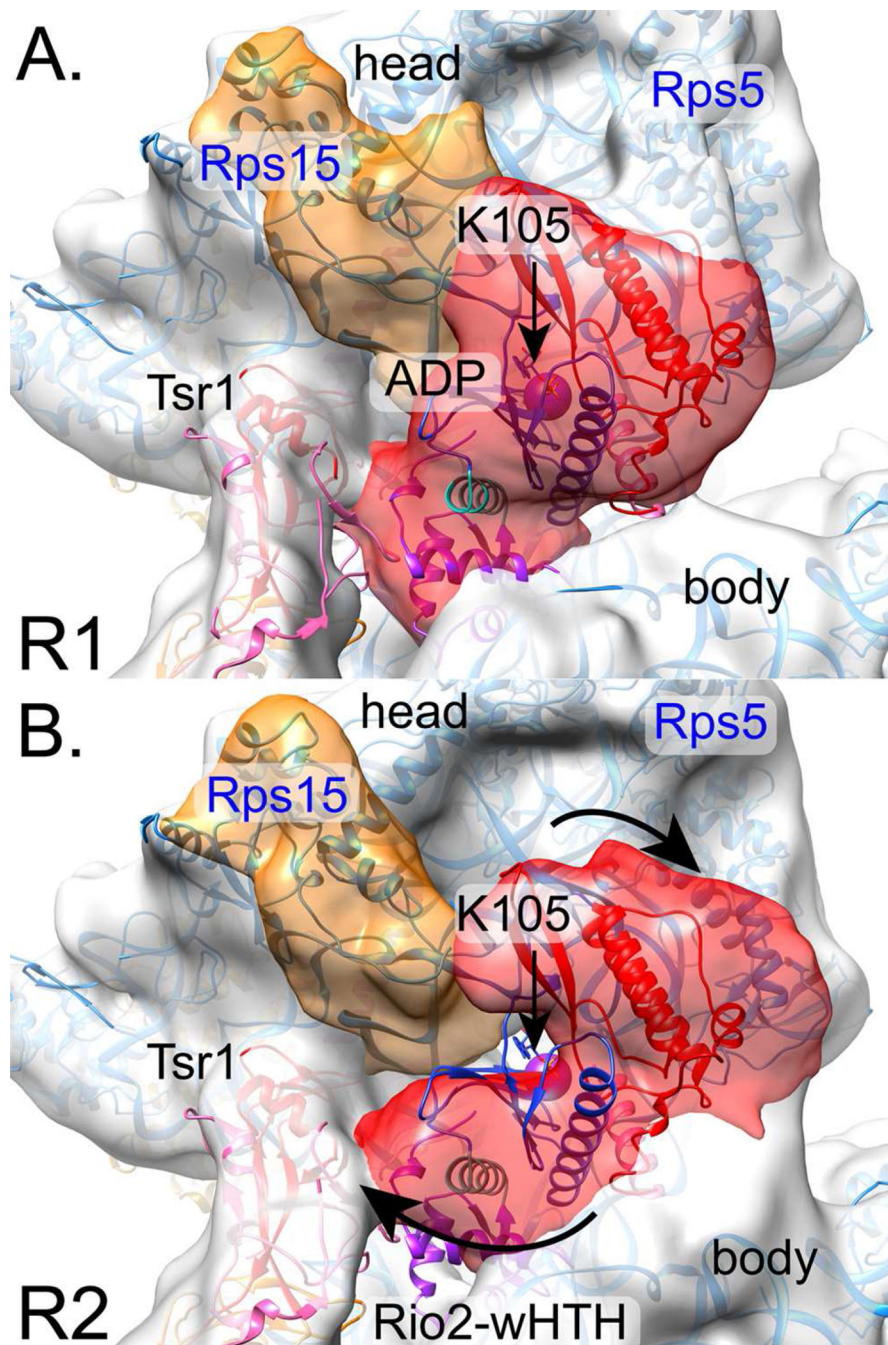


Fig. 3. Rio2 adopts multiple conformations, bound between the head and the body

A. Subclass R1 shows Rio2 as an extended, wedge-shaped density (red), with a small lobe close to Tsr1^{IV} and a larger lobe bridging the head and body. The X-ray crystal structure of *C. thermophilum* Rio2 homolog (PDB 4GYG, (Ferreira-Cerca et al., 2012)) fits well, in light of C-terminal sequence differences between *ct*Rio2 and *sc*Rio2. Rio2 is colored per Ferreira-Cerca et al., 2012: the N-terminal wHTH is purple, the N lobe is blue, the C-lobe is in red, and two linker helices are turquoise or pink. A purple sphere represents K105's Ca atom position.

B. Subclass R2 reveals density corresponding to Rio2 as a U-shape (red) into which the X-ray crystal structure of *ct*Rio2 does not fit well. To fit the density, the wHTH would need to rotate towards Tsr1^{IV}, while the C-lobe would need to rotate away from its central ATP binding pocket and towards the pre-40S body. Arrows demark these suggested movements. See also Figure S2.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

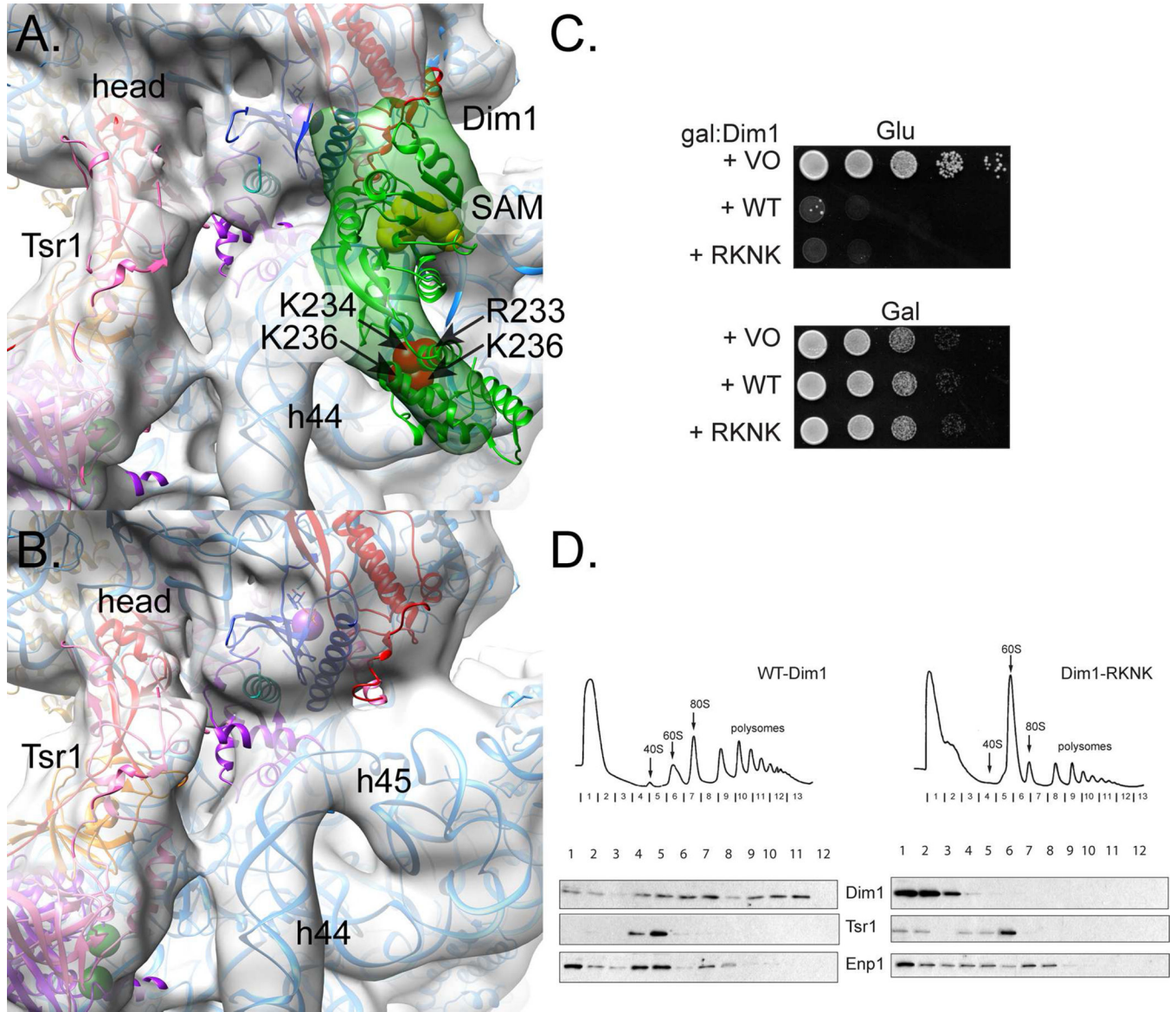


Fig. 4. Dim1 binding joins the head to the body

A. Local 3D classification of Dim1 reveals a single class with crescent-shaped density (green) that corresponds well to the X-ray crystal structure of human Dim1 (PDB 1ZQ9, (Dong et al., 2005). By orienting its N-terminus to Rio2, R233-K236 (four red spheres represent each C α atom position), which are together responsible for binding pre-40S, point to rRNA in the body of the small subunit. Yellow spheres for each SAM atom mark the active site.

B. A second class emerged from 3D classification that lacked Dim1, suggesting this AF is not integral to the integrity of the pre-40S subunit.

C. Growth of galactose-inducible/glucose-repressible Dim1 (GAL::Dim1) yeast cells carrying an empty vector (VO), or vectors coding wild-type (WT) or Dim1_RKNK is compared on glucose and galactose-containing plates.

D. Sucrose gradients of cell extracts from GAL::Dim1 cells transformed with vectors carrying WT-Dim1 or Dim1_RKNK, grown in glucose for 16 h. Shown are absorbance profiles at 254 nm and Western blots for Dim1, Enp1 and Tsr1. See also Figures S2 and S3.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

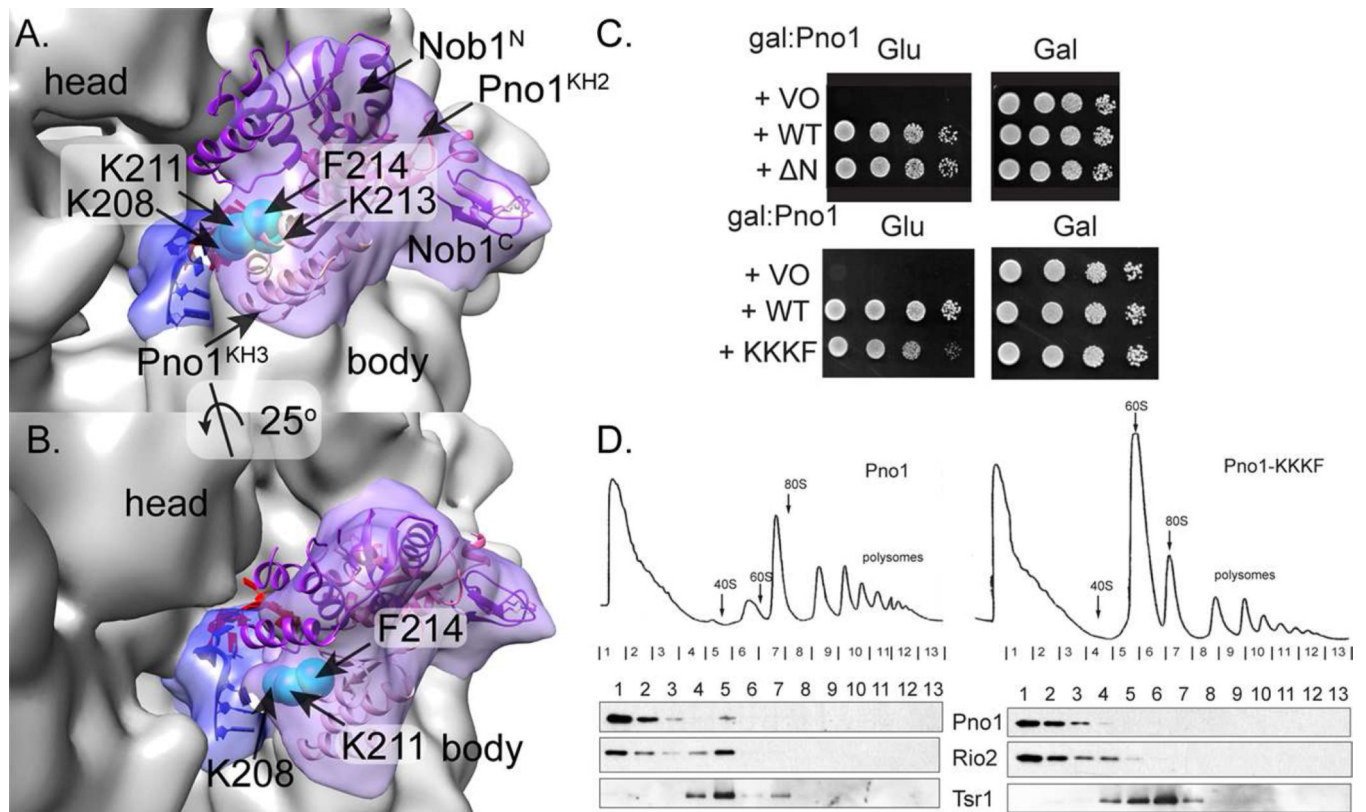


Fig. 5. The platform does not form a strongly pre-formed active site for ITS1 cleavage

A. A large domain near the head likely corresponds to the main Nob1 domain because it corresponds in size and shape to the core *P. horikoshii* homolog Nob1 structure whereas a smaller density at the periphery could represent its smaller C-terminal domain (purple ribbon, PDB 2LCQ, (Veith et al., 2012)). Nob1 sits on a C-shaped string of two Pno1 KH domains, modeled from the RNA-bound *P. horikoshii* KH domains (PDB 3AEV, (Jia et al., 2010)). KH2 is dark pink, KH3 is light pink, and bound RNA is blue. Each Ca atom of four amino acids mutagenized in Pno1 is a cyan sphere. The unaccounted density is blue and we modeled six ssRNA nucleotides from the Pno1 co-crystal structure to show the strong correlation between the shape of the empty density and this proposed assignment.

B. Our proposed position of the 3' end of the rRNA (blue) is not far from the mature position of those nucleotides (red).

C. Growth of galactose-inducible/glucose-repressible Pno1 (GAL::Pno1) yeast cells carrying an empty vector (VO), wild-type (WT), N Pno1 (top), or Pno1_KKKF (bottom) is compared on plates with glucose or galactose.

D. Sucrose gradients of cell extracts from GAL::Pno1 cells transformed with vectors carrying WT-Pno1 or Pno1_KKKF, grown in glucose for 16 h. Shown are absorbance profiles at 254 nm and Western blots for Pno1, Rio2 and Tsr1. See also Figure S5.

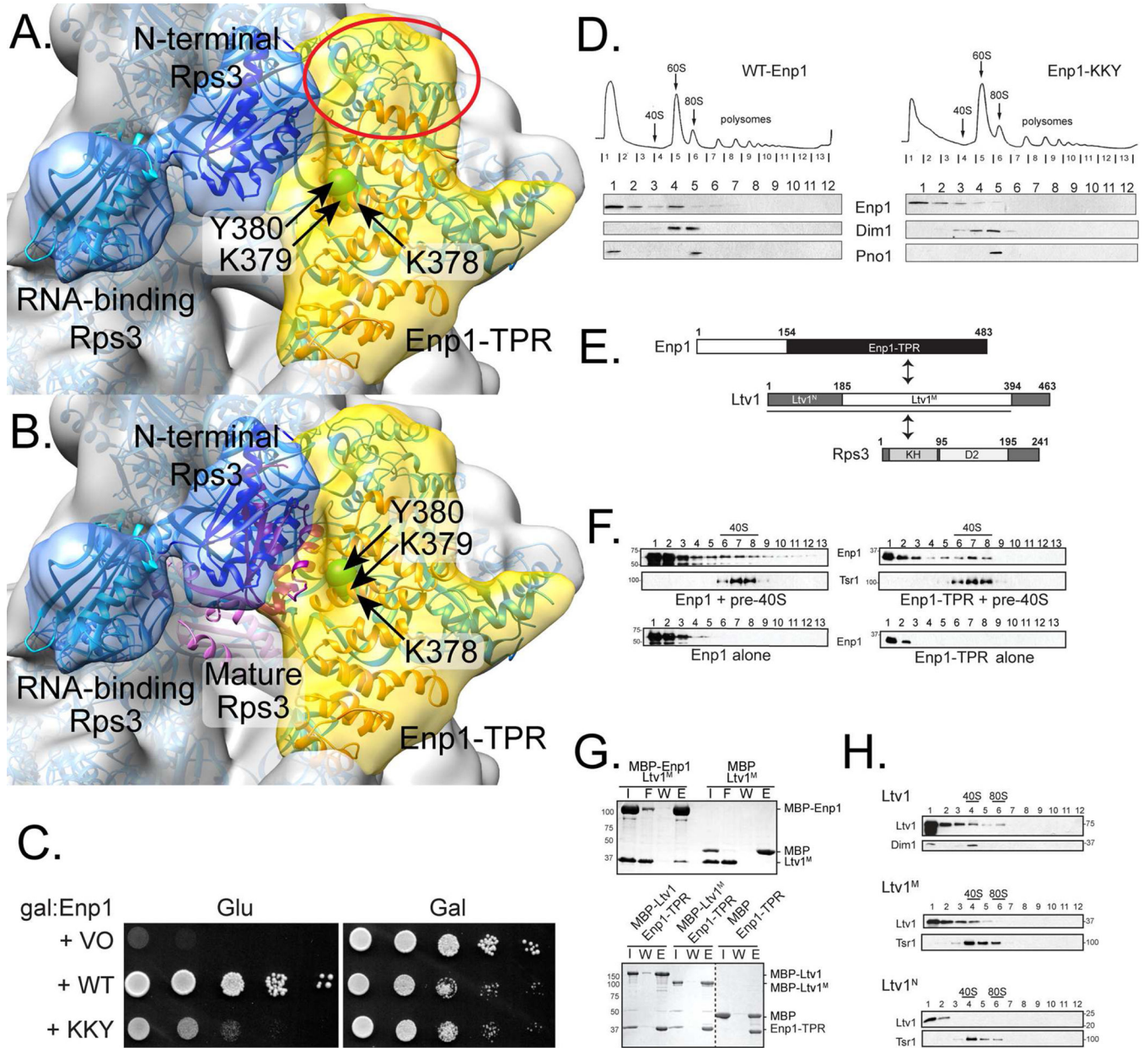


Fig. 6. Rps3 binds peripherally in pre-40S

A. Rps3 does not sit in its final position in pre-40S but can be visualized as a bow-shaped density bound at the edge of the beak (blue, modeled from Rps3 in the mature yeast 40S subunit structure PDB 4V88, (Ben-Shem et al., 2011)). Ltv1 and Enp1 are not structurally well characterized, but they likely fit in the remaining beak density (yellow, red oval) such that the Enp1-TPR motifs (orange ribbon, modeled in Phyre (Kelley et al., 2015)), which contain the essential amino acids K378-380 (a green sphere marks the position of each amino acid's C α atom), form the pre-40S binding site that bridges from the head to the body.

B. As in A, but showing the mature position of Rps3 from PDB 4V88, Ben-Shem et al., 2011 (pink ribbon).

- C. Growth of galactose-inducible/glucose-repressible Enp1 (GAL::Enp1) yeast cells carrying an empty vector (VO), or vectors coding wild-type (WT) or Enp1_KKY, is compared on glucose or galactose-containing plates.
- D. Sucrose gradients of cell extracts from GAL::Enp1 cells transformed with vectors carrying WT-Enp1 or Enp1_KKY, grown in glucose for 16 h. Shown are absorbance profiles at 254 nm and Western blots for Enp1, Dim1, and Pno1.
- E. Schematic representation of the interactions in the Enp1/Ltv1/Rps3 complex. Ltv1 core fragment (Ltv1^M) binds Rps3 and Enp1. Enp1-TPR binds Ltv1^M. Rps3's N-terminal KH domain interacts with Ltv1. Rps3 binds Ltv1^N and Ltv1^M.
- F. Enp1-TPR is sufficient for binding pre-40S. Western blot analyses of gradient fractions from Ltv1 ribosomes reconstituted with wild-type Enp1 or Enp1-TPR. Position of pre-40S, determined by absorbance at 254 nm, is indicated. Enp1 or Enp1-TPR migrated in top fractions of the gradients without ribosomes.
- G. Enp1 (top) and Enp1-TPR (bottom) bind to Ltv1^M. Coomassie-stained SDS-PAGE of protein binding assays. The fractions shown are I, input; W, the final wash; and E, eluted. The dotted line represents a lane of the gel that was irrelevant to the experiment so digitally deleted.
- H. Ltv1^M is sufficient for binding to pre-ribosomes. Sucrose gradients of total cell extracts from Ltv1 cells expressing WT-Ltv1, Ltv1^N or Ltv1^M. Western blots for Ltv1 and Dim1 are shown and the position of 40S and 80S ribosomes based on absorbance profiles is marked. See also Figures S6.

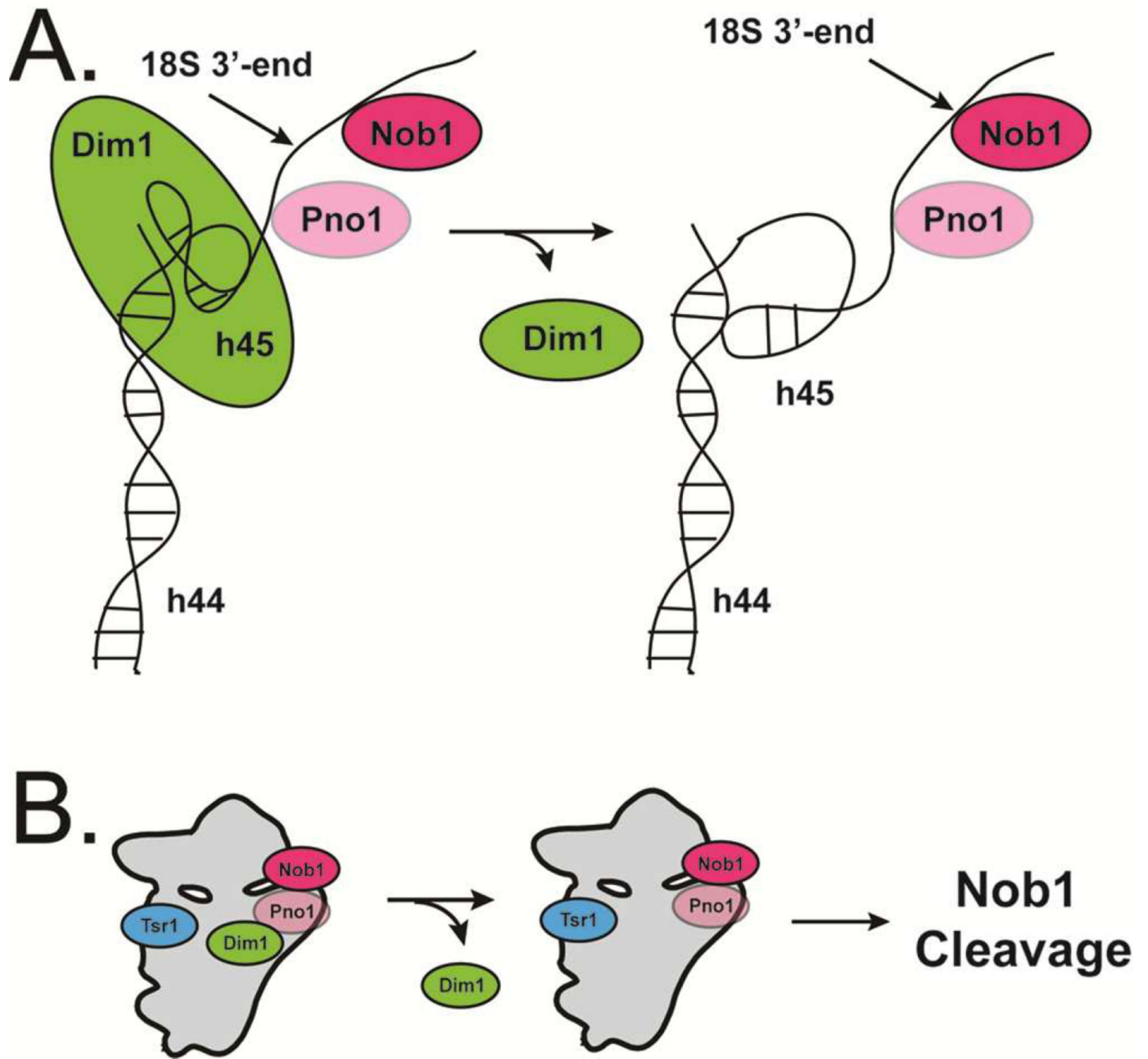


Fig. 7. A model for Dim1-regulated 18S rRNA maturation in pre-40S

In Dim1-containing pre-40S, h45 position is disrupted but upon Dim1 release, h45 repositions and the 3' end of the rRNA threads through pre-40S to the Nob1 active site.