Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Cryo-EM Analysis of Human
Small Ribosomal Subunit Biogenesis

Michael Franz Ameismeier

aus

Regensburg, Deutschland

2020

List of publications

Publication 1 – Ameismeier and Cheng *et al.* 2018

<u>Visualizing late states of human 40S ribosomal subunit maturation</u>
Michael Ameismeier[‡], Jingdong Cheng[‡], Otto Berninghausen and Roland Beckmann
Nature 558, 249-253, doi:10.1038/s41586-018-0193-0 (2018).

Publication 2 – Ameismeier *et al.* 2020

<u>Structural basis for the final steps of human 40S ribosome maturation</u>
Michael Ameismeier, Ivo Zemp, Jasmin van den Heuvel, Otto Berninghausen, Ulrike Kutay and Roland Beckmann
Nature 587, 683-687, doi: 10.1038/s41586-020-2929-x (2020)

Publication 3 – Thoms, Buschauer and Ameismeier *et al.* 2020

<u>Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2</u>
Matthias Thoms[‡], Robert Buschauer[‡], **Michael Ameismeier**[‡], Lennart Koepke, Timo Denk, Maximilian Hirschenberger, Hanna Kratzat, Manuel Hayn, Timur Mackens-Kiani, Jingdong Cheng, Jan H. Straub, Christina M. Stuerzel, Thomas Froehlich, Otto Berninghausen, Thomas Becker, Frank Kirchhoff, Konstantin M. J. Sparrer and Roland Beckmann
Science 369, 1249-1255, doi:10.1126/science.abc8665 (2020).

Publication 4 – Kratzat and Mackens-Kiani *et al.* 2020

<u>Structural inventory of native ribosomal ABCE1-43S pre-initiation complexes</u>
Hanna Kratzat[‡], Timur Mackens-Kiani[‡], **Michael Ameismeier**, Jingdong Cheng, Estelle Dacheux, Abdelkader Namane, Otto Berninghausen, Micheline Fromont-Racine, Thomas Becker and Roland Beckmann
bioRxiv, 2020.2007.2009.194902 (2020).
In revision at EMBO

Publication 5 – Cheng *et al.* 2020

<u>90S pre-ribosome transformation into the primordial 40S subunit</u>
Jingdong Cheng[‡], Benjamin Lau[‡], Giuseppe La Venuta, **Michael Ameismeier**, Otto Berninghausen, Ed Hurt and Roland Beckmann
Science 369, 1470-1476, doi: 10.1126/science.abb4119

[‡] These authors contributed equally

Contribution report

The work presented in this thesis has been conducted under the supervision of Prof. Dr. Roland Beckmann in his group at the Gene Center, Ludwig-Maximilian-University in Munich between September 2016 and August 2020. Besides listed authors, several people have been important to the successful publication of the results. Please be referred to the end of the document for further acknowledgments.

Publication 1 – Ameismeier and Cheng *et al.* 2018

In this study we have analyzed the formation of small ribosomal subunits in human cells using cryo-electron microscopy and single particle analysis. Based on five native pre-40S structures and one reconstituted particle we proposed a sequence of events that shape the 40S particle. I generated stable cell lines, performed the affinity purification of native complexes, prepared cryo-EM grids and assisted in sample screening. Furthermore, I conducted data processing of all native complexes and, together with co-first author Dr. Jingdong Cheng, built molecular models. I contributed to data analysis and interpretation, as well as manuscript writing and figure preparation.

Publication 2 – Ameismeier *et al.* 2020

In this study we have continued our analysis of 40S maturation in human cells and extended the sequence of maturation events by a set of five distinctive intermediates. We described very late events and showed the central functions of ribosome biogenesis factors RIOK1, PNO1 and NOB1, as well as discovered novel factors eIF1AD and LRRC47. Here, I conducted all experiments required for native complex purification and sample preparation. I assisted in cryo-EM data collection and performed all data processing and model building. In addition, I contributed to data analysis, data interpretation, and manuscript writing and prepared most figures.

Publication 3 – Thoms, Buschauer and Ameismeier *et al.* 2020

This study describes the structure and effect of SARS-CoV-2 Nsp1 bound to the small ribosomal subunit. Nsp1 blocks the mRNA entry tunnel, leading to a complete shutdown of translation and suppression of cellular anti-viral responses. Here, I assisted in sample preparation, performed data analysis of all native complexes, and aided to molecular modelling. Furthermore, I contributed to data interpretation, figure preparation and manuscript writing.

Publication 4 – Kratzat and Mackens-Kiani *et al.* 2020

This publication describes an inventory of native human and yeast translation initiation complexes containing the ribosome splitting factor ABCE1 and provides high-resolution insight into the structure of the human initiation factor complexes eIF2 and eIF3. Here, I have prepared the samples of human initiation complexes and conducted the analysis of cryo-EM data.

Publication 5 – Cheng *et al.* 2020

This study analyzes the transition from 90S to early 40S assembly intermediates in yeast. In contrast to previous models, the presented structures indicate a successive shedding of assembly factor modules instead of the joint release as a 5' ETS particle. Here, I assisted in manuscript writing and figure preparation.

Summary

Ribosomes are key components of the cellular machinery which carries out protein synthesis. These highly sophisticated macromolecular complexes consist of two subunits, each comprising ribosomal ribonucleic acid (rRNA) and proteins. Due to their central biological role and complex structure, formation of new ribosomal subunits is a long and complicated process. It starts in the nucleolus with the transcription of rRNA and finishes in the cytoplasm. Several hundred protein and RNA factors are involved along this path and ensure correct modification, folding and processing of rRNA, as well as incorporation of ribosomal proteins. This process has been extensively studied in many different organisms ranging from bacteria to higher eukaryotes, while most insights were gained from the model organism *Saccharomyces cerevisiae*. However, human ribosome biogenesis has been shown to be different in various ways.

The results presented in this cumulative dissertation provided the first detailed insights into biogenesis intermediates of the human small ribosomal subunit. In two publications, a total of ten structurally distinct assembly intermediates were described, which were extracted from human cell culture via affinity purification and subsequently solved by cryo-electron microscopy.

The first publication covers compositional and conformational changes that occur at a late nuclear and early cytoplasmic stage. Five structures of native 40S intermediates, termed state A to E, showed a stepwise maturation of mainly the 3' major domain and the position of ribosomal assembly factors RRP12, PNO1, ENP1, LTV1, TSR1, NOB1, RIOK2, RIOK1, BUD23 and TRMT112. The earliest structure, state A, still lacked ribosomal proteins uS2, uS5 and eS21 for a complete 40S body, as well as the endonuclease NOB1. Also, helices 35 − 40 of the 3' major domain were coordinated in an immature position by RRP12 and the decoding center was occupied by the methyltransferase BUD23 and its binding partner TRMT112. In state B, helices 35 − 37 have moved to their designated positions, which allowed the recruitment of NOB1, uS2, uS5 and eS21. Concomitantly, BUD23 and TRMT112 have dissociated from the particle. Release of RRP12 and a 140° movement of helix 39 and 40 marked the transition to state C, during which the atypical kinase RIOK2 has also bound to the intersubunit side. In state D, uS3, eS10, uS10 and uS14 have bound to the beak at a position previously occupied by ENP1 and the N-terminal part of LTV1. Finally, in state E, a so far unidentified factor had replaced the C-terminus of RIOK2 within the mRNA entry channel. Throughout the described sequence, the 3' end of the 18S rRNA precursor was tightly coordinated by PNO1 and NOB1. This kept the designated endonucleolytic cleavage site at distance from the active center of NOB1, which prevented a premature processing at site 3.

The second publication extended the series described in the first by five additional structures of small subunit assembly intermediates. Here, we provided a detailed description of the final steps in 18S rRNA precursor processing. The first new structure, state F1, shows the kinase RIOK1 bound to the top of helix 44 in an outward facing position. While PNO1 and NOB1 are still bound, TSR1 and LTV1 have left the particle during the transition from state E. In addition, novel factor LRRC47 has bound to the intersubunit side of the particle, with its C-terminal domain occupying partially the position of TSR1. In a two-step sequence, RIOK1 repositioned inward and finally settled at the decoding center, similar to RIOK2. During this transition, a second novel factor, eIF1AD has bound next to RIOK1. Also, helix 44 has finally matured and PNO1 had been displaced from the particle. This allowed NOB1 to reposition its PIN domain and active

center above the substrate. In state G, an endonuclease deficient mutant of NOB1 is thus trapped in its active conformation, providing insights into the final processing step of 18S rRNA. Finally, two additional assembly intermediates show the sequential release of NOB1, which is replaced by ribosomal protein eS26, and LRRC47. Only the release of assembly factors RIOK1 and eIF1AD remains for 40S maturation to complete.

In conclusion, high-resolution structures obtained by cryo-electron microscopy enabled unprecedented insight into late 40S maturation and offered a wealth of information on the assembly factors involved. Several comprehensive principles, such as coordination of large-scale rRNA rearrangements by biogenesis factors, were derived and allowed drawing parallels to 90S and 60S assembly pathways. Together, the presented results provide the foundation for future work, both for further experimental elucidation of assembly factor function and for structural analysis of other human pre-ribosomal particles. Finally, in the light of an increasing number of ribosome assembly associated diseases, termed ribosomopathies, and the central role of ribosome formation in several types of cancer, high-resolution information on maturation intermediates enables for the first time the development of a targeted, structure-driven intervention strategy.

List of abbreviations

AF	Assembly factor
ATP	Adenosine triphosphate
CP	Central protuberance
CRACK	UV crosslinking and complementary DNA analysis
cryo-EM	Cryo-electron microscopy
DC	Decoding center
DNA	Deoxyribonucleic acid
ES	Expansion segments
ETS	External transcribed spacers
FISH	Fluorescence *in situ* hybridization
GAC	GTPase-associated center
GDP	Guanosine diphosphate
GTP	Guanosine triphosphate
IC	Initiation complex
IGS	Intergenic spacer
ISS	Intersubunit space
ITS	Internal transcribed spacer
LR-NES	Leucine-rich NES
LSU	Large subunit
MAP	Mitogen-activated protein
NES	Nuclear export signal
NOR	Nucleolar organizing region
NPC	Nuclear pore complex
ORF	Open reading frame
PIC	Pre-initiation complex
PTC	Peptidyl transferase center
RNA	Ribonucleic acid
mRNA	Messenger RNA
rRNA	Ribosomal RNA
snoRNA	Small nucleolar RNA
snoRNP	Small nucleolar ribonucleoprotein
snRNA	Small nuclear RNA
tRNA	Transfer RNA
RP	Ribosomal protein
RQC	Ribosome-associated quality control
SPA	Single-particle analysis
SSU	Small subunit
UPR	Unfolded protein response
UV	Ultraviolet

# Table of contents

List of figures

1 Introduction

Across the three domains of life, the flow of genetic information follows the 'central dogma of molecular biology' (Crick, 1970; Crick, 1958). With some modifications to its original meaning, this principle summarizes the processes that are fundamental to all biological systems: replication, transcription and translation (see Berg et al., 2019 for a general overview). During replication, two identical copies of deoxyribonucleic acids (DNA) strands are produced from a single entity, thereby duplicating the information stored within these large molecules and allowing a propagation of such information to the next generation of cells. During transcription certain DNA segments termed genes are then transcribed into a related type of macromolecules called RNA. Depending on the gene, transcription can result in a variety of different RNA molecules that can be categorized as either coding or non-coding. Coding RNA molecules are called messenger RNA (mRNA) and contain information on a protein's amino acid sequence. Non-protein-coding or non-coding RNAs, on the other hand, do not serve as a template for protein synthesis. This category rather comprises a large diversity of RNA molecules with varying lengths and functions, including transfer (tRNA), ribosomal, small nuclear (snRNA) and small nucleolar RNA (snoRNA). Finally, during translation, a new protein is produced by the ribosome. One amino acid after another is brought to the ribosome by their tRNA and added to a growing polypeptide chain. By matching the right tRNA to nucleotide triplets within the protein encoding segment (open reading frame, ORF), called codons (Crick et al., 1961), the mRNA is gradually decoded.

1.1 Overview of ribosome structure and function

Ribosomes are macromolecular complexes and the central component to protein synthesis in all biological systems. They consist of two structural and functional subunits that together catalyze the stepwise assembly of proteins (Figure 1A). Each subunit contains ribosomal proteins (RPs) and RNAs, with their overall number, structure and size being dependent on the organism (Korobeinikova et al., 2012). Ribosomes and their subunits are commonly described by their respective sedimentation coefficient, which is determined by the particle's size, density, and shape and is measure in Svedberg (S) units. The prokaryotic ribosomes from bacteria and archaea have a Svedberg value of 70S and consist of a small 30S and a large 50S subunit. In eukaryotes, higher values of 80S for the fully assembled ribosome and 40S and 60S for the small and large subunit, respectively, reflect an increase in size during evolution (Fox, 2010). In addition to cytosolic ribosomes, eukaryotic cells can contain a variety of additional types within their organelles, such as a 55S mitochondrial (Greber et al., 2015) or a 70S chloroplast ribosome (Graf et al., 2017). In accordance with the endosymbiotic theory, these ribosomes have evolved from their prokaryotic ancestors.

Each ribosomal subunit can be further dissected into its protein and RNA components. For example, in *Escherichia coli*, a common bacterial model organism, the small 30S subunit contains a 16S rRNA and 21 ribosomal proteins, while the large 50S subunit is made of a 5S and 23S rRNA and 33 proteins (Kaczanowska and Ryden-Aulin, 2007). Archaeal ribosomes resemble those of bacterial origin in terms of overall (30S and 50S) and rRNA (16S, 5S and 23S) size, but already contain some additional, specific RPs (Armache et al., 2013). And finally, containing a fourth rRNA and several specific proteins, eukaryotic 80S ribosomes are larger than their prokaryotic counterpart (Melnikov et al., 2012). In humans, the small 40S subunit consists of the 18S rRNA and 33 proteins, while the large 60S subunit contains the 28S, 5S and 5.8S rRNA, as well as

47 proteins (Anger et al., 2013). In contrast to archaea, overall ribosome composition is well conserved in eukaryotes, with only a few exceptions, such as the lack of eL28 in the yeast *Saccharomyces cerevisiae* or an additional acidic ribosomal phosphoprotein (P3) in plants (Lecompte et al., 2002; Szick et al., 1998). Differences also exist in the length of individual rRNAs, resulting in a 25S rRNA in the large subunit of *S. cerevisiae* and 28S rRNA in those from humans.
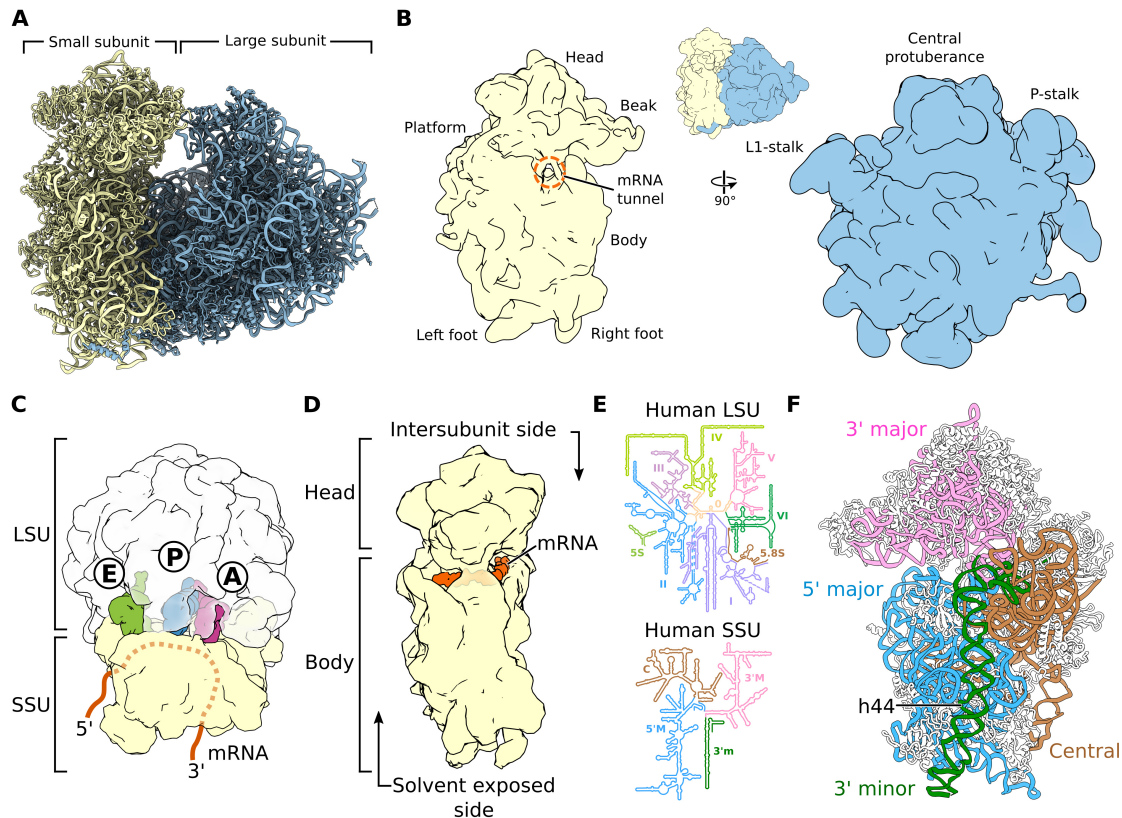
Figure 1 | Overview of structural features of the ribosome. (A) Model representation of a human 80S ribosome with the small subunit colored in yellow and the large subunit in blue (PDB 6Y0G). (B) Volume of a human 80S ribosome (EMDB 10668) and both subunits in their classical views. Major structural features are labelled. (C) Surface representation of an 80S ribosome (PDB 6TNU and 6R5Q) with A, P, and E-site tRNA. mRNA path indicated in orange. (D) Surface representation of a human small subunit with the bound mRNA highlighted in orange (PDB 6Y0G). (E) Secondary structure diagram of the human large (top) and small (bottom) subunit with all domains colored and labelled (adapted from http://apollo.chemistry.gatech.edu/RibosomeGallery). (F) Model representation of the human 40S with the 18S rRNA colored according to its domains. rRNA h44 of the 3' minor domain labelled.

Several additional structural hallmarks of the ribosome are distinguishable in the subunit's classical views (Figure 1B). The SSU can be divided into a main body and its head (Korn, 1980). The body contains a left and right foot, as well as the platform, while the head displays a protrusion called the beak. The LSU is often depicted in the 'crown view', where the central protuberance (CP), L1-stalk and P-stalk (L7/L12-stalk in prokaryotes) are visible as characteristic features. Both subunits have also several rRNA helices, called expansion segments (ES) that often extend from the particle. The length of these rRNA helices are species

dependent and account for most of the rRNA size differences between bacterial and eukaryotic ribosomes (Wilson and Doudna Cate, 2012).

Despite such differences in size and composition, the ribosomal core is remarkably well conserved. Several studies over the last decades have elucidated the structure of bacterial, archaeal and eukaryotic ribosomes or their subunits using X-ray crystallography (Ban et al., 1998; Ban et al., 2000; Schluenzen et al., 2000; Tocilj et al., 1999; Wimberly et al., 2000) and cryo-electron microscopy (cryo-EM) and single-particle analysis (SPA; Agrawal et al., 1996; Anger et al., 2013; Frank et al., 1995; Greber et al., 2012a; Khatter et al., 2015; Khusainov et al., 2016). In fact, due to their size, ease of purification and high biological relevance, ribosomes and their subunits have been a popular sample for structure determination and method development well over the past 40 years and countless publications cover diverse aspects of ribosome function (see Brown and Shao, 2018 for a review). The resulting models have revealed an overall conservation of functional sites within the small and large subunit. The small subunit (SSU) contains the mRNA binding site, a cavity that runs perpendicular to the long side of the complex and links both faces of the particle (see Figure 1, C and D). A part of the mRNA tunnel forms the decoding center (DC), where tRNA molecules are matched to each codon during translation, effectively decoding the mRNA sequence and translating the genetic code into a chain of amino acids. In assembled ribosomes, the small subunit is bound to the large subunit (LSU) at their intersubunit sides by several distinct bridging elements (Ben-Shem et al., 2011; Liu and Fredrick, 2016). In the LSU and close to the decoding center of the SSU is the peptidyl transferase center (PTC), the catalytic core of the ribosome, where every amino acid is added from its tRNA to the growing peptide chain. Early high resolution structural data has shown that the catalytic center is composed of rRNA segments with the nearest protein residing approximately 18 angstroms (Å) away (Ban et al., 2000; Hansen et al., 2002; Nissen et al., 2000; Schluenzen et al., 2000; Schmeing et al., 2002). The ribosome is therefore a prominent example of a ribozyme, an enzyme whose catalytic function is mediated by RNA moieties. From the PTC, the nascent polypeptide chain follows a path through the LSU of approximately 100 Å called the exit tunnel, before finally exiting the ribosome at the solvent exposed side (Frank et al., 1995; Voss et al., 2006). Three different tRNA binding sites, called A, P and E-site, are located within the intersubunit space (ISS; Agrawal et al., 1996). The A-site marks the entry position of aminoacyl-tRNAs during the elongation phase and is the position where the cognate tRNA is paired via its anticodon to the codon on the mRNA. The P-site contains the peptidyl-tRNA with the growing polypeptide chain and the E-site a deacylated tRNA, which is ready to be released from the ribosome.

Based on three-dimensional structures of the ribosome, a secondary structure map of each rRNA can be generated, highlighting the interaction between rRNAs and their domains (Figure 1E). The rRNA of the large subunit is divided into six distinct domains (domains I to VI), which are all clustered around a central domain 0 (Ban et al., 2000). The rRNA of the small subunit can be segmented into four parts: the 5' and central domains, which make up a large portion of the body; the 3' major domain, which constitutes the head; and the 3' minor domain, of which a large part forms the elongated helix 44 (h44) that runs across the intersubunit side (Figure 1F; Schluenzen et al., 2000).

Fully matured ribosomal RNA contains several nucleosides whose ribose or base moiety has been chemically modified. Indeed, with more than 2% of altered residues, rRNA is the second most modified type of RNA in eukaryotes after tRNAs (Sloan et al., 2017). While over 160 different kinds of RNA

modification have been described so far (Boccaletto et al., 2018), ribosomes contain a rather limited set. The most common modifications in eukaryotes are the attachment of a methyl group to the 2' hydroxyl group of a ribose moiety and the isomerization of uridine to pseudouridine (Ψ; Sloan et al., 2017). These modifications are introduced, with a few exceptions, by a multi-protein complex guided by a snoRNA, which in its entirety is termed small nucleolar ribonucleoprotein (snoRNP). Eukaryotic snoRNPs can be classified into two sets: box C/D snoRNPs, which are responsible for most 2'-*O*-methylations and box H/ACA snoRNPs, which carry out pseudouridylation (Watkins and Bohnsack, 2012). Besides the snoRNA, box C/D snoRNPs contain the proteins Snu13 (15.5 K in humans), Nop56, Nop58 and Nop1 (fibrillarin in humans), of which the latter is the methyltransferase (Aittaleb et al., 2003; Singh et al., 2008). In case of box H/ACA snoRNPs, associated proteins are Nhp2, Nop10, Gar1 and the pseudouridine synthase Cbf5 (dyskerin in humans; Lafontaine et al., 1998a). In addition to snoRNA guided complexes, a few stand-alone enzymes are known to introduce rRNA base modifications at specific positions. In yeast, six of such bases have been identified on the small and large subunit each, comprising mono- or di-methylated pyrimidine and purine rings and acetylated or aminocarboxypropylated pyrimidine rings (see Sharma and Lafontaine, 2015 for a review). While all six modifications of the small subunit are highly conserved in humans, only three of the base methylations are present in human large ribosomal subunits (Sloan et al., 2017). Interestingly, many modifications are found around functionally important sites and are required for optimal translational efficiency (Baxter-Roshek et al., 2007; Liang et al., 2007, 2009). Although deletion of a single or even multiple modifications has no effect on cell viability and ribosome function, mutations in or lack of core snoRNP proteins result in severe growth defects (Gautier et al., 1997; Schimmang et al., 1989; Tollervey et al., 1991; Tollervey et al., 1993). Modified residues are therefore important for the overall stability by contributing with facilitated base stacking or additional hydrogen bonds.

1.2   The eukaryotic translation cycle

In prokaryotes and eukaryotes alike, protein synthesis occurs in four steps: translation initiation, elongation, termination, and recycling. During initiation, components of the translational machinery are assembled, and the starting position is identified on the mRNA. In the elongation phase the gene is subsequently translated codon by codon into a polypeptide chain. Once a stop codon is reached, protein synthesis is terminated and the new protein released, before the ribosome and all remaining factors are recycled (for reviews see Dever and Green, 2012; Hinnebusch and Lorsch, 2012; Jackson et al., 2010; Rodnina, 2018).

In eukaryotes translation initiation is more complicated than its bacterial counterpart and requires the concerted action of at least 12 initiation factors (reviewed in Hinnebusch, 2017; Hinnebusch and Lorsch, 2012; Jackson et al., 2010). A defining feature of the canonical process thereof is the scanning mechanism by which the mRNA sequence is probed base by base from the 5' to the 3' end to find the first start codon, most often the triplet AUG (Hinnebusch, 2014). During initiation, a 43S pre-initiation complex (PIC) forms first, containing the small ribosomal subunit and the trimeric complex, which consists of eIF2, the initiator tRNA (Met-tRNA$_i^{Met}$) and guanosine triphosphate (GTP; Kapp and Lorsch, 2004; Kolitz and Lorsch, 2010). Formation of the 43S PIC is promoted by initiation factors eIF1, eIF1A, eIF3 and eIF5 (Algire et al., 2002; Asano et al., 2001; Majumdar et al., 2003). The fully assembled 43S PIC subsequently engages with mRNA which was preactivated by the tripartite initiation factor complex eIF4F, consisting of eIF4A, eIF4E and

eIF4G (Hinnebusch and Lorsch, 2012), which leads to the formation of the 48S initiation complex (IC; Korneeva et al., 2000; LeFebvre et al., 2006; Methot et al., 1996; Mitchell et al., 2010). The mRNA is subsequently scanned base by base for a matching start codon. In most cases, this is an AUG triplet embedded in the so-called Kozak sequence (purine at position -3 and guanine following the AUG; Kozak, 1986; Pisarev et al., 2006). In a last step, eIF2-GDP and eIF5 are released from the PIC and the 60S subunit is recruited to the complex by eIF1A and the GTPase eIF5B (Acker et al., 2006; Pestova et al., 2000). This requires the interaction of eIF5B-GTP with the C-terminal DIDDI-motif of eIF1A (Marintchev et al., 2003; Olsen et al., 2003). Subunit joining subsequently triggers GTP hydrolysis by eIF5B, resulting in its release from the complex and conformational changes that ultimately lead to the dissociation of eIF1A (Acker et al., 2009; Lee et al., 2002).

Once the formation of the 80S IC is completed, translation elongation can start. The decoding of mRNA is a three-step mechanism: first, cognate aminoacyl-tRNA is brought to the complex by the eukaryotic elongation factor 1A (eEF1A; Jakobsson et al., 2018). Upon correct codon recognition at the A-site, eEF1A is released, allowing the aminoacyl-tRNA to fully accommodate (Gromadski et al., 2006; Liu et al., 2015; Lorsch and Herschlag, 1998; Pape et al., 1999). Next, the tRNA acceptor stem repositions to the PTC and the nascent polypeptide chain is transferred onto the new amino acid (Schmeing et al., 2005a; Schmeing et al., 2005b; Voorhees et al., 2009). And finally, following peptidyl transfer, the peptidyl-tRNA is translocated into the P-site, which vacates the A-site for a new cycle (Cornish et al., 2008; Frank and Agrawal, 2000). The elongation cycle with tRNA delivery and recognition, peptidyl transferase reaction and finally mRNA and tRNA translocation usually occurs until a suitable stop codon is reached.

In most cases, no tRNA matches any of the three potential stop codons UAA, UAG and UGA. Instead, stop codons are recognized by termination factors, leading to translation termination and the release of the newly synthesized protein (Brenner et al., 1965; Caskey et al., 1968). In eukaryotes, two termination (or release) factors, eRF1 and eRF3, are required for stop codon recognition and peptide release. Following stop codon recognition by eRF1 and GTP hydrolysis by eRF3, conformational changes places a conserved GGQ motif of eRF1 within the PTC, which results in a nucleophilic attack on the peptidyl-tRNA ester bond by a coordinated water (Jin et al., 2010; Matheisl et al., 2015). This leads to hydrolysis and release of the nascent peptide.

Following translation termination and the release of the nascent protein, post-termination ribosomes still contain mRNA, a deacylated tRNA at the P-site and at least release factor eRF1. Therefore, before reengagement of these components can occur in a new cycle of translation, subunit dissociation is required. In termination or pre-recycling complexes the ATP-binding cassette protein ABCE1 (Rli1 in yeast) binds the 80S ribosome at the intersubunit space near the GTPase-associated center (GAC), a common binding site of translational GTPases such as EF-G, eEF2 and eRF3 (Preis et al., 2014; Shao et al., 2016). However, despite extensive structural insights, the exact mechanism of subsequent ribosome splitting by ABCE1 remains unclear (Becker et al., 2012; Franckenberg et al., 2012; Heuer et al., 2017a; Preis et al., 2014). Upon subunit dissociation, mRNA and P-site tRNA remain bound to the small subunit and are released in a final step by translation initiation factors eIF1A, eIF1 and eIF3 with its subunit eIF3j (Pisarev et al., 2007; Pisarev et al., 2010).

1.3   A ribosome's central role in biological processes

As the main component of the translational machinery, the ribosome has a central role in protein homeostasis (proteostasis) and is integrated into a complex regulatory system that orchestrates the selective production, folding, translocation and degradation of proteins required by the cell under any internal and environmental circumstances (Balch et al., 2008; Diaz-Villanueva et al., 2015). Correct and controlled protein synthesis is an essential part thereof and protein translation can be altered specifically or globally. Often posttranslational modifications such as phosphorylation regulate key player activity in translation (Sonenberg and Hinnebusch, 2009). A well-studied example is the phosphorylation of Ser51 of eIF2, which leads to the sequestering of GEF eIF2B and reduced TC formation. Four serine kinases are able to carry out phosphorylation under different kinds of cellular stress: (I) GCN2 during amino acid starvation, (II) PERK in context of the unfolded protein response (UPR) pathway, (III) PKR at the presence of double stranded RNA and (IV) HRI during heme deprivation (Dever, 1997; Shao et al., 2001; Su et al., 2008; Sudhakar et al., 2000; Vattem et al., 2001; Zhang et al., 2006). Translation initiation is further regulated by blocking access of eIF4E to the 5' cap of mRNAs through interactions with eIF4E binding proteins (4E-BPs). Phosphorylation of 4E-BP1 by the central nutrient sensor mammalian target of rapamycin (mTOR) reduces its eIF4E binding capacity and thus increases cellular translation initiation rate (Gingras et al., 1999; Raught and Gingras, 1999; Showkat et al., 2014). These are just two prominent examples of highly complex regulatory mechanisms embedding protein translation into the cellular response to internal or external steady-state perturbation (Dever et al., 2016).

Apart from the rather systemic responses, ribosomes are also central to individual mRNA and protein quality control mechanisms to prevent aberrant protein formation. Different defects can lead to ribosome stalling on the mRNA, eventually requiring rescue. Examples include mRNAs with premature stop codons, truncated mRNAs, unresolved mRNA secondary structures or damages, as well as shortage of aminoacylated tRNA (see recent reviews Karamyshev and Karamysheva, 2018; Simms et al., 2017). A multitude of factors are able to detect such defects and resolve translational arrest in a process called ribosome-associated quality control (RQC), often leading to mRNA, tRNA or nascent peptide degradation (Brandman et al., 2012; Frischmeyer et al., 2002; Klauer and van Hoof, 2012; Su et al., 2019; van Hoof et al., 2002). RQC is complemented by co-translational quality control mechanisms, which surveil the nascent peptide exiting the ribosome. Various factors regulate nascent peptide modifications, folding, translocation or degradation (see Kramer et al., 2009; Pechmann et al., 2013 for reviews).

Finally, ribosome production is an expensive process and thus itself under tight control. Several cellular signaling pathways directly feed into its regulation. Almost all types of stress a cell can endure, including low or high temperature, nutrient deprivation, oxidative and radiation damage, heavy metals or changes in osmolarity affect ribosome biogenesis at various levels (Piazzi et al., 2019).

1.4   Overview of ribosome formation

Ribosomes are usually required to be efficiently and reliably generated at high numbers and a surprising amount of effort is put into their production. Proliferating yeast cells, for example, have been estimated to produce more than 4000 ribosomal subunits per minute, HeLa cells around 7500 (Lewis and Tollervey, 2000; Warner, 1999). These estimates lead to some staggering numbers in rapidly growing yeast cells: 60%

of overall transcription accounts for rRNA production, 50% of RNA polymerase II (Pol II) transcription and 90% of mRNA splicing are devoted to ribosomal proteins, and every nuclear pore is required to import about 1000 RPs and export 25 subunits per minute. On top, there is the need for processing factors, snoRNPs, transport, quality control mechanisms, and more. This renders ribosome biogenesis as one of the most energy demanding cellular processes (Granneman and Tollervey, 2007; Warner, 1999).

In eukaryotic cells, ribosome biogenesis follows an overall conserved path, best described in the model organism *Saccharomyces cerevisiae*. A long rRNA precursor (47S in humans, 35S in *S. cerevisiae*) is transcribed by RNA polymerase I (Pol I) within the nucleolus, which later yields the 18S, 5.8S and 28S (25S in yeast) rRNA after endo- and exonucleolytic processing. Along the way from the nucleolus to the cytoplasm, this primary transcript is folded, modified, and processed. Meanwhile, the 5S rRNA is produced by RNA polymerase III (Pol III; Dieci et al., 2007). Ribosome biogenesis thus requires the joint action of all three RNA polymerases. In addition, an ever-increasing number of assembly factors (AFs) is involved at various stages of ribosome formation, many of which display enzymatic functions, including ATPases, GTPases, endo- or exonucleases, methyltransferases and RNA helicases (Fromont-Racine et al., 2003; Henras et al., 2008).

Finally, ribosome biogenesis is a complex process of sequential and sometimes parallel occurring structural changes in the maturing particles. At the start, transcription of the primary rRNA precursor occurs and the initial pre-ribosome forms. This particle, called 90S or small subunit processome (Dragon et al., 2002; Grandi et al., 2002), then gives rise to the small and large subunit precursors after rRNA cleavage. Maturation pathways of the two subunits thus separate and the particles independently mature. Ribosome biogenesis is therefore usually divided into the 90S, pre-60S, and pre-40S related parts (see Woolford and Baserga, 2013 as an example) and this thesis follows this scheme. Figure 2 provides an overview of the entire ribosome maturation pathway.

1.5 Early steps in eukaryotic ribosome biogenesis

Eukaryotic ribosome biogenesis starts in the nucleus. Despite their lack of further membrane-enclosed subcompartments, nuclei are not homogenous and readily form specialized areas called organelles (Nunes and Moretti, 2017). The first to be discovered and usually the most prominent is the nucleolus, a dense area within the nucleus, where ribosomal rRNA is transcribed by RNA polymerase I (Birnstiel et al., 1966; Ritossa and Spiegelman, 1965). Nucleoli of higher eukaryotes form around nucleolar organizing regions (NOR), transcriptionally active sites of rDNA gene repeats (Colley et al., 2000). In humans, approximately 200 – 600 rDNA genes encode for the 47S polycistronic rRNA precursor, clustered in tandem arrays on the short arms of the five acrocentric chromosomes (chromosomes 13, 14 ,15, 21 and 22; Gonzalez and Sylvester, 1995; Tantravahi et al., 1976). Between individuals and amongst cell types this rDNA copy number can vary extensively (Gibbons et al., 2015; Xu et al., 2017). In contrast, *Saccharomyces cerevisiae* (budding yeast) has a single rDNA locus on chromosome XII, which contains 100 – 200 copies of 35S rDNA genes, interspaced by reversely oriented 5S rDNA segments (Warner, 1989). While 5S rRNA transcription is nucleolar in *S. cerevisiae* due to their physical coupling to the 35S rDNA repeats, in humans and many other eukaryotes, 5S rRNA is encoded in separate tandem repeat segments and therefore its transcription is not necessarily nucleolar, but often occurs in close proximity (Haeusler and Engelke, 2006; Vierna et al., 2013).
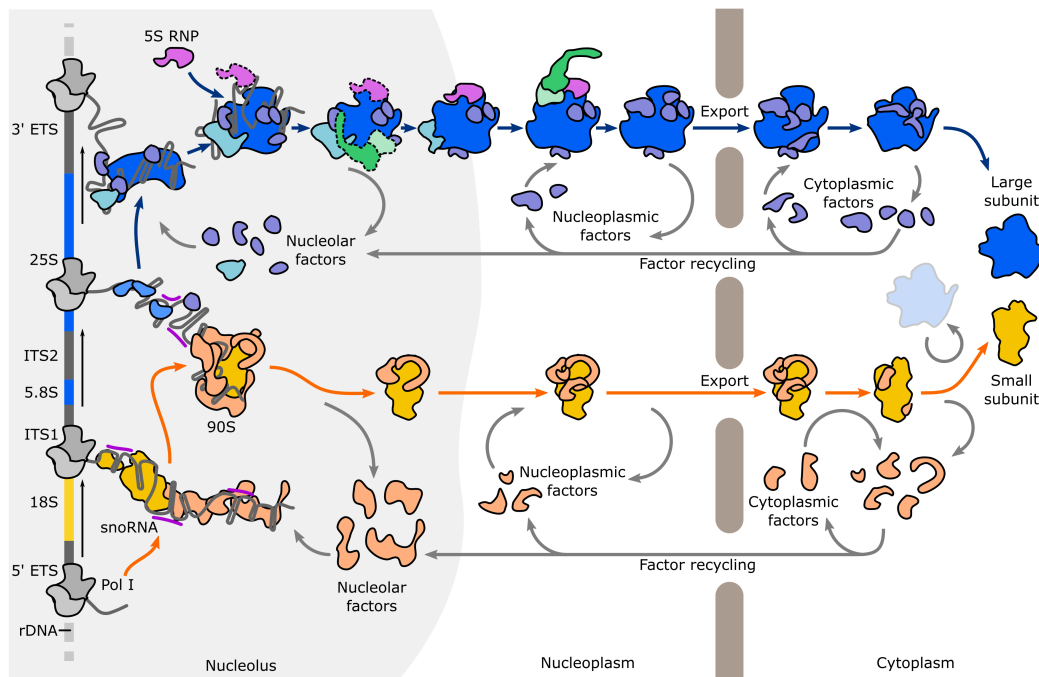
Figure 2 | Schematic overview of the ribosome biogenesis pathway. Nascent primary rRNA transcript of Pol I is bound by nucleolar AFs forming the 90S pre-ribosome. Following rRNA cleavage, the large (blue) and small subunit precursors (yellow) separately mature in the nucleolus, nucleoplasm and finally cytoplasm. A multitude of AFs associate and dissociate at various timepoints during maturation. Numbers, shape, and binding position of assembly factors displayed are exemplary. Adapted from Kressler, 2017.

Transcription of the first large rRNA precursor by Pol I marks the start in ribosome formation and is thus a key target in many regulatory pathways, including mTOR and mitogen-activated protein (MAP) kinase signaling (Grummt, 2003). Active transcription of rDNA tandem repeats has been famously visualized by analyzing nuclear chromatin spreads with electron microscopy (Miller and Beatty, 1969). After dissociation of purified nucleoli, several rRNA side branches of growing lengths with terminal knobs are visible protruding from a central DNA strand, which results in the characteristic Christmas tree-like appearance. Images of such "Miller spreads" revealed important details, including the head-to-tail orientation of rDNA repeats, irregularities in non-transcribed intergenic spacer (IGS) lengths, the order of 18S to 28S transcription, or the co-transcriptional formation and even release of the first rRNA processing complex (Miller and Beatty, 1969; Mougey et al., 1993; Osheim et al., 2004; Reeder et al., 1976).

Within the primary transcript, the rRNA segments are separated by internal transcribed spacers 1 and 2 (ITS1 and ITS2) and flanked by external transcribed spacers at both ends (5' ETS and 3' ETS, see Figure 3A). Due to the 5' to 3' direction of transcription, the first rRNA segments to emerge are the 5' ETS, followed by the 18S rRNA and ITS1. Co-transcriptional compaction of the growing RNA and incorporation of RPs and assembly factors leads to the formation of the 90S particle or SSU processome (Dragon et al., 2002; Grandi et al., 2002), which harbors the 18S rRNA and is therefore the first precursor of the small ribosomal subunit. Early stages in ribosome formation can be described by summarizing the complex events into three main

tasks which need to be accomplished: (I) initial endo- and exonucleolytic processing of rRNA; (II) chemical modification of selected nucleotides; and (III) reduction of conformational freedom of the pre-rRNA by the guided formation of secondary structures and incorporation of RPs and assembly factors (Woolford and Baserga, 2013).

Early pre-ribosomal RNA processing steps

With an overall conserved architecture of the primary transcript in eukaryotes, similar rRNA processing mechanisms have been described in different eukaryotic species. One reoccurring scheme thereof is the multi-step removal of excessive RNA, in contrast to an arguably simpler single endonucleolytic cleavage step (Henras et al., 2015). In the model organism *S. cerevisiae*, the 3' end of the primary transcript is usually generated co-transcriptionally by the endonuclease Rnt1, which is required for efficient termination of Pol I transcription (El Hage et al., 2008; Kufel et al., 1999; Reeder et al., 1999). While detectable amounts of the full 35S pre-rRNA in yeast suggested potential post-transcriptional processing (Udem and Warner, 1972), co-transcriptional release of the 5' ETS, as well as cleavage within the ITS1 have later been described (Kos and Tollervey, 2010; Osheim et al., 2004; Veinot-Drebot et al., 1988). Further experiments have shown that early 5' ETS and ITS1 processing events can indeed occur both co- or post-transcriptionally and the order of these processing events is not fixed, however, co-transcriptional processing is more common in growing cells (reviewed in Fernandez-Pevida et al., 2015). Three sites play a role during the initial phase of ribosome formation and their cleavage all occur within the 90S pre-ribosomal particle (Venema and Tollervey, 1999). First, processing at sites A_0 and A_1 within the 5' ETS allows its removal from the particle and the formation of the mature 5' end of 18S rRNA (Beltrame et al., 1994; Hughes and Ares, 1991). Following cleavage, the 5' ETS is quickly degraded by the nuclear exosome (de la Cruz et al., 1998). While the endonuclease of A_0 has not yet been identified, Utp24 is thought to be responsible for A_1 processing (Tomecki et al., 2015; Wells et al., 2016). Interestingly, Utp24 is also the putative catalytic enzyme for processing at site A_2, which lies within the ITS1 and marks the main cleavage event leading to the separation of small and large subunit rRNA precursor segments (Wells et al., 2016).

Pre-ribosomal RNA processing is generally conserved between yeast and humans, although differences exist in timing of individual cleavage events and the occurrence of alternative pathways (Henras et al., 2015; Tomecki et al., 2017). The human 47S primary transcript is larger than its yeast counterpart and contains an additional processing site A' at the 5' end. Cleavage at site A' occurs very early and is thought to be the only co-transcriptional processing that occurs in mammals (Lazdins et al., 1997). In addition, cleavage at A' seems to be independent of the other 5' ETS processing events and only requires the U3 snoRNP and UTP-A, two of the 90S factor modules, as well as the exonuclease XRN2 (Kass et al., 1987; Sloan et al., 2014; Sloan et al., 2013; Vance et al., 1985). The remaining processing sites are comparable to yeast, with A0, 1, E and 2 in humans corresponding to A_0, A_1, A_2 and A_3 in yeast (Henras et al., 2015). Their order of processing, however, is different. While in yeast, 5' ETS processing (A_0 and A_1) usually occurs before ITS1 cleavage at site A_2, two alternative pathways exist in humans, one of which separates the SSU and LSU precursors at site 2 before 5' ETS shortening (see Figure 3B; Cerezo et al., 2019; Henras et al., 2015). This is generally assumed to be the major pathway, because its 30S and 21S rRNA intermediates are usually more abundant (Cerezo et al., 2019; Henras et al., 2015; Preti et al., 2013). The ribonucleoprotein RNase for mitochondrial processing (MRP), consisting of seven proteins and one noncoding RNA, has recently been identified as the

endonuclease in humans (Goldfarb and Cech, 2017). The complex is similar to its yeast homolog, which has been shown to mediate A_3 cleavage (Mattijssen et al., 2010). Efficient site 2 processing requires RPs uL24 and eL15, as well as the assembly factors BOP1, NOL12, RBM28, and PES1, which are homologous to yeast Erb1, Rrp17, Nop4 and Nop7, respectively. In yeast, these factors are involved in site A_3 processing and subsequent exonucleolytic trimming (see section 1.6; Henras et al., 2015; Lapik et al., 2004; Oeffinger et al., 2009; Preti et al., 2013; Sahasranaman et al., 2011; Sloan et al., 2013). Taken together, this highlights the similarities between site 2 and A_3 processing in humans and yeast, respectively. Subsequent cleavage at site A0, presumably by the endonuclease hUTP23, results in the 26S rRNA intermediate, which is then processed in a second step by hUTP24 to form 21S rRNA (Cerezo et al., 2019; Tomecki et al., 2015; Wells et al., 2016; Wells et al., 2017). In an alternative pathway, 5' ETS processing occurs first, leading to the short-lived 43S rRNA after A0 and to the 41S intermediate after site 1 cleavage. 41S rRNA is then separated into the small and large subunit precursor rRNAs by site 2 or site E processing (Henras et al., 2015; Tomecki et al., 2017).

Early pre-ribosomal RNA modification

As briefly introduced above, ribosomal RNA is extensively modified by box H/ACA and box C/D snoRNPs. Considering the requirement of rRNA accessibility during modification, it is not surprising that guided ribose methylation has also been shown to occur co-transcriptionally in yeast for the SSU and at least in part for LSU rRNA (Birkedal et al., 2015; Greenberg and Penman, 1966; Kos and Tollervey, 2010). In fact, only 2'-O-methyl-adenosine Am100 of the 18S rRNA is not co-transcriptionally modified, while six nucleotides of the 25S rRNA (A817, G867, A876, A2256, U2421 and A2640) appeared to be modified post-transcriptionally (Birkedal et al., 2015). Interestingly, only the U3 snoRNP, which has no nucleotide modification activity, is stably associated with purified 90S particles (Barandun et al., 2017; Kornprobst et al., 2016; Sun et al., 2017), indicating that most modifications occur rapidly during transcription and before the nascent rRNA is compacted into the 90S pre-ribosome. Similar results have been obtained for humans, although some late-acting snoRNPs exist which depend on the helicase activity of DDX21 (Sloan et al., 2015). Notably, with the acetyltransferase Kre33 (NAT10 in humans) and methyltransferase Emg1 also two of the stand-alone modifying enzymes are stably bound to the 90S particle (Cheng et al., 2019; Cheng et al., 2017; Liu and Thiele, 2001).

Compaction of pre-ribosomal RNA into the 90S particle

During transcription, the nascent rRNA precursor is quickly bound by snoRNPs, assembly factors and ribosomal proteins and numerous factors involved in this process have been identified by genetic, biochemical, and structural assays. Association of RNA and protein to the early transcript initiates the correct formation of secondary and tertiary structures and incorporation of ribosomal proteins (Woolford and Baserga, 2013). Many of the early-acting assembly factors are organized in modules which bind the rRNA in a concerted and hierarchical manner (Chaker-Margot et al., 2015; Hunziker et al., 2016; Kass et al., 1990; Kressler et al., 2017; Pérez-Fernández et al., 2007) and several recent cryo-EM structures of the 90S pre-ribosome have provided detailed insights into their location and function (Barandun et al., 2017; Chaker-Margot et al., 2017; Cheng et al., 2019; Cheng et al., 2017; Kornprobst et al., 2016; Sun et al., 2017).
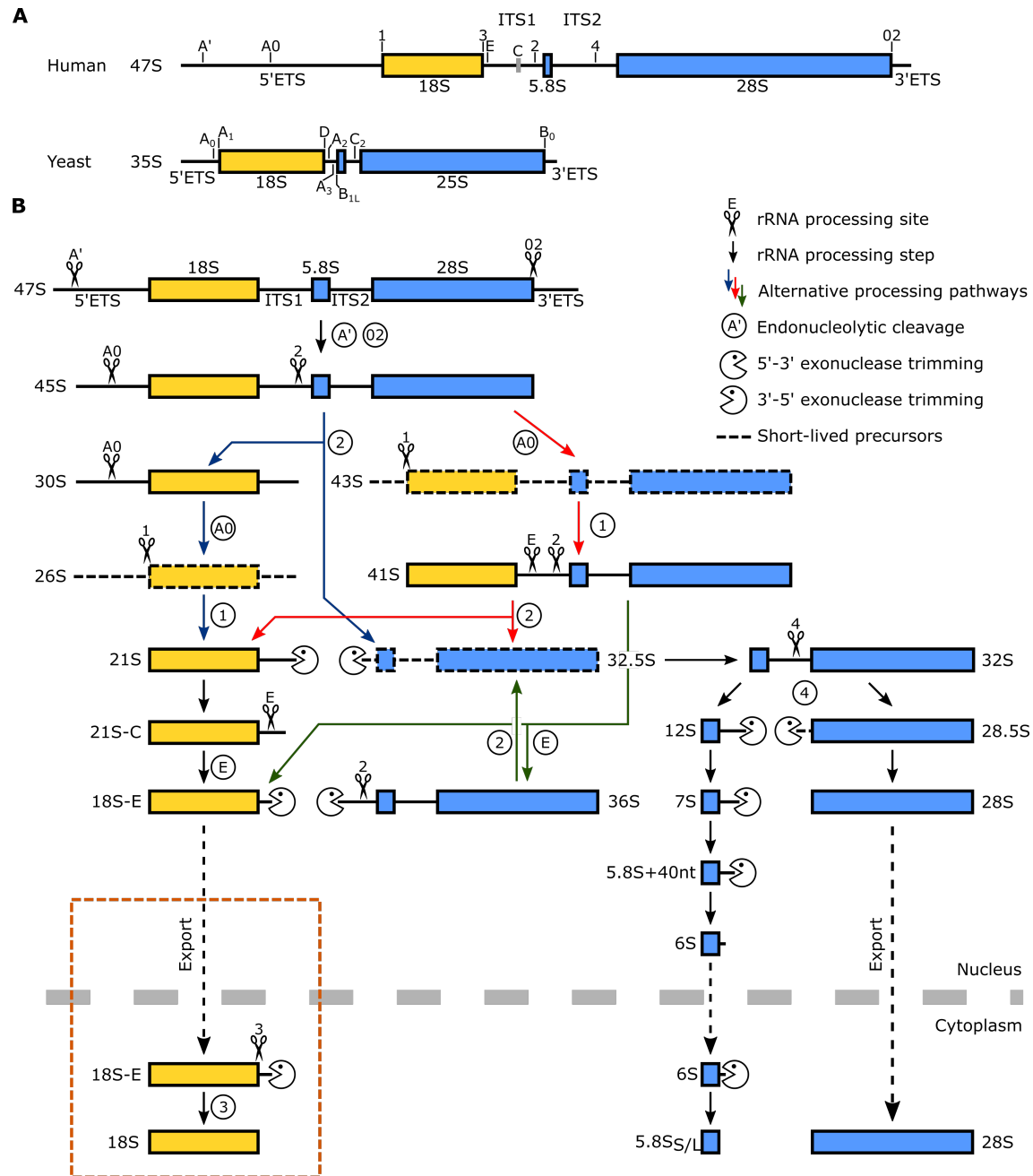
Figure 3 | Pre-ribosomal RNA processing in yeast and humans. (A) Schematic overview of the primary transcript in humans and *S. cerevisiae*. All endonucleolytic cleavage sites and rRNA segments are labelled. The grey box labelled 'C' corresponds to a highly conserved region within the ITS1. Drawing to scale. (B) Summary of all major rRNA processing pathways in humans. The primary 47S transcript is gradually processed in a series of endo- and exonucleolytic cleavage steps. Colored arrows indicate alternative processing pathways and the orange box highlights the focus of research presented in this thesis. Drawing not to scale. Adapted from Henras, 2015.

Initially, the emerging 5' ETS is recognized by the heteroheptameric UTP-A subcomplex, which in turn is required for the association of the U3 snoRNP and UTP-B module as well as the Mpp10 complex (Pérez-Fernández et al., 2007). Binding of UTP-A to the first three helices of the 5' ETS leads to the formation of the 90S particle base, followed by association of the hexameric UTP-B complex, which binds to 5' ETS helices H3 – H5 and forms a substantial part of the 90S body (Cheng et al., 2017; Hunziker et al., 2016). Together with the 5' end of the nascent transcript, both assembly factor modules act as a scaffold for recruitment of further factors, including the U3 snoRNP, the Mpp10 complex and a few single proteins (Barandun et al., 2017; Chaker-Margot et al., 2015; Cheng et al., 2017). The U3 snoRNA contains conserved sequence motifs and binds, in addition to the U3-specific protein Rrp9, with Nop1, Nop56, Nop5 and Snu13 the common core of box C/D snoRNPs (Wormsley et al., 2001; Zhang et al., 2013). However, U3 snoRNP does not act as a 2'-*O*-methyltransferase, but serves as a central rRNA chaperone instead, binding several sites within the 5' ETS and 18S rRNA and thereby tethering them in close proximity (Dutca et al., 2011; Hughes, 1996). Besides assisting in long-range interactions, U3 snoRNA also prevents formation of the central pseudoknot, a conserved structural feature of the mature 40S (Sharma and Tollervey, 1999). U3 snoRNA further interacts with Mpp10, which binds to UTP-B and 5' ETS rRNA via its associated proteins Imp3 and Imp4 at the core of the 90S particle (Barandun et al., 2017; Cheng et al., 2017; Wormsley et al., 2001).

Following the 5' ETS the 18S rRNA is transcribed. Using a series of 3'-truncated RNAs, a multitude of assembly factors has been identified that bind to each of the four domains in a seemingly sequential manner (Chaker-Margot et al., 2015; Zhang et al., 2016). Interestingly, while the order of rRNA binding is presumably as described and follows transcription, recent high-resolution structures of early 90S particles show a reversed integration of the 5' domain and the rest of the 18S rRNA. The central, 3' major and 3' minor domains appear already stably incorporated, but the 5' domain is still disordered and requires the Kre33-Enp2-Brf2-Lcp5 module for maturation (Cheng et al., 2019). Within the fully assembled 90S, many proteins bind to the four 18S rRNA domains and the 5' end of the ITS1, including multi-protein modules like the UTP-C complex, the Noc4-Nop14 module or the central Bms1-Rcl1 dimer (Cheng et al., 2019; Kuhn et al., 2009; Lin et al., 2013). In addition to a large number of assembly factors, several ribosomal proteins of the small subunit are also already incorporated at early stages (Cheng et al., 2017; Sun et al., 2017).

After the formation of the complex 90S particle, pre-rRNA processing and release occurs. Disentanglement of the rRNA-U3 hybrid requires the RNA helicase Dhr1 (DHX37 in humans; Choudhury et al., 2019; Sardana et al., 2015; Zhu et al., 2016), while the exosome and interacting factors are necessary for the release and degradation of the 5'ETS (Mitchell, 2010; Thoms et al., 2015). Cleavage at site A_2 leads to the separation of the small and large subunit precursor RNAs, which are subsequently separately processed.

1.6   Formation of the large subunit

Despite the presence of mainly SSU RPs and assembly factors within the 90S particle, subunit separation seems to be coordinated with the transcription of at least parts of the LSU rRNA precursor and is dependent on bridging elements like Rrp5 (Hierlmeier et al., 2013; Khoshnevis et al., 2019; Lebaron et al., 2013). Following A_2 cleavage in yeast, the released $27SA_2$ rRNA is subsequently split and shortened by a series of endo- and exonucleolytic processing steps (Henras et al., 2015). Interestingly, further 5' end processing by cleavage at site A_3 or its alternative B_{1L} is coupled to the maturation of the 3' end (Kufel et al., 1999). In the main pathway, $27SA_2$ is cut at site A_3 by MRP and the resulting $27SA_3$ is trimmed by exonucleases Rat1 and

Rrp17 to form the mature 5' end of 5.8S rRNA in 27SB (Henry et al., 1994; Johnson, 1997; Oeffinger et al., 2009; Schmitt and Clayton, 1993). Next, cleavage within the ITS2 at site C_2 by Las1 separates the 5.8S and 25S precursors and leads to the formation of the 7S and 26S pre-rRNA (Fromm et al., 2017; Gasse et al., 2015). 7S rRNA is subsequently processed at the 3' end by sequential exonucleolytic trimming activity of the exosome, Rrp6 and Ngl2, of which the latter occurs within the cytosol after particle export (Allmang et al., 1999; Briggs et al., 1998; Faber et al., 2002; Mitchell et al., 1996; Thomson and Tollervey, 2010). Also, final processing of 26S rRNA within the nucleus by Rat1 leads to the mature 25S rRNA (Geerlings et al., 2000). Processing of the large subunit precursor RNAs is largely conserved and similar steps lead to the formation of the human 5.8S and 28S rRNA (Aubert et al., 2018; Preti et al., 2013; Tafforeau et al., 2013).

Maturation of the large subunit is a complex process that requires numerous auxiliary factors, takes place within three different cellular compartments and can be divided into three main stages: (I) formation of the initial precursor particle in the nucleolus, (II) maturation within the nucleoplasm and export licensing, and (III) final maturation steps within the cytosol (Woolford and Baserga, 2013). Recently, several structures of distinct nuclear and cytoplasmic 60S precursors have been solved by cryo-EM, which have shed light on main maturation events (Barrio-Garcia et al., 2016; Greber et al., 2012b; Kargas et al., 2019; Kater et al., 2017; Ma et al., 2017; Malyutin et al., 2017; Sanghai et al., 2018; Zhou et al., 2019). Interestingly, in contrast to the SSU maturation, rRNA domains of the large subunit mature in a more interdependent manner, presumably due to their highly intertwined structures (Ben-Shem et al., 2011). Also, 60S biogenesis displays less distinct AF modules and rather involves a multitude of single proteins, many of which interact on the particle. Throughout the maturation process, reoccurring concepts emerge. In some cases, AF bind to specific areas to block local maturation and their release is required to proceed. Also, maturation of entire substructures is sometimes initialized at immature positions before major structural rearrangements lead to their repositioning. And finally, steps which require input of energy in form of ATP or GTP hydrolysis might represent irreversible checkpoints.

Initial formation of the LSU precursor in the nucleolus

The first steps in 60S assembly take place upon A_2 cleavage and $27SA_2$ release within the nucleolus. Early precursor particles were shown to contain several RNA binding proteins, RNA helicases, and few snoRNAs, in addition to the $27SA_2$ rRNA, indicating ongoing modification and folding (Dez and Tollervey, 2004; Hierlmeier et al., 2013; Lebaron et al., 2013; Rodríguez-Galán et al., 2013). As mentioned above, processing at site A_3 requires a matured 3' end and binding of the ribosomal protein uL3 (Gamalinda et al., 2014). Following 5' maturation of the designated 5.8S rRNA and association of further ribosomal proteins and assembly factors (Gamalinda et al., 2014; Kressler et al., 2008), stable incorporation of 25S rRNA domains I, II and parts of domain VI were shown. With many of their interacting ribosomal proteins already in their final position, this led to the formation of the core and solvent exposed side of the particle. These early pre-60S intermediates also contain a distinct structural feature called 'foot', which comprises the ITS2 rRNA and its associated AFs (Kater et al., 2017; Sanghai et al., 2018). Proceeding from the initial particle, step by step, functional features of the mature ribosome are assembled, including the PTC, GAC and the exit tunnel (Kargas et al., 2019; Zhou et al., 2019).

However, before the particle is released from the nucleolus, further structural rearrangements are required. Initially, the L1 stalk is kept in an immature conformation by a set of early-acting AFs. Removal of the Ytm1-Erb1 dimer by Rea1, also known as Mdn1, is necessary for the dissociation of a set of proteins that interact with the meandering N-terminal tail of Erb1, which ultimately allows repositioning of the L1 stalk (Kater et al., 2017; Sahasranaman et al., 2011; Tang et al., 2008). Also, although associated to the particle, the 5S RNP is not yet stably incorporated in early particles and the premature CP presumably forms prior to transit to the nucleoplasm (Kater et al., 2017; Kressler et al., 2008). Finally, processing of the ITS2 at site C_2 by the newly associated endonuclease Las1 occurs. This, however, might not be a prerequisite for further maturation, as 60S-like particles have been shown to form translating ribosomal complexes after impaired C_2 cleavage (Sarkar et al., 2017).

Nucleoplasmic maturation and export competence

Cryo-EM structures of 60S precursors have revealed further structural remodeling within the nucleoplasm. Early nucleoplasmic particles already contain the majority of ribosomal proteins and large parts of the core have matured, while several functionally important areas like the exit tunnel, ISS or CP are still decorated with assembly factors (Barrio-Garcia et al., 2016; Bradatsch et al., 2012; Leidig et al., 2014; Wu et al., 2016). Following C_2 cleavage, both resulting rRNA segments are trimmed within the nucleoplasm. Via the adapter protein Nop53, the 3′-5′ RNA helicase Mtr4 and the associated exosome are recruited to the pre-60S particle, leading to the shortening of the 3′ end of 7S rRNA by Rrp44 and Rrp6 and the dismantling of the 'foot' structure (Fromm et al., 2017; Schuller et al., 2018; Thoms et al., 2015). The 26S rRNA precursor is first phosphorylated at the 5′ end by the kinase Grc3 and then processed by the exonuclease Rat1 and its activating partner Rai1 (Braglia et al., 2010; Gasse et al., 2015; Xiang et al., 2009). In addition to the removal of ITS2, pre-60S particles undergo a major restructuring at the CP. The 5S RNP, consisting of 5S rRNA and ribosomal proteins uL18 and uL5, is initially integrated into the nascent particle in a premature conformation (Leidig et al., 2014). Release of assembly factors Rpf2 and Rrs1, which bind between the 5S RNP and rRNA helices of the future CP, appears central to the subsequent rotation of 5S rRNA by approximately 180° (Asano et al., 2015; Barrio-Garcia et al., 2016; Kharde et al., 2015; Leidig et al., 2014; Zhang et al., 2007). 5S RNP rotation and stabilization of the CP requires Rea1 and assembly factors Sda1 and the Rix1-Ipi1-Ipi3 complex. Then, ATP and GTP hydrolysis by Rea1 and Nog2, respectively, results in the release of Rsa4, Sda1, Nog2, and the large Rea1-Rix1-Ipi1-Ipi3 module, ultimately allowing export adapter Nmd3 to bind to the previously occupied site (Barrio-Garcia et al., 2016; Matsuo et al., 2014; Ulbrich et al., 2009). Nmd3 interacts with the central export factor Crm1, also known as exportin 1 (XPO1) and thus aids together with Ran-GTP to the directionality of the passage through the nuclear pore complex (NPC; Askjaer et al., 1999; Ho et al., 2000). Furthermore, the assembly factor Arx1, which binds at the end of the exit tunnel, directly interacts with the phenylalanine-glycine (FG)-rich nucleoporins located at the center of the NPC and thus assists in shuttling through the NPC (Bradatsch et al., 2007; Hung et al., 2008). Finally, several additional factors have been proposed to facilitate nuclear export, including the armadillo (ARM)-like fold containing protein Rrp12 (Oeffinger et al., 2004). Displacement of Mrt4 by Yvh1 at the site of the yet to be formed P-stalk might also still occur in the nucleus, enabling association of the heterodimeric export factor module Mex67-Mtr2, although recent results suggest a cytoplasmic association of Yvh1 (Altvater et al., 2012; Greber, 2016; Lo et al., 2009; Sarkar et al., 2016).

Final maturation in the cytosol

Following the export into the cytosol, processing of 6S rRNA by Ngl2 and removal of the remaining biogenesis factors completes 60S formation (Thomson and Tollervey, 2010). This final maturation process occurs sequentially and several steps are driven by ATP or GTP hydrolysis (Kressler et al., 2010; Lo et al., 2010). Initially, the AAA-like ATPase Drg1 is recruited to the pre-60S particle by Rlp24, leading to the replacement of Rlp24 with ribosomal protein eL24 and, in addition, to the loss of the assembly factors Nog1, Nug1 and Nsa2 (Altvater et al., 2012; Pertschy et al., 2007). Incorporation of eL24 allows subsequent recruitment of Rei1, Jjj1 and the HSP70-like ATPase Ssa, which are necessary for Arx1 release (Hung and Johnson, 2006; Lo et al., 2010; Meyer et al., 2010). In parallel, Yvh1 is replaced by RP uL10 at the P-stalk and completion of this structural feature, as well as Arx1 release is a prerequisite for further progression (Lebreton et al., 2006; Lo et al., 2010). Ribosomal protein uL16 is incorporated next and the remaining factors Nmd3 and Tif6 are released, which requires the activity of GTPases Lsg1 and Efl1, as well as auxiliary proteins Sqt1 and Sdo1 (Lo et al., 2010; Senger et al., 2001; Weis et al., 2015).

1.7   Formation of the small subunit

Cleavage within the ITS1, which separates the large and small subunit rRNA precursors, marks the start of the final 40S maturation phase. In stark contrast to the 90S, nuclear and cytoplasmic pre-40S particles contain only a relatively small set of AFs and large portions of the particle are already assembled (Klinge and Woolford, 2019; Strunk et al., 2011). During the transition from 90S to early 40S precursors, release of all major assembly factor modules, as well as the separated 5' ETS is required for the reorganization of the particle's rRNA domains. Within the 90S the 5' domain is largely matured, the remaining central, 3' major and 3' minor domains, however, still need to reposition for integration (Barandun et al., 2017; Chaker-Margot et al., 2017; Cheng et al., 2017; Klinge and Woolford, 2019; Kornprobst et al., 2016). Little is known about this transition process, but it is generally assumed that loss of the majority of AFs together with the 5' ETS occurs in form of a combined 5' ETS particle that is similar in composition to the particle at the beginning of 90S formation (Allmang et al., 2000; Barandun et al., 2018; Chaker-Margot, 2018; Pena et al., 2017). With the loss of the 5' ETS scaffold most other factors also dissociate and only a handful of assembly factors remain, including Dim1, Rrp12, Enp1, and Pno1 (also known as Dim2) in yeast (Campbell and Karbstein, 2011; Schafer et al., 2003). Additional proteins, such as the endonuclease Nob1, GTPase-like protein Tsr1, the ATPase Rio2 and Ltv1 then join the complex before the particle reaches export competence.

Final rRNA processing steps during late 40S maturation

Besides conformational changes in rRNA and integration of the remaining ribosomal proteins, removal of the ITS1 remnant occurs during the last stage of 40S maturation. In yeast, the final 18S precursor is the 20S rRNA, which stems from the initial transcript after A_1 and A_2 cleavage. It is exported into the cytoplasm unprocessed, where it is shortened by a single endonucleolytic cleavage step at site D, yielding the 3' end of mature 18S rRNA (Fatica et al., 2003; Fatica et al., 2004; Pertschy et al., 2009). In humans, however, SSU rRNA maturation is far more complex. As mentioned previously, the major pathway entails site 2 processing within the ITS1 prior to 5' ETS shortening, leading to the abundant 30S rRNA intermediate. Subsequent trimming of the 5' end by site A0 and 1 cleavage results in the 21S rRNA (Henras et al., 2015). While the 5'

end has thus matured, 21S still contains approximately 940 nucleotides of the ITS1 at its 3' end (Sloan et al., 2013). It is then gradually shortened by a series of endo- and exonucleolytic processing steps. First, RRP6, the catalytic component of the nuclear exosome, trims the 3' end by removing approximately 300 nucleotides, before it stops at a highly conserved region approximately 635 nucleotides into the ITS1 (Carron et al., 2011; Preti et al., 2013; Sloan et al., 2013). This leads to the 21S-C rRNA precursor, which is then cleaved at site E by hUTP24. This can occur at two different sites, leaving 78 or 81 nucleotides of ITS1 bound to the 18S 3' end (Preti et al., 2013). The resulting 18S-E rRNA is then further shortened by the exonuclease poly(A)-specific ribonuclease (PARN) to 24 – 40 nucleotides (Ishikawa et al., 2017; Montellese et al., 2017). PARN, which is also involved in processing of various other coding and noncoding RNAs (Martinez et al., 2000; Moon et al., 2015), acts still within the nucleus and is required for efficient particle export and final rRNA maturation (Montellese et al., 2017). The poly(A) polymerase PAPD5 has been suggested to add untemplated adenosines to 18S-E to facilitate substrate recognition by PARN (Montellese et al., 2017). Finally, the shortened 18S-E rRNA precursor is exported into the cytoplasm, where the last cleavage by NOB1 occurs (Zemp et al., 2009). Interestingly, cytoplasmic precursors of 18S rRNA have been shown to contain to some extent additional uridines at their 3' end, which suggests further intermediate processing including poly-uridylation and potentially subsequent exonucleolytic trimming (Montellese et al., 2017; Preti et al., 2013). Whether this is required for 18S maturation or signals for rRNA degradation remains to be elucidated.

Nucleoplasmic maturation and particle export

Following the disassembly of the 90S pre-ribosome, the emerging small subunit precursor is readily prepared for nuclear export and quickly transferred into the cytoplasm (Schafer et al., 2003). Similar to large subunit precursors, pre-40S particles require the transport receptor Crm1 and the small protein Ran in complex with GTP for nuclear export (Moy and Silver, 1999, 2002). Crm1 is a ring-shaped, HEAT-repeat containing protein, which is able to bind a large variety of cargo complexes via a loosely conserved nuclear export signal sequence (NES; Dong et al., 2009). It can then pass the core of the nuclear pore complex together with its cargo by transiently interacting with FG repeat containing nucleoporins, before GTP hydrolysis by Ran leads to the dissociation of the complex in the cytoplasm (Koyama and Matsuura, 2010). As previously mentioned, during 60S biogenesis the essential assembly factor Nmd3 serves as an adapter protein and mediates the interaction with transport receptor Crm1 via its C-terminal NES (Ho et al., 2000). So far, no comparable NES-containing protein has been found on 40S precursors. Instead, various non-essential assembly factors are thought to contribute to an efficient nuclear pore passage.

With Ltv1 and Pno1 in yeast, and RIOK2 in humans, three of the pre-40S associated assembly factors were initially predicted to contain NES sequences and have therefore been suggested as potential Crm1 adapter proteins (Seiser et al., 2006; Vanrobays et al., 2008; Zemp et al., 2009). However, closer inspection of Ltv1 and Pno1 later indicated separate, Crm1 independent effects on 40S biogenesis and export. The putative NES of Pno1 does not match a refined version of the leucine-rich (LR)-NES, which was defined based on Crm1 crystal structures (Guttler et al., 2010; Merwin et al., 2014) and mutation or deletion of the NES sequence of Ltv1 did not cause nuclear retention of Ltv1 and a cytoplasmic enrichment of 20S rRNA precursors (Fassio et al., 2010). Interestingly, nuclear accumulation of uS3 and uS5, two ribosomal proteins that are exported as part of the pre-40S complex, after Ltv1 mutation was attributed to secondary effects

after failed recycling of assembly factors during the cytosolic stage of 40S maturation (Fassio et al., 2010). Similar ambiguous results were obtained for Rio2 in yeast, when nuclear retention of uS5 and cytoplasmic accumulation of 20S precursors were observed in mutant Rio2 strains (Schafer et al., 2003; Vanrobays et al., 2003). Interaction of human RIOK2 with CRM1, however, was shown to occur *in vitro* via a functional, C-terminal NES and a contribution of RIOK2 to export efficiency was also shown *in vivo* (Zemp et al., 2009). Additionally, a crystal structure of human CRM1 bound with the NES signal sequence of RIOK2 allowed the identification of important interacting residues (Fung et al., 2015). Another protein which copurifies with pre-40S particles and shuttles between the nucleus and cytoplasm is Slx9 and its human homolog FAM207A (also known as C21orf70; Faza et al., 2012; Wyler et al., 2011). Yeast Slx9 does not interact with Crm1 directly. Instead it is able to bind Rio2 and form a trimeric complex with RanGTP, which then in turn can recruit the transport receptor Crm1 to the NES of Rio2 (Fischer et al., 2015). Crm1 association to the maturing pre-40S ribosome therefore appears to occur in a stepwise, non-canonical manner. Similarly, yeast protein Yrb2 has been shown to be required for efficient pre-40S export and to bind Crm1 and RanGTP *in vitro* (Faza et al., 2012; Maurer et al., 2001; Moy and Silver, 2002). RanBP3, the human homologue of Yrb2 binds CRM1 and stabilizes CRM1-cargo complexes with certain NES-containing proteins (Englmeier et al., 2001; Lindsay et al., 2001). Finally, human PDCD2L and its paralog PDCD2 have recently been shown to shuttle between nucleus and cytoplasm in a CRM1 dependent manner. Both contain functional NES sequences and are able to bind pre-40S complexes (Landry-Voyer et al., 2016). Together, these findings indicate that pre-40S particles do not rely on a single essential, NES-bearing assembly factor capable of efficient Crm1 recruitment.

Several proteins have been reported to assist in NPC passage by directly interacting with FG-rich nucleoporins. Similar to the pre-60S export, the Mex67-Mtr2 heterodimer is able to bind 40S precursors and support their export in yeast (Faza et al., 2012; Yao et al., 2008; Yao et al., 2007). Mex67 and Mtr2 dimerize with their NTF2-like domains and interact with pre-40S and pre-60S particles via loop insertions within these domains (Faza et al., 2012; Fribourg and Conti, 2003; Yao et al., 2007). Both proteins are also able to simultaneously bind pre-ribosomal particles and FG-rich regions of nucleoporins, thus facilitating nuclear export (Grant et al., 2002; Strasser et al., 2000; Suyama et al., 2000). Similarly, karyopherin-like protein Rrp12 has been shown to interact with both, small and large subunit precursors in yeast (Oeffinger et al., 2004). While Rrp12 seems non-essential for pre-60S maturation, pre-40S particles require Rrp12 for efficient nuclear export, due to its capability to bind FG-rich nucleoporins and RanGTP (Moriggi et al., 2014; Oeffinger et al., 2009). In addition, Rrp12 is involved in early 90S maturation and remains bound to the emerging 40S precursor until its release within the cytosol. Interestingly, after Rrp12 depletion pre-ribosomal particles lack the assembly factor Ltv1, indicating a potential role of Rrp12 in Ltv1 recruitment (Moriggi et al., 2014).

Apart from the described assembly factors involved in export, a few additional proteins interact with the particle within the nucleus. The methyltransferase Bud23 methylates G1575 of the 18S rRNA and has been shown to be required for 40S precursors to reach export competence. Interestingly, G1575 methylation is not essential and Bud23 presence rather than its enzymatic activity is vital for normal 40S maturation (White et al., 2008). Bud23 depends on its binding partner Trm112 for stability and interacts with RNA helicase Dhr1 prior to A2 cleavage (Figaro et al., 2012; Letoquart et al., 2014; Sardana et al., 2014). Bud23 and Trm112 are evolutionary conserved in eukaryotes and the human homologs BUD23 (also known as

WBSCR22) and TRMT112 act accordingly during human 40S biogenesis (Haag et al., 2015; Ounap et al., 2013; Ounap et al., 2015; Zorbas et al., 2015). Dim1, a second methyltransferase binding late pre-40S particles in yeast, is responsible for a conserved dimethylation of bases A1781 and A1782 close to the decoding site and exhibits a similar requirement for its physical presence rather than enzymatic activity as Bud23 (Lafontaine et al., 1994; Lafontaine et al., 1995). Unlike Bud23, however, yeast Dim1 and its human homolog DIMT1L act at different stages of 40S maturation. While yeast Dim1 remains bound to the pre-40S particle and shuttles into the cytosol where dimethylation occurs, human DIMT1L exerts its function within the nucleus and dissociates prior to particle export (Brand et al., 1977; Zorbas et al., 2015). Depletion of the casein kinase homolog Hrr25 has been shown to inhibit 40S export in yeast, although the mechanism remains unknown (Schafer et al., 2006). Both Hrr25 and the isoforms δ and ε of its human homolog casein kinase 1 (CK1) are required for cytoplasmic 40S maturation as described below (Mitterer et al., 2016; Zemp et al., 2014). Finally, the GTPase-like protein Tsr1 binds to pre-40S particles before nuclear export. A crystal structure of yeast Tsr1 revealed structural similarities to translational GTPases, such as EF-Tu, however, it lacks several active sites residues, rendering it incapable of GTP binding and hydrolysis (McCaughan et al., 2016). While Tsr1 is required for 20S rRNA processing in yeast, depletion leads to 20S accumulation within the cytosol, which suggests a non-essential role in particle export (Leger-Silvestre et al., 2004) .

Final steps in 40S maturation

Following the export of the 40S precursor into the cytoplasm, release of the remaining assembly factors and final rRNA processing occurs. The exact order of events and mechanisms involved, however, remain poorly understood. One of the first proteins to be released is likely Rrp12, because it is present only in a minor fraction of cytoplasmic pre-40S particles (Moriggi et al., 2014). In humans, release of RRP12 requires the atypical kinase RIOK2 (Wyler et al., 2011) and overall conservation suggests a similar situation in yeast. Rio2 and its human homolog are both essential for the cytoplasmic maturation of the small ribosomal subunit and mutations or deletions lead to 20S and 18S-E processing defects, respectively (Geerlings et al., 2003; Vanrobays et al., 2003; Zemp et al., 2009). Structural analysis of related Rio2 kinases from *Chaetomium thermophilum* and *Archaeoglobus fuldigus* have revealed a kinase fold similar to the canonical eukaryotic protein kinase (ePK) domain, however, lacking important features such as the activation loop and substrate interface (Ferreira-Cerca et al., 2012; LaRonde-LeBlanc and Wlodawer, 2004). In addition to an observed phosphoaspartate intermediate and an occluded active site of Rio2 on the ribosome, this suggests an ATPase rather than kinase function of Rio2 during 40S maturation (Ferreira-Cerca et al., 2012). Interestingly, physical presence instead of ATP hydrolyzation by RIOK2 is required for ENP1 dissociation in humans, whereas LTV1, PNO1 and NOB1 release depend on its enzymatic activity (Zemp et al., 2009).

Probing of yeast U1191 aminocarboxypropylation by the transferase Tsr3 indicated dissociation of Ltv1 and Enp1 precedes that of Tsr1 and Rio2, which in turn is followed by Pno1 and Nob1 (Hector et al., 2014). Ltv1 forms a subcomplex with Enp1 and ribosomal protein uS3 at the particle's beak and is important for the coordinated integration of uS3 (Mitterer et al., 2016; Schafer et al., 2006). Phosphorylation of Ltv1 by the kinase Hrr25 in yeast and CK1δ or CK1ε in humans is required for its subsequent release (Ghalei et al., 2015; Zemp et al., 2014). Interestingly, Ltv1 binding partner Enp1 is also phosphorylated in yeast and humans and the kinase activity of CK1δ/ε is required for ENP1 dissociation (Ghalei et al., 2015; Zemp et al., 2014). Because lack of CK1δ/ε in humans leads to a stalled recycling of the remaining assembly factors ENP1, LTV1,

RRP12, RIOK2, PNO1 and NOB1, phosphorylation and release of LTV1 appears critical for later 40S maturation events (Zemp et al., 2014).

Maturation of the beak region results in a pre-40S particle poised for final rRNA processing. These SSU precursors still contain at least ribosomal assembly factors Tsr1, Rio2, Pno1, and Nob1, as well as the methyltransferase Dim1 in yeast. Subsequent mechanisms leading to endonucleolytic cleavage of 20S and 18S-E rRNA precursors, respectively, differ greatly between yeast and humans. In yeast, Nob1 activation is thought to occur in 80S-like complexes and to require binding and enzymatic activity of several additional factors (Lebaron et al., 2012; Strunk et al., 2012; Turowski et al., 2014). Involvement of the large subunit in SSU maturation is described as a quality control step, ensuring structural integrity and functionality of nascent 40S subunits (Strunk et al., 2012). Formation of 80S-like complexes is promoted by the GTPase eIF5B, also known as Fun12, which also serves in subunit joining during canonical translation initiation (Hinnebusch, 2014; Lebaron et al., 2012; Pestova et al., 2000). Disruption of eIF5B GTPase activity or its interaction with the large subunit reduces 20S processing efficiency (Lebaron et al., 2012). Following 80S-like formation, the ATPase Fap7 is thought to bind and lead to the release of the assembly factors Dim1 and Tsr1, before dissociating after ATP hydrolysis (Ghalei et al., 2017; Strunk et al., 2012). Similar to Fap7, the exact timing and mode of action of another late-acting assembly factor, Rio1, remains uncertain. Rio1 is the second protein from the RIO kinase family functioning in ribosome biogenesis after Rio2. It contains a similar central nucleotide-binding domain as Rio2 but lacks its N-terminal winged helix domain (Ferreira-Cerca et al., 2014). Like Rio2, Rio1 likely serves as an ATPase, whose activity is required for 40S maturation and essential for yeast cell viability (Angermayr et al., 2002; Ferreira-Cerca et al., 2014; Turowski et al., 2014; Vanrobays et al., 2001). Depletion of Rio1 or overexpression of a mutant form leads to enrichment of assembly factors Dim1, Tsr1, Fap7, Nob1 and Pno1 on 80S-like particles, indicative of a role of Rio1 in the release of these factors (Belhabich-Baumas et al., 2017; Ferreira-Cerca et al., 2014). Interestingly, ATP binding rather than hydrolysis is required for Rio1 mediated stimulation of site D cleavage by Nob1, which occurs in 80S-like particles lacking Dim1, Tsr1 and Fap7. Finally, overlapping binding sites of Pno1 and Nob1 to the 3' end of 20S rRNA, as well as a lower abundance of Pno1 in cleavage competent 80S particles suggest a release of Pno1 prior to Nob1 activation and final rRNA processing (Turowski et al., 2014).

The most striking difference when comparing this pathway with the final steps of human 40S maturation is the lack of evidence for the requirement of a potential 80S-like complex formation. While the human homolog of Fap7, hCINAP, has also been shown to be involved in 18S-E processing, and its ATPase activity stimulates NOB1 mediated cleavage of 18S-E rRNA *in vitro* (Bai et al., 2016), neither eIF5B nor 60S subunits have been demonstrated to be required for 40S biogenesis. In fact, the sole assembly factor known to interact with the final 40S precursors is RIOK1, the human homolog of yeast atypical kinase Rio1. It has been shown to act after RIOK2 and stimulate 18S-E rRNA processing. In addition, expression of a kinase-dead mutant leads to cytoplasmic accumulation of PNO1 and NOB1, but not ENP1, LTV1 and RIOK2, indicative of an active role of RIOK1 in PNO1 and NOB1 recycling from 40S particles (Widmann et al., 2012). How RIOK1 triggers NOB1 activity, however, and whether other factors are involved remains unclear.

Cryo-EM structures of cytoplasmic 40S precursors

Early cryo-EM data has revealed the location of several biogenesis factors at low resolution in yeast and humans, including Enp1/ENP1 and Ltv1/LTV1, GTPase-like protein Tsr1/TSR1, the methyltransferase Dim1, the atypical kinase Rio2/RIOK2, as well as the endonuclease Nob1/NOB1 and its binding partner Pno1/PNO1 (Figure 4; Larburu et al., 2016; Strunk et al., 2011). These insights complemented results obtained by ultraviolet (UV) crosslinking and complementary DNA analysis (CRACK; Granneman et al., 2010; Turowski et al., 2014), which allows mapping of protein binding sites onto the 18S rRNA precursor. Within late 40S precursors, these biogenesis factors occupy important positions: Enp1 and Ltv1 form a complex at the particle's beak, preventing the accommodation of ribosomal proteins uS3, uS10, eS10 and uS14 at their mature position. Tsr1 and Rio2 are located at the intersubunit side of the particle, where they block the A- and P-sites, the binding position of translational GTPases and accommodation of the large subunit in its canonical position. In addition, both factors interact with an immature h44, where the top segment of h44 is shifted away from its final position, thus disrupting the decoding site (Strunk et al., 2011). Finally, the assembly factor Pno1, as well as the endonuclease Nob1 have been identified at the platform, close the 3' end of the 18S rRNA and the 5' end of ITS1.

Many of the nuclear and some cytoplasmic factors are missing in the described complexes and, so far, no structure of a nuclear pre-40S has been solved. Over the years, several large-scale screens have identified hundreds of proteins involved in human ribosome formation. While most factors appear conserved across species, many have been found lacking homologs in yeast or displaying different or additional functions (Badertscher et al., 2015; Farley-Barnes et al., 2018; Tafforeau et al., 2013; Wild et al., 2010). Structural analysis of human pre-ribosomal complexes is therefore an important addition to ongoing experimental efforts.
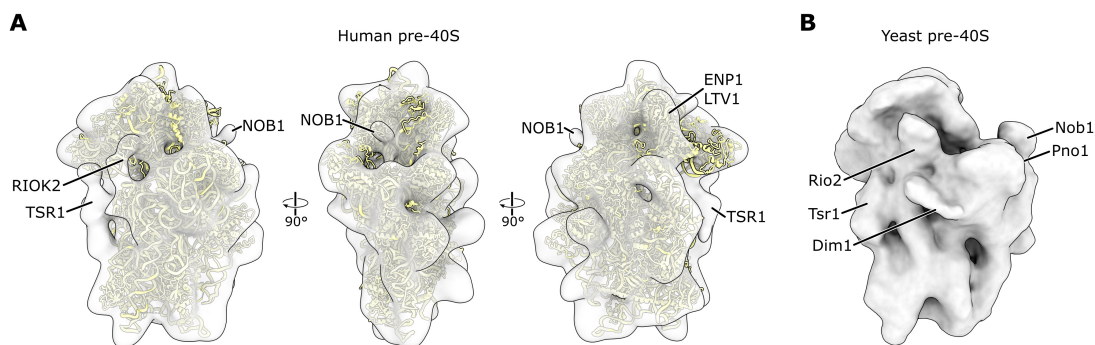
Figure 4 | Cryo-EM volumes of a human and yeast pre-40S particle. (A) Overlay of a low-resolution cryo-EM volume of a human 40S precursor (EMDB 3300) and a model of the mature subunit (PDB 6Y0G) show additional densities assigned to various biogenesis factors. (B) Cryo-EM volume of a yeast 40S precursor with biogenesis factors labelled (EMDB 1927).

2 Aims of this thesis

Eukaryotic ribosome biogenesis is a complex process and, so far, best understood in the model organism *Saccharomyces cerevisiae*. Decades of research have provided vast information on yeast rRNA processing, compaction and modification, and the recruitment of ribosomal proteins. As outlined above, however, major differences exist between yeast and human assembly mechanisms and many systems remain poorly understood. In the light of an increasing list of human diseases related to ribosome biogenesis, termed ribosomopathies (Aspesi and Ellis, 2019; Farley-Barnes et al., 2019; Kampen et al., 2020), an in-depth understanding of human biogenesis pathways is therefore desirable.

This dissertation aimed at providing insights into maturation principles of the human small ribosomal subunit using cryo-electron microscopy and single-particle analysis. To obtain the first high-resolution structure of a 40S precursor, a suitable expression system was required which would allow purification of human pre-ribosomal complexes at sufficient quantity and quality for cryo-EM structure determination. Therefore, based on methods previously described (Wyler et al., 2011) stable human cell lines were generated, which expressed selected ribosome biogenesis factors upon tetracycline induction. Using a range of pre-40S AFs, a system was set up to tag and purify native complexes, which were subsequently analyzed by cryo-EM.

Utilizing the ribosomal biogenesis factor PNO1, which binds to a broad range of possible pre-40S intermediates, as bait for affinity purification, as well as recent developments in computational analysis of cryo-EM data, several distinct precursor states were targeted simultaneously for publication 1 (Ameismeier et al., 2018). High-resolution data on one or more 40S precursors was to be compared to similar ongoing projects by colleagues working with the yeast system. Furthermore, data from several states would provide answers to fundamental questions regarding the functions of specific human assembly factors and the degree of conservation of maturation principles between yeast and humans.

The results of publication 1 provided detailed information on late nuclear and early cytoplasmic maturation events. This first set of higher resolved 40S assembly intermediates was to be expanded in a second publication (Ameismeier et al., 2020). The aim was to further dissect the release mechanisms of TSR1, PNO1 and NOB1 during final 40S processing steps and analyze the compositional and conformational changes of the pre-40S particle required for the activation of the endonuclease NOB1. In support of structural data, immunofluorescence, and fluorescence *in situ* hybridization (FISH) experiments were performed, adding to our understanding of the function of the described SSU assembly factors.

3 Discussion and Outlook

In the two publications included in this dissertation, a total of ten distinct native small ribosomal subunit precursors were described and ordered into a logical sequence based on compositional differences and existing functional data (Figure 5; Ameismeier et al., 2018; Ameismeier et al., 2020). Together they provided the first high-resolution structures of human ribosome biogenesis precursors and allowed a detailed description of late 40S maturation principles. In a first phase, structural rearrangements in rRNA and a stepwise incorporation of late assembling ribosomal proteins result in the completion of the 3' major domain. Then, rearrangements in rRNA segments and a shift of the catalytically active domain of NOB1 occurs, activating the endonuclease and triggering site 3 processing. Furthermore, during maturation several factors bind and obstruct the tRNA binding sites, keeping the A- and P-site in an immature conformation or block access of translation initiation components. In total, 13 assembly factors were observed coordinating the stepwise maturation of the small ribosomal subunit. Six of them (TSR1, ENP1, LTV1, RIOK2, PNO1 and NOB1) have been previously observed at low resolution (Larburu et al., 2016; Strunk et al., 2011). Four factors (RRP12, RIOK1, BUD23 and TRMT112) have not been visualized before, but their roles in 40S maturation have been described (Letoquart et al., 2014; Moriggi et al., 2014; Widmann et al., 2012); yet, two are entirely novel factors (LRRC47 and eIF1AD) and one remains unidentified (factor X).

Maturation of the 3' major domain and head formation

In yeast, disassembly of the 90S pre-ribosome leads to a particle termed Dis-C, which displays a near-mature 5' and central, but partially flexible 3' major and 3' minor domains. In addition, this nuclear 40S maturation intermediate contains several remaining 90S-associated assembly factors, including the RNA helicase Dhr1 and GTPase Bms1, which is later replaced by Tsr1 prior to nuclear export (Cheng et al., 2017; Heuer et al., 2017b; Scaiola et al., 2018; Cheng et al., in press). Assuming a similar maturation pathway of the human small ribosomal subunit, the so far earliest precursor described in the sequence, state A, appears slightly more matured (Ameismeier et al., 2018; Cheng et al., in press). Large segments of the head have already formed and are likely stabilized in their position through interactions with the AF TSR1. TSR1 binds with its GTPase-like domains to an area on the 40S which is also utilized by several translation factors and, in addition, contacts the head with domain IV, possibly contributing to a more stable head conformation. A similar interaction was also observed in yeast pre-40S particles (Heuer et al., 2017b; Scaiola et al., 2018). With ribosomal proteins uS2, uS5, and eS21 yet to bind, state A misses a central part of the body. In addition, the endonuclease NOB1, which carries out the final rRNA processing step at site 3 has not yet associated. Because all three ribosomal proteins are required for nuclear export (Ferreira-Cerca et al., 2005; O'Donohue et al., 2010) and NOB1 has also been shown to be incorporated into the particle in the nucleoplasm (Zemp et al., 2009), state A is the sole precursor in the described sequence that can be unambiguously assigned to the nucleoplasm. Since no export particle has been described so far, it remains unknown how incorporation of these ribosomal proteins is linked to export competence.
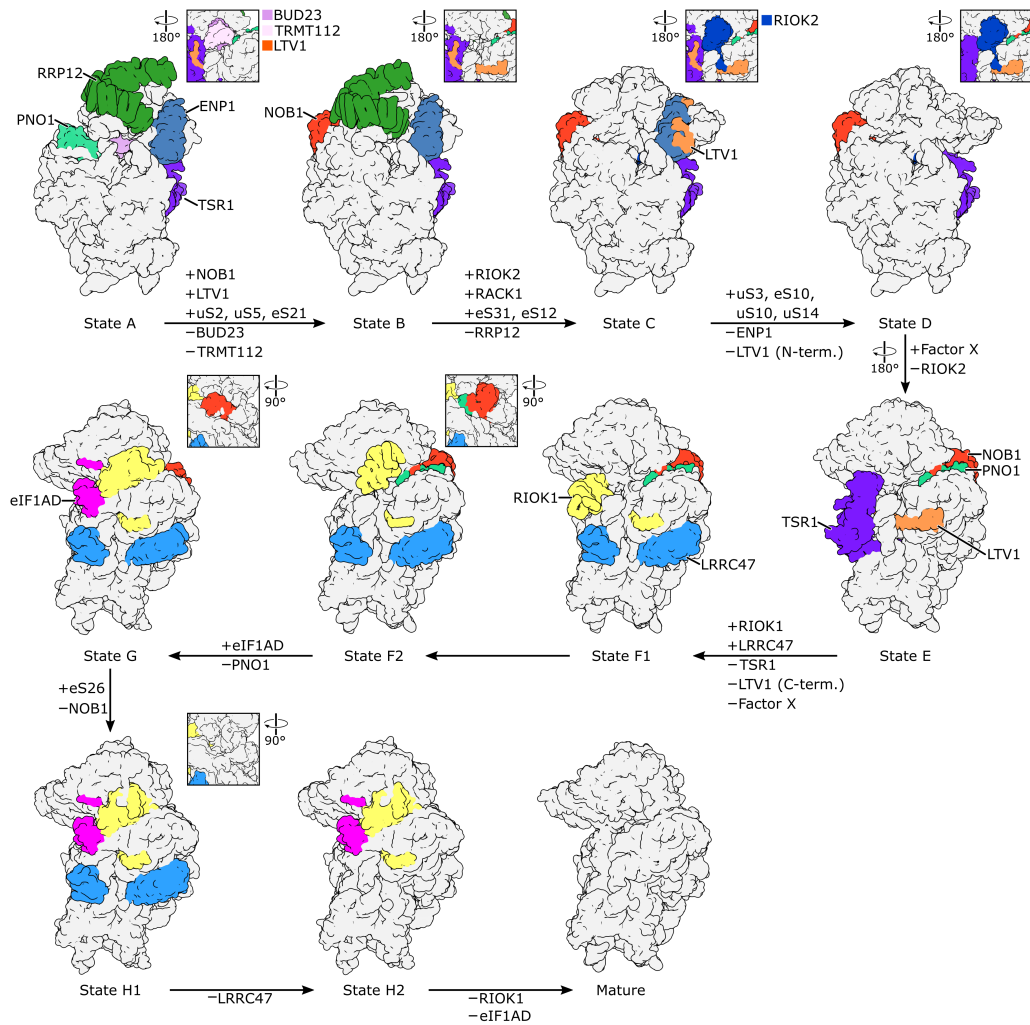
Figure 5 | Summary of late 40S maturation events. Surface representations of ten pre-40S states and the mature subunit visualize the step-wise assembly of the 3' major domain (top row), maturation of the intersubunit side and activation of NOB1 (middle row), and finally the release of remaining assembly factors (bottom row). Biogenesis factors are colored and labelled accordingly. Small inlays show the intersubunit side (states A – D) and platform (states F1 – H1), respectively.

During the transition from state A to D, several large-scale rearrangements of rRNA segments occur in a sequential manner (see Figure 6A and Ameismeier et al., 2018). Initially, rRNA helices h35 – h37 and the three-way junction consisting of h38, h39 and h40 are located at an immature position and are coordinated by the nuclear export factor RRP12. In state B, h35 – h37 have dissociated from the N-terminal half of RRP12 and relocated to their mature position. In a second step, RRP12 dissociates and enables a 140° rotation of the h38 – h40 segment. Notably, ribosomal protein eS17 had already associated to the immature three-way junction and accompanied its movement during the transition to state C. In state D, ribosomal proteins uS3, eS10, uS10 and uS14 have replaced ENP1 at the beak, which resulted in a shift of h34 and the formation of the central linker between h34, h35 and h38 (Ameismeier et al., 2018; Heuer et al., 2017b). Such pre-assembly of large sections of rRNA and subsequent repositioning is a common concept

during ribosome biogenesis. Similar events occur during the formation of the large subunit, for example when the L1-stalk and the central protuberance are initially constructed at an immature orientation and only later incorporated at their designated position (Kater et al., 2020; Kater et al., 2017; Leidig et al., 2014; Wu et al., 2016). During 60S biogenesis such major restructuring events require energy expenditure and are considered important checkpoints during assembly progression, because they are often linked to changes in subcellular localization and thus contribute to the directionality of the process (Barrio-Garcia et al., 2016; Greber, 2016; Kater et al., 2017). So far, it remains unknown what triggers the observed conformational changes in pre-40S intermediates and whether energy input is required to enable structural rearrangements. With RIOK2 and CK1δ/ε two kinases play a role in late nucleoplasmic and early cytoplasmic maturation steps and could account for energy input. Both factors have been shown to be required for efficient RRP12 release, which suggests an active role in driving the 3' major maturation (Wyler et al., 2011; Zemp et al., 2014). However, in case of RIOK2, physical presence rather than ATP hydrolysis seems to be important for maturation progression (Zemp et al., 2009). And how phosphorylation of ENP1 and LTV1 by CK1 links to the release of RRP12 and conformational changes in rRNA, or whether CK1 exerts an additional function remains also unclear. Further work is therefore required to untangle events occurring shortly before and after nuclear export.
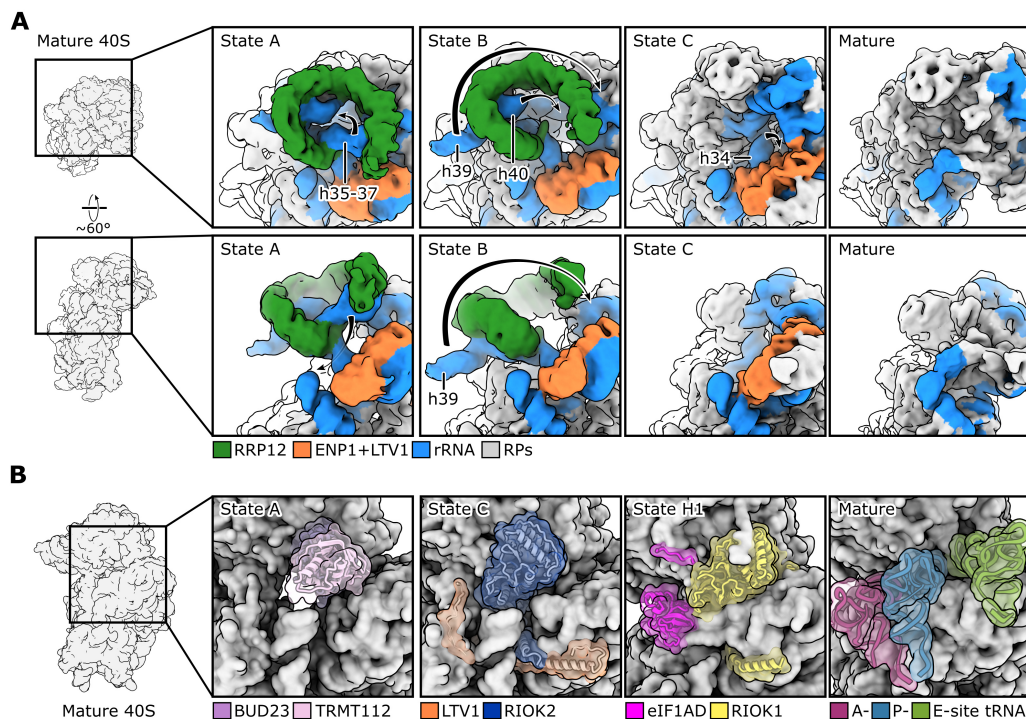
Figure 6 | Maturation of the 3' major and 3' minor domains. (A) Cryo-EM volumes of State A - C and the mature 40S showing rRNA rearrangements during assembly progression. rRNA and AFs colored and moving rRNA segments labeled. (B) Surface representation of states A, C, and H1, and the mature 40S with AFs bound to the designated tRNA binding sites shown additionally in cartoon representation and are color accordingly. Putative tRNA binding highlighted by docking an A-, P- and E-site tRNA to the mature 40S (PDBs 6TNU and 6R5Q).

Maturation of the 3' minor domain and formation of tRNA binding sites

During maturation, several factors bind and obstruct the 40S intersubunit side, preventing the formation of the tRNA binding area and the association of the large subunit (Figure 6B; Ameismeier et al., 2018; Heuer et al., 2017b; Scaiola et al., 2018; Strunk et al., 2011; Ameismeier et al., 2020). In all states prior to the release of PNO1 and the concomitant activation of NOB1, rRNA helix 44 is kept in an immature conformation, in which the top segment is shifted outwards. Thus, the linker region between h44, h45 and h28, which is a central component of the A and P sites, remains disordered (Ameismeier et al., 2018). A similar situation is also observed in yeast cytoplasmic 40S precursors (Heuer et al., 2017b; Scaiola et al., 2018; Strunk et al., 2011). *In-vitro* reconstitution experiments have shown that incubation of mature 40S subunits with AFs PNO1 and NOB1 is sufficient to reverse h44 maturation (Ameismeier et al., 2018), suggesting a functional link between h44 position and AF presence. Ribosome biogenesis factors like TSR1, LTV1 and RIOK2 probe the immature conformation of h44 by inserting terminal helices underneath the shifted segment of h44 (Ameismeier et al., 2018; Heuer et al., 2017b; Scaiola et al., 2018). This likely stabilizes the immature conformation and prevents premature h44 relocation due to steric hinderance. Thus, maturation events at the intersubunit side are presumably linked to final 18S-E processing.

Additional factors that block the designated tRNA binding sites are the methyltransferase BUD23, the RIOK2-related kinase RIOK1 and the initiation factor homolog eIF1AD (Ameismeier et al., 2020). In state A, the heterodimer BUD23-TRMT112 occupies an area surrounding the substrate G1639 between the A- and P-site (Ameismeier et al., 2018). In both yeast and humans, BUD23 has been shown to be required for efficient nuclear maturation of 40S precursors and that its physical presence rather than its catalytic activity is essential for particle export (Haag et al., 2015; White et al., 2008; Zorbas et al., 2015). However, the effect of BUD23 association on assembly progression remains unclear. Reports that Bud23 interacts with DEAH-box RNA helicase Dhr1 in yeast suggest that BUD23 has already exerted its essential function during the transition from a late 90S to the early pre-40S state (Sardana et al., 2013; Sardana et al., 2014). BUD23 is thus one of several assembly factors which play a functional role in ribosome maturation besides their catalytical activity. For example, physical presence of the yeast methyltransferase Dim1 and its human homolog DIMT1L, respectively, has been shown to be sufficient for assembly progression of the small subunit. Likewise, the methyltransferase Spb1 has been implied to coordinate central rRNA folding events during large subunit formation (Kater et al., 2020; Kater et al., 2017; Lafontaine et al., 1998b; Zorbas et al., 2015).

In later cytoplasmic stages, atypical kinases RIOK2 and RIOK1 replace BUD23 and coordinate either the immature or mature h44, respectively. In addition, both factors contain terminal extensions, which are inserted into the mRNA entry tunnel, forming a connection between both sides of the particle (Ameismeier et al., 2018; Ameismeier et al., 2020). Replacement of the C-terminal helices of RIOK2 by two helices of the unidentified factor X could indicate that the mRNA tunnel is used to coordinate maturation events on the solvent exposed side with those on the intersubunit side.

In summary, throughout the observed maturation process, the important tRNA binding site is kept in an immature conformation or shielded by assembly factors. Binding or release of each factor then coincides with major maturation events at the head or platform, respectively.

Activation of NOB1 and final 18S-E rRNA processing

The final rRNA processing step in 40S assembly is performed by the endonuclease NOB1 in both yeast and humans (Fatica et al., 2003; Sloan et al., 2019; Zhang et al., 2005). Although NOB1 binds to the particle already within the nucleus, access to its substrate at the site 3 cleavage site is restricted during late nucleoplasmic and most cytoplasmic stages (Ameismeier et al., 2018; Fatica et al., 2003; Pertschy et al., 2009). Following NOB1 incorporation, the 3' end of 18S rRNA and few nucleotides of the adjacent ITS1 remain stably enclosed between PNO1 and NOB1. Thus, similar to a situation observed in yeast 90S pre-ribosomes, where the endonuclease Utp24 is initially blocked from accessing its substrate, the PIN domain of NOB1, harboring the active site, is kept at a distance of roughly 30 Å from the rRNA cleavage site (Ameismeier et al., 2018; Ameismeier et al., 2020). The observed, inactive conformation of NOB1 explains previous chemical crosslinking results in yeast, where Nob1 preferentially interacted with h40, approximately 33 Å away from its substrate (Granneman et al., 2010; Turowski et al., 2014). Interestingly, association of the endonuclease appears different in the model organism *S. cerevisiae*. So far, Nob1 is lacking in all high-resolution structures and experimental data from both CRAC and chemical probing indicate a mostly flexible and solvent accessible 3' end (Granneman et al., 2010; Heuer et al., 2017b; Scaiola et al., 2018). Although Nob1 has been observed in a low-resolution cryo-EM structure (Strunk et al., 2011) and discrepancies in Nob1 occupancy might be accounted for by differences in sample preparation, a less robust interaction with the 40S precursor seems likely. Whether this also translates into a different mode of Nob1 activation remains to be elucidated.

In humans, activation of the endonuclease appears tightly controlled by the kinase RIOK1, which initially binds to the still immature h44 before accommodating at the matured tRNA binding sites together with eIF1AD (Figure 7A; Ameismeier et al., 2020). Both factors are essential for PNO1 release and 18S-E processing, as shown by AF recycling defects and ITS1 accumulation within the cytosol (Widmann et al., 2012; Ameismeier et al., 2020). Binding of ATP by the kinase RIOK1, which has been shown to be critical for 18S-E processing *in vitro* (Turowski et al., 2014), might contribute to conformational changes in RIOK1, allowing the kinase to settle further into the cleft between 40S body and head. Assisted by the novel biogenesis factor eIF1AD, this might drive the relocation of h44 and h28 and the formation of the linker region and tRNA binding sites. Displacement of h28 by the flexible loop of RIOK1 likely removes PNO1 from the particle and thereby allows the rotation of the NOB1 PIN domain (Figure 7B). Helix 28 thus links the completion of the functionally important tRNA binding area to final 18S maturation. Activation of the endonuclease NOB1 appears similar to A_1 processing in yeast, in which formation of the rRNA helix 1 leads to rRNA rearrangements and positioning of the A1 site substrate close to the Utp24 endonuclease (Cheng et al., in press).

Interestingly, ribosomal protein eS26 has been shown to be required for efficient 18S-E and 20S processing in humans and yeast, respectively (O'Donohue et al., 2010; Schutz et al., 2014). Following NOB1 release, eS26 is the ultimate ribosomal protein to stably incorporate into the newly synthesized 40S and replaces PNO1 and NOB1 in coordinating the 3' end of the 18S rRNA (Ameismeier et al., 2020; Anger et al., 2013). So far, structural data presented in publication 2 failed to explain a functional role of eS26 during site 3 processing, because it was observed at its canonical position and exclusively after rRNA cleavage occurred
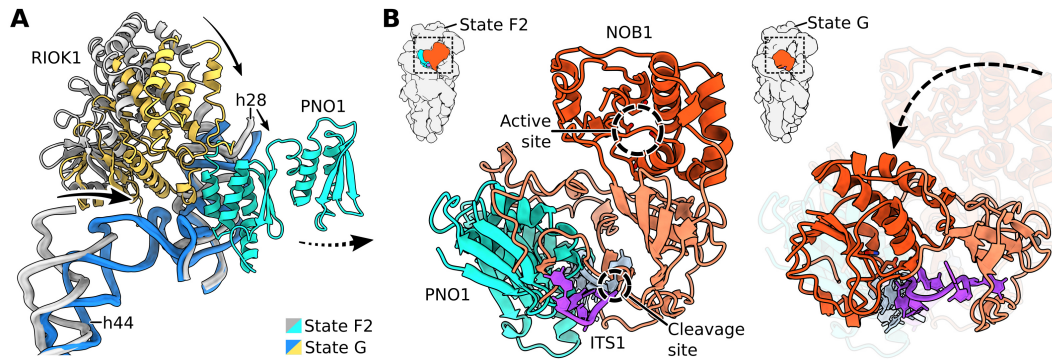
Figure 7 | Structural rearrangements during NOB1 activation. (A) Model representation of rRNA segments h44, h28 and h45, as well as RIOK1 and PNO1 in state F2 and G. Accommodation of RIOK1 and concomitant maturation of h44 and h28 removes PNO1 from the particle. Arrows show the direction of movement. Dashed arrow indicates the removal of PNO1. (B) Model representation of PNO1, ITS1 and NOB1 in state F2 (left) and G (right). Release of PNO1 leads to repositioning of the PIN domain (red) above the substrate. Active site of NOB1 and site 3 cleavage site marked with a dashed circle.

(Ameismeier et al., 2020). While an active role of eS26 in NOB1 activation cannot be excluded, its depletion could also indirectly affect ribosome biogenesis. In yeast, eS26 has previously been suggested to be imported into the nucleus by members of the importin system and escorted to the 90S pre-ribosome by the chaperone Tsr2 (Schutz et al., 2014). Binding of eS26 to early 40S precursors is reported to be dependent on interactions with ribosomal protein uS11 and the ATPase Fap7 (Pena et al., 2016) and deletion of either Fap7 or Tsr2 prevents efficient 20S rRNA processing in the cytoplasm (Granneman et al., 2005; Schutz et al., 2014). This indicates that eS26 might also be required for correct nuclear maturation and deleterious effects are delayed until the cytoplasmic stage in its absence. Further work is therefore needed to elucidate the role of ribosomal proteins during late 40S assembly, especially considering contradicting results showing a lack of eS26 on 90S particles (Sturm et al., 2017).

Final maturation in humans and yeast

In yeast, an extensive quality control mechanism has been proposed to play a central role in the final phase of 40S maturation, which involves binding of the 60S subunit and a translation-like cycle (Lebaron et al., 2012; Strunk et al., 2012). This model was developed based on by a variety of observations: first, several assembly factors and the immature 20S rRNA accumulate in 80S-like particles after inhibiting the normal maturation pathway. For example, deletion of the ATPase Fap7 was shown to lead to 80S-like particles containing AFs Dim1, Tsr1, Enp1, Nob1 and Pno1 (Ghalei et al., 2015; Granneman et al., 2005; Strunk et al., 2012) and mutation of Rio1 and disruption of its ATPase activity can trap Nob1, Tsr1, Pno1 and Fap7 on 80S-like complexes (Belhabich-Baumas et al., 2017). Similarly, deletion of Rio1 results in 80S complexes containing late 40S AFs and immature 20S rRNA (Ferreira-Cerca et al., 2014). Furthermore, particles purified via Rio1 contain, besides 20S rRNA and Nob1, also the large subunit and are competent of site D cleavage *in vitro* (Turowski et al., 2014). Such accumulation of aberrant complexes containing both pre-40S and 60S indicates that at least at some stage during 40S maturation subunit joining can occur. Second, several factors that are known to interact with 80S ribosomes also affect 40S assembly. For example, efficient rRNA processing requires the GTPase Fun12, also known as eIF5B, which mediates subunit joining

during translation initiation (Hinnebusch, 2014; Lebaron et al., 2012). This is supported by the observation that particles purified via Fun12 contain the endonuclease Nob1 and the 20S pre-rRNA (Garcia-Gomez et al., 2014). Furthermore, deletion of either Dom34 or Rli1 leads to accumulation of Nob1, Pno1, Enp1 and Dim1 in 80S-like complexes (Strunk et al., 2011) and particles purified with Rli1 also co-purify 20S rRNA (Soudet et al., 2010). Rli1 and Dom34 are both factors which bind to 80S ribosomes and separate the ribosomal subunits after translation (Dever and Green, 2012). Finally, mutations in the ribosomal protein uL3 of the large subunit suppresses the GTPase activity of Fun12 and leads to 20S processing defects (Garcia-Gomez et al., 2014).

A structure of such a pseudo-80S complex solved by cryo-EM was recently published on bioRxiv (Rai et al., 2020). Following Fap7 deletion, particles purified via the AF Tsr1 contain the methyltransferase Dim1, Pno1 and to some extent Nob1. The obtained structure showed a displaced large subunit with Dim1 wedged into the intersubunit space, however, lacked density for the endonuclease. In addition, while retaining its N-terminal helix below rRNA h44, Tsr1 has rotated away from the small subunit. Notably, large parts of the platform have unfolded with a repositioning of h24 and the loss of ribosomal proteins eS1 and uS11. But so far, it is unclear how dissociation of already matured segments aids in maturation of the 40S subunit. In summary, despite numerous indications of the formation of a pseudo-80S at least at some stage during SSU maturation, one important question remains. Is the formation of a pseudo-80S complex required for proper assembly of all 40S subunits or does binding of the large subunit occur only in a subpopulation of pre-40S particles? Accumulation of 80S-like complexes is often only observed after mutation or deletion of important factors and can therefore also be attributed to the disruption of a separate control mechanism, which would normally rescue only low abundant, off-pathway intermediates. Further work is therefore required to establish pseudo-80S formation as a universal and essential step during canonical 40S maturation.

So far, no comparable mechanism has been described which would monitor the quality of nascent human 40S subunits in an 80S-like complex. In publication 2, structures are presented which include two novel factors, both of which interact with the particle before and after cleavage (Montellese et al., 2020; Ameismeier et al., 2020). Interestingly, both eIF1AD and LRRC47 have no homologous proteins in yeast. While no functional role could yet be assigned experimentally to LRRC47, presence of the protein effectively prevents association of the large subunit in its canonical position, similar to the immature h44 in all preceding states (Ameismeier et al., 2018). Furthermore, eIF1AD, a homolog of the translation initiation factor eIF1A, has been shown to be required for 40S assembly and might serve as a potential quality control factor by monitoring the future binding site of eIF1A (Ameismeier et al., 2020). The novel biogenesis factor would thereby constitute another addition to a long list of proteins which serve as placeholders during ribosome formation (Espinar-Marchena et al., 2017). Together, this suggests an alternative route to final rRNA processing and small ribosomal subunit maturation as described in yeast. Non-conserved factors eIF1AD and LRRC47 might therein replace the function of the large subunit in driving structural rearrangements which result in the activation of NOB1. However, further experimental validation of the role of eIF1AD and its interplay with RIOK1 in PNO1 release and NOB1 activation is required.

In conclusion, the work presented in this thesis provides the first high-resolution structures of human pre-ribosomal particles and enables a detailed description of the factors involved. Based on the observation of

ten conformationally and compositionally distinct 40S precursors, general maturation principles were derived and put into the context of pre-existing functional data. Together, this provides the basis for future experiments designed to further elucidate the role of the assembly factors described here and, in addition, can serve as the foundation for structural analysis of other human pre-ribosomal particles, including the 90S and pre-60S complexes. Considering a multitude of human diseases linked to ribosome formation, including ribosomopathies and several types of cancer (Kampen et al., 2020; Mills and Green, 2017), high-resolution data might enable the development of clinical intervention strategies by structure-based drug design.

Acknowledgements

This dissertation would not have been possible without the many people who accompanied me over the years, and I am extremely grateful for the support I have experienced along the way. First and foremost, I would like to thank Roland for his supervision and the opportunity to work in his lab. I cannot express how much I enjoyed working in the productive and friendly environment he accomplishes to create in his group. I would also like to thank Prof. Karl-Peter Hopfner for being my second advisor and Prof. Karl-Peter Hopfner, Prof. Klaus Förstemann, and Dr. Gregor Witte for their part in my thesis committee and the defense. Thank you also to Prof. Julian Stingele and Dr. Dietmar Martin for their role as 5th and 6th advisor. I would like to thank QBM for their funding of my PhD.

I would like to thank all collaborators with whom I had the pleasure to work. A big shout-out to Prof. Ulrike Kutay, Dr. Ivo Zemp and Jasmin van den Heuvel for excellent functional data, their openness about their work and the ease of our collaboration. I would also like to thank Konstantin Sparrer and his team at the Ulm university for their part in our SARS-CoV-2 paper.

My biggest gratitude is due to all the people in the lab that made those 4 years truly amazing. A big thank you to all past and current members and the efforts they put into our social activities! I won't forget the fun we had at Oktoberfest, Starkbierfest (at least those parts I still remember), Halloween, Christmas, spontaneous grill and paper parties and so much more. A huge THANK YOU especially to our technicians, Andrea, Heidi, and Joanna! I am absolutely, dead certain that without you nothing I have accomplished would have been possible. You are the foundation this lab is built on and I cannot thank you enough for your support. The same holds true for the other side of the road: thank you Otto, Susi and Charlotte for all you did in teaching me how to operate the microscopes (don't touch the emergency shutdown knob… got it now!), screening all my dirty samples and providing overall excellent EM services. I would also like to thank colleagues who helped me out during my projects: Sarah for her help at the beginning of my PhD, Jingdong for his magic model building, as well as several plasmids and cell lines, and Matthias and Robert for everything they did during 6 weeks of speedy paper writing, the awesome support you provided and the fun we had as office mates! A huge thank you also to Thomas for all the support, scientific discussions, and fun we had in and outside the lab (with the occasional Whisky). We should definitely go and play Badminton at some point!

A very special thank you is due to Lukas, Kathi and Hanna. Without you my PhD might have been quicker, but also a lot less fun. I will miss the midday cooking sessions, rooftop coffee breaks and evening parties with you more than anything else. Lukas, we have been studying/working together now since the first day of our Bachelor. Thank you for all you did in the last 10 years.

Finally, I would like to devote the last words of my thesis to my family, friends, and wife. Thank you for everything.

References

Acker, M.G., Shin, B.S., Dever, T.E., and Lorsch, J.R. (2006). Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. J Biol Chem *281*, 8469-8475.

Acker, M.G., Shin, B.S., Nanda, J.S., Saini, A.K., Dever, T.E., and Lorsch, J.R. (2009). Kinetic analysis of late steps of eukaryotic translation initiation. J Mol Biol *385*, 491-506.

Agrawal, R.K., Penczek, P., Grassucci, R.A., Li, Y., Leith, A., Nierhaus, K.H., and Frank, J. (1996). Direct visualization of A-, P-, and E-site transfer RNAs in the Escherichia coli ribosome. Science *271*, 1000-1002.

Aittaleb, M., Rashid, R., Chen, Q., Palmer, J.R., Daniels, C.J., and Li, H. (2003). Structure and function of archaeal box C/D sRNP core proteins. Nat Struct Biol *10*, 256-263.

Algire, M.A., Maag, D., Savio, P., Acker, M.G., Tarun, S.Z., Jr., Sachs, A.B., Asano, K., Nielsen, K.H., Olsen, D.S., Phan, L.*, et al.* (2002). Development and characterization of a reconstituted yeast translation initiation system. RNA *8*, 382-397.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J *18*, 5399-5410.

Allmang, C., Mitchell, P., Petfalski, E., and Tollervey, D. (2000). Degradation of ribosomal RNA precursors by the exosome. Nucleic Acids Res *28*, 1684-1691.

Altvater, M., Chang, Y., Melnik, A., Occhipinti, L., Schutz, S., Rothenbusch, U., Picotti, P., and Panse, V.G. (2012). Targeted proteomics reveals compositional dynamics of 60S pre-ribosomes after nuclear export. Mol Syst Biol *8*, 628.

Ameismeier, M., Cheng, J., Berninghausen, O., and Beckmann, R. (2018). Visualizing late states of human 40S ribosomal subunit maturation. Nature *558*, 249-253.

Ameismeier, M., Zemp, I., van den Heuvel, J., Berninghausen, O., Kutay, U., and Beckmann, R. (2020). Structural basis for the final steps of human 40S ribosome maturation. Nature *587*, 683-687.

Anger, A.M., Armache, J.P., Berninghausen, O., Habeck, M., Subklewe, M., Wilson, D.N., and Beckmann, R. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80-85.

Angermayr, M., Roidl, A., and Bandlow, W. (2002). Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression. Mol Microbiol *44*, 309-324.

Armache, J.P., Anger, A.M., Marquez, V., Franckenberg, S., Frohlich, T., Villa, E., Berninghausen, O., Thomm, M., Arnold, G.J., Beckmann, R.*, et al.* (2013). Promiscuous behaviour of archaeal ribosomal proteins: implications for eukaryotic ribosome evolution. Nucleic Acids Res *41*, 1284-1293.

Asano, K., Phan, L., Valasek, L., Schoenfeld, L.W., Shalev, A., Clayton, J., Nielsen, K., Donahue, T.F., and Hinnebusch, A.G. (2001). A multifactor complex of eIF1, eIF2, eIF3, eIF5, and tRNA(i)Met promotes initiation complex assembly and couples GTP hydrolysis to AUG recognition. Cold Spring Harb Symp Quant Biol *66*, 403-415.

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

Askjaer, P., Bachi, A., Wilm, M., Bischoff, F.R., Weeks, D.L., Ogniewski, V., Ohno, M., Niehrs, C., Kjems, J., Mattaj, I.W.*, et al.* (1999). RanGTP-regulated interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase. Mol Cell Biol *19*, 6276-6285.

Aspesi, A., and Ellis, S.R. (2019). Rare ribosomopathies: insights into mechanisms of cancer. Nat Rev Cancer *19*, 228-238.

Aubert, M., O'Donohue, M.F., Lebaron, S., and Gleizes, P.E. (2018). Pre-Ribosomal RNA Processing in Human Cells: From Mechanisms to Congenital Diseases. Biomolecules *8*.

Badertscher, L., Wild, T., Montellese, C., Alexander, L.T., Bammert, L., Sarazova, M., Stebler, M., Csucs, G., Mayer, T.U., Zamboni, N.*, et al.* (2015). Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells. Cell Rep *13*, 2879-2891.

Bai, D., Zhang, J., Li, T., Hang, R., Liu, Y., Tian, Y., Huang, D., Qu, L., Cao, X., Ji, J.*, et al.* (2016). The ATPase hCINAP regulates 18S rRNA processing and is essential for embryogenesis and tumour growth. Nat Commun *7*, 12310.

Balch, W.E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008). Adapting proteostasis for disease intervention. Science *319*, 916-919.

Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassucci, R.A., Sweet, R., Frank, J., Moore, P.B., and Steitz, T.A. (1998). A 9 A resolution X-ray crystallographic map of the large ribosomal subunit. Cell *93*, 1105-1115.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science *289*, 905-920.

Barandun, J., Chaker-Margot, M., Hunziker, M., Molloy, K.R., Chait, B.T., and Klinge, S. (2017). The complete structure of the small-subunit processome. Nat Struct Mol Biol *24*, 944-953.

Barandun, J., Hunziker, M., and Klinge, S. (2018). Assembly and structure of the SSU processome-a nucleolar precursor of the small ribosomal subunit. Curr Opin Struct Biol *49*, 85-93.

Barrio-Garcia, C., Thoms, M., Flemming, D., Kater, L., Berninghausen, O., Bassler, J., Beckmann, R., and Hurt, E. (2016). Architecture of the Rix1-Rea1 checkpoint machinery during pre-60S-ribosome remodeling. Nat Struct Mol Biol *23*, 37-44.

Baxter-Roshek, J.L., Petrov, A.N., and Dinman, J.D. (2007). Optimization of ribosome structure and function by rRNA base modification. PLoS One *2*, e174.

Becker, T., Franckenberg, S., Wickles, S., Shoemaker, C.J., Anger, A.M., Armache, J.P., Sieber, H., Ungewickell, C., Berninghausen, O., Daberkow, I.*, et al.* (2012). Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. Nature *482*, 501-506.

Belhabich-Baumas, K., Joret, C., Jady, B.E., Plisson-Chastang, C., Shayan, R., Klopp, C., Henras, A.K., Henry, Y., and Mougin, A. (2017). The Rio1p ATPase hinders premature entry into translation of late pre-40S pre-ribosomal particles. Nucleic Acids Res *45*, 10824-10836.

Beltrame, M., Henry, Y., and Tollervey, D. (1994). Mutational analysis of an essential binding site for the U3 snoRNA in the 5' external transcribed spacer of yeast pre-rRNA. Nucleic Acids Res *22*, 4057-4065.

Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 A resolution. Science *334*, 1524-1529.

Berg, J.M., Tymoczko, J.L., Gatto, G.J., and Stryer, L. (2019). Biochemistry, Ninth edition. edn (New York: W.H. Freeman/Macmillan Learning).

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2015). Profiling of ribose methylations in RNA by high-throughput sequencing. Angew Chem Int Ed Engl *54*, 451-455.

Birnstiel, M.L., Wallace, H., Sirlin, J.L., and Fischberg, M. (1966). Localization of the ribosomal DNA complements in the nucleolar organizer region of Xenopus laevis. Natl Cancer Inst Monogr *23*, 431-447.

Boccaletto, P., Machnicka, M.A., Purta, E., Piatkowski, P., Baginski, B., Wirecki, T.K., de Crecy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A.*, et al.* (2018). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res *46*, D303-D307.

Bradatsch, B., Katahira, J., Kowalinski, E., Bange, G., Yao, W., Sekimoto, T., Baumgartel, V., Boese, G., Bassler, J., Wild, K.*, et al.* (2007). Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. Mol Cell *27*, 767-779.

Bradatsch, B., Leidig, C., Granneman, S., Gnadig, M., Tollervey, D., Bottcher, B., Beckmann, R., and Hurt, E. (2012). Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. Nat Struct Mol Biol *19*, 1234-1241.

Braglia, P., Heindl, K., Schleiffer, A., Martinez, J., and Proudfoot, N.J. (2010). Role of the RNA/DNA kinase Grc3 in transcription termination by RNA polymerase I. EMBO Rep *11*, 758-764.

Brand, R.C., Klootwijk, J., Van Steenbergen, T.J., De Kok, A.J., and Planta, R.J. (1977). Secondary methylation of yeast ribosomal precursor RNA. Eur J Biochem *75*, 311-318.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.-W., Zhou, S., King, D., Shen, P.S., Weibezahn, J.*, et al.* (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042-1054.

Brenner, S., Stretton, A.O., and Kaplan, S. (1965). Genetic code: the 'nonsense' triplets for chain termination and their suppression. Nature *206*, 994-998.

Briggs, M.W., Burkard, K.T., and Butler, J.S. (1998). Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J Biol Chem *273*, 13255-13263.

Brown, A., and Shao, S. (2018). Ribosomes and cryo-EM: a duet. Curr Opin Struct Biol *52*, 1-7.

Campbell, M.G., and Karbstein, K. (2011). Protein-protein interactions within late pre-40S ribosomes. PLoS One *6*, e16194.

Carron, C., O'Donohue, M.F., Choesmel, V., Faubladier, M., and Gleizes, P.E. (2011). Analysis of two human pre-ribosomal factors, bystin and hTsr1, highlights differences in evolution of ribosome biogenesis between yeast and mammals. Nucleic Acids Res *39*, 280-291.

Caskey, C.T., Tompkins, R., Scolnick, E., Caryk, T., and Nirenberg, M. (1968). Sequential translation of trinucleotide codons for the initiation and termination of protein synthesis. Science *162*, 135-138.

Cerezo, E., Plisson-Chastang, C., Henras, A.K., Lebaron, S., Gleizes, P.E., O'Donohue, M.F., Romeo, Y., and Henry, Y. (2019). Maturation of pre-40S particles in yeast and humans. Wiley Interdiscip Rev RNA *10*, e1516.

Chaker-Margot, M. (2018). Assembly of the small ribosomal subunit in yeast: mechanism and regulation. RNA *24*, 881-891.

Chaker-Margot, M., Barandun, J., Hunziker, M., and Klinge, S. (2017). Architecture of the yeast small subunit processome. Science *355*.

Chaker-Margot, M., Hunziker, M., Barandun, J., Dill, B.D., and Klinge, S. (2015). Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis. Nat Struct Mol Biol *22*, 920-923.

Cheng, J., Baßler, J., Fischer, P., Lau, B., Kellner, N., Kunze, R., Griesel, S., Kallas, M., Berninghausen, O., Strauss, D.*, et al.* (2019). Thermophile 90S Pre-ribosome Structures Reveal the Reverse Order of Co-transcriptional 18S rRNA Subdomain Integration. Molecular cell.

Cheng, J., Kellner, N., Berninghausen, O., Hurt, E., and Beckmann, R. (2017). 3.2-A-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. Nat Struct Mol Biol *24*, 954-964.

Cheng, J., Lau, B., La Venuta, G., Ameismeier, M., Berninghausen, O., Hurt, E., and Beckmann, R. 90S pre-ribosome transformation into the primordial 40S subunit. *Science, in press*

Choudhury, P., Hackert, P., Memet, I., Sloan, K.E., and Bohnsack, M.T. (2019). The human RNA helicase DHX37 is required for release of the U3 snoRNP from pre-ribosomal particles. RNA Biol *16*, 54-68.

Colley, A., Beggs, J.D., Tollervey, D., and Lafontaine, D.L. (2000). Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. Mol Cell Biol *20*, 7238-7246.

Cornish, P.V., Ermolenko, D.N., Noller, H.F., and Ha, T. (2008). Spontaneous intersubunit rotation in single ribosomes. Mol Cell *30*, 578-588.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Crick, F.H. (1958). On protein synthesis. Symp Soc Exp Biol *12*, 138-163.

Crick, F.H., Barnett, L., Brenner, S., and Watts-Tobin, R.J. (1961). General nature of the genetic code for proteins. Nature *192*, 1227-1232.

de la Cruz, J., Kressler, D., Tollervey, D., and Linder, P. (1998). Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. EMBO J *17*, 1128-1140.

Dever, T.E. (1997). Using GCN4 as a reporter of eIF2 alpha phosphorylation and translational regulation in yeast. Methods *11*, 403-417.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Dever, T.E., Kinzy, T.G., and Pavitt, G.D. (2016). Mechanism and Regulation of Protein Synthesis in Saccharomyces cerevisiae. Genetics *203*, 65-107.

Dez, C., and Tollervey, D. (2004). Ribosome synthesis meets the cell cycle. Curr Opin Microbiol *7*, 631-637.

Diaz-Villanueva, J.F., Diaz-Molina, R., and Garcia-Gonzalez, V. (2015). Protein Folding and Mechanisms of Proteostasis. Int J Mol Sci *16*, 17193-17230.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. Trends Genet *23*, 614-622.

Dong, X., Biswas, A., Suel, K.E., Jackson, L.K., Martinez, R., Gu, H., and Chook, Y.M. (2009). Structural basis for leucine-rich nuclear export signal recognition by CRM1. Nature *458*, 1136-1141.

Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., Mitchell, B.M., Porwancher, K.A., Wehner, K.A., Wormsley, S., Settlage, R.E., Shabanowitz, J., Osheim, Y.*, et al.* (2002). A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature *417*, 967-970.

Dutca, L.M., Gallagher, J.E., and Baserga, S.J. (2011). The initial U3 snoRNA:pre-rRNA base pairing interaction required for pre-18S rRNA folding revealed by in vivo chemical probing. Nucleic Acids Res *39*, 5164-5180.

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

Englmeier, L., Fornerod, M., Bischoff, F.R., Petosa, C., Mattaj, I.W., and Kutay, U. (2001). RanBP3 influences interactions between CRM1 and its nuclear protein export substrates. EMBO Rep *2*, 926-932.

Espinar-Marchena, F.J., Babiano, R., and de la Cruz, J. (2017). Placeholder factors in ribosome biogenesis: please, pave my way. Microbial Cell *4*, 144-168.

Faber, A.W., Van Dijk, M., Raue, H.A., and Vos, J.C. (2002). Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in Saccharomyces cerevisiae. RNA *8*, 1095-1101.

Farley-Barnes, K.I., McCann, K.L., Ogawa, L.M., Merkel, J., Surovtseva, Y.V., and Baserga, S.J. (2018). Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. Cell Rep *22*, 1923-1934.

Farley-Barnes, K.I., Ogawa, L.M., and Baserga, S.J. (2019). Ribosomopathies: Old Concepts, New Controversies. Trends Genet *35*, 754-767.

Fassio, C.A., Schofield, B.J., Seiser, R.M., Johnson, A.W., and Lycan, D.E. (2010). Dominant mutations in the late 40S biogenesis factor Ltv1 affect cytoplasmic maturation of the small ribosomal subunit in Saccharomyces cerevisiae. Genetics *185*, 199-209.

Fatica, A., Oeffinger, M., Dlakic, M., and Tollervey, D. (2003). Nob1p is required for cleavage of the 3' end of 18S rRNA. Mol Cell Biol *23*, 1798-1807.

Fatica, A., Tollervey, D., and Dlakic, M. (2004). PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. RNA *10*, 1698-1701.

Faza, M.B., Chang, Y., Occhipinti, L., Kemmler, S., and Panse, V.G. (2012). Role of Mex67-Mtr2 in the nuclear export of 40S pre-ribosomes. PLoS Genet *8*, e1002915.

Fernandez-Pevida, A., Kressler, D., and de la Cruz, J. (2015). Processing of preribosomal RNA in Saccharomyces cerevisiae. Wiley Interdiscip Rev RNA *6*, 191-209.

Ferreira-Cerca, S., Kiburu, I., Thomson, E., LaRonde, N., and Hurt, E. (2014). Dominant Rio1 kinase/ATPase catalytic mutant induces trapping of late pre-40S biogenesis factors in 80S-like ribosomes. Nucleic Acids Res *42*, 8635-8647.

Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H., and Milkereit, P. (2005). Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. Mol Cell *20*, 263-275.

Ferreira-Cerca, S., Sagar, V., Schafer, T., Diop, M., Wesseling, A.M., Lu, H., Chai, E., Hurt, E., and LaRonde-LeBlanc, N. (2012). ATPase-dependent role of the atypical kinase Rio2 on the evolving pre-40S ribosomal subunit. Nat Struct Mol Biol *19*, 1316-1323.

Figaro, S., Wacheul, L., Schillewaert, S., Graille, M., Huvelle, E., Mongeard, R., Zorbas, C., Lafontaine, D.L., and Heurgue-Hamard, V. (2012). Trm112 is required for Bud23-mediated methylation of the 18S rRNA at position G1575. Mol Cell Biol *32*, 2254-2267.

Fischer, U., Schauble, N., Schutz, S., Altvater, M., Chang, Y., Faza, M.B., and Panse, V.G. (2015). A non-canonical mechanism for Crm1-export cargo complex assembly. Elife *4*.

Fox, G.E. (2010). Origin and Evolution of the Ribosome. Cold Spring Harbor Perspectives in Biology *2*, a003483-a003483.

Franckenberg, S., Becker, T., and Beckmann, R. (2012). Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. Curr Opin Struct Biol *22*, 786-796.

Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature *406*, 318-322.

Frank, J., Zhu, J., Penczek, P., Li, Y., Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., Lata, R.K., and Agrawal, R.K. (1995). A model of protein synthesis based on cryo-electron microscopy of the E. coli ribosome. Nature *376*, 441-444.

Fribourg, S., and Conti, E. (2003). Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. EMBO Rep *4*, 699-703.

Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrerio, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science *295*, 2258-2261.

Fromm, L., Falk, S., Flemming, D., Schuller, J.M., Thoms, M., Conti, E., and Hurt, E. (2017). Reconstitution of the complete pathway of ITS2 processing at the pre-ribosome. Nature communications *8*, 1787-1787.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene *313*, 17-42.

Fung, H.Y., Fu, S.C., Brautigam, C.A., and Chook, Y.M. (2015). Structural determinants of nuclear export signal orientation in binding to exportin CRM1. Elife *4*.

Gamalinda, M., Ohmayer, U., Jakovljevic, J., Kumcuoglu, B., Woolford, J., Mbom, B., Lin, L., and Woolford, J.L. (2014). A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. Genes & Development *28*, 198-210.

Garcia-Gomez, J.J., Fernandez-Pevida, A., Lebaron, S., Rosado, I.V., Tollervey, D., Kressler, D., and de la Cruz, J. (2014). Final pre-40S maturation depends on the functional integrity of the 60S subunit ribosomal protein L3. PLoS Genet *10*, e1004205.

Gasse, L., Flemming, D., and Hurt, E. (2015). Coordinated ribosomal ITS2 RNA processing by the Las1 complex integrating endonuclease, polynucleotide kinase, and exonuclease activities. Molecular cell *60*, 808-815.

Gautier, T., Berges, T., Tollervey, D., and Hurt, E. (1997). Nucleolar KKE/D repeat proteins Nop56p and Nop58p interact with Nop1p and are required for ribosome biogenesis. Mol Cell Biol *17*, 7088-7098.

Geerlings, T.H., Faber, A.W., Bister, M.D., Vos, J.C., and Raue, H.A. (2003). Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20 S Pre-rRNA in Saccharomyces cerevisiae. J Biol Chem *278*, 22537-22545.

Geerlings, T.H., Vos, J.C., and Raue, H.A. (2000). The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-->3' exonucleases. RNA *6*, 1698-1703.

Ghalei, H., Schaub, F.X., Doherty, J.R., Noguchi, Y., Roush, W.R., Cleveland, J.L., Stroupe, M.E., and Karbstein, K. (2015). Hrr25/CK1delta-directed release of Ltv1 from pre-40S ribosomes is necessary for ribosome assembly and cell growth. J Cell Biol *208*, 745-759.

Ghalei, H., Trepreau, J., Collins, J.C., Bhaskaran, H., Strunk, B.S., and Karbstein, K. (2017). The ATPase Fap7 Tests the Ability to Carry Out Translocation-like Conformational Changes and Releases Dim1 during 40S Ribosome Maturation. Mol Cell *67*, 990-1000 e1003.

Gibbons, J.G., Branco, A.T., Godinho, S.A., Yu, S., and Lemos, B. (2015). Concerted copy number variation balances ribosomal DNA dosage in human and mouse genomes. Proc Natl Acad Sci U S A *112*, 2485-2490.

Gingras, A.C., Gygi, S.P., Raught, B., Polakiewicz, R.D., Abraham, R.T., Hoekstra, M.F., Aebersold, R., and Sonenberg, N. (1999). Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev *13*, 1422-1437.

Goldfarb, K.C., and Cech, T.R. (2017). Targeted CRISPR disruption reveals a role for RNase MRP RNA in human preribosomal RNA processing. Genes Dev *31*, 59-71.

Gonzalez, I.L., and Sylvester, J.E. (1995). Complete sequence of the 43-kb human ribosomal DNA repeat: analysis of the intergenic spacer. Genomics *27*, 320-328.

Graf, M., Arenz, S., Huter, P., Donhofer, A., Novacek, J., and Wilson, D.N. (2017). Cryo-EM structure of the spinach chloroplast ribosome reveals the location of plastid-specific ribosomal proteins and extensions. Nucleic Acids Res *45*, 2887-2896.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D.*, et al.* (2002). 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Mol Cell *10*, 105-115.

Granneman, S., Nandineni, M.R., and Baserga, S.J. (2005). The putative NTPase Fap7 mediates cytoplasmic 20S pre-rRNA processing through a direct interaction with Rps14. Mol Cell Biol *25*, 10352-10364.

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

Granneman, S., and Tollervey, D. (2007). Building ribosomes: even more expensive than expected? Curr Biol *17*, R415-417.

Grant, R.P., Hurt, E., Neuhaus, D., and Stewart, M. (2002). Structure of the C-terminal FG-nucleoporin binding domain of Tap/NXF1. Nat Struct Biol *9*, 247-251.

Greber, B.J. (2016). Mechanistic insight into eukaryotic 60S ribosomal subunit biogenesis by cryo-electron microscopy. RNA *22*, 1643-1662.

Greber, B.J., Bieri, P., Leibundgut, M., Leitner, A., Aebersold, R., Boehringer, D., and Ban, N. (2015). Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. Science *348*, 303-308.

Greber, B.J., Boehringer, D., Godinic-Mikulcic, V., Crnkovic, A., Ibba, M., Weygand-Durasevic, I., and Ban, N. (2012a). Cryo-EM structure of the archaeal 50S ribosomal subunit in complex with initiation factor 6 and implications for ribosome evolution. J Mol Biol *418*, 145-160.

Greber, B.J., Boehringer, D., Montellese, C., and Ban, N. (2012b). Cryo-EM structures of Arx1 and maturation factors Rei1 and Jjj1 bound to the 60S ribosomal subunit. Nat Struct Mol Biol *19*, 1228-1233.

Greenberg, H., and Penman, S. (1966). Methylation and processing of ribosomal RNA in HeLa cells. J Mol Biol *21*, 527-535.

Gromadski, K.B., Daviter, T., and Rodnina, M.V. (2006). A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. Mol Cell *21*, 369-377.

Grummt, I. (2003). Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. Genes Dev *17*, 1691-1702.

Guttler, T., Madl, T., Neumann, P., Deichsel, D., Corsini, L., Monecke, T., Ficner, R., Sattler, M., and Gorlich, D. (2010). NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. Nat Struct Mol Biol *17*, 1367-1376.

Haag, S., Kretschmer, J., and Bohnsack, M.T. (2015). WBSCR22/Merm1 is required for late nuclear pre-ribosomal RNA processing and mediates N7-methylation of G1639 in human 18S rRNA. RNA *21*, 180-187.

Haeusler, R.A., and Engelke, D.R. (2006). Spatial organization of transcription by RNA polymerase III. Nucleic Acids Res *34*, 4826-4836.

Hansen, J.L., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2002). Structural insights into peptide bond formation. Proc Natl Acad Sci U S A *99*, 11670-11675.

Hector, R.D., Burlacu, E., Aitken, S., Le Bihan, T., Tuijtel, M., Zaplatina, A., Cook, A.G., and Granneman, S. (2014). Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. Nucleic Acids Res *42*, 12138-12154.

Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A., and Gleizes, P.E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdiscip Rev RNA *6*, 225-242.

Henras, A.K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A., and Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci *65*, 2334-2359.

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

Heuer, A., Gerovac, M., Schmidt, C., Trowitzsch, S., Preis, A., Kötter, P., Berninghausen, O., Becker, T., Beckmann, R., and Tampé, R. (2017a). Structure of the 40S-ABCE1 post-splitting complex in ribosome recycling and translation initiation. Nature structural & molecular biology.

Heuer, A., Thomson, E., Schmidt, C., Berninghausen, O., Becker, T., Hurt, E., and Beckmann, R. (2017b). Cryo-EM structure of a late pre-40S ribosomal subunit from Saccharomyces cerevisiae. Elife *6*.

Hierlmeier, T., Merl, J., Sauert, M., Perez-Fernandez, J., Schultz, P., Bruckmann, A., Hamperl, S., Ohmayer, U., Rachel, R., Jacob, A*., et al.* (2013). Rrp5p, Noc1p and Noc2p form a protein module which is part of early large ribosomal subunit precursors in S. cerevisiae. Nucleic acids research *41*, 1191-1210.

Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem *83*, 779-812.

Hinnebusch, A.G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. Trends Biochem Sci *42*, 589-611.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb Perspect Biol *4*.

Ho, J.H., Kallstrom, G., and Johnson, A.W. (2000). Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. J Cell Biol *151*, 1057-1066.

Hughes, J.M. (1996). Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA. J Mol Biol *259*, 645-654.

Hughes, J.M., and Ares, M., Jr. (1991). Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. EMBO J *10*, 4231-4239.

Hung, N.J., and Johnson, A.W. (2006). Nuclear recycling of the pre-60S ribosomal subunit-associated factor Arx1 depends on Rei1 in Saccharomyces cerevisiae. Mol Cell Biol *26*, 3718-3727.

Hung, N.J., Lo, K.Y., Patel, S.S., Helmke, K., and Johnson, A.W. (2008). Arx1 is a nuclear export receptor for the 60S ribosomal subunit in yeast. Mol Biol Cell *19*, 735-744.

Hunziker, M., Barandun, J., Petfalski, E., Tan, D., Delan-Forino, C., Molloy, K.R., Kim, K.H., Dunn-Davies, H., Shi, Y., Chaker-Margot, M.*, et al.* (2016). UtpA and UtpB chaperone nascent pre-ribosomal RNA and U3 snoRNA to initiate eukaryotic ribosome assembly. Nat Commun *7*, 12090.

Ishikawa, H., Yoshikawa, H., Izumikawa, K., Miura, Y., Taoka, M., Nobe, Y., Yamauchi, Y., Nakayama, H., Simpson, R.J., Isobe, T.*, et al.* (2017). Poly(A)-specific ribonuclease regulates the processing of small-subunit rRNAs in human cells. Nucleic Acids Res *45*, 3437-3447.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.

Jakobsson, M.E., Malecki, J., and Falnes, P.O. (2018). Regulation of eukaryotic elongation factor 1 alpha (eEF1A) by dynamic lysine methylation. RNA Biol *15*, 314-319.

Jin, H., Kelley, A.C., Loakes, D., and Ramakrishnan, V. (2010). Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. Proc Natl Acad Sci U S A *107*, 8593-8598.

Johnson, A.W. (1997). Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. Mol Cell Biol *17*, 6122-6130.

Kaczanowska, M., and Ryden-Aulin, M. (2007). Ribosome biogenesis and the translation process in Escherichia coli. Microbiol Mol Biol Rev *71*, 477-494.

Kampen, K.R., Sulima, S.O., Vereecke, S., and De Keersmaecker, K. (2020). Hallmarks of ribosomopathies. Nucleic Acids Res *48*, 1013-1028.

Kapp, L.D., and Lorsch, J.R. (2004). GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. J Mol Biol *335*, 923-936.

Karamyshev, A.L., and Karamysheva, Z.N. (2018). Lost in Translation: Ribosome-Associated mRNA and Protein Quality Controls. Front Genet *9*, 431.

Kargas, V., Castro-Hartmann, P., Escudero-Urquijo, N., Dent, K., Hilcenko, C., Sailer, C., Zisser, G., Marques-Carvalho, M.J., Pellegrino, S., Wawiórka, L.*, et al.* (2019). Mechanism of completion of peptidyltransferase centre assembly in eukaryotes. eLife *8*.

Kass, S., Craig, N., and Sollner-Webb, B. (1987). Primary processing of mammalian rRNA involves two adjacent cleavages and is not species specific. Mol Cell Biol *7*, 2891-2898.

Kass, S., Tyc, K., Steitz, J.A., and Sollner-Webb, B. (1990). The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. Cell *60*, 897-908.

Kater, L., Mitterer, V., Thoms, M., Cheng, J., Berninghausen, O., Beckmann, R., and Hurt, E. (2020). Construction of the Central Protuberance and L1 Stalk during 60S Subunit Biogenesis. Mol Cell.

Kater, L., Thoms, M., Barrio-Garcia, C., Cheng, J., Ismail, S., Ahmed, Y.L., Bange, G., Kressler, D., Berninghausen, O., Sinning, I*., et al.* (2017). Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. Cell *171*, 1599-1610 e1514.

Kharde, S., Calviño, F.R., Gumiero, A., Wild, K., and Sinning, I. (2015). The structure of Rpf2-Rrs1 explains its role in ribosome biogenesis. Nucleic acids research *43*, 7083-7095.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640-645.

Khoshnevis, S., Liu, X., Dattolo, M.D., and Karbstein, K. (2019). Rrp5 establishes a checkpoint for 60S assembly during 40S maturation. RNA *25*, 1164-1176.

Khusainov, I., Vicens, Q., Bochler, A., Grosse, F., Myasnikov, A., Menetret, J.F., Chicher, J., Marzi, S., Romby, P., Yusupova, G*., et al.* (2016). Structure of the 70S ribosome from human pathogen Staphylococcus aureus. Nucleic Acids Res *44*, 10491-10504.

Klauer, A.A., and van Hoof, A. (2012). Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. Wiley Interdiscip Rev RNA *3*, 649-660.

Klinge, S., and Woolford, J.L., Jr. (2019). Ribosome assembly coming into focus. Nat Rev Mol Cell Biol *20*, 116-131.

Kolitz, S.E., and Lorsch, J.R. (2010). Eukaryotic initiator tRNA: finely tuned and ready for action. FEBS Lett *584*, 396-404.

Korn, A.P. (1980). Dark field electron microscopy of the 30S subunit of the E. coli ribosome. Ultramicroscopy *5*, 513-520.

Korneeva, N.L., Lamphear, B.J., Hennigan, F.L., and Rhoads, R.E. (2000). Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. J Biol Chem *275*, 41369-41376.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Kos-Braun, I., Kos, M., Thoms, M., Berninghausen, O., Beckmann, R*., et al.* (2016). Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. Cell *166*, 380-393.

Korobeinikova, A.V., Garber, M.B., and Gongadze, G.M. (2012). Ribosomal proteins: structure, function, and evolution. Biochemistry (Mosc) *77*, 562-574.

Kos, M., and Tollervey, D. (2010). Yeast pre-rRNA processing and modification occur cotranscriptionally. Molecular cell *37*, 809-820.

Koyama, M., and Matsuura, Y. (2010). An allosteric mechanism to displace nuclear export cargo from CRM1 and RanGTP by RanBP1. EMBO J *29*, 2002-2013.

Kozak, M. (1986). Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc Natl Acad Sci U S A *83*, 2850-2854.

Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat Struct Mol Biol *16*, 589-597.

Kressler, D., Hurt, E., and Baßler, J. (2017). A Puzzle of Life: Crafting Ribosomal Subunits. Trends in Biochemical Sciences.

Kressler, D., Hurt, E., and Bassler, J. (2010). Driving ribosome assembly. Biochimica et biophysica acta *1803*, 673-683.

Kressler, D., Roser, D., Pertschy, B., and Hurt, E. (2008). The AAA ATPase Rix7 powers progression of ribosome biogenesis by stripping Nsa1 from pre-60S particles. The Journal of cell biology *181*, 935-944.

Kufel, J., Dichtl, B., and Tollervey, D. (1999). Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. RNA *5*, 909-917.

Kuhn, H., Hierlmeier, T., Merl, J., Jakob, S., Aguissa-Toure, A.H., Milkereit, P., and Tschochner, H. (2009). The Noc-domain containing C-terminus of Noc4p mediates both formation of the Noc4p-Nop14p submodule and its incorporation into the SSU processome. PLoS One *4*, e8370.

Lafontaine, D., Delcour, J., Glasser, A.L., Desgres, J., and Vandenhaute, J. (1994). The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18 S rRNA is essential in yeast. J Mol Biol *241*, 492-497.

Lafontaine, D., Vandenhaute, J., and Tollervey, D. (1995). The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. Genes Dev *9*, 2470-2481.

Lafontaine, D.L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M., and Tollervey, D. (1998a). The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. Genes Dev *12*, 527-537.

Lafontaine, D.L., Preiss, T., and Tollervey, D. (1998b). Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? Mol Cell Biol *18*, 2360-2370.

Landry-Voyer, A.M., Bilodeau, S., Bergeron, D., Dionne, K.L., Port, S.A., Rouleau, C., Boisvert, F.M., Kehlenbach, R.H., and Bachand, F. (2016). Human PDCD2L Is an Export Substrate of CRM1 That Associates with 40S Ribosomal Subunit Precursors. Mol Cell Biol *36*, 3019-3032.

Lapik, Y.R., Fernandes, C.J., Lau, L.F., and Pestov, D.G. (2004). Physical and functional interaction between Pes1 and Bop1 in mammalian ribosome biogenesis. Mol Cell *15*, 17-29.

Larburu, N., Montellese, C., O'Donohue, M.F., Kutay, U., Gleizes, P.E., and Plisson-Chastang, C. (2016). Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. Nucleic Acids Res *44*, 8465-8478.

LaRonde-LeBlanc, N., and Wlodawer, A. (2004). Crystal structure of A. fulgidus Rio2 defines a new family of serine protein kinases. Structure *12*, 1585-1594.

Lazdins, I.B., Delannoy, M., and Sollner-Webb, B. (1997). Analysis of nucleolar transcription and processing domains and pre-rRNA movements by in situ hybridization. Chromosoma *105*, 481-495.

Lebaron, S., Schneider, C., van Nues, R.W., Swiatkowska, A., Walsh, D., Bottcher, B., Granneman, S., Watkins, N.J., and Tollervey, D. (2012). Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. Nat Struct Mol Biol *19*, 744-753.

Lebaron, S., Segerstolpe, A., French, S.L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A.L., Wieslander, L., and Tollervey, D. (2013). Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. Mol Cell *52*, 707-719.

Lebreton, A., Saveanu, C., Decourty, L., Rain, J.C., Jacquier, A., and Fromont-Racine, M. (2006). A functional network involved in the recycling of nucleocytoplasmic pre-60S factors. J Cell Biol *173*, 349-360.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D., and Poch, O. (2002). Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acids Res *30*, 5382-5390.

Lee, J.H., Pestova, T.V., Shin, B.S., Cao, C., Choi, S.K., and Dever, T.E. (2002). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. Proc Natl Acad Sci U S A *99*, 16689-16694.

LeFebvre, A.K., Korneeva, N.L., Trutschl, M., Cvek, U., Duzan, R.D., Bradley, C.A., Hershey, J.W., and Rhoads, R.E. (2006). Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. J Biol Chem *281*, 22917-22932.

Leger-Silvestre, I., Milkereit, P., Ferreira-Cerca, S., Saveanu, C., Rousselle, J.C., Choesmel, V., Guinefoleau, C., Gas, N., and Gleizes, P.E. (2004). The ribosomal protein Rps15p is required for nuclear exit of the 40S subunit precursors in yeast. EMBO J *23*, 2336-2347.

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

Letoquart, J., Huvelle, E., Wacheul, L., Bourgeois, G., Zorbas, C., Graille, M., Heurgue-Hamard, V., and Lafontaine, D.L. (2014). Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. Proc Natl Acad Sci U S A *111*, E5518-5526.

Lewis, J.D., and Tollervey, D. (2000). Like attracts like: getting RNA processing together in the nucleus. Science *288*, 1385-1389.

Liang, X.H., Liu, Q., and Fournier, M.J. (2007). rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. Mol Cell *28*, 965-977.

Liang, X.H., Liu, Q., and Fournier, M.J. (2009). Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. RNA *15*, 1716-1728.

Lin, J., Lu, J., Feng, Y., Sun, M., and Ye, K. (2013). An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme. PLoS Biol *11*, e1001669.

Lindsay, M.E., Holaska, J.M., Welch, K., Paschal, B.M., and Macara, I.G. (2001). Ran-binding protein 3 is a cofactor for Crm1-mediated nuclear protein export. J Cell Biol *153*, 1391-1402.

Liu, P.C., and Thiele, D.J. (2001). Novel stress-responsive genes EMG1 and NOP14 encode conserved, interacting proteins required for 40S ribosome biogenesis. Mol Biol Cell *12*, 3644-3657.

Liu, Q., and Fredrick, K. (2016). Intersubunit Bridges of the Bacterial Ribosome. J Mol Biol *428*, 2146-2164.

Liu, W., Chen, C., Kavaliauskas, D., Knudsen, C.R., Goldman, Y.E., and Cooperman, B.S. (2015). EF-Tu dynamics during pre-translocation complex formation: EF-Tu.GDP exits the ribosome via two different pathways. Nucleic Acids Res *43*, 9519-9528.

Lo, K.Y., Li, Z., Bussiere, C., Bresson, S., Marcotte, E.M., and Johnson, A.W. (2010). Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. Mol Cell *39*, 196-208.

Lo, K.Y., Li, Z., Wang, F., Marcotte, E.M., and Johnson, A.W. (2009). Ribosome stalk assembly requires the dual-specificity phosphatase Yvh1 for the exchange of Mrt4 with P0. J Cell Biol *186*, 849-862.

Lorsch, J.R., and Herschlag, D. (1998). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. Biochemistry *37*, 2180-2193.

Ma, C., Wu, S., Li, N., Chen, Y., Yan, K., Li, Z., Zheng, L., Lei, J., Woolford, J.L., Jr., and Gao, N. (2017). Structural snapshot of cytoplasmic pre-60S ribosomal particles bound by Nmd3, Lsg1, Tif6 and Reh1. Nat Struct Mol Biol *24*, 214-220.

Majumdar, R., Bandyopadhyay, A., and Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. J Biol Chem *278*, 6580-6587.

Malyutin, A.G., Musalgaonkar, S., Patchett, S., Frank, J., and Johnson, A.W. (2017). Nmd3 is a structural mimic of eIF5A, and activates the cpGTPase Lsg1 during 60S ribosome biogenesis. The EMBO journal *36*, 854-868.

Marintchev, A., Kolupaeva, V.G., Pestova, T.V., and Wagner, G. (2003). Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: a new interaction between old partners. Proc Natl Acad Sci U S A *100*, 1535-1540.

Martinez, J., Ren, Y.G., Thuresson, A.C., Hellman, U., Astrom, J., and Virtanen, A. (2000). A 54-kDa fragment of the Poly(A)-specific ribonuclease is an oligomeric, processive, and cap-interacting Poly(A)-specific 3' exonuclease. J Biol Chem *275*, 24222-24230.

Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015). Structure of a human translation termination complex. Nucleic Acids Res *43*, 8615-8626.

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

Mattijssen, S., Welting, T.J., and Pruijn, G.J. (2010). RNase MRP and disease. Wiley Interdiscip Rev RNA *1*, 102-116.

Maurer, P., Redd, M., Solsbacher, J., Bischoff, F.R., Greiner, M., Podtelejnikov, A.V., Mann, M., Stade, K., Weis, K., and Schlenstedt, G. (2001). The nuclear export receptor Xpo1p forms distinct complexes with NES transport substrates and the yeast Ran binding protein 1 (Yrb1p). Mol Biol Cell *12*, 539-549.

McCaughan, U.M., Jayachandran, U., Shchepachev, V., Chen, Z.A., Rappsilber, J., Tollervey, D., and Cook, A.G. (2016). Pre-40S ribosome biogenesis factor Tsr1 is an inactive structural mimic of translational GTPases. Nat Commun *7*, 11789.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nat Struct Mol Biol *19*, 560-567.

Merwin, J.R., Bogar, L.B., Poggi, S.B., Fitch, R.M., Johnson, A.W., and Lycan, D.E. (2014). Genetic analysis of the ribosome biogenesis factor Ltv1 of Saccharomyces cerevisiae. Genetics *198*, 1071-1085.

Methot, N., Song, M.S., and Sonenberg, N. (1996). A region rich in aspartic acid, arginine, tyrosine, and glycine (DRYG) mediates eukaryotic initiation factor 4B (eIF4B) self-association and interaction with eIF3. Mol Cell Biol *16*, 5328-5334.

Meyer, A.E., Hoover, L.A., and Craig, E.A. (2010). The cytosolic J-protein, Jjj1, and Rei1 function in the removal of the pre-60 S subunit factor Arx1. J Biol Chem *285*, 961-968.

Miller, O.L., Jr., and Beatty, B.R. (1969). Visualization of nucleolar genes. Science *164*, 955-957.

Mills, E.W., and Green, R. (2017). Ribosomopathies: There's strength in numbers. Science *358*.

Mitchell, P. (2010). Rrp47 and the function of the Sas10/C1D domain. Biochem Soc Trans *38*, 1088-1092.

Mitchell, P., Petfalski, E., and Tollervey, D. (1996). The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. Genes Dev *10*, 502-513.

Mitchell, S.F., Walker, S.E., Algire, M.A., Park, E.H., Hinnebusch, A.G., and Lorsch, J.R. (2010). The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. Mol Cell *39*, 950-962.

Mitterer, V., Murat, G., Rety, S., Blaud, M., Delbos, L., Stanborough, T., Bergler, H., Leulliot, N., Kressler, D., and Pertschy, B. (2016). Sequential domain assembly of ribosomal protein S3 drives 40S subunit maturation. Nat Commun *7*, 10336.

Montellese, C., Montel-Lehry, N., Henras, A.K., Kutay, U., Gleizes, P.E., and O'Donohue, M.F. (2017). Poly(A)-specific ribonuclease is a nuclear ribosome biogenesis factor involved in human 18S rRNA maturation. Nucleic Acids Res *45*, 6822-6836.

Montellese, C., van den Heuvel, J., Ashiono, C., Dorner, K., Melnik, A., Jonas, S., Zemp, I., Picotti, P., Gillet, L.C., and Kutay, U. (2020). USP16 counteracts mono-ubiquitination of RPS27a and promotes maturation of the 40S ribosomal subunit. Elife *9*.

Moon, D.H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., Tai, A.K., and Agarwal, S. (2015). Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. Nat Genet *47*, 1482-1488.

Moriggi, G., Nieto, B., and Dosil, M. (2014). Rrp12 and the Exportin Crm1 participate in late assembly events in the nucleolus during 40S ribosomal subunit biogenesis. PLoS Genet *10*, e1004836.

Mougey, E.B., O'Reilly, M., Osheim, Y., Miller, O.L., Jr., Beyer, A., and Sollner-Webb, B. (1993). The terminal balls characteristic of eukaryotic rRNA transcription units in chromatin spreads are rRNA processing complexes. Genes Dev *7*, 1609-1619.

Moy, T.I., and Silver, P.A. (1999). Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. Genes Dev *13*, 2118-2133.

Moy, T.I., and Silver, P.A. (2002). Requirements for the nuclear export of the small ribosomal subunit. J Cell Sci *115*, 2985-2995.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920-930.

Nunes, V.S., and Moretti, N.S. (2017). Nuclear subcompartments: an overview. Cell Biol Int *41*, 2-7.

O'Donohue, M.F., Choesmel, V., Faubladier, M., Fichant, G., and Gleizes, P.E. (2010). Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. J Cell Biol *190*, 853-866.

Oeffinger, M., Dlakic, M., and Tollervey, D. (2004). A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. Genes Dev *18*, 196-209.

Oeffinger, M., Zenklusen, D., Ferguson, A., Wei, K.E., El Hage, A., Tollervey, D., Chait, B.T., Singer, R.H., and Rout, M.P. (2009). Rrp17p is a eukaryotic exonuclease required for 5' end processing of Pre-60S ribosomal RNA. Mol Cell *36*, 768-781.

Olsen, D.S., Savner, E.M., Mathew, A., Zhang, F., Krishnamoorthy, T., Phan, L., and Hinnebusch, A.G. (2003). Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. EMBO J *22*, 193-204.

Osheim, Y.N., French, S.L., Keck, K.M., Champion, E.A., Spasov, K., Dragon, F., Baserga, S.J., and Beyer, A.L. (2004). Pre-18S ribosomal RNA is structurally compacted into the SSU processome prior to being cleaved from nascent transcripts in Saccharomyces cerevisiae. Molecular cell *16*, 943-954.

Ounap, K., Kasper, L., Kurg, A., and Kurg, R. (2013). The human WBSCR22 protein is involved in the biogenesis of the 40S ribosomal subunits in mammalian cells. PLoS One *8*, e75686.

Ounap, K., Leetsi, L., Matsoo, M., and Kurg, R. (2015). The Stability of Ribosome Biogenesis Factor WBSCR22 Is Regulated by Interaction with TRMT112 via Ubiquitin-Proteasome Pathway. PLoS One *10*, e0133841.

Pape, T., Wintermeyer, W., and Rodnina, M. (1999). Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. EMBO J *18*, 3800-3807.

Pechmann, S., Willmund, F., and Frydman, J. (2013). The ribosome as a hub for protein quality control. Mol Cell *49*, 411-421.

Pena, C., Hurt, E., and Panse, V.G. (2017). Eukaryotic ribosome assembly, transport and quality control. Nat Struct Mol Biol *24*, 689-699.

Pena, C., Schutz, S., Fischer, U., Chang, Y., and Panse, V.G. (2016). Prefabrication of a ribosomal protein subcomplex essential for eukaryotic ribosome formation. Elife *5*.

Pérez-Fernández, J., Román, A., De Las Rivas, J., Bustelo, X.R., and Dosil, M. (2007). The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. Molecular and cellular biology *27*, 5414-5429.

Pertschy, B., Saveanu, C., Zisser, G., Lebreton, A., Tengg, M., Jacquier, A., Liebminger, E., Nobis, B., Kappel, L., van der Klei, I.*, et al.* (2007). Cytoplasmic recycling of 60S preribosomal factors depends on the AAA protein Drg1. Molecular and cellular biology *27*, 6581-6592.

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

Pestova, T.V., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., and Hellen, C.U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature *403*, 332-335.

Piazzi, M., Bavelloni, A., Gallo, A., Faenza, I., and Blalock, W.L. (2019). Signal Transduction in Ribosome Biogenesis: A Recipe to Avoid Disaster. Int J Mol Sci *20*.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286-299.

Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U., and Pestova, T.V. (2006). Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. Genes Dev *20*, 624-636.

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Mol Cell *37*, 196-210.

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D.E., Berninghausen, O., Green, R., Becker, T., and Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. Cell Rep *8*, 59-65.

Preti, M., O'Donohue, M.F., Montel-Lehry, N., Bortolin-Cavaille, M.L., Choesmel, V., and Gleizes, P.E. (2013). Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. Nucleic Acids Res *41*, 4709-4723.

Rai, J., Parker, M.D., Huang, H., Choy, S., Ghalei, H., Johnson, M.C., Karbstein, K., and Stroupe, M.E. (2020). An opened pre-40S platform enables temporal regulation of Fap7. bioRxiv, 617910.

Raught, B., and Gingras, A.C. (1999). eIF4E activity is regulated at multiple levels. Int J Biochem Cell Biol *31*, 43-57.

Reeder, R.H., Guevara, P., and Roan, J.G. (1999). Saccharomyces cerevisiae RNA polymerase I terminates transcription at the Reb1 terminator in vivo. Mol Cell Biol *19*, 7369-7376.

Reeder, R.H., Higashinakagawa, T., and Miller, O., Jr. (1976). The 5' leads to 3' polarity of the Xenopus Ribosomal RNA precursor molecule. Cell *8*, 449-454.

Ritossa, F.M., and Spiegelman, S. (1965). Localization of DNA Complementary to Ribosomal Rna in the Nucleolus Organizer Region of Drosophila Melanogaster. Proc Natl Acad Sci U S A *53*, 737-745.

Rodnina, M.V. (2018). Translation in Prokaryotes. Cold Spring Harb Perspect Biol *10*.

Rodríguez-Galán, O., García-Gómez, J.J., and de la Cruz, J. (2013). Yeast and human RNA helicases involved in ribosome biogenesis: current status and perspectives. Biochimica et biophysica acta *1829*, 775-790.

Sahasranaman, A., Dembowski, J., Strahler, J., Andrews, P., Maddock, J., and Woolford, J.L. (2011). Assembly of Saccharomyces cerevisiae 60S ribosomal subunits: role of factors required for 27S pre-rRNA processing. The EMBO journal *30*, 4020-4032.

Sanghai, Z.A., Miller, L., Molloy, K.R., Barandun, J., Hunziker, M., Chaker-Margot, M., Wang, J., Chait, B.T., and Klinge, S. (2018). Modular assembly of the nucleolar pre-60S ribosomal subunit. Nature *556*, 126-126.

Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., Marcotte, E.M., Tollervey, D., Correll, C.C., and Johnson, A.W. (2015). The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. PLoS Biol *13*, e1002083.

Sardana, R., White, J.P., and Johnson, A.W. (2013). The rRNA methyltransferase Bud23 shows functional interaction with components of the SSU processome and RNase MRP. RNA *19*, 828-840.

Sardana, R., Zhu, J., Gill, M., and Johnson, A.W. (2014). Physical and functional interaction between the methyltransferase Bud23 and the essential DEAH-box RNA helicase Ecm16. Mol Cell Biol *34*, 2208-2220.

Sarkar, A., Pech, M., Thoms, M., Beckmann, R., and Hurt, E. (2016). Ribosome-stalk biogenesis is coupled with recruitment of nuclear-export factor to the nascent 60S subunit. Nat Struct Mol Biol *23*, 1074-1082.

Sarkar, A., Thoms, M., Barrio-Garcia, C., Thomson, E., Flemming, D., Beckmann, R., and Hurt, E. (2017). Preribosomes escaping from the nucleus are caught during translation by cytoplasmic quality control. Nature structural & molecular biology.

Scaiola, A., Pena, C., Weisser, M., Bohringer, D., Leibundgut, M., Klingauf-Nerurkar, P., Gerhardy, S., Panse, V.G., and Ban, N. (2018). Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit. EMBO J.

Schafer, T., Maco, B., Petfalski, E., Tollervey, D., Bottcher, B., Aebi, U., and Hurt, E. (2006). Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. Nature *441*, 651-655.

Schafer, T., Strauss, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003). The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J *22*, 1370-1380.

Schimmang, T., Tollervey, D., Kern, H., Frank, R., and Hurt, E.C. (1989). A yeast nucleolar protein related to mammalian fibrillarin is associated with small nucleolar RNA and is essential for viability. EMBO J *8*, 4015-4024.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F.*, et al.* (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. Cell *102*, 615-623.

Schmeing, T.M., Huang, K.S., Kitchen, D.E., Strobel, S.A., and Steitz, T.A. (2005a). Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. Mol Cell *20*, 437-448.

Schmeing, T.M., Huang, K.S., Strobel, S.A., and Steitz, T.A. (2005b). An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. Nature *438*, 520-524.

Schmeing, T.M., Seila, A.C., Hansen, J.L., Freeborn, B., Soukup, J.K., Scaringe, S.A., Strobel, S.A., Moore, P.B., and Steitz, T.A. (2002). A pre-translocational intermediate in protein synthesis observed in crystals of enzymatically active 50S subunits. Nat Struct Biol *9*, 225-230.

Schmitt, M.E., and Clayton, D.A. (1993). Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in Saccharomyces cerevisiae. Mol Cell Biol *13*, 7935-7941.

Schuller, J.M., Falk, S., Fromm, L., Hurt, E., and Conti, E. (2018). Structure of the nuclear exosome captured on a maturing preribosome. Science, eaar5428-eaar5428.

Schutz, S., Fischer, U., Altvater, M., Nerurkar, P., Pena, C., Gerber, M., Chang, Y., Caesar, S., Schubert, O.T., Schlenstedt, G.*, et al.* (2014). A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly. Elife *3*, e03473.

Seiser, R.M., Sundberg, A.E., Wollam, B.J., Zobel-Thropp, P., Baldwin, K., Spector, M.D., and Lycan, D.E. (2006). Ltv1 is required for efficient nuclear export of the ribosomal small subunit in Saccharomyces cerevisiae. Genetics *174*, 679-691.

Senger, B., Lafontaine, D.L., Graindorge, J.S., Gadal, O., Camasses, A., Sanni, A., Garnier, J.M., Breitenbach, M., Hurt, E., and Fasiolo, F. (2001). The nucle(ol)ar Tif6p and Efl1p are required for a late cytoplasmic step of ribosome synthesis. Mol Cell *8*, 1363-1373.

Shao, J., Grammatikakis, N., Scroggins, B.T., Uma, S., Huang, W., Chen, J.J., Hartson, S.D., and Matts, R.L. (2001). Hsp90 regulates p50(cdc37) function during the biogenesis of the activeconformation of the heme-regulated eIF2 alpha kinase. J Biol Chem *276*, 206-214.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Sharma, K., and Tollervey, D. (1999). Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing. Mol Cell Biol *19*, 6012-6019.

Sharma, S., and Lafontaine, D.L.J. (2015). 'View From A Bridge': A New Perspective on Eukaryotic rRNA Base Modification. Trends Biochem Sci *40*, 560-575.

Showkat, M., Beigh, M.A., and Andrabi, K.I. (2014). mTOR Signaling in Protein Translation Regulation: Implications in Cancer Genesis and Therapeutic Interventions. Mol Biol Int *2014*, 686984.

Simms, C.L., Thomas, E.N., and Zaher, H.S. (2017). Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip Rev RNA *8*.

Singh, S.K., Gurha, P., and Gupta, R. (2008). Dynamic guide-target interactions contribute to sequential 2'-O-methylation by a unique archaeal dual guide box C/D sRNP. RNA *14*, 1411-1423.

Sloan, K.E., Bohnsack, M.T., Schneider, C., and Watkins, N.J. (2014). The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. RNA *20*, 540-550.

Sloan, K.E., Knox, A.A., Wells, G.R., Schneider, C., and Watkins, N.J. (2019). Interactions and activities of factors involved in the late stages of human 18S rRNA maturation. RNA Biol *16*, 196-210.

Sloan, K.E., Leisegang, M.S., Doebele, C., Ramírez, A.S., Simm, S., Safferthal, C., Kretschmer, J., Schorge, T., Markoutsa, S., Haag, S.*, et al.* (2015). The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21. Nucleic acids research *43*, 553-564.

Sloan, K.E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G.J., and Watkins, N.J. (2013). Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. J Cell Biol *200*, 577-588.

Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.D., Lafontaine, D.L.J., and Bohnsack, M.T. (2017). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol *14*, 1138-1152.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Soudet, J., Gelugne, J.P., Belhabich-Baumas, K., Caizergues-Ferrer, M., and Mougin, A. (2010). Immature small ribosomal subunits can engage in translation initiation in Saccharomyces cerevisiae. EMBO J *29*, 80-92.

Strasser, K., Bassler, J., and Hurt, E. (2000). Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. J Cell Biol *150*, 695-706.

Strunk, B.S., Loucks, C.R., Su, M., Vashisth, H., Cheng, S., Schilling, J., Brooks, C.L., 3rd, Karbstein, K., and Skiniotis, G. (2011). Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. Science *333*, 1449-1453.

Strunk, B.S., Novak, M.N., Young, C.L., and Karbstein, K. (2012). A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. Cell *150*, 111-121.

Sturm, M., Cheng, J., Bassler, J., Beckmann, R., and Hurt, E. (2017). Interdependent action of KH domain proteins Krr1 and Dim2 drive the 40S platform assembly. Nat Commun *8*, 2213.

Su, Q., Wang, S., Gao, H.Q., Kazemi, S., Harding, H.P., Ron, D., and Koromilas, A.E. (2008). Modulation of the eukaryotic initiation factor 2 alpha-subunit kinase PERK by tyrosine phosphorylation. J Biol Chem *283*, 469-475.

Su, T., Izawa, T., Thoms, M., Yamashita, Y., Cheng, J., Berninghausen, O., Hartl, F.U., Inada, T., Neupert, W., and Beckmann, R. (2019). Structure and function of Vms1 and Arb1 in RQC and mitochondrial proteome homeostasis. Nature *570*, 538-542.

Sudhakar, A., Ramachandran, A., Ghosh, S., Hasnain, S.E., Kaufman, R.J., and Ramaiah, K.V. (2000). Phosphorylation of serine 51 in initiation factor 2 alpha (eIF2 alpha) promotes complex formation between eIF2 alpha(P) and eIF2B and causes inhibition in the guanine nucleotide exchange activity of eIF2B. Biochemistry *39*, 12929-12938.

Sun, Q., Zhu, X., Qi, J., An, W., Lan, P., Tan, D., Chen, R., Wang, B., Zheng, S., Zhang, C.*, et al.* (2017). Molecular architecture of the 90S small subunit pre-ribosome. Elife *6*.

Suyama, M., Doerks, T., Braun, I.C., Sattler, M., Izaurralde, E., and Bork, P. (2000). Prediction of structural domains of TAP reveals details of its interaction with p15 and nucleoporins. EMBO Rep *1*, 53-58.

Szick, K., Springer, M., and Bailey-Serres, J. (1998). Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes. Proc Natl Acad Sci U S A *95*, 2378-2383.

Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Molecular cell *51*, 539-551.

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

Tantravahi, R., Miller, D.A., Dev, V.G., and Miller, O.J. (1976). Detection of nucleolus organizer regions in chromosomes of human, chimpanzee, gorilla, orangutan and gibbon. Chromosoma *56*, 15-27.

Thoms, M., Thomson, E., Baßler, J., Gnädig, M., Griesel, S., and Hurt, E. (2015). The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. Cell *162*, 1029-1038.

Thomson, E., and Tollervey, D. (2010). The final step in 5.8S rRNA processing is cytoplasmic in Saccharomyces cerevisiae. Mol Cell Biol *30*, 976-984.

Tocilj, A., Schlunzen, F., Janell, D., Gluhmann, M., Hansen, H.A., Harms, J., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., *et al.* (1999). The small ribosomal subunit from Thermus thermophilus at 4.5 A resolution: pattern fittings and the identification of a functional site. Proc Natl Acad Sci U S A *96*, 14252-14257.

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M., and Hurt, E.C. (1991). The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. EMBO J *10*, 573-583.

Tollervey, D., Lehtonen, H., Jansen, R., Kern, H., and Hurt, E.C. (1993). Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. Cell *72*, 443-457.

Tomecki, R., Labno, A., Drazkowska, K., Cysewski, D., and Dziembowski, A. (2015). hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. RNA Biol *12*, 1010-1029.

Tomecki, R., Sikorski, P.J., and Zakrzewska-Placzek, M. (2017). Comparison of preribosomal RNA processing pathways in yeast, plant and human cells - focus on coordinated action of endo- and exoribonucleases. FEBS Lett *591*, 1801-1850.

Turowski, T.W., Lebaron, S., Zhang, E., Peil, L., Dudnakova, T., Petfalski, E., Granneman, S., Rappsilber, J., and Tollervey, D. (2014). Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. Nucleic Acids Res *42*, 12189-12199.

Udem, S.A., and Warner, J.R. (1972). Ribosomal RNA synthesis in Saccharomyces cerevisiae. J Mol Biol *65*, 227-242.

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

van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262-2264.

Vance, V.B., Thompson, E.A., and Bowman, L.H. (1985). Transfection of mouse ribosomal DNA into rat cells: faithful transcription and processing. Nucleic Acids Res *13*, 7499-7513.

Vanrobays, E., Gelugne, J.P., Gleizes, P.E., and Caizergues-Ferrer, M. (2003). Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in Saccharomyces cerevisiae. Mol Cell Biol *23*, 2083-2095.

Vanrobays, E., Gleizes, P.E., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Caizergues-Ferrer, M., and Gelugne, J.P. (2001). Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein. EMBO J *20*, 4204-4213.

Vanrobays, E., Leplus, A., Osheim, Y.N., Beyer, A.L., Wacheul, L., and Lafontaine, D.L. (2008). TOR regulates the subcellular distribution of DIM2, a KH domain protein required for cotranscriptional ribosome assembly and pre-40S ribosome export. RNA *14*, 2061-2073.

Vattem, K.M., Staschke, K.A., and Wek, R.C. (2001). Mechanism of activation of the double-stranded-RNA-dependent protein kinase, PKR: role of dimerization and cellular localization in the stimulation of PKR phosphorylation of eukaryotic initiation factor-2 (eIF2). Eur J Biochem *268*, 3674-3684.

Veinot-Drebot, L.M., Singer, R.A., and Johnston, G.C. (1988). Rapid initial cleavage of nascent pre-rRNA transcripts in yeast. J Mol Biol *199*, 107-113.

Venema, J., and Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. Annu Rev Genet *33*, 261-311.

Vierna, J., Wehner, S., Honer zu Siederdissen, C., Martinez-Lage, A., and Marz, M. (2013). Systematic analysis and evolution of 5S ribosomal DNA in metazoans. Heredity (Edinb) *111*, 410-421.

Voorhees, R.M., Weixlbaumer, A., Loakes, D., Kelley, A.C., and Ramakrishnan, V. (2009). Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. Nat Struct Mol Biol *16*, 528-533.

Voss, N.R., Gerstein, M., Steitz, T.A., and Moore, P.B. (2006). The geometry of the ribosomal polypeptide exit tunnel. J Mol Biol *360*, 893-906.

Warner, J.R. (1989). Synthesis of ribosomes in Saccharomyces cerevisiae. Microbiol Rev *53*, 256-271.

Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. Trends Biochem Sci *24*, 437-440.

Watkins, N.J., and Bohnsack, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley interdisciplinary reviews RNA *3*, 397-414.

Weis, F., Giudice, E., Churcher, M., Jin, L., Hilcenko, C., Wong, C.C., Traynor, D., Kay, R.R., and Warren, A.J. (2015). Mechanism of eIF6 release from the nascent 60S ribosomal subunit. Nature structural & molecular biology *22*, 914-919.

Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J., and Schneider, C. (2016). The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. Nucleic Acids Res *44*, 5399-5409.

Wells, G.R., Weichmann, F., Sloan, K.E., Colvin, D., Watkins, N.J., and Schneider, C. (2017). The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain. Nucleic Acids Res *45*, 4796-4809.

White, J., Li, Z., Sardana, R., Bujnicki, J.M., Marcotte, E.M., and Johnson, A.W. (2008). Bud23 methylates G1575 of 18S rRNA and is required for efficient nuclear export of pre-40S subunits. Mol Cell Biol *28*, 3151-3161.

Widmann, B., Wandrey, F., Badertscher, L., Wyler, E., Pfannstiel, J., Zemp, I., and Kutay, U. (2012). The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. Mol Biol Cell *23*, 22-35.

Wild, T., Horvath, P., Wyler, E., Widmann, B., Badertscher, L., Zemp, I., Kozak, K., Csucs, G., Lund, E., and Kutay, U. (2010). A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. PLoS Biol *8*, e1000522.

Wilson, D.N., and Doudna Cate, J.H. (2012). The structure and function of the eukaryotic ribosome. Cold Spring Harb Perspect Biol *4*.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M., Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. Nature *407*, 327-339.

Woolford, J.L., Jr., and Baserga, S.J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics *195*, 643-681.

Wormsley, S., Samarsky, D.A., Fournier, M.J., and Baserga, S.J. (2001). An unexpected, conserved element of the U3 snoRNA is required for Mpp10p association. RNA *7*, 904-919.

Wu, S., Tutuncuoglu, B., Yan, K., Brown, H., Zhang, Y., Tan, D., Gamalinda, M., Yuan, Y., Li, Z., Jakovljevic, J.*, et al.* (2016). Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. Nature *534*, 133-137.

Wyler, E., Zimmermann, M., Widmann, B., Gstaiger, M., Pfannstiel, J., Kutay, U., and Zemp, I. (2011). Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. RNA *17*, 189-200.

Xiang, S., Cooper-Morgan, A., Jiao, X., Kiledjian, M., Manley, J.L., and Tong, L. (2009). Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. Nature *458*, 784-788.

Xu, B., Li, H., Perry, J.M., Singh, V.P., Unruh, J., Yu, Z., Zakari, M., McDowell, W., Li, L., and Gerton, J.L. (2017). Ribosomal DNA copy number loss and sequence variation in cancer. PLoS Genet *13*, e1006771.

Yao, W., Lutzmann, M., and Hurt, E. (2008). A versatile interaction platform on the Mex67-Mtr2 receptor creates an overlap between mRNA and ribosome export. EMBO J *27*, 6-16.

Yao, W., Roser, D., Köhler, A., Bradatsch, B., Bassler, J., and Hurt, E. (2007). Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. Molecular cell *26*, 51-62.

Zemp, I., Wandrey, F., Rao, S., Ashiono, C., Wyler, E., Montellese, C., and Kutay, U. (2014). CK1delta and CK1epsilon are components of human 40S subunit precursors required for cytoplasmic 40S maturation. J Cell Sci *127*, 1242-1253.

Zemp, I., Wild, T., O'Donohue, M.F., Wandrey, F., Widmann, B., Gleizes, P.E., and Kutay, U. (2009). Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. J Cell Biol *185*, 1167-1180.

Zhang, F., Hamanaka, R.B., Bobrovnikova-Marjon, E., Gordan, J.D., Dai, M.S., Lu, H., Simon, M.C., and Diehl, J.A. (2006). Ribosomal stress couples the unfolded protein response to p53-dependent cell cycle arrest. J Biol Chem *281*, 30036-30045.

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

Zhang, L., Lin, J., and Ye, K. (2013). Structural and functional analysis of the U3 snoRNA binding protein Rrp9. RNA *19*, 701-711.

Zhang, L., Wu, C., Cai, G., Chen, S., and Ye, K. (2016). Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. Genes Dev *30*, 718-732.

Zhang, Y., Ni, J., Zhou, G., Yuan, J., Ren, W., Shan, Y., Tang, W., Yu, L., and Zhao, S. (2005). Cloning, expression and characterization of the human NOB1 gene. Mol Biol Rep *32*, 185-189.

Zhou, D., Zhu, X., Zheng, S., Tan, D., Dong, M.Q., and Ye, K. (2019). Cryo-EM structure of an early precursor of large ribosomal subunit reveals a half-assembled intermediate. Protein Cell *10*, 120-130.

Zhu, J., Liu, X., Anjos, M., Correll, C.C., and Johnson, A.W. (2016). Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. Mol Cell Biol *36*, 965-978.

Zorbas, C., Nicolas, E., Wacheul, L., Huvelle, E., Heurgue-Hamard, V., and Lafontaine, D.L. (2015). The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. Mol Biol Cell *26*, 2080-2095.

Publications

Publications discussed in this dissertation:

Publication 1 – Ameismeier and Cheng et al. 2018

Publication 2 – Ameismeier et al. 2020

Publications not discussed in this dissertation:

Publication 3 – Thoms, Buschauer and Ameismeier et al. 2020

Publication 4 – Kratzat and Mackens-Kiani et al. 2020

Publication 5 – Cheng et al. 2020

ARTICLE

Visualizing late states of human 40S ribosomal subunit maturation

Michael Ameismeier[1,2], Jingdong Cheng[1,2], Otto Berninghausen[1] & Roland Beckmann[1]*

The formation of eukaryotic ribosomal subunits extends from the nucleolus to the cytoplasm and entails hundreds of assembly factors. Despite differences in the pathways of ribosome formation, high-resolution structural information has been available only from fungi. Here we present cryo-electron microscopy structures of late-stage human 40S assembly intermediates, representing one state reconstituted in vitro and five native states that range from nuclear to late cytoplasmic. The earliest particles reveal the position of the biogenesis factor RRP12 and distinct immature rRNA conformations that accompany the formation of the 40S subunit head. Molecular models of the late-acting assembly factors TSR1, RIOK1, RIOK2, ENP1, LTV1, PNO1 and NOB1 provide mechanistic details that underlie their contribution to a sequential 40S subunit assembly. The NOB1 architecture displays an inactive nuclease conformation that requires rearrangement of the PNO1-bound 3′ rRNA, thereby coordinating the final rRNA folding steps with site 3 cleavage.

Ribosomes are universally conserved macromolecular complexes that translate mRNA into protein. In eukaryotes, they consist of a small 40S and a large 60S subunit that together comprise four rRNAs and about 80 ribosomal proteins. During their assembly, a multitude of over 200 *trans*-acting ribosome biogenesis factors (RBFs) ensures correct cleavage, modification and folding of rRNA and concomitant incorporation of ribosomal proteins[1–3].

In eukaryotes, ribosome biogenesis starts in the nucleolus with the transcription of three of the four rRNAs as a primary transcript[4]. After dissociation of the large subunit precursor, pre-40S particles are exported to the cytoplasm, where final maturation processes occur[5]. Many proteins are involved in the later stages of human small subunit maturation, several of which we observe in this work, including the methyltransferase BUD23 (also known as WBSCR22) together with TRMT112, the armadillo (ARM)-like protein RRP12, structural proteins ENP1 (also known as BYSL) and LTV1, the endonuclease NOB1 with its binding partner PNO1, GTPase-like protein TSR1, as well as the atypical kinases RIOK1 and RIOK2[6–8].

High-resolution structures of pre-40S particles from fungi such as *Saccharomyces cerevisiae* have recently been published[9,10]. Despite differences in rRNA processing[11] and the composition and function of RBFs[12,13], however, human ribosomal precursors have been only described at low resolution so far[14], limiting the information to overall positioning of RBFs.

Here we report cryo-electron microscopy (cryo-EM) reconstructions of several small ribosomal subunit precursors that provide detailed insights into late human 40S maturation principles. As a reoccurring concept, we found that assembly factors stabilize rRNA segments in distinct immature conformations before allowing their accommodation into their respective mature positions. Thereby, accurate step-wise and sequential maturation of rRNA from 5′ to 3′ is ensured, and premature engagement of 40S precursors in 80S formation and erroneous translation is prevented.

Cryo-EM analysis of the human pre-40S ribosome

To understand better the maturation process of human 40S subunits, we aimed at structurally analysing native complexes, purified with PNO1 as bait (states A–E), and reconstituted pre-40S particles in vitro (state R). Single particle cryo-EM analysis (Extended Data Fig. 1) resulted in a series of pre-40S structures that represent six different middle to late assembly states (Fig. 1a and Extended Data Fig. 2a, b).

In state A, we found the RBFs RRP12, ENP1 and PNO1 together with the methyltransferases BUD23 and TRMT112. State B lacks BUD23 and TRMT112, but comprises LTV1 and ribosomal proteins uS2, uS5 and eS21, as well as the endonuclease NOB1. From state C on, RRP12 is dissociated, and its place is occupied by RACK1. In addition, the kinase RIOK2, and ribosomal proteins eS12 and eS31 have bound, revealing remarkable conservation when compared with highly similar pre-40S subunits from yeast[9,10] (Extended Data Fig. 2d). In state D, ENP1 and LTV1 have been replaced by ribosomal proteins uS3, uS10, eS10 and uS14. Finally, an unassigned protein (factor X) is present in state E, partially occupying the binding site of the C-terminal helix of RIOK2.

The atypical kinase RIOK1 has also been described to have a role in the final steps of pre-40S maturation, together with PNO1 and NOB1[7]. We did not observe such a state in our native pull-outs and, therefore, attempted to reconstitute it in vitro using purified mature 40S subunits and recombinant proteins (state R). We found that PNO1 and RIOK1 have displaced ribosomal proteins eS26 and eL41, and that NOB1, PNO1 and the pre-18S rRNA in state R adopt a highly similar conformation as in state E, including an immature helix 44 (h44) and a shifted 3′ end.

Among all pre-40S reconstructions, two (states C and R) could be refined to average resolutions below 4 Å. A full model of state C was built (Fig. 1b and Extended Data Fig. 2c), including biogenesis factors TSR1, NOB1, PNO1, ENP1 and RIOK2, as well as parts of LTV1, and a short segment of RIOK1 was built in state R (Extended Data Fig. 2e). The average resolution of the other states was between 4 and 4.5 Å, with considerably lower local resolution in flexible areas. It was, however, sufficient to allow for unambiguous rigid body fitting of our models and assignment of RRP12, BUD23 and TRMT112 to extra densities in states A and B (Fig. 1b).

The maturation process of the 3′ major domain

The maturation process of the central region around helices h35–h40 is dominated by large shifts of rRNA, in which RRP12 has a crucial
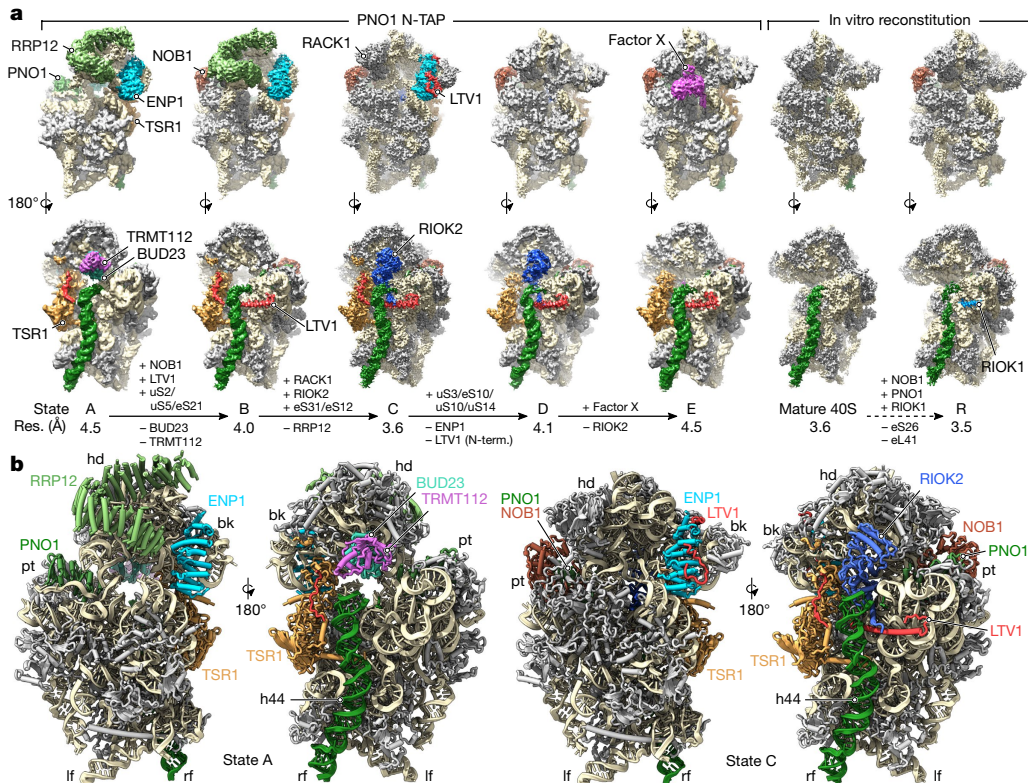
Fig. 1 | Structural analysis of human pre-40S particles. a, Cryo-EM reconstructions of five native pre-40S particles, obtained via N-terminal tandem affinity purification (N-TAP), and the mature 40S with a precursor-like complex reconstituted in vitro (see Methods). Names, average resolution and changes in protein composition are stated below. **b**, Models of state A (left) and C (right) 40S precursors with ribosomal biogenesis factors and h44 highlighted (grey: ribosomal proteins, light brown: rRNA). bk, beak; hd, head; lf, left foot; pt, platform; rf, right foot.

role (Fig. 2 and Extended Data Fig. 3a). In state A, it coordinates h35–h37 and the region around the three-way-junction of h38–h39–h40 in their immature position, and with h41 and uS9 parts of the head (Fig. 2a, top). The flipping of helices h35–h37 downwards by about 90° then allows binding of the uS2–uS5–eS21 cluster together with NOB1 (Fig. 2a, middle and Extended Data Fig. 3b). Once RRP12 is released, h39 rotates almost 140° around the h34 axis to interact with h41, while h40 and eS17 accommodate close to their mature position in state C (Fig. 2a, bottom, and Fig. 2b). With the movement of these central rRNA helices, the interface is then formed for binding of RACK1. In yeast, Rrp12 has been shown to facilitate export of pre-40S particles through direct interaction with nucleoporins and the small GTPase Gsp1p[15]. In combination with our results, this suggests that RRP12 could serve in a quality control step, in which correct head formation and subsequent export are coupled.

Despite major rRNA rearrangements, helices h34, h39 and h40 are yet to settle into their mature position (Fig. 2c), which depends on maturation events within the beak (Extended Data Fig. 3c). In our earlier particles, ENP1 and LTV1 occupy the central position on the beak, preventing the accommodation of the uS3–uS10–eS10–uS14 cluster and the mature positioning of h34, similar to what has been observed in yeast[9]. After the release of ENP1 and LTV1 from the beak in state D, ribosomal proteins uS3, uS10, eS10 and uS14 settle into their respective sites, and this coincides with a movement of h34 closer to its mature position (Extended Data Fig. 3d).

The maturation process of the decoding centre

We observed an immature decoding centre in all native complexes and, to our surprise, also in native 40S particles treated with PNO1, NOB1 and RIOK1 at high molar excess (20×; Extended Data Fig. 4a). Similar to yeast[9], h44 is shifted outward by approximately 25 Å in states A–E and 11 Å in state R (Fig. 3a). This displacement prevents correct folding of the connecting region between h44 and h28 or h45, and therefore formation of the A and P sites. The biogenesis factors TSR1, LTV1 and RIOK2 probe this immature conformation through distinct terminal helices that bind underneath h44, probably contributing to its stabilization (Fig. 3b).

TSR1, an inactive mimic of translational GTPases[16], interacts with the pre-40S body near rRNA helices h3–h5, h15 and h44 via domains I and III, with its N-terminal helix passing beneath h44 up to h12 (Fig. 3b and Extended Data Fig. 4b), as well as with the head via its domain

Fig. 2 | RRP12 binding and sequential pre-18S rRNA rearrangements during head formation. a, Movement of h35–h37 and h39 during early head formation. Volumes filtered at 6 Å. **b, c**, Detailed view on rRNA rearrangements. **b**, The transition from state B to C is described by a 140° movement of h39 and h40. **c**, Final maturation requires repositioning of h34, h39 and h40.
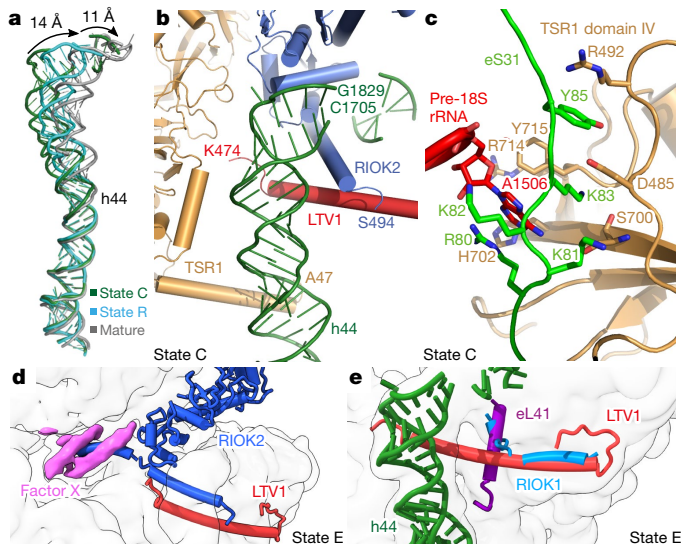
Fig. 3 | Positioning of h44 and maturation of the decoding centre.
a, Comparison of h44 in its immature and mature position. **b**, Location of terminal helices of TSR1, LTV1 and RIOK2 between displaced h44 and the body. **c**, Interactions between TSR1 and h34 with the N-terminal extension of eS31. **d**, **e**, Clashes of biogenesis factor RIOK2 with unassigned factor X (**d**), RIOK1 with the C terminus of LTV1, and RIOK1 with eL41 (**e**).

IV, as observed in yeast[9]. In addition, we observe distinct interactions between TSR1, pre-18S rRNA and eS31, which is recruited in state C. Together with domain IV of TSR1, the shifted N terminus of eS31 (Extended Data Fig. 4c) coordinates a flipped out A1506 of the junction region of h32, h33 and h34, stabilizing h34 in an immature state after the rearrangement of the h34–h40 region (Fig. 3c). Supported by an insertion in domain II, which binds to uS13 and uS19 in the 3′ major domain, domain IV and eS31 might therefore serve as a link between maturation events within the beak and the intersubunit side.

Another connection could represent LTV1, of which we observed two interacting segments (Fig. 1a). In addition to the N-terminal part that binds ENP1 similarly to published structures[17], its C terminus forms a long helix that stretches for about 60 Å across the intersubunit side and ends beneath h44 (Fig. 3b), a position largely occupied by Dim1 in late yeast 40S precursors[18]. Unlike in yeast, however, human dimethylation by DIMT1 is nuclear and, therefore, DIMT1 is not involved in the final maturation steps in the cytoplasm[8].

The kinase RIOK2 is observed in states C and D, binding to the decoding centre between the head and the platform as previously described[9]. In addition, we found its C terminus forming an elongated helix, which passes beneath h44 and is deeply buried within the tunnel that connects the solvent and intersubunit side (Fig. 3b). Surprisingly, in state E, factor X occupies parts of this tunnel with two helices, and superimposition of both states reveals overlapping binding sites (Fig. 3d).

Apart from RIOK2, BUD23–TRMT112 and the kinase domain of RIOK1 also bind to the neck and platform of pre-40S particles (Extended Data Fig. 4a, d). Furthermore, the C terminus of RIOK1 binds between h27, h44 and h45, a site previously occupied by LTV1 and later by eL41 in mature 40S particles (Fig. 3e). This is consistent with data suggesting that the C terminus of Rio1 is crucial for its binding to the ribosome in yeast[19]. Furthermore, the clash of RIOK1 with LTV1 and RIOK2 could explain why the sole presence of RIOK1 and not its ATPase activity is sufficient for their release[7].

Coordination of site 3 cleavage by PNO1 and NOB1
PNO1 has a decisive role in selectively binding and stabilizing h28 and h44 in an immature conformation in all states (Fig. 4 and Extended Data Fig. 5a). Superimposition of state C with mature 40S reveals two major clashes: First, the last helix of the K homology (KH) 2 domain of PNO1 would clash with h28 (Fig. 4a and Extended Data Fig. 5b, c). This retains
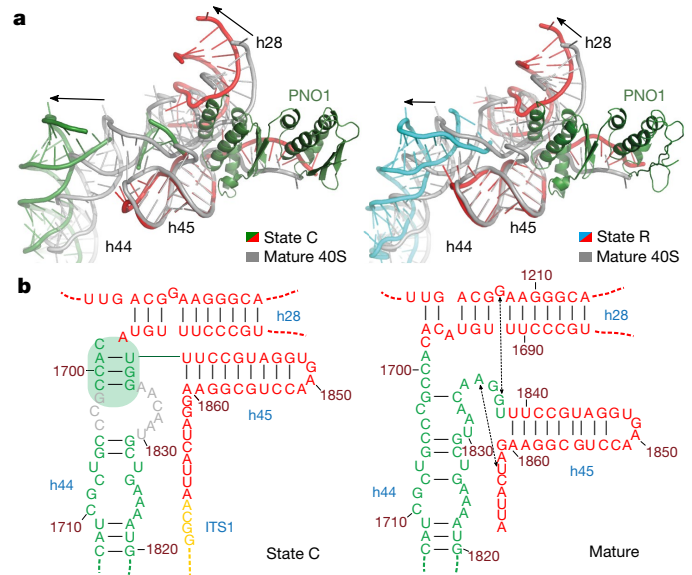
Fig. 4 | PNO1 interaction determines state of the decoding centre.
a, Conformation of helices h28, h44 and h45 in native state C (left) and the reconstituted particle (state R, right) compared to their mature position (grey). Displacements stabilized by the presence of PNO1 are indicated (arrows). **b**, Immature base pairing as a result of shifted helices h28 and h44. Summary of base interactions in state C and the mature 40S, with a focus on h28, h44 and h45. Base pairings are indicated by black lines and the relevant base stacking interactions by dashed lines. Immature helix highlighted in green. For an extended view, see Extended Data Fig. 6. ITS1, internal transcribed spacer.

h28 in a backward position and prevents stacking between G1207 and G1837 (Extended Data Fig. 5d). Second, PNO1 would also clash with the connecting region between h44 and h45, prohibiting its mature fold while breaking another stacking between A1835 and A1863. As a result, A1863 is flipped out and can be recognized and bound by PNO1 (Extended Data Fig. 5e). Concomitant with the displacement of h44, novel immature base pairing is formed between nucleotides 1699–1701 and 1836–1838 (Fig. 4b and Extended Data Fig. 5f). Notably, as observed in state R, incubation of fully matured 40S subunits with recombinant PNO1 (together with NOB1 and RIOK1) was sufficient to reverse the final maturation step by dissociating eS26 and forcing the rRNA to unfold partially (Fig. 4a, right, and Extended Data Fig. 5b).

Besides stabilizing the central rRNA region in a distinct immature state and consistent with previous results[20], we found PNO1 directly interacting with NOB1, which is absent in early pre-ribosomal particles[6]. Structurally, NOB1 can be divided into four parts: the PilT N-terminus (PIN) endonuclease domain, the insertion domain, the zinc-finger domain and the C-terminal domain. The PIN and zinc-finger domains form a stable core of NOB1 (Extended Data Fig. 5g) and mediate its interaction with ribosomal proteins uS2 and eS17 (Fig. 5a). The insertion domain binds to the KH1 domain of PNO1 involving residues Phe128 and Trp208 (Fig. 5b), and the C-terminal part of NOB1 forms a large loop, which interacts with the hydrophobic groove between the two KH domains of PNO1 (Fig. 5c).

In contrast to state R, we can trace several bases into the density for the internal transcribed spacer 1 (ITS1) in all native states (Extended Data Fig. 5h), which is a clear indication that site 3 cleavage has yet to happen. When compared with mature 40S, the 3′ end appears shifted from its mature position and is bound by PNO1 and NOB1. The C-terminal domain of NOB1 contains several positively charged residues that stabilize the backbone of the 3′ end. In addition, Glu120 of the insertion and Arg398 of the C-terminal extension hold the ultimate nucleotide before site 3, A1869, in a tight pocket. In parallel, PNO1 binds the 3′ end and the ITS1 region via residues within its two KH domains, with Arg182, Ser235 and Tyr238 forming hydrogen bonds with G1862 and A1863, contributing to a sequence
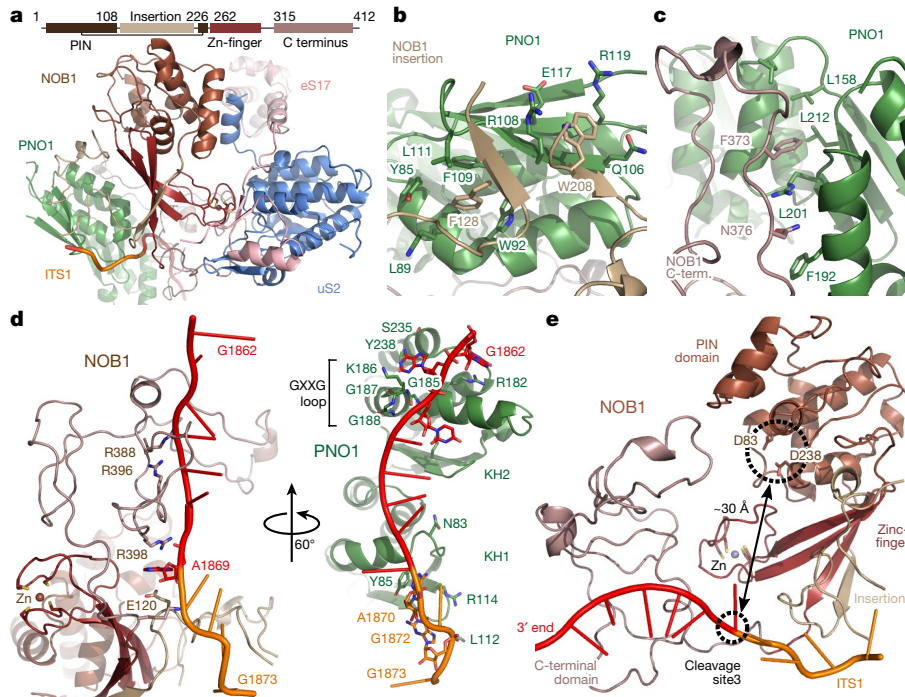
Fig. 5 | NOB1 and PNO1 coordinate site 3 in an inactive conformation. **a**, Domain architecture and overall structure of NOB1 shows binding of NOB1 by PNO1, uS2 and eS17. **b**, **c**, Molecular detail of the interface of the insertion domain (**b**) and the C-terminal domain (**c**) of NOB1 with PNO1. **d**, Coordination of the pre-18S rRNA 3′ end by NOB1 and PNO1. **e**, Insertion and C terminus of NOB1 shield the site 3 cleavage site from its catalytic centre (dashed circles).

specificity in binding (Fig. 5d and Extended Data Fig. 5h), essentially resembling the situation observed in yeast[9]. Importantly, the RNA around cleavage site 3 is captured at a large distance (approximately 30 Å) from the rigid catalytic centre of the nuclease NOB1 by both, PNO1 and NOB1 itself (Fig. 5e). Movement of the ITS1 towards the catalytic centre is prevented by the insertion and extension domain of NOB1. As a consequence, the PNO1–NOB1 complex is kept in a catalytically silent conformation along the observed maturation process. Clearly, a large conformational change, and probably release of PNO1, will be required to trigger the cleavage of site 3 by releasing

the RNA substrate from its confinement. This would be in agreement with previous hypotheses suggesting that, in the presence of Rio1, cleavage at site D in yeast (site 3 in humans) by Nob1 is facilitated after Pno1 dissociation[21].

Conclusion

Taken together, late maturation of human small ribosomal subunit occurs in a sequential manner (Fig. 6): The largely mature 40S body (state A) is further completed by the incorporation of several ribosomal proteins, allowing for recruitment of NOB1 as observed in state
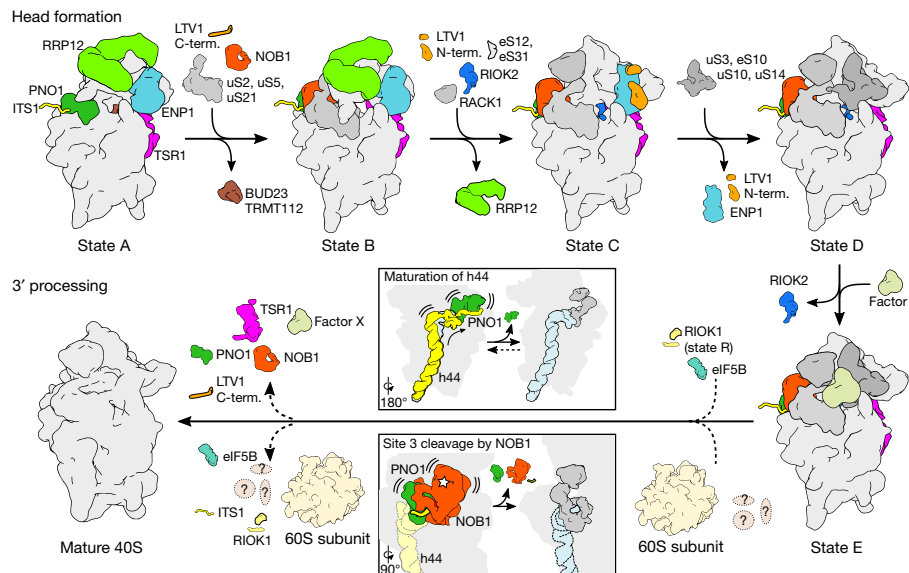
Fig. 6 | A model of stepwise maturation of the human 40S ribosomal subunit. Illustrations of distinct human 40S precursors during head formation and 3′ processing. Transitions in RBF composition not shown in this work are displayed with dotted arrows. Movement of h44 and cleavage of site 3 by NOB1 is shown in inserts with focus on PNO1 (green), h44 (yellow and blue), NOB1 (orange) and uS26 and uS28 (grey). The 60S model is based on EMDB-5592.

B. At the same time, the head and beak are formed by going through a sequence of distinct immature rRNA conformations, which is guided by interactions with BUD23, TRMT112, RRP12, ENP1, LTV1, TSR1 and RIOK2, and eventually results in export competence. Once in the cytoplasm, several factors have been suggested to have a role in eukaryotes in triggering the final maturation (adopting native conformation, factor release, site 3 cleavage), including RIOK1, eIF5B (Fun12 in yeast) and mature 60S subunits[22]. Interestingly, eIF5B as well as mature 60S subunits would clash at several sites with our state E, which may contribute to driving these last structural rearrangements (Extended Data Fig. 7). Finally, formation of the decoding region is inevitably accompanied by dissociation or repositioning of PNO1, which, in turn, may facilitate the cleavage of 18S-E rRNA at site 3 and thus concludes 40S formation.

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at https://doi.org/10.1038/s41586-018-0193-0.

1. Woolford, J. L. Jr & Baserga, S. J. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* **195**, 643–681 (2013).
2. Zemp, I. & Kutay, U. Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Lett.* **581**, 2783–2793 (2007).
3. Tschochner, H. & Hurt, E. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol.* **13**, 255–263 (2003).
4. Phipps, K. R., Charette, J. & Baserga, S. J. The small subunit processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip. Rev. RNA* **2**, 1–21 (2011).
5. Rouquette, J., Choesmel, V. & Gleizes, P. E. Nuclear export and cytoplasmic processing of precursors to the 40S ribosomal subunits in mammalian cells. *EMBO J.* **24**, 2862–2872 (2005).
6. Wyler, E. et al. Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *RNA* **17**, 189–200 (2011).
7. Widmann, B. et al. The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. *Mol. Biol. Cell* **23**, 22–35 (2012).
8. Zorbas, C. et al. The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol. Biol. Cell* **26**, 2080–2095 (2015).
9. Heuer, A. et al. Cryo-EM structure of a late pre-40S ribosomal subunit from *Saccharomyces cerevisiae*. *eLife* **6**, e30189 (2017).
10. Scaiola, A. et al. Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit. *EMBO J.* **37**, e98499 (2018).
11. Henras, A. K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A. & Gleizes, P. E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **6**, 225–242 (2015).
12. Tafforeau, L. et al. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of pre-rRNA processing factors. *Mol. Cell* **51**, 539–551 (2013).
13. Badertscher, L. et al. Genome-wide RNAi screening identifies protein modules required for 40S subunit synthesis in human cells. *Cell Reports* **13**, 2879–2891 (2015).
14. Larburu, N. et al. Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. *Nucleic Acids Res.* **44**, 8465–8478 (2016).
15. Oeffinger, M., Dlakic, M. & Tollervey, D. A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. *Genes Dev.* **18**, 196–209 (2004).
16. McCaughan, U. M. et al. Pre-40S ribosome biogenesis factor Tsr1 is an inactive structural mimic of translational GTPases. *Nat. Commun.* **7**, 11789 (2016).
17. .Sun, Q. et al. Molecular architecture of the 90S small subunit pre-ribosome. *eLife* **6**, (2017).
18. Strunk, B. S. et al. Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* **333**, 1449–1453 (2011).
19. Ferreira-Cerca, S., Kiburu, I., Thomson, E., LaRonde, N. & Hurt, E. Dominant Rio1 kinase/ATPase catalytic mutant induces trapping of late pre-40S biogenesis factors in 80S-like ribosomes. *Nucleic Acids Res.* **42**, 8635–8647 (2014).
20. Woolls, H. A., Lamanna, A. C. & Karbstein, K. Roles of Dim2 in ribosome assembly. *J. Biol. Chem.* **286**, 2578–2586 (2011).
21. Turowski, T. W. et al. Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. *Nucleic Acids Res.* **42**, 12189–12199 (2014).
22. Strunk, B. S., Novak, M. N., Young, C. L. & Karbstein, K. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* **150**, 111–121 (2012).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Molecular cloning. An N-terminal Strep-Flag tag was added to the multiple cloning site of the commercial vector pcDNA5/FRT/TO (Invitrogen). The region coding for PNO1 was amplified from human cDNA (Amsbio) using KOD DNA polymerase (Merck) and inserted into the modified pcDNA5/FRT/TO_StFLAG vector with BamHI and XhoI. DNA coding for human RIOK1, PNO1 and NOB1 was amplified from a HEK293T cell reverse transcription cDNA library and cloned into a modified pET 28a vector (SUMO tag).

Generation of cell lines. The generation of cell lines that stably express Strep-Flag-tagged PNO1 was adapted from previous work[23]. In brief, HEK Flp-In 293 T-Rex cells (Invitrogen) were grown to 70% confluency and transfected with pcDNA5/FRT/TO_StFLAG-PNO1 and pOG44 (Invitrogen) using Lipofectamine 3000 (Thermo Scientific). Selection was performed following the manufacturer's protocols with 150 µg ml^{-1} hygromycin B (Thermo Scientific). Cells were maintained in DMEM (Thermo Scientific), containing 10% FBS, 100 µg ml^{-1} hygromycin B, 10 µg ml^{-1} blasticidin, and 1× penicillin/streptomycin and GlutaMAX (Thermo Scientific). Cell lines were not authenticated or tested for mycoplasma contamination.

Native complex preparation. Native 40S precursors were purified from a stable HEK293 T-Rex Flp-In cell line as previously described[6]. Twenty-four hours before collection, PNO1 expression was induced with 1.6 µg ml^{-1} tetracycline. Cells were collected in 0.025% trypsin/EDTA (Thermo Scientific), pelleted for 7 min at 1,800g and 4 °C and washed once with PBS. Cells were then lysed for 30 min on ice using lysis buffer (10 mM HEPES pH 7.6, 100 mM KCl, 2 mM MgCl$_2$, 1 mM dithiothreitol (DTT), 0.5 mM NaF, 0.1 mM Na$_2$V$_3$O$_4$, 0.5% NP-40 substitute, and 1× protease inhibitor (Sigma Aldrich). The lysate was cleared for 15 min at 4,400g and 4 °C before transfer to equilibrated StrepTactin XT affinity beads (IBA Lifesciences), followed by a 2 h incubation in an overhead rotator. Beads were collected in small columns and washed four times with buffer A (10 mM HEPES pH 7.6, 100 mM KCl, 2 mM MgCl$_2$, 1 mM DTT, 0.5 mM NaF, 0.1 mM Na$_2$V$_3$O$_4$ and 1× protease inhibitor). Bound material was then eluted five times with 30 min incubation in buffer A containing 50 mM D-biotin (Roth). Combined eluates were transferred to equilibrated anti-Flag affinity beads (Sigma-Aldrich), incubated for 2 h, followed by five washing steps with buffer A and elution of purified complexes by five times incubation for 30 min with 0.2 mg ml^{-1} 3×Flag peptide (Sigma Aldrich) in buffer A. Finally, the combined eluates were concentrated on 300 kDa molecular mass cut-off filters (Sartorius) and analysed for protein concentration on a NanoDrop photometer (Thermo Scientific). Nikkol was added to a final concentration of 0.05% to improve ice quality after grid preparation. Samples were then analysed by polyacrylamide gel electrophoresis on 4–12% Bis-Tris gradient gels (NuPAGE, Thermo Scientific), stained with SimplyBlue (Thermo Scientific) and bands were cut out and identified separately in-house via mass spectrometry.

Reconstitution of pre-40S complexes. For recombinant expression, NOB1, PNO1 and RIOK1 plasmids were transformed into *Escherichia coli* Bl21 strains and induced with 0.1 mM IPTG at 15 °C. Cells were lysed in 50 mM Tris pH 8.0, 300 mM NaCl, the lysate subjected to Ni-NTA affinity chromatography and proteins eluted overnight by on-column cleavage with Ulp1 at 4 °C. Samples were purified further using ion exchange and size exclusion chromatography with a final buffer condition of 20 mM Tris-HCl pH 8.0, 300 mM NaCl, and 5 mM DTT.

Human 40S subunits were purified from a HEK293T cell line. Cells were collected and disrupted using lysis buffer (20 mM HEPES pH 7.4, 100 mM potassium acetate, 7.5 mM magnesium acetate$_2$, 1 mM DTT, 10 µg cycloheximide, 0.5% NP-40 and 1× protease inhibitor) and 80S ribosomes pelleted using a sucrose cushion (50 mM Tris pH 8.0, 500 mM potassium acetate, 25 mM magnesium acetate, 1 mM DTT, 1 M sucrose and 0.1% Nikkol). The resuspended ribosomes were stored in subunit separation buffer (50 mM HEPES pH 7.4, 500 mM KCl, 2 mM MgCl$_2$ and 2 mM DTT). To obtain separated 40S subunits, resuspended ribosomes were treated with puromycin at a final concentration of 1 mM, first 15 min on ice and then 10 min at 37 °C. After the treatment, the mixture was applied onto a 10–40% sucrose gradient and run overnight in a SW 40 rotor at 49,500g for 18 h. The 40S peak was collected and changed to storage buffer conditions (20 mM HEPES pH 7.4, 100 mM potassium acetate, 2.5 mM magnesium acetate and 2 mM DTT).

For the reconstitution of pre-40S complexes, purified small subunits (approximately 0.1 µM) were mixed with 2 µM RIOK1, 2 µM PNO1 and 2 µM NOB1 in storage buffer on ice for 30 min. Complexes were stabilized by crosslinking by adding glutaraldehyde to a final concentration of 0.5% (v/v). After 10 min incubation at room temperature, grids were prepared as described.

Electron microscopy and image processing. Pre-coated (2 nm) R3/3 holey carbon supported copper grids (Quantifoil) were glow discharged at 0.2 hPa for 20 s. Then, 3.5 µl of sample was directly applied onto each grid, blotted for 2–3 s at 4 °C and plunge frozen in liquid ethane using a Vitrobot Mark IV (FEI Company).

Grids were screened for ice quality and cryo-EM data acquired on a Titan Krios transmission electron microscope (FEI Company) at 300 kV under low-dose conditions (10 frames at about 2.5 e$^-$ Å^{-2}) with a nominal pixel size of 1.084 Å per pixel on the object scale using the semi-automated software EM-TOOLS (TVIPS). A total of 17,345 and 11,977 micrographs of the native PNO1-pullout were collected on a Falcon II direct electron detector (datasets 1 and 2) at nominal defocus ranges from −1.0 to −2.5 µm. In total, 2,663 and 8,115 micrographs were collected of the two datasets 'mature 40S' (dataset 3) and 'reconstituted pre-40S' (dataset 4). Original image stacks were aligned, summed and drift-corrected using MotionCor2[24]. Contrast-transfer-function parameters and resolution were estimated for each micrograph using CTFFIND4[25] and Gctf[26], respectively. Micrographs with an estimated resolution below 4 Å and astigmatism below 5% were manually screened for contamination or carbon rupture. Then, 11,356, 10,870, 2,523 and 7,753 micrographs were submitted to automated particle picking using Gautomatch (https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) and human 40S (EMDB-5592) as reference, resulting in 1,154,906, 1,637,274, 407,657 and 959,348 picked particles. These were then reference-free 2D-classified and classes individually 3D refined as shown in Extended Data Fig. 1 using Relion (version 2.1-beta-1)[27,28]. Datasets 1 and 2 were individually processed before all particles of corresponding states were combined and further analysed (see Extended Data Fig. 1d).

Model building and refinement. In general, for the molecular model building of ribosome biogenesis factors, we initially performed secondary structure prediction (PSIPRED[29]) and, whenever crystal structures of homologues were available, we used 3D structure prediction by using SWISS-MODEL[30]. A full model was built for state C, while working models were prepared for all other states.

In class C, the *Saccharomyces cerevisiae* crystal structure (PDB code 5IW7) was used as a reference for TSR1, followed by manual refinement and de novo building of the missing extensions (47–75, 392–405 and 464–485) into the density in Coot[31]. For RIOK2, the *Chaetomium thermophilum* crystal structure (PDB code 4GYG) was used to generate a homology model, which was then rigid body fitted into its density and manually adjusted, before building de novo its C-terminal helices (494–544). The PNO1 structure is based on the *Pyrococcus horikoshii* crystal structure (PDB code 3AEV) with some minor adjustments to account for structural differences. For NOB1, the *P. horikoshii* NMR solution structure (PDB: code 2LCQ) was used as a reference to build a starting model for the PIN domain, followed by a manual refinement. The zinc-finger, insertion and C-terminal extension were all built de novo (94–233 and 256–412). The *S. cerevisiae* crystal structure (PDB code 5WWO) was used as a starting model for ENP1 and parts of LTV1. ENP1 and the N terminus of LTV1 were then manually adjusted and built into the map in Coot, respectively. The C-terminal helix of LTV1 (414–417) was built de novo. The 18S rRNA and associated ribosomal proteins were modelled using the human 40S ribosome (PDB code 5A2Q) as a reference, followed by rigid body fitting and manual adjustment. Immature h34 and h44, the linker region between h44 and h45, as well as the 3′ end and ITS1 were built into their respective densities. The final model was real space refined at a resolution of 3.6 Å with secondary structure restraints for proteins and RNA, generated by ProSMART[32] and LIBG[33], using PHENIX[34] and REFMAC5[35]. Final model evaluation was done with MolProbity[36]. Overfitting statistics were calculated by a random displacement of atoms in the model followed by a refinement with REFMAC5 against one of the half maps. Finally, Fourier shell correlation curves are calculated between the volume of the refined model and both half maps using Relion.

For state A, our model of state C was fitted into the density, before removing or adjusting parts to account for conformational and compositional differences. Namely, NOB1, LTV1, RACK1, RIOK2, as well as uS2, uS5, eS12, eS21, eS31 and parts of eS17 were removed from the model. A secondary structure model of RRP12 obtained in state B was placed in the respective density. Models for immature rRNA helices h34 and h39–40 were used from a model of state B and manually adjusted to their slightly different conformation. Helices h35–h37 from state C were rigid body fitted to additional densities enclosed by RRP12. A homology model of BUD23 and TRMT112 was generated based on PDB code 4QTU and rigid body placed into its respective density.

For state B, the model for state C was rigid body docked into the density. Well-resolved areas within the body were then Phenix real space refined, while less resolved areas within the head were solely manually checked. To account for structural differences, RACK1 and RIOK2, as well as eS12 and eS31 were removed from the structure. Flipped out 18S rRNA regions h34, h39 and h40 from state C were rigid body fitted together with the N-terminal part of eS17 into the focus-refined map in Chimera. RRP12 was modelled on a secondary structure level by manually placing poly-alanine alpha-helices in rod shaped densities.

For state D and E, rRNA and ribosomal proteins from the mature state, as well as NOB1, PNO1, RIOK2, LTV1 and TSR1 from state C were rigid body fitted into their respective density using Chimera.

A model for the mature 40S (PDB code 5A2Q) was fitted into and Phenix real space refined against the final mature 40S volume.

Finally, 18S rRNA and ribosomal proteins of the mature state were fitted into the reconstituted 40S precursor volume (state R). NOB1 and PNO1 were used from our model of state C and for RIOK1, only the C terminus which shows high resolution was de novo built, while the kinase domain was left out.
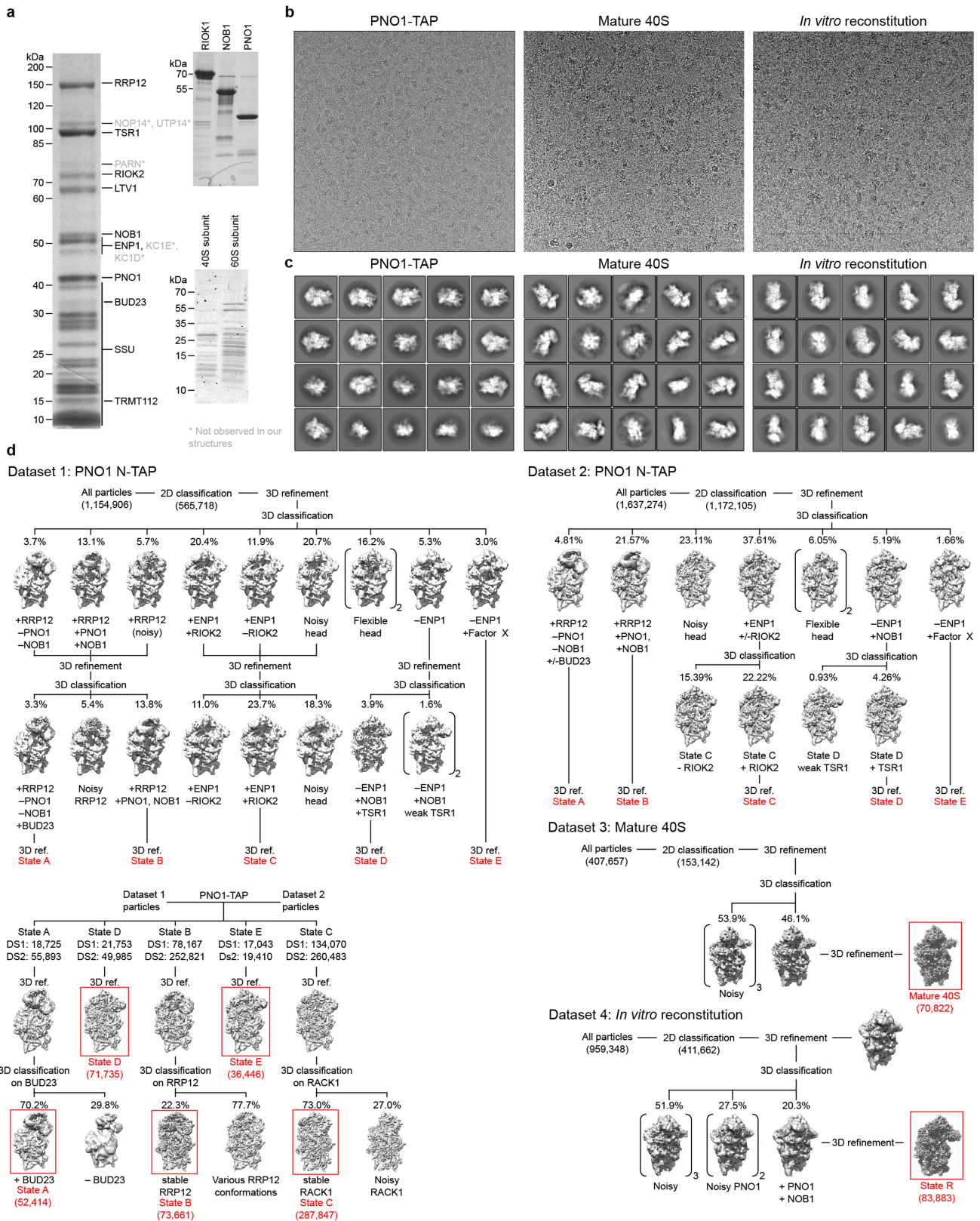
Maps and models were visualized and figures created with the PyMOL Molecular Graphics System (Version 1.7.4, Schrödinger, LLC), ChimeraX[37] and UCSF Chimera.

Statistics and reproducibility. Purification and sample preparation of native pre-40S complexes was done twice ($n = 2$) with equal results. Cryo-EM data from two different grids has been collected with similar results (see Extended Data Fig. 1d). Reconstitution of 40S precursors and respective cryo-EM data collection has been done once ($n = 1$). No statistical analysis has been applied throughout the work.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability statement. All cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes EMD-4349, EMD-4348, EMD-4337, EMD-4350, EMD-4351, EMD-4353 and EMD-4352. The atomic model of state C and all working models have been deposited in the Protein Data Bank (PDB) under accessions 6G4W, 6G4S, 6G18, 6G51, 6G53, 6G5I and 6G5H.

23. Glatter, T., Wepf, A., Aebersold, R. & Gstaiger, M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol. Syst. Biol.* **5**, 237 (2009).
24. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
25. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
26. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
27. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
28. Kimanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**, e18722 (2016).
29. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349–W357 (2013).
30. Biasini, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–W258 (2014).
31. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
32. Nicholls, R. A., Fischer, M., McNicholas, S. & Murshudov, G. N. Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr. D* **70**, 2487–2499 (2014).
33. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).
34. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
35. Vagin, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D* **60**, 2184–2195 (2004).
36. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
37. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

a

b

PNO1-TAP | Mature 40S | *In vitro* reconstitution

c

PNO1-TAP | Mature 40S | *In vitro* reconstitution

d

Dataset 1: PNO1 N-TAP

Dataset 2: PNO1 N-TAP

Dataset 3: Mature 40S

Dataset 4: *In vitro* reconstitution

Extended Data Fig. 1 | Sample preparation and cryo-EM analysis of the pre-40S ribosome. a, Coomassie stained SDS–PAGE analysis of samples used in this work with bands labelled as identified by mass spectrometry. RBFs that we do not observe in our structures are marked with an asterisk. 60S bands are shown for comparison (see Supplementary Fig. 1 for source data). **b**, Representative micrographs from three datasets, low-pass filtered at 15 Å. **c**, Representative 2D classes of three datasets showing various orientations of pre-40S particles. **d**, Cryo-EM data processing scheme for all datasets with final volumes highlighted in red. Classes that could not be further refined or sorted are labelled flexible or noisy. For PNO1 N-TAP, two datasets were collected and individually processed. Particles of respective states were combined and further analysed (see Methods). Complex purification was done three times independently with the same results. Cryo-EM data collection and analysis was done twice for native complexes with similar results. Mass spectrometry analysis, pre-40S reconstitution and data collection for datasets 3 and 4 were done once. DS, dataset.
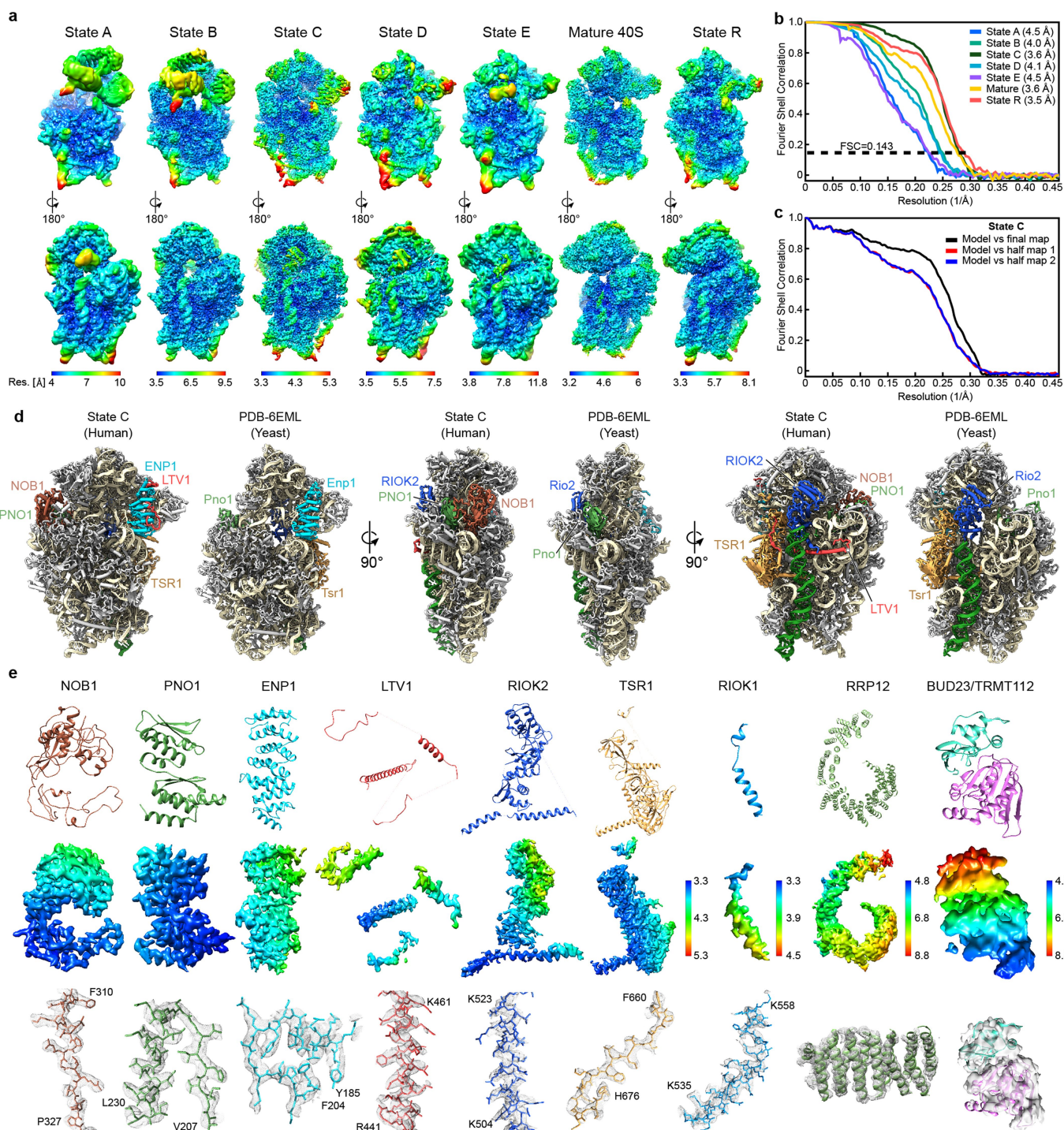
Extended Data Fig. 2 | Local resolution, refinement and model statistics. a, Local resolution distribution as estimated by Relion ranging from approximately 3 Å in well-resolved areas to 12 Å in more flexible parts. Colouring according to scale bars. **b**, Fourier shell correlation (FSC) plot with average resolutions according to the 'gold standard' (FSC = 0.143) stated in the legend. **c**, FSC plot of the state C model against cryo-EM maps as calculated by REFMAC5 (see Supplementary Data 1 for source data). **d**, Structural comparison between state C and a pre-40S particle from yeast (PDB code 6EML). Assembly factors coloured as in Fig. 1a. **e**, Cartoon representations of models of RBFs (top) are shown together with their respective density (mid). Volumes are coloured according to local resolution, which ranges from 3 Å in more rigid areas to 9 Å in flexible parts. RRP12, BUD23 and TRMT112 are less resolved, which only permitted placing of dummy helices and rigid body fitting of homology models, respectively. Examples of well resolved areas are depicted below.
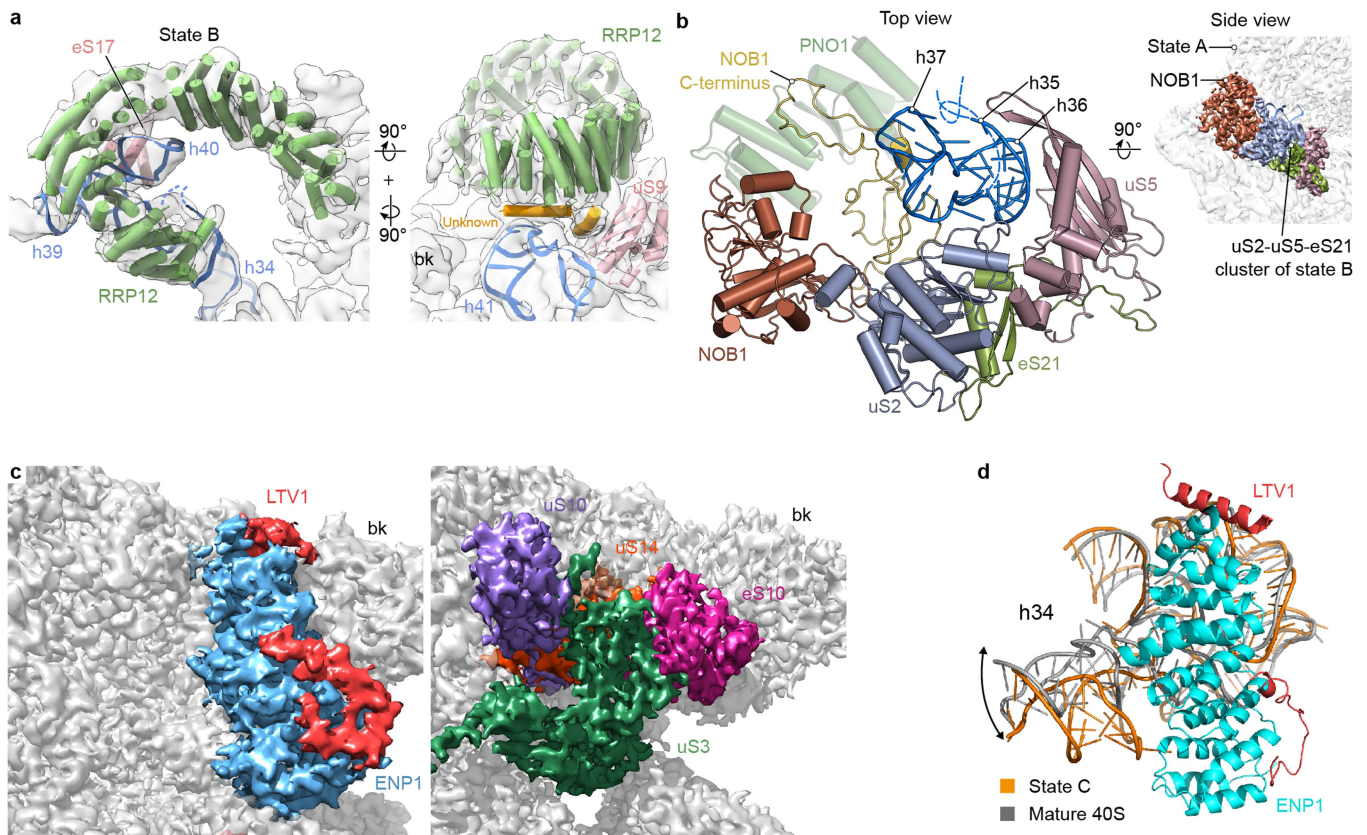
Extended Data Fig. 3 | Structural details of assembly factors and beak formation. a, Positioning of RRP12 in state B shows interaction with eS17 and uS9, as well as rRNA h34, h39, h40 and h41. Two additional helices of unknown identity bridge RRP12 and the head. Poly-alanine helices are placed in respective densities. **b,** Recruitment of NOB1 and the uS2–uS5–eS21 cluster after h35–h37 flipping; eS17 is omitted from this view. **c,** Close-up view of the beak region before (left) and after (right) replacement of ENP1-LTV1 with the uS3–eS10–uS20 cluster. **c,** Movement of h34 during maturation.
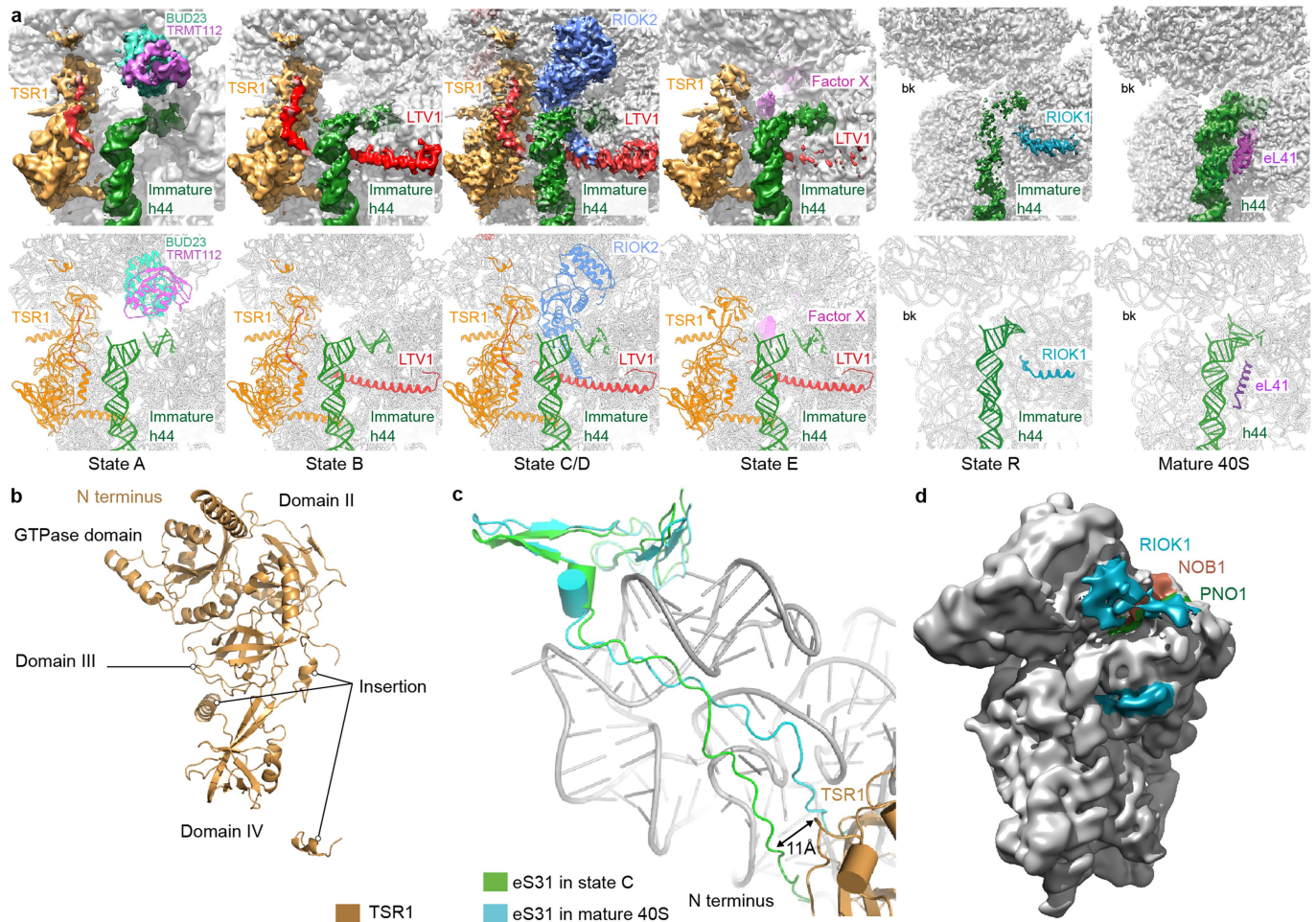
Extended Data Fig. 4 | The maturation of the decoding centre.
a, Close-up views of the decoding centre in density (top) and model (bottom) representation showing the position of the biogenesis factors close to h44. **b**, Domain arrangement of TSR1. Insertions and extensions of TSR1 are labelled. **c**, Conformational change of the N terminus of eS31 during maturation. The alignment was based on the C-terminal zinc-finger domain. **d**, Overall structure of the reconstituted particle low-pass filtered at 12 Å, showing the approximate positioning of RIOK1 (blue).
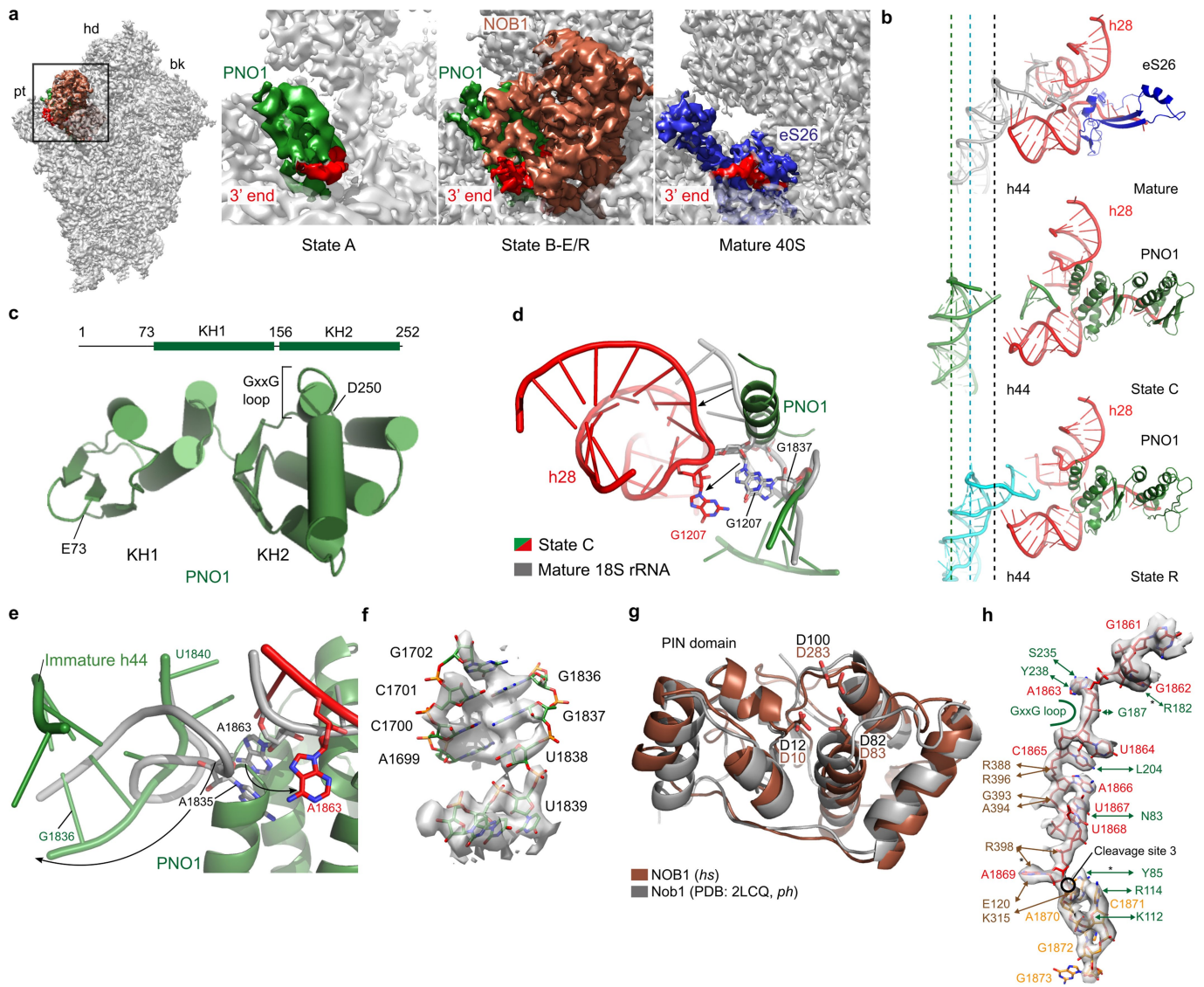
Extended Data Fig. 5 | PNO1, NOB1 and coordination of the 3′ rRNA end. a, The 3′ end region (box) throughout the maturation process. PNO1 and NOB1 bind similarly in state B–E and R. **b,** Cartoon representation of eS26 and PNO1 with parts of the pre-18S rRNA at different stages in maturation. h44 is shifted in state C. Displacement of eS26 by PNO1 in state R leads to a partially shifted h44. **c,** Overall structure of PNO1 with its two KH domains. **d, e,** Detailed view on the effects of PNO1 binding. Clashing of PNO1 with h28 and the linker region between h44 and h45

leads to a disruption of the base stacking between G1207 and G1837 (**d**) and A1835 and A1863 (**e**). **f,** Residues involved in novel base pairing with their respective electron density in state C. **g,** Comparison of the PIN domain of human (*hs*) NOB1 in state C and the NMR structure of Nob1 from *Pyrococcus horikoshii* (*ph*; PDB code 2LCQ), with their conserved active site residues highlighted. **h,** Summary of residues involved in binding of the rRNA 3′ end (asterisk indicates stacking). Electron density of state C surrounding the 3′ end is shown.
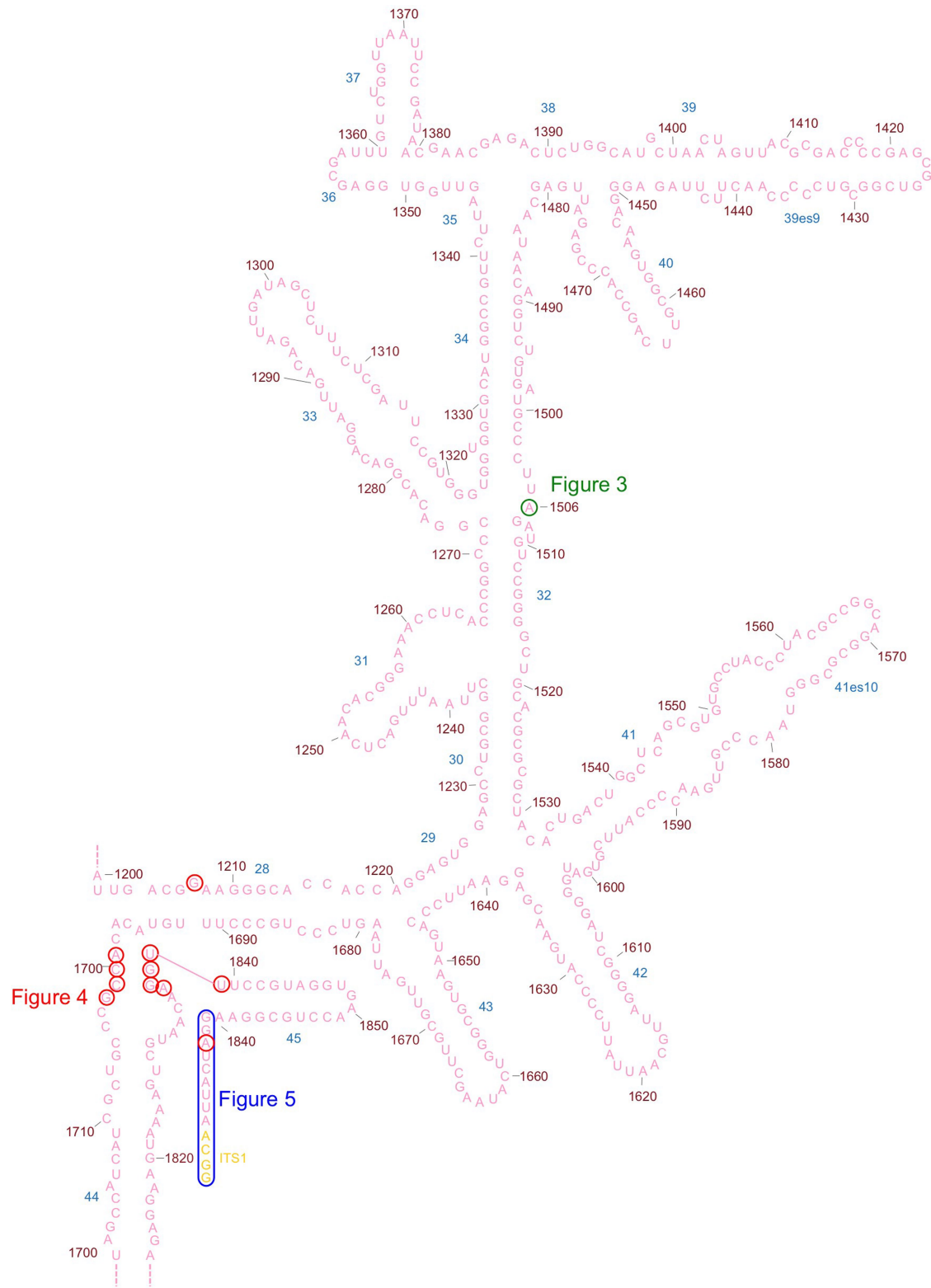
Extended Data Fig. 6 | 2D diagram of the pre-18S rRNA head region of state C. Extended secondary structure diagram of the pre-18S rRNA head region of state C. Residues mentioned throughout the text are highlighted. Related to Fig. 4b.
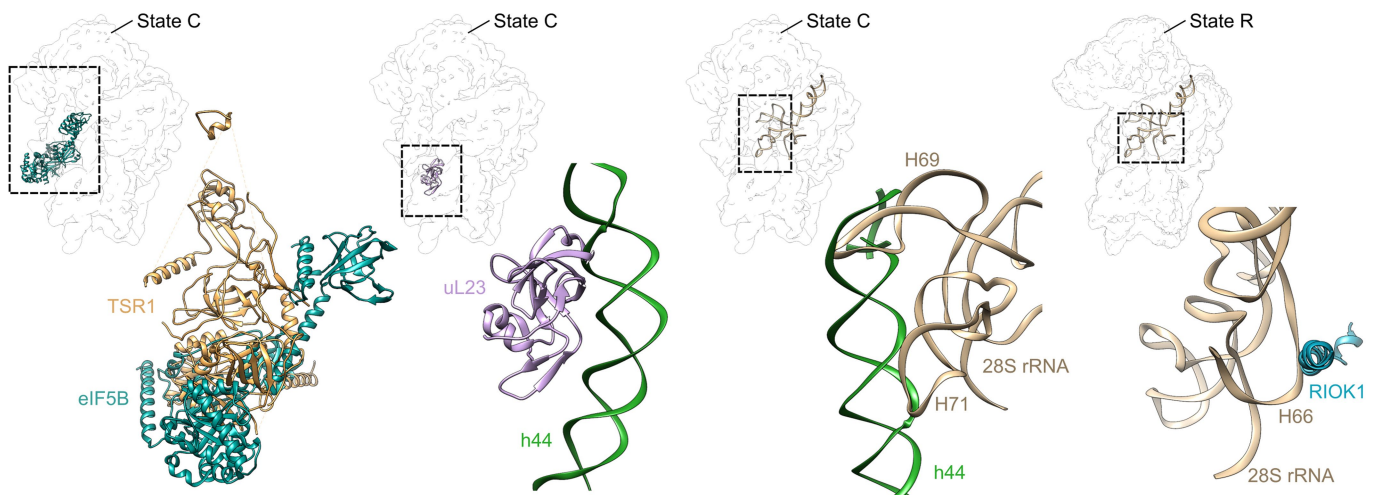
Extended Data Fig. 7 | 60S subunits components clash with immature h44. Overview over clash sites of large subunit components with pre-40S particles. Left, initiation factor eIF5B, which is potentially involved in formation of the 80S-like ribosome complex, occupies a similar binding site as TSR1. Middle, ribosomal protein uL23 and helices H69 and H71 of 28S rRNA clash with h44 in its immature position. Right, finally helix H66 of 28S rRNA clashes with a RIOK1 helix. A model of an 80S ribosome with eIF5B (PDB code 4UJD) was aligned to pre-40S states, and the factors are shown as overlays together with one of the 40S precursors.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	PNO1-TAP (EMD-4437) (PDB-6G18)	State R (EMD-4353) (PDB-6G5I)	Mature 40S (EMD-4352) (PDB-6G5H)
DATA COLLECTION			
Magnification	129,151	129,151	129,151
Voltage (kV)	300	300	300
Electron exposure (e⁻ Å⁻²)	25	25	25
Defocus range (μm)	-0.4 to -3.2	-0.7 to -3.3	-0.7 to -3.8
Pixel size (Å)	1.084	1.084	1.084
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	1,737,823	959,348	407,657
Final particle images (no.)	287,847 (state C)	83,883	70,822
Map resolution (Å)	3.6 (state C)	3.5	3.6
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	3.3-7.5	3.3-6.8	3.4-8.6
MODEL REFINEMENT	**State C**		
Refinement			
Initial model used	5A2Q		
Model resolution (Å)	3.6		
FSC threshold	0.5		
Map sharpening B factor (Å²)	-210.9		
Model resolution range (Å)	3.3-7.5		
Model composition			
Non-hydrogen atoms	83825		
Protein residues	6060		
RNA bases	1658		
Ligands	2		
B factors (Å²)	141.1		
R.m.s. deviations			
Bond lengths (Å)	0.0068		
Bond angles (°)	1.19		
Validation			
Molprobity score	2.46		
Clashscore, all atoms	6.28		
Poor rotamers (%)	7.77		
Ramachandran plot			
Favoured (%)	93.33		
Allowed (%)	5.85		
Outliers (%)	0.82		
Validation (RNA)			
Correct sugar puckers (%)	96.14		
Good backbone conformations (%)	67.37		

Summary of relevant parameters used during cryo-EM data collection and processing. Refinement and validation statistics are provided for the molecular model of state C.

nature research

Corresponding author(s):   Roland Beckmann

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	EM-Tools
Data analysis	MotionCor2, CTFFIND4 v4.1.10, Gctf v1.06, Relion 2.1-beta-1, Gautomatch v0.56, PSIPRED v3.3, SWISS-MODEL, ProSMART v0.856/LIBG/ REFMAC5 from CCP4 v7.0, Phenix 1.12 (includes MolProbity), Chimera 1.12, Pymol 1.8, Coot 0.8.8, ChimeraX v0.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All cryo-EM density maps have been deposited in the EM Data Bank under the accession codes EMD-4349, EMD-4348, EMD-4337, EMD-4350, EMD-4351,

EMD-4353 and EMD-4352. The atomic model of state C and all working models have been deposited in the Protein Data Bank under PDB 6G4W, 6G4S, 6G18, 6G51, 6G53, 6G5I and 6G5H.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were applied to pre-determine sample sizes
Data exclusions	No data were excluded from the analysis
Replication	All biochemical and biological experiments were successfully and reliably reproduced. Cryo-EM data collection and data analysis was performed once.
Randomization	No randomization was required for the reported experiments
Blinding	Blinding was not required for the reported experiments

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293 T-REx Flp-In cell line: Thermo Fischer Scientific
Authentication	Cells were not authenticated
Mycoplasma contamination	The commercial cell line used in this work was not tested for mycoplasma contamination, however we have a stringent testing policy with all foreign cell lines brought to the laboratory to keep the laboratory mycoplasma free.
Commonly misidentified lines (See ICLAC register)	No cell lines used throughout the work are listed in the ICLAC database

Article

Structural basis for the final steps of human 40S ribosome maturation

Michael Ameismeier[1], Ivo Zemp[2], Jasmin van den Heuvel[2,3], Matthias Thoms[1], Otto Berninghausen[1], Ulrike Kutay[2] & Roland Beckmann[1✉]

Eukaryotic ribosomes consist of a small 40S and a large 60S subunit that are assembled in a highly coordinated manner. More than 200 factors ensure correct modification, processing and folding of ribosomal RNA and the timely incorporation of ribosomal proteins[1,2]. Small subunit maturation ends in the cytosol, when the final rRNA precursor, 18S-E, is cleaved at site 3 by the endonuclease NOB1[3]. Previous structures of human 40S precursors have shown that NOB1 is kept in an inactive state by its partner PNO1[4]. The final maturation events, including the activation of NOB1 for the decisive rRNA-cleavage step and the mechanisms driving the dissociation of the last biogenesis factors have, however, remained unresolved. Here we report five cryo-electron microscopy structures of human 40S subunit precursors, which describe the compositional and conformational progression during the final steps of 40S assembly. Our structures explain the central role of RIOK1 in the displacement and dissociation of PNO1, which in turn allows conformational changes and activation of the endonuclease NOB1. In addition, we observe two factors, eukaryotic translation initiation factor 1A domain-containing protein (EIF1AD) and leucine-rich repeat-containing protein 47 (LRRC47), which bind to late pre-40S particles near RIOK1 and the central rRNA helix 44. Finally, functional data shows that EIF1AD is required for efficient assembly factor recycling and 18S-E processing. Our results thus enable a detailed understanding of the last steps in 40S formation in human cells and, in addition, provide evidence for principal differences in small ribosomal subunit formation between humans and the model organism *Saccharomyces cerevisiae*.

Ribosomes are the central components of the cellular machinery that translates messenger RNA into protein. In eukaryotes, these complexes consist of a small 40S and a large 60S subunit, which harbour the decoding and catalytic activities, respectively. Together, both subunits comprise four rRNAs and about 80 ribosomal proteins. Ribosome biogenesis is coordinated by more than 200 *trans*-acting protein and RNA factors[1,2], and starts in the nucleolus with the transcription of a single primary transcript onto which these factors and ribosomal proteins are assembled to form an early pre-ribosome. Endonucleolytic cleavage leads to its separation into small and large subunit precursors[3], which further mature within the nucleus and are subsequently exported into the cytoplasm, where final assembly steps occur. Small subunit formation ends with the cleavage of 18S-E, the last 18S rRNA precursor, by NOB1 at site 3 and the release of the remnants of the flanking internal transcribed spacer 1 (ITS1), as well as all remaining biogenesis factors[1,5].

We previously described the structures of several nuclear and early cytoplasmic 40S precursors leading up to the final and decisive endonucleolytic processing step[4]. Therefore, it remained unclear how the conformational maturation of the central region in the 40S subunit is coupled to NOB1 activation. In yeast, final cleavage by Nob1 has been suggested to occur in 80S-like complexes and requires binding and enzymatic activity of several additional factors, including the translational GTPase eIF5B and the ATPase Fap7[6–8]. Involvement of the large subunit in 40S maturation is thought of as a quality-control step, ensuring structural integrity and functionality of nascent 40S subunits[6]. The 80S-like formation does not appear to occur in humans and, to our knowledge, no comparable quality control mechanism has been described.

In this study, we report cryo-electron microscopy (cryo-EM) structures of 40S precursors, which provide insights into the mechanism of NOB1 activation and the finalization of 40S assembly. They show that RIOK1, an atypical kinase related to the earlier-acting RIOK2[9,10], serves as a central coordinator of conformational and compositional changes in 40S assembly intermediates. Together with the biogenesis factor EIF1AD, rearrangements of RIOK1 couple the conformational maturation of the 18S rRNA with the displacement of PNO1 and allow for the activation of the endonuclease NOB1.

Cryo-EM structures of 40S precursors

For the structural analysis of these late transitions, we used catalytically inactive mutants of RIOK1 and NOB1 to purify several native 40S

[1]Gene Center Munich, Department of Biochemistry, University of Munich, Munich, Germany. [2]Institute of Biochemistry, Department of Biology, ETH Zurich, Zurich, Switzerland. [3]Molecular Life Sciences Ph.D. Program, Zurich, Switzerland. ✉e-mail: beckmann@genzentrum.lmu.de
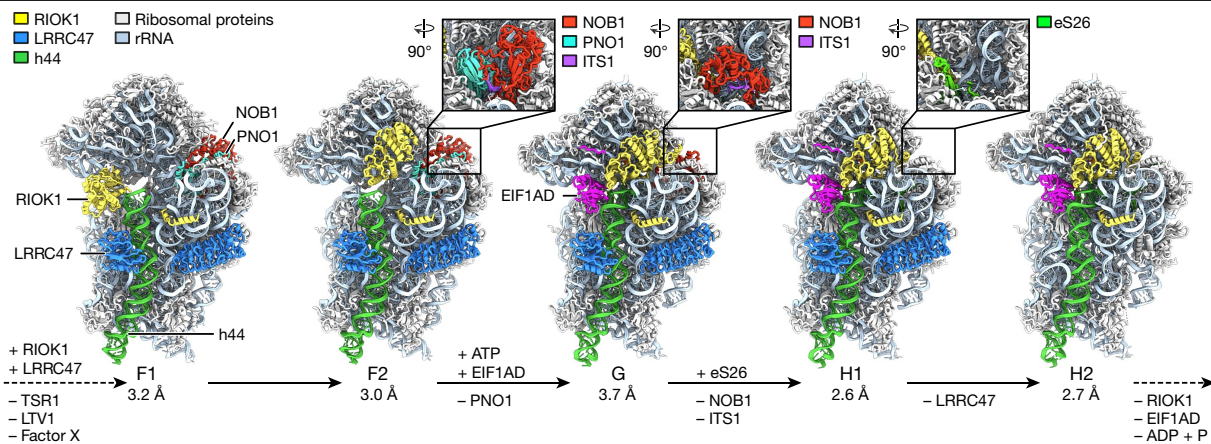
Fig. 1 | Structures of human pre-40S ribosomal intermediates during final processing steps. Ribosomal biogenesis factors NOB1, PNO1, RIOK1, LRRC47 and EIF1AD, ribosomal protein eS26 and rRNA segments ITS1 and h44 are named and highlighted in colour (grey: ribosomal proteins, light blue: 18S-E rRNA). Assigned names, overall resolution and changes in protein and nucleotide composition for each state are displayed at the bottom. Magnified regions show the conformational and compositional changes around ITS1. P, phosphate.

precursors from human cells (Extended Data Fig. 1a). We determined their structures at an overall resolution of 2.6 to 3.7 Å using ensemble single-particle cryo-EM (Extended Data Figs. 1, 2, Extended Data Table 1) and placed the particles in a logical sequence on the basis of conformational and compositional differences (Fig. 1). Resulting states F1–H2[4] are subsequent to our previously described sequence and represent the events just before and after the final 18S rRNA processing step. State F1 is characterized by the loss of the ribosome biogenesis factors (RBFs) TSR1, LTV1 and the unidentified factor X, as well as the presence of LRRC47 and RIOK1. RIOK1 repositions in state F2, before settling into its final conformation in state G.

The transition into state G is accompanied by binding of EIF1AD, concomitant with structural rearrangements of rRNA, displacement of PNO1 and the subsequent activation of NOB1. After cleavage at site 3 and release of NOB1 and the remains of ITS1, ribosomal protein eS26 (also known as S26) can be observed in its canonical position at the platform in state H1. Finally, loss of LRRC47 in state H2 concludes the sequence, with only the release of RIOK1 and EIF1AD remaining for human 40S maturation to be completed.

RIOK1 coordinates final maturation steps

On the basis of our structures, RIOK1 appears central to conformational rearrangements and the activation of NOB1. RIOK1, which is the founding member of the RIO protein family, displays three different binding modes to the particle in states F1, F2 and G, with its position at the decoding centre in state G resembling that of previously bound RIOK2[4] (Extended Data Fig. 3a). In state F1, two C-terminal helices, required for binding to the particle[11], have displaced the C terminus of the RBF LTV1, resulting in its release[9] (Extended Data Fig. 3b). The N terminus of RIOK1 remains flexible in the first two states, as it is observed only as fuzzy density, but has already replaced the two helices of factor X within the mRNA entry tunnel (Fig. 2a, Extended Data Fig. 3c). Initially, RIOK1 is bound with its RIO domain between rRNA helix 18 (h18) and h34, close to the flexible top region of h44, in an outward facing position, partially occupying the position of TSR1 in earlier particles (Figs. 1, 2a, Extended Data Fig. 3d). A short stretch, N-terminal to the RIO domain, augments the central, 5-stranded β-sheet (β') of RIOK1 and, together with a short α-helix (α'), contributes to binding of rRNA h34 (Fig. 2a). Notably, in state F2, this interaction is lost and RIOK1 rotates by approximately 90° to position the kinase domain close to the decoding centre, while the N terminus stays fixed within the tunnel (Fig. 1, Extended Data Fig. 3c). In these initial states, RIOK1 appears

to be in an apo form, devoid of a nucleotide within its active site, and several regions, including the functionally important flexible loop[11,12], remain disordered (Fig. 2a). Following accommodation of RIOK1 at the decoding centre (Fig. 2b) and association of EIF1AD, however, RIOK1 is ATP-bound throughout the rest of the sequence (states G–H2 in Figs. 1, 2c, d). This is consistent with previous data indicating that ATP binding is sufficient for rRNA cleavage[8], although ATP hydrolysis appears to increase processing efficiency[8,9].

At its final position in states G–H2, RIOK1 coordinates the hypermodified nucleotide 1248, which is converted during 40S maturation to 1-methyl-3-α-amino-α-carboxypropyl pseudouridine ($m^1acp^3\Psi$) in a three-step mechanism[13] (Extended Data Fig. 3e, f). This nucleotide is
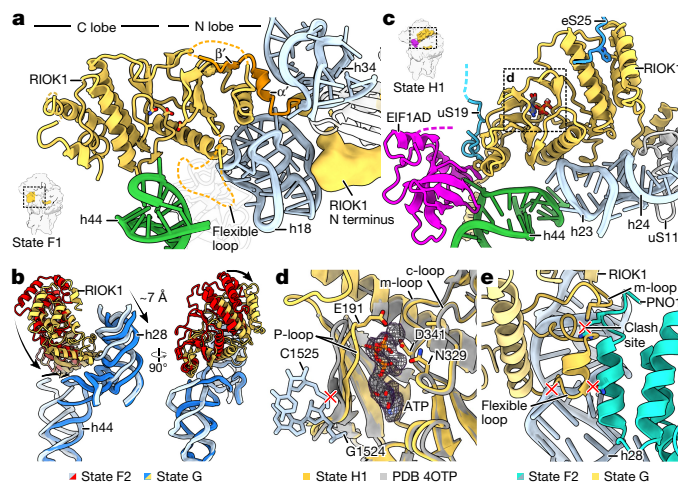
Fig. 2 | Conformational rearrangements of RIOK1 coordinate late steps in 40S maturation. a, Cartoon representation of RIOK1 bound to state F1 with secondary structure elements α' and β' and rRNA segments labelled. The N-terminal stretch that binds h18 and h34 is highlighted in orange. Putative RIOK1 N-terminus density was Gaussian filtered at 1.7σ. **b,** Superposition of RIOK1, h28 and h44 in states F2 and G. **c,** Cartoon representation of ATP-bound RIOK1 bound to the decoding centre in state H1 and its interactions with rRNA, EIF1AD, uS19 and eS25. **d,** Electron density and model of Mg^{2+}-ATP bound to the active site of RIOK1 in state H1. Superposition of ADP-bound human RIOK1 (Protein Data Bank (PDB): 4OTP) shown in grey. **e,** Superposition of RIOK1 in state G and h28 and PNO1 in state F2 with clashes of RIOK1 m-loop and flexible loop. Clash sites are marked with red crosses.

located close to the P-site and a loss of its modifications has recently been shown to be a major feature of several cancer types[14]. Notably, aminocarboxypropylation by TSR3 occurs during the final stages of 40S formation and is required for efficient 20S and 18S-E processing in yeast and humans, respectively[13,15]. Therefore, RIOK1 activity is likely to be dependent on a late and functionally important modification that may serve as a quality checkpoint.

Concomitantly and facilitated by the dissociation of terminal helices of TSR1 and LTV1[4,16,17], the maturation of the intersubunit surface has further progressed by a rotation and translation of rRNA h44 into its mature position in state G (Figs. 1, 2b). The immature base pairing between A1699–G1702 and G1836–U1839 observed in preceding states has been resolved and the region connecting h44 with h45 and h28 has formed. Furthermore, h28, the linker between the main body and head of the 40S, is shifted by approximately 7 Å by the flexible loop of RIOK1 (Fig. 2b, e). Together with overlapping positions of the m-loop and flexible loop of RIOK1 in state G and the C-terminal helix of PNO1 in preceding states, these structural changes in rRNA conformation result in the release of PNO1 and ultimately in the activation of NOB1 (Fig. 1). Although eS26 is known to be required for efficient processing of 18S-E rRNA[18], eS26 accommodation was not observed after PNO1 release in state G, because its canonical binding site is still blocked by the relocated PIN domain of NOB1.

Local resolution in F1 and F2 particles prevented a detailed comparison between the apo and ATP-bound form; however, conformational changes in RIOK1 are highly likely to govern this transition. In Rio1 of *Archaeoglobus fulgidus*, nucleotide binding has been shown to induce conformational changes in the flexible loop, as well as a contraction of the RIO domain around the nucleotide binding site[12]. ATP binding by human RIOK1 may, in a similar way, enable settling of RIOK1 into its final position accompanied by rRNA remodelling and, eventually, activation of NOB1. Moreover, comparison of the ATP-bound RIOK1 active site with a crystal structure containing ADP and a phospho-aspartate intermediate showed a considerable relocation of the P-loop (Fig. 2d). Similar conformational changes during ATP hydrolysis could explain the subsequent release of the RIOK1 kinase domain from the particle, as rRNA residues G1524 and C1525 would clash with the P-loop in an ADP-bound state (Fig. 2d). Thus, ATP binding and hydrolysis by RIOK1 are central in the last steps of 40S maturation.

Characterization of EIF1AD and LRRC47

Consistent with recent work[19], we identified two factors, EIF1AD and LRRC47, acting alongside RIOK1 in late human 40S maturation (Figs. 1, 3a, b), both of which have no homologues in the yeast *S. cerevisiae*. To our knowledge, EIF1AD and LRRC47 have not been identified as ribosome biogenesis factors in high-throughput RNA-mediated interference (RNAi) screens, probably owing to the choice of experimental readout[20,21]. LRRC47 consists of two major domains, an N-terminal leucine-rich repeat and a C-terminal phenylalanyl-tRNA synthetase or B3/4 domain (Fig. 3a, Extended Data Fig. 4a). Both parts were bound to states F1–H1, albeit with different rigidity. The N-terminal leucine-rich repeat was present stoichiometrically and invariably bound to rRNA helices h11, h20, h22 and h24, as well as uS15 below the platform, potentially acting as an anchor initiating pre-40S binding. This is consistent with a low-resolution cryo-EM reconstruction of purified pre-40S complexes using N-terminally tagged LRRC47 as bait, which reveals simultaneous binding of TSR1 and the N-terminal domain (Extended Data Fig. 4b). The C-terminal domain, however, was both flexibly and sub-stoichiometrically bound. The best-resolved part comprised residues 408–442, which formed an α-helix–loop–α-helix motif, bound to rRNA h5 and h13–h14 and occupied the position of the formerly bound N-terminal helix of TSR1[4] (Fig. 3c, d). Unlike TSR1, however, both LRRC47 domains could accommodate the conformational changes of rRNA h44 during maturation (Extended Data Fig. 4c).

Notably, both domains of LRRC47 would clash with a 60S subunit in its 80S conformation, thus suggesting a potential anti-association activity.

The second factor, EIF1AD, is a paralogue of eukaryotic translation initiation factor 1A (eIF1A)[22], and is essential in multiple human cell lines[23]. Although it is well conserved from fungi to humans, it is absent in *S. cerevisiae*. When compared to the position of eIF1A on the mature 40S subunit, EIF1AD showed a similar binding mode on the pre-40S particle (Figs. 1, 3b, Extended Data Fig. 4d). A conserved tryptophan residue (W62) coordinated the flipped-out base A1825 of rRNA h44 and additionally interacted with residues F171 and I142 of RIOK1 and uS19 (also known as S15), respectively (Fig. 3e). The main oligonucleotide- or oligosaccharide-binding fold was slightly shifted relative to the 40S particle compared with eIF1A in yeast translation initiation complexes, and the N terminus of EIF1AD was bound between rRNA h18 and h34 (Extended Data Fig. 4e). In contrast to eIF1A, however, EIF1AD lacks the C-terminal IDDI motif required for interaction with eIF5B, which is essential to trigger 60S subunit joining for 80S formation[24] (Extended Data Fig. 4f). Instead, a conserved stretch of EIF1AD is bound to rRNA h30, uS13 and uS19, with the C terminus looping over the RIOK1 kinase domain (Fig. 3b, Extended Data Fig. 4f, g) next to the N terminus of eS25 (Fig. 2c, Extended Data Fig. 4g).

To analyse their functional relevance, we depleted EIF1AD and LRRC47, as well as RIOK2 as a positive control[25], from HeLa cells using RNAi and monitored the distribution of the 40S assembly factors ENP1 (also known as bystin), TSR1, RIOK1, NOB1 and PNO1 by immunofluorescence (Fig. 4a, Extended Data Fig. 5). LRRC47 knockdown had no effect on recycling of TSR1, PNO1 and NOB1 after release from pre-40S particles or on ITS1 localization, indicating a non-essential function in 40S maturation (Extended Data Fig. 5). Cells lacking EIF1AD, however, displayed cytoplasmic enrichment and enhanced nuclear exclusion of RIOK1, PNO1 and NOB1, similar to control cells after RIOK2 depletion. This shows that EIF1AD is required for efficient release of PNO1, NOB1 and RIOK1 from pre-40S particles, consistent with the observed direct contact to RIOK1 and an effect on RIOK1 function. Unlike RIOK2, EIF1AD is not required for TSR1 and ENP1 release, as indicated by a retained primarily nuclear localization of both factors. This indicates that EIF1AD acts downstream of RIOK2 and after the release of TSR1 and ENP1 from the particle, in accordance with our structural data. In addition to the defect in RBF recycling, cells lacking EIF1AD show an increased ITS1 signal in the cytosol, highlighting the contribution of EIF1AD to the final steps of 18S rRNA processing (Fig. 4b). On the basis
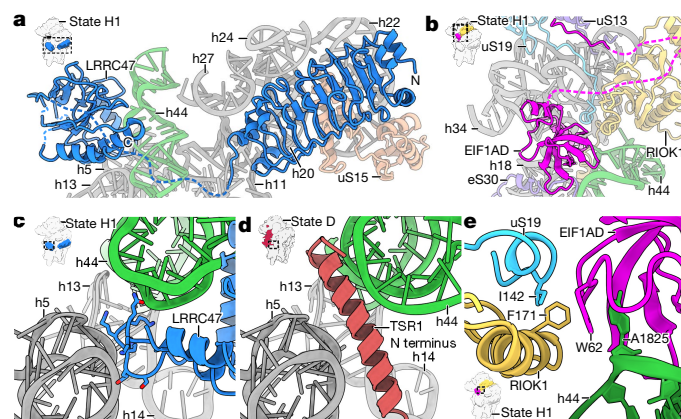
Fig. 3 | Association of LRRC47 and EIF1AD with 40S subunit precursors.
a, **b**, Overview of LRRC47 (**a**) and EIF1AD (**b**) binding sites to the pre-40S particle. **c**, Cartoon representation of rRNA helices and the LRRC47 helix–loop–helix motif with interacting residues shown as sticks. **d**, Model of the N-terminal helix of TSR1 occupying the same binding site in state D (PDB: 6G51). **e**, Detailed view of the interaction of EIF1AD with RIOK1, uS19 and rRNA helix h44.

a

| siCtrl | siRIOK2 | siEIF1AD | siPCtrl | siPEIF1AD |

PNO1

NOB1 + LMB

RIOK1

TSR1

ENP1

EIF1AD

b

Cy5-5'ITS1
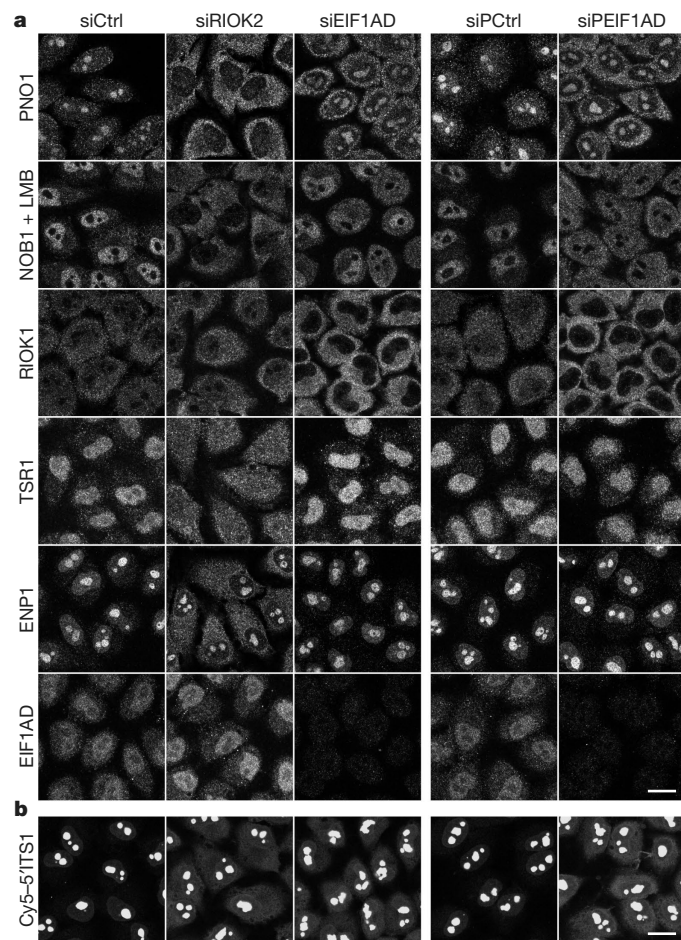
Fig. 4 | EIF1AD depletion leads to late cytoplasmic 40S maturation defects.
a, Immunofluorescence analysis of the indicated RBFs in HeLa cells after
RNAi-mediated depletion of EIF1AD (siEIF1AD) or RIOK2 (siRIOK2) using single
(si) or pooled (siP) siRNAs (see Methods). siCtrl, control. Scale bar, 20 μm. LMB,
leptomycin B. **b**, Fluorescence in situ hybridization (FISH) analysis of the
experiment in **a**, revealing cytoplasmic accumulation of 18S-E pre-rRNA upon
EIF1AD depletion. FISH images were processed in parallel, using a gamma
correction of 1.5. All experiments were performed in biological triplicates (*n* = 3).

of our structural and functional data, we therefore suggest that EIF1AD
can stabilize the RIOK1 kinase in its locked-in, ATP-bound state, thereby
contributing to the conformational changes of the pre-rRNA and the
ultimate activation of NOB1.

Activation of the endonuclease NOB1

The release of PNO1 has previously been shown to occur directly before
the final 3′ end cleavage by Nob1 in yeast[8]. Notably, we were unable
to purify and detect complexes that contained NOB1 without PNO1
when using wild-type (WT) RIOK1 (RIOK1(WT)) (data not shown) or
RIOK1(D324A) as bait. This indicated a rather fast activation of NOB1
upon PNO1 displacement, followed by cleavage at site 3 and release
of NOB1 and ITS1. The nuclease-deficient mutant NOB1(D10N)[26] was
trapped in a monomeric and active conformation, engaging its sub-
strate RNA in about 5% of the particles (Fig. 1, state G, Extended Data
Fig. 1e, f).

State G revealed a conformational switch of NOB1: the PIN domain,
which harbours the active site, rotated relative to its invariant zinc
finger domain by about 55° (Fig. 5a, Extended Data Fig. 6a), reflecting
a structural and functional independence that had previously been
indicated biochemically[27]. The insertion and C-terminal extension of
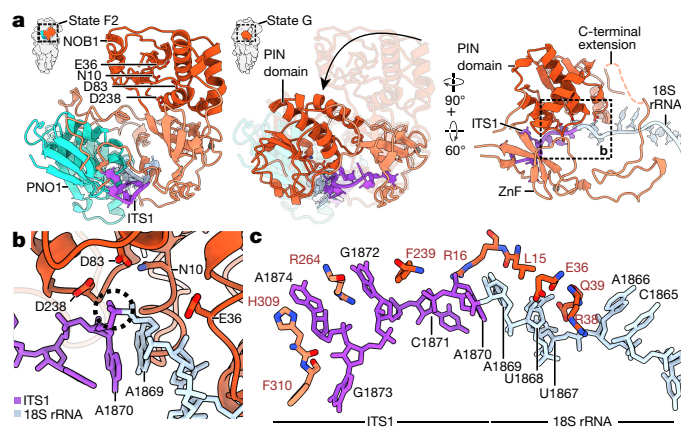
Fig. 5 | Active conformation of NOB1 for site 3 cleavage. a, Cartoon
representation of NOB1(D10N) bound to the platform in inactive (state F2)
and active (state G) conformations, with the active-site residues labelled.
Movement of the PIN domain towards ITS1 after PNO1 release is indicated by an
arrow. **b**, Detailed view of the active site of NOB1(D10N) with its substrate.
Dashed circle marks site 3 cleavage site. **c**, Stick representation of NOB1
residues that interact with the ITS1 and 3′ end. ZnF, zinc finger domain.

NOB1, which had been bound to PNO1 in states F1 and F2, were now flex-
ible and only partially resolved. In addition, a conformational shift in
the rRNA 3′ end–ITS1 segment resulted in the precise accommodation
of the site 3 cleavage site between A1869 and A1870 within the active
site of the nuclease (Fig. 5b, Extended Data Fig. 6b). NOB1 residues
D83, D238, E36 and the mutated N10 are located in close proximity and
can be expected to mediate cleavage in a wild-type background. NOB1
has recently been shown to bind rRNA segments around the 3′ end
and ITS1 in both yeast and human[27,28]. While we saw extensive interac-
tion between NOB1 and the 18S-E rRNA around site 3, no base-specific
interaction was apparent. Instead, binding of the rRNA was mediated
through several π–π interactions between nucleobases and amino acid
side chains, including a tight coordination of G1872 and A1874 of the
ITS1 by H309, R264 and F239 of NOB1 (Fig. 5c, Extended Data Fig. 6c).
Such close interaction with the ITS1 may result in the joint release of
a NOB1–ITS1 complex after site 3 cleavage without the requirement
for further 18S rRNA remodelling. The processing of 18S-E rRNA and
dissociation of NOB1 and ITS1 coincides with incorporation of the last
ribosomal protein, eS26 (state H1 in Fig. 1, Extended Data Fig. 6a), which
has been shown to be required for efficient processing[18]. This concludes
the maturation process of the 18S rRNA 3′ end region.

Conclusion

Our data provides detailed insights into the final steps of human 40S
maturation. In accordance with functional data[9], RIOK1 exerts a central
role in coordinating rRNA rearrangements with PNO1 release and NOB1
activation, which also requires EIF1AD. The influence of this transla-
tion initiation paralogue on maturation might ensure that the final
maturation of 40S subunits can only occur if they are competent for
eIF1A binding and thus translation initiation, which would fit into the
reoccurring concept of 'placeholder factors' in ribosome biogenesis
(reviewed in ref. [29]). Unlike in the model organism *S. cerevisiae*, in which
final rRNA processing has been proposed to require formation of an
80S-like particle in a quality-control step[6], involvement of the large
subunit in site 3 processing has not been shown in human cells[5]. As a
mimic of translation initiation factor eIF1A, the closely related protein
EIF1AD might therefore serve as a potential quality-control factor in
an alternative proofreading step. This seems consistent with an inhibi-
tion of 60S binding by LRRC47 due to steric hindrance at the subunit
interface.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2929-x.

1. Woolford, J. L. Jr & Baserga, S. J. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* **195**, 643–681 (2013).
2. Bohnsack, K. E. & Bohnsack, M. T. Uncovering the assembly pathway of human ribosomes and its emerging links to disease. *EMBO J.* **38**, e100278 (2019).
3. Henras, A. K., Plisson-Chastang, C., O'Donohue, M.-F., Chakraborty, A. & Gleizes, P.-E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **6**, 225–242 (2015).
4. Ameismeier, M., Cheng, J., Berninghausen, O. & Beckmann, R. Visualizing late states of human 40S ribosomal subunit maturation. *Nature* **558**, 249–253 (2018).
5. Cerezo, E. et al. Maturation of pre-40S particles in yeast and humans. *Wiley Interdiscip. Rev. RNA* **10**, e1516 (2019).
6. Strunk, B. S., Novak, M. N., Young, C. L. & Karbstein, K. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* **150**, 111–121 (2012).
7. Lebaron, S. et al. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.* **19**, 744–753 (2012).
8. Turowski, T. W. et al. Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. *Nucleic Acids Res.* **42**, 12189–12199 (2014).
9. Widmann, B. et al. The kinase activity of human Rio1 is required for final steps of cytoplasmic maturation of 40S subunits. *Mol. Biol. Cell* **23**, 22–35 (2012).
10. Angermayr, M., Roidl, A. & Bandlow, W. Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression. *Mol. Microbiol.* **44**, 309–324 (2002).
11. Ferreira-Cerca, S., Kiburu, I., Thomson, E., LaRonde, N. & Hurt, E. Dominant Rio1 kinase/ATPase catalytic mutant induces trapping of late pre-40S biogenesis factors in 80S-like ribosomes. *Nucleic Acids Res.* **42**, 8635–8647 (2014).
12. Laronde-LeBlanc, N., Guszczynski, T., Copeland, T. & Wlodawer, A. Structure and activity of the atypical serine kinase Rio1. *FEBS J.* **272**, 3698–3713 (2005).
13. Meyer, B. et al. Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. *Nucleic Acids Res.* **44**, 4304–4316 (2016).
14. Babaian, A. et al. Loss of m1acp3Ψ ribosomal RNA modification is a major feature of cancer. *Cell Rep.* **31**, 107611 (2020).
15. Hector, R. D. et al. Snapshots of pre-rRNA structural flexibility reveal eukaryotic 40S assembly dynamics at nucleotide resolution. *Nucleic Acids Res.* **42**, 12138–12154 (2014).
16. Heuer, A. et al. Cryo-EM structure of a late pre-40S ribosomal subunit from *Saccharomyces cerevisiae*. *eLife* **6**, e30189 (2017).
17. Scaiola, A. et al. Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit. *EMBO J.* **37**, e98499 (2018).
18. O'Donohue, M. F., Choesmel, V., Faubladier, M., Fichant, G. & Gleizes, P. E. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *J. Cell Biol.* **190**, 853–866 (2010).
19. Montellese, C. et al. USP16 counteracts mono-ubiquitination of RPS27a and promotes maturation of the 40S ribosomal subunit. *eLife* **9**, e54435 (2020).
20. Badertscher, L. et al. Genome-wide RNAi screening identifies protein modules required for 40S subunit synthesis in human cells. *Cell Rep.* **13**, 2879–2891 (2015).
21. Farley-Barnes, K. I. et al. Diverse regulators of human ribosome biogenesis discovered by changes in nucleolar number. *Cell Rep.* **22**, 1923–1934 (2018).
22. Yu, J. & Marintchev, A. Comparative sequence and structure analysis of eIF1A and eIF1AD. *BMC Struct. Biol.* **18**, 11 (2018).
23. Bertomeu, T. et al. A high-resolution genome-wide CRISPR/Cas9 viability screen reveals structural features and contextual diversity of the human cell-essential proteome. *Mol. Cell. Biol.* **38**, e00302-17 (2018).
24. Acker, M. G., Shin, B. S., Dever, T. E. & Lorsch, J. R. Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. *J. Biol. Chem.* **281**, 8469–8475 (2006).
25. Zemp, I. et al. Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *J. Cell Biol.* **185**, 1167–1180 (2009).
26. Pertschy, B. et al. RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.* **284**, 35079–35091 (2009).
27. Sloan, K. E., Knox, A. A., Wells, G. R., Schneider, C. & Watkins, N. J. Interactions and activities of factors involved in the late stages of human 18S rRNA maturation. *RNA Biol.* **16**, 196–210 (2019).
28. Lamanna, A. C. & Karbstein, K. Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proc. Natl Acad. Sci. USA* **106**, 14259–14264 (2009).
29. Espinar-Marchena, F. J., Babiano, R. & Cruz, J. Placeholder factors in ribosome biogenesis: please, pave my way. *Microb. Cell* **4**, 144–168 (2017).

Article

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Generation of cell lines

Cell lines expressing RIOK1(D324A), NOB1(D10N) or LRRC47 were generated as previously described[4,30]. In brief, regions coding for RIOK1, NOB1 and LRRC47 were amplified from human cDNA (Amsbio) using KOD DNA polymerase (Merck). An internal Strep-Flag tag following amino acid 496 was added to RIOK1 before insertion into a pcDNA5/FRT/TO vector. In case of NOB1 and LRRC47, the PCR product was inserted into a pcDNA5/FRT/TO vector that contained an N-terminal Strep-Flag tag. RIOK1(D324A) and NOB1(D10N) were generated using site-directed mutagenesis by PCR. HEK Flp-In 293 T-Rex cells (Invitrogen) were transfected with 0.5 μg of the target plasmid and 4.5 μg pOG44 (Invitrogen) in 10-cm cell-culture dishes at approximately 70% confluence, using 20 μg polyethyleneimine (PEI). Selection was performed with 150 μg ml^{-1} hygromycin B (Thermo Scientific) and cells were subsequently maintained in DMEM (Thermo Scientific), containing 10% FBS, 100 μg ml^{-1} hygromycin B, 10 μg ml^{-1} blasticidin, and 1× penicillin/streptomycin and GlutaMAX (Thermo Scientific). Commercial cell lines have not been authenticated and tested negative for mycoplasma contamination.

Native complex preparation

Native pre-ribosomal complexes were purified as previously described[4]. In brief, cell lines were grown to approximately 80% confluence, before protein expression was induced for 24 h with 1.6 μg ml^{-1} tetracycline. Cells from 20 to 40 15-cm cell culture dishes were collected in 0.025% trypsin–EDTA (Thermo Scientific), washed once in 1× phosphate-buffered saline (PBS) and pelleted at 1,600g at 4 °C. Cells were then lysed in wash buffer (20 mM HEPES pH 7.6, 150 mM potassium acetate, 5 mM MgCl$_2$, 1 mM DTT, 0.5 mM NaF, 0.1 mM Na$_3$VO$_4$, 1× protease inhibitor (Sigma Aldrich)) containing 0.5% NP-40 substitute for 30 min on an overhead shaker at 4 °C. The lysate was cleared by centrifugation at 4,000g for 15 min at 4 °C and incubated with 4 ml of equilibrated StrepTactin XT affinity beads (IBA Life Sciences) for 2 h. Beads were subsequently collected in a small column, washed 5 times with 4 ml wash buffer and eluted 7 times by incubating the beads with 4 ml wash buffer containing 50 mM D-biotin (Roth) for 30 min. Eluates were combined and incubated with 100 μl of equilibrated anti-Flag affinity beads (Sigma-Aldrich) for 2 h at 4 °C. In case of RIOK1 and LRRC47, the first purification step using the Strep tag was skipped and the cleared cell lysate was directly incubated with anti-Flag affinity beads. The beads were then washed four times with 1 ml wash buffer, before the bound material was eluted six times by incubation with 100 μl of 20 mM HEPES pH 7.6, 150 mM potassium acetate, 5 mM MgCl$_2$, 1 mM DTT, 0.05% Nikkol and 0.2 mg ml^{-1} 3× Flag peptide (Sigma Aldrich) for 30 min. Finally, the combined eluates were concentrated on 300-kDa molecular mass cut-off filters (Sartorius). The concentration was estimated on a NanoDrop photometer (Thermo Scientific) and the samples analysed by polyacrylamide gel electrophoresis on 4–12% Bis-Tris gradient gels (NuPAGE, Thermo Scientific), which were stained with SimplyBlue (Thermo Scientific) and subsequently cut into segments, which were individually analysed in-house by mass spectrometry.

Electron microscopy and image processing

Copper grids with holey carbon support films (R3/3, Quantifoil) and a 2 nm pre-coated continuous carbon layer were glow discharged at 2.1 × 10^{-1} mbar for 20 s. A 3.5 μl sample was applied to the grid in a Vitrobot Mark IV (FEI Company), blotted for 2 s after 45 s of incubation at 4 °C and plunge frozen in liquid ethane. Grids were screened for sample quality and data collection was then performed on a Titan Krios at 300 kV. In case of RIOK1(D324A), 7712 movies were collected with a nominal pixel size of 1.059 Å and defocus of 0.5–2.5 μm using a K2 Summit direct electron detector under low-dose conditions (48 frames at approximately 1 e$^-$ Å^{-2}) using EPU (FEI Company). For NOB1(D10N) and LRRC47, a Titan Krios, equipped with a Falcon II direct electron detector was used at 300 kV under low-dose conditions (10 frames at approximately 2.5 e$^-$ Å^{-2}) using EM-TOOLS (TVIPS). For NOB1(D10N), 14,397 and 14,548 movies were collected at a pixel size of 1.084 Å and 0.846 Å, respectively, at a defocus range of 1.0–3.0 μm (dataset 2 and 3). Gain-corrected movie frames were motion corrected and summed with MotionCor2[31] and contrast-transfer-function (CTF) parameters were determined with CTFFIND4 and Gctf[32]. After manual screening for image quality, 7,365 (RIOK1(D324A)), 14,391 and 12,948 (NOB1(D10N)) micrographs were used for automated particle picking by Gautomatch (https://www2.mrc-lmb.cam.ac.uk/research/locally-developed-software/zhang-software/#gauto). Totals of 1,822,874 (RIOK1(D324A)), 1,744,341 and 1,946,432 (NOB1(D10N)) particle images were then sorted using reference-free classification and 3D refinement in Relion 3.0[33,34] (see Extended Data Fig. 1). After sorting, particles from each state were separately subjected to Bayesian polishing and CTF refinement. Final reconstructions were then automatically B-factor sharpened with Relion and their local resolution estimated. In case of state F2, a local refinement with a mask covering the head of the 40S and RIOK1 (Extended Data Fig. 2d) was performed.

As micrographs of the two NOB1(D10N) datasets were recorded at a different magnification, and therefore a nominal pixel size, particle images had to be rescaled before combining them with those of the first dataset. To do this, particle images from both datasets were extracted without rescaling and their respective 3D reconstructions were compared. Using Chimera, the relative pixel size of the second dataset was determined as 0.846 Å. Then, particles from the second dataset were re-extracted in a box size of 492 and rescaled in Relion to 384, effectively changing the pixel size to 1.084 Å. Finally, the particle images were cut at a box size of 360 using the Relion image handler command and subsequently combined with the images from the other dataset.

In case of LRRC47, 10,083 movies were collected at a nominal pixel size of 1.084 Å, which were processed as described above and subsequently manually screened for image quality. Gautomatch was used to pick 967,247 particles, which were then 2D classified in Relion 3.1. However, owing to low sample and grid quality, as well as a strong preferential orientation, 3D refinement in Relion failed to provide a reliable result. The particles were therefore iteratively 3D classified in cryoSPARC[35] using the heterogenous refinement program, until 113,317 particles yielded a 6.5 Å reconstruction.

Model building and refinement

The molecular model of the mature human 40S (PDB: 6G5H) was used together with the crystal structure of RIOK1 (PDB: 4OTP) as a starting model for state H1. All ribosomal proteins and RNA were manually checked, missing parts were built and all residues real-space refined in Coot[36,37]. At the given resolution of 2.6 Å, known post-transcriptional modifications of rRNA that were identifiable (as listed in table 1 in ref. [38]) were added (Extended Data Fig. 2e, f). The model of RIOK1 was rigid body fitted into its density and subsequently manually extended by the N terminus, the flexible loop, as well as several C-terminal helices. The bound nucleotide was identified as ATP, coordinated by two magnesium ions. The NMR structure (PDB: 2DGY) was used to build EIF1AD, where C- and N-terminal segments had to be adjusted. Furthermore, a short stretch across uS13 and uS19 was identified as the C terminus of EIF1AD. Phyre 2[39] was used to generate homology models of LRRC47, which were then used as starting models. The N-terminal leucine-rich repeat motif was built in its entirety, while the C-terminal domain was less well resolved in all states, allowing only two helices to be properly built. Except for these helices, the homology model for the C-terminal domain was rigid-body fitted and trimmed to poly-alanine. The complete model for state H1 was real-space refined in Phenix[40] with the

parameter 'nonbonded weight' set to 500. The molecular model of state H2 was derived from H1 by deleting LRRC47, followed by real-space refinement in Phenix.

State G was built by a combination of the models of state H1 and sections of our previously published state C (PDB: 6G18). Ribosomal protein eS26 was replaced with a model of NOB1, where the PIN domain was manually rotated, the remaining sections adjusted to the density and the slightly relocated 3′ end and ITS1 was built. All residues not resolved in this state were removed and the model real-space refined in Phenix.

Finally, models of state F1 and F2 were based on the model of state H1, combined with segments of state C (PDB: 6G18), including the immature h44, as well as biogenesis factors PNO1, NOB1 and the relocated ITS1. EIF1AD was deleted from both models and RIOK1 rigid-body fitted into its respective density. All additional helices of RIOK1 not resolved in their states were removed and, in case of state F1, the additional β-strand and N-terminal stretch manually built. Owing to the highly flexible 40S head in state F2, the position of RIOK1 was adjusted in a locally refined map that covered the head region only. Both models were then real-space refined in Phenix.

Maps and models were visualized and figures created with the PyMOL Molecular Graphics System (v.2.3.2, Schrödinger) and ChimeraX[41].

Antibodies

Antibodies targeting PNO1 (dilution 1:2,000), NOB1 (1:5,000), ENP1 (1:15,000), RIOK1 (1:8,000) and TSR1 (1:1,000) have been previously described[9,25,42]. Anti-EIF1AD (20528-1-AP; immunofluorescence: 1:150; western blot: 1:1,000) and anti-LRRC47 (23217-1-AP; 1:1,500) were purchased from Proteintech, anti-beta-actin from Sigma Aldrich (A1987; 1:20,000; Extended Data Fig. 5a) and from Santa Cruz (sc-47778; 1:1,000; Extended Data Fig. 5b). Goat anti-mouse Alexa Fluor 488 (A-11001; 1:300) secondary antibody for immunofluorescence, goat anti-mouse Alexa Fluor Plus 680 (A32729; 1:10,000) and goat anti-rabbit Alexa Fluor Plus 800 (A32735; 1:10,000) secondary antibodies for western blotting were purchased from Invitrogen.

RNA-mediated interference

The siRNA oligos siCtrl (Allstars negative control, 1027281), siEIF1AD (5′-ACCGCAGACAGUAUCAUGAGA-3′) and siLRRC47 (5′-CCGCGUUUGU UUGGCCGGUUU-3′) were purchased from Qiagen and siRIOK2 (5′-GGAUCUUGGAUAUGUUUAA-3′) was purchased from Microsynth. Control, EIF1AD and LRRC47 siPOOLs (siP) were obtained from siTOOLs Biotech. Transfections of siRNAs into HeLa cells was performed with INTERFERin (Polyplus transfection), using siRNA oligonucleotides and siPOOLs at 10 and 5 nM final concentration, respectively. After 72 h of RNAi, cells were fixed with 4% PFA and analysed by immunofluorescence or FISH as previously described[25], or lysed in in SDS–PAGE sample buffer for western blot analysis. LMB was purchased from LC Laboratories.

Statistics and reproducibility

No statistical analysis has been applied to this work. All attempts at reproducing the purification of ribosomal subunit precursors were successful. Cellular assays were performed with at least of biological triplicates ($n = 3$).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank under the accession codes EMD-11517, EMD-11518, EMD-11519, EMD-11520 and EMD-11521. The atomic models have been deposited in the Protein Data Bank under accession codes 6ZXD, 6ZXE, 6ZXF, 6ZXG and 6ZXH.

30. Wyler, E. et al. Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *RNA* **17**, 189–200 (2011).
31. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
32. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
33. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
34. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
35. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
36. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
37. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. D* **71**, 136–153 (2015).
38. Taoka, M. et al. Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res.* **46**, 9289–9298 (2018).
39. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
40. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
41. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
42. Zemp, I. et al. CK1δ and CK1ε are components of human 40S subunit precursors required for cytoplasmic 40S maturation. *J. Cell Sci.* **127**, 1242–1253 (2014).
43. van Tran, N. et al. The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Res.* **47**, 7719–7733 (2019).

Extended Data Fig. 1 | Sample preparation and cryo-EM data analysis.
a, SDS−PAGE analysis of native pre-40S complexes purified with RIOK1(D324) and NOB1(D10N). Identified protein bands are labelled. For gel source data, see Supplementary Fig. 1. **b**, Representative micrographs from the three data sets. Scale bar 50 nm **c**, Subset of 2D averages of extracted particles after initial 2D classification. **d**, **e**, Summarized classification scheme of RIOK1(D324A) (**d**) and NOB1(D10N) (**e**). Particles of final states, marked in orange, were subjected to CTF parameter refinement and Bayesian polishing before the last 3D refinement. **f**, Particles of state G from both NOB1(D10N) data sets were combined and classified. Final volume is marked in orange. **g**, Summarized classification scheme of LRRC47.
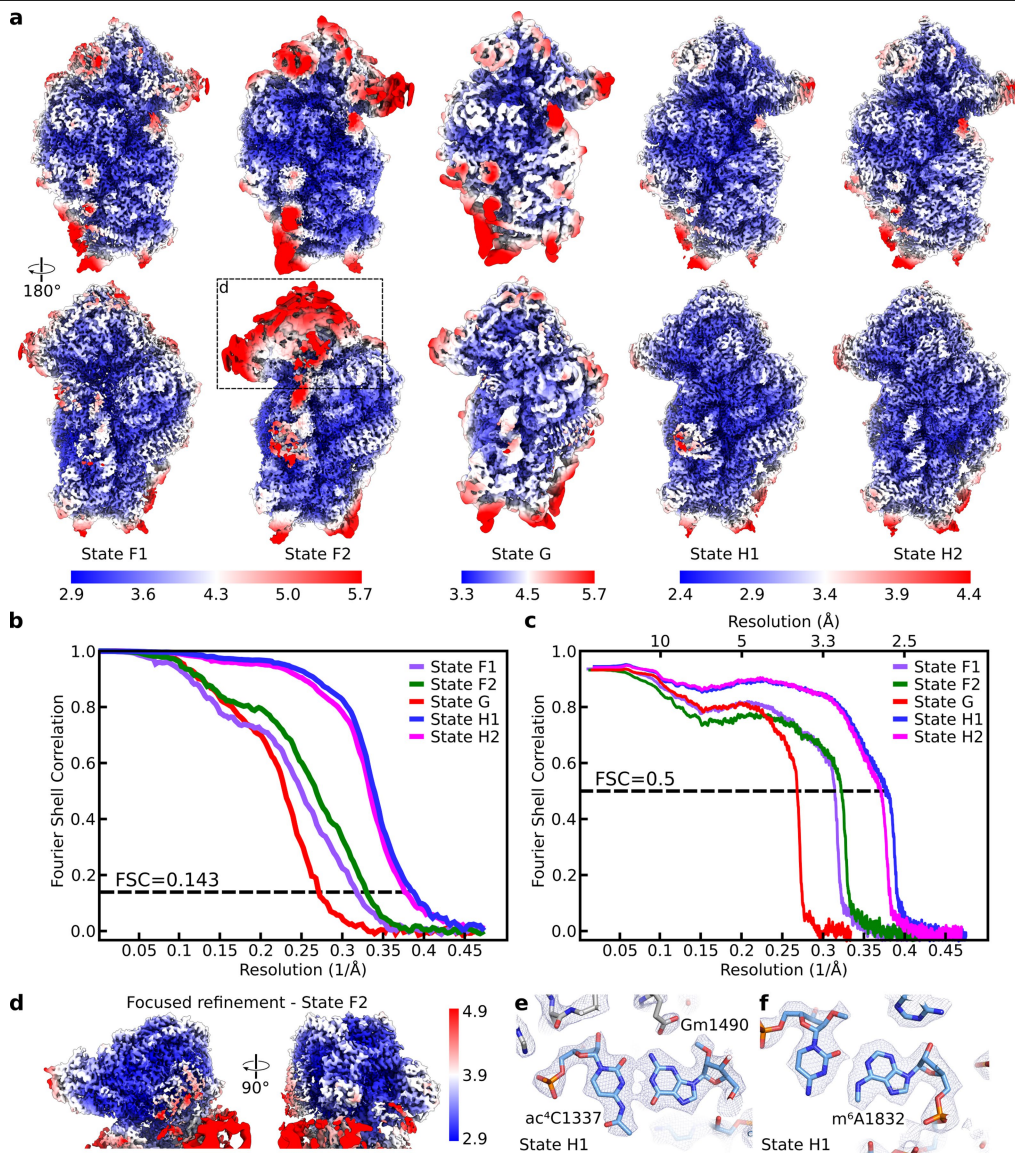
Extended Data Fig. 2 | Local resolution, refinement and model statistics.
a, Local resolution distribution of states F1 – H2 with their respective colour grading scheme as estimated by Relion. **b**, Fourier shell correlation (FSC) curve for all states. Average resolution values as stated in Fig. 1 are calculated according to the 'gold standard' at FSC = 0.143. **c**, FSC plot of the models against their volume as provided by Phenix. **d**, Local resolution estimation of the 'head' region of state F2 after focused refinement. **e**, **f**, Model and cryo-EM density of state H1 around the post-transcriptionally modified rRNA residues C1337, G1490 (**e**), and A1832 (see ref. [43].) (**f**).
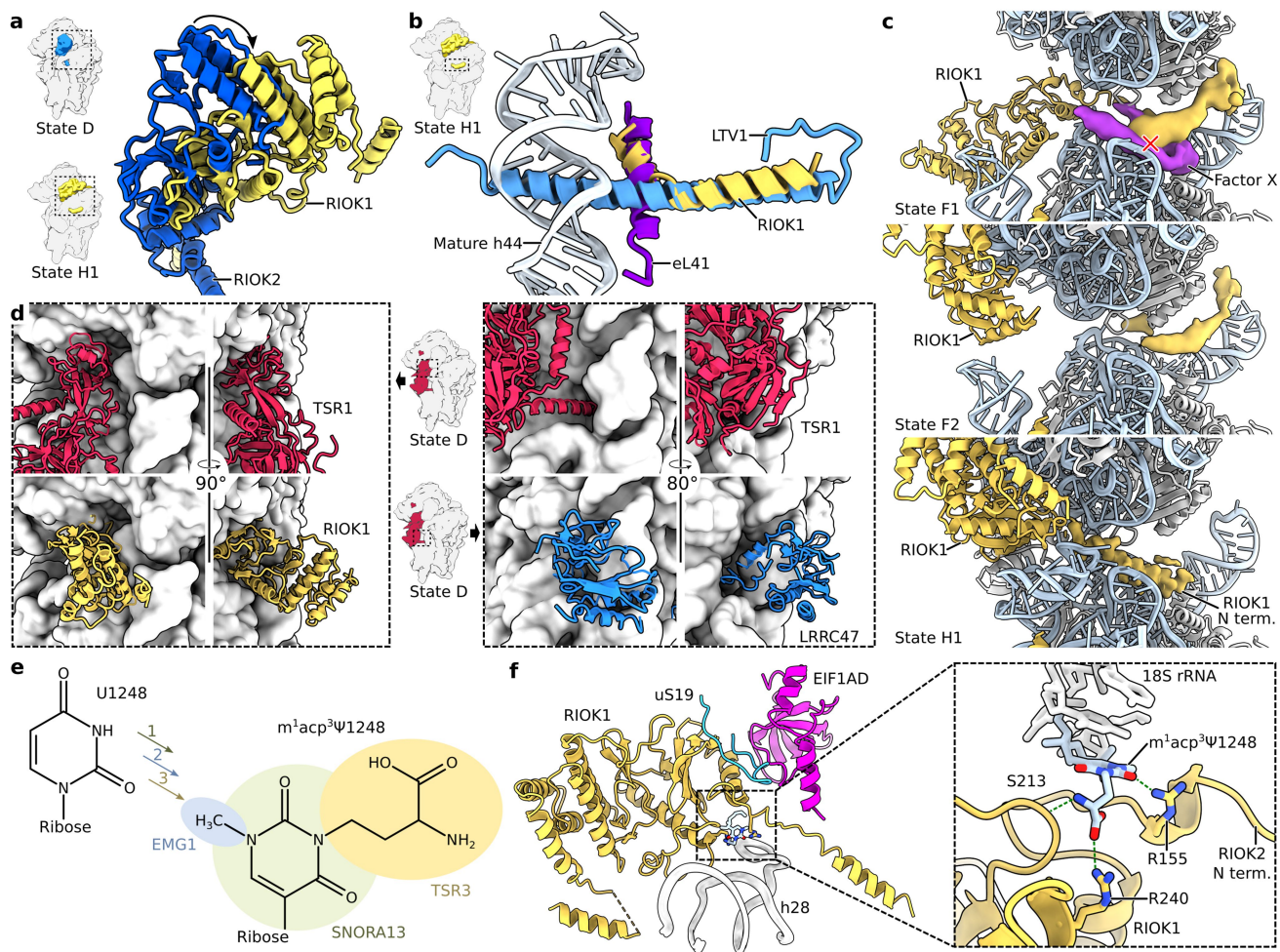
Extended Data Fig. 3 | Details on RIOK1 binding to the pre-40S particle.
a, Superposition of models of RIOK1 and RIOK2 after alignment of state D (PDB-6G51) and H1 highlight the overlapping binding site at the decoding centre and the rotation of their central RIO domain by approximately 15°.
b, Cartoon representation of eL41 (PDB-6G5H), as well as the C terminus of LTV1 (PDB-6G51) and RIOK1 relative to the top of the matured h44 in state H1. Overlaps in binding sites highlight mutually exclusive binding. **c**, Models of state F1, F2 and H1 with focus on RIOK1 position. Unaccounted density within the mRNA entry tunnel in states F1 and F2 (yellow) is likely the flexible N terminus of RIOK1, which overlaps with the two helices of factor X present in state E (marked with a red cross, see ref. [4].). **d**, Surface representation of state D (PDB-6G51) with RBF TSR1 in cartoon representation. Models of RIOK1 (left) and LRRC47 (right) of state F1 after alignment of the particles emphasize the overlaps in binding sites around h44. **e**, Post-transcriptional modification of U1248 leads to formation of 1-methyl-3-α-amino-α-carboxypropyl pseudouridine ($m^1acp^3\Psi$). Modifying enzymes and their contribution to the structure are indicated by colours. **f**, Cartoon representation showing the coordination of $m^1acp^3\Psi1248$ by RIOK1.
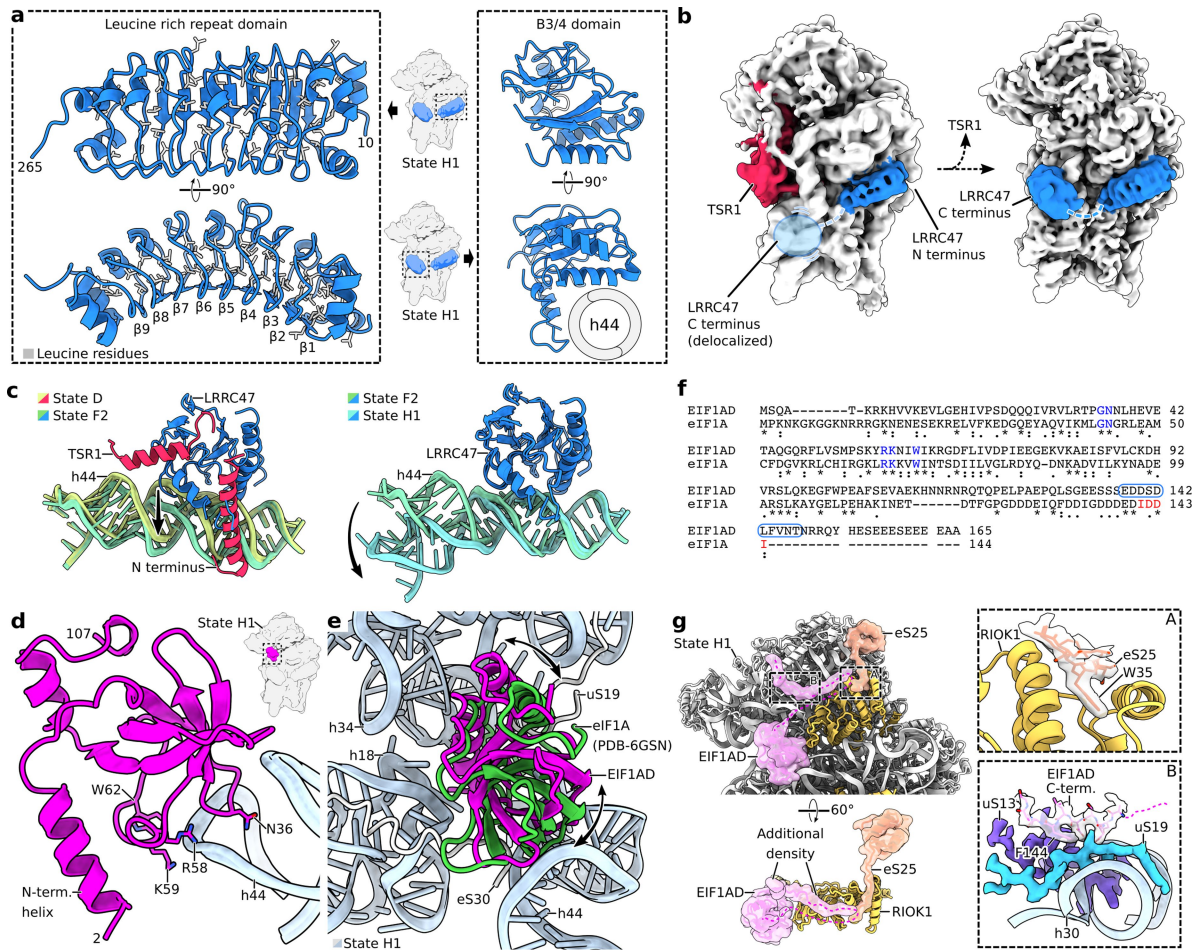
Extended Data Fig. 4 | Structural details of novel factors LRRC47 and EIF1AD. a, Overall structure of the two domains of LRRC47. Leucine residues, secondary structure elements and position of h44 are highlighted. **b**, Low-resolution cryo-EM reconstruction of a sample using N-terminally tagged LRRC47 (left) and state H1 filtered at 7 Å (right). The leucine-rich domain of LRRC47 (blue) binds simultaneously with TSR1, while the C-terminal domain remains delocalized and is therefore not visible. **c**, Models of LRRC47, h44 and parts of TSR1 show the conformational changes in h44 that accompany the transition between the states. A central part of h44 moves after release of TSR1, enabled by the lack of its N-terminal helix (left). The B3/4 domain of LRRC47 would clash with both displayed helices of TSR1. LRRC47 continues to bind in an almost unchanged position after maturation of h44 (right). **d**, Structure of

EIF1AD with its N-terminal helix and residues N36, R58, K59 and W62 labelled. **e**, Cartoon representation of rRNA segments and ribosomal proteins surrounding EIF1AD in state H1. Model of yeast eIF1A (PDB-6GSN) after alignment of a pre-48S translation initiation complex to the pre-40S particle shows a shifted binding location. **f**, Sequence alignment of human EIF1AD and eIF1A. Conserved residues that bind to rRNA are coloured blue and the IDDI motif of eIF1A in red. The conserved C-terminal stretch that binds to uS13 and uS19 is marked with a blue box. **g**, Model of state H1 with cryo-EM volume of EIF1AD and eS25 Gaussian filtered at 1.5 standard deviations. Additional density extends from well resolved parts of the C terminus of EIF1AD. Detailed views on eS25 N terminus and EIF1AD C terminus are shown in boxes A and B.
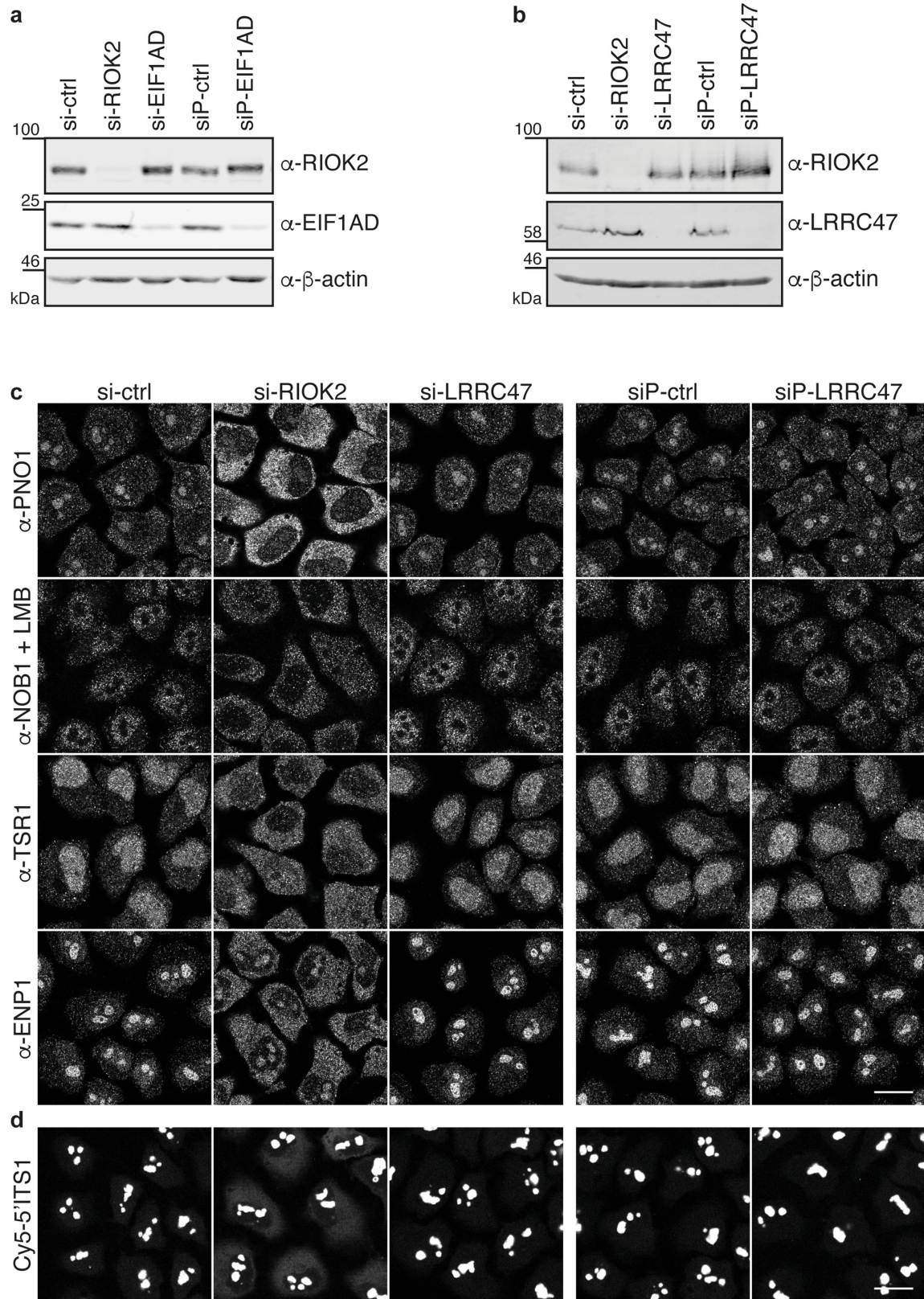
Extended Data Fig. 5 | EIF1AD but not LRRC47 depletion affects late 40S subunit maturation. a, Western blot analysis of the experiment shown in Fig. 4 confirming the effectiveness of siRNA treatments for EIF1AD and RIOK2. **b**, Western blot analysis for the experiment in **c** confirming the depletion of LRRC47 or RIOK2 upon siRNA treatment. For gel source data of **a** and **b**, see Supplementary Fig. 1. **c**, Immunofluorescence analysis of HeLa cells treated with siRNAs against LRRC47 or RIOK2 using antibodies against the indicated RBFs. For immunofluorescence analysis of NOB1, cells were treated with 20 nM leptomycin B (LMB) for 90 min. Note that only RIOK2 but not LRRC47 depletion leads to cytoplasmic recycling defects of the tested RBFs. Scale bar, 20 μm. **d**, FISH analysis of experiment in **b**, revealing cytoplasmic accumulation of 18S-E pre-rRNA upon RIOK2 but not LRRC47 depletion. FISH pictures were processed in parallel, using a gamma correction of 1.5. All experiments were done in triplicates ($n = 3$).
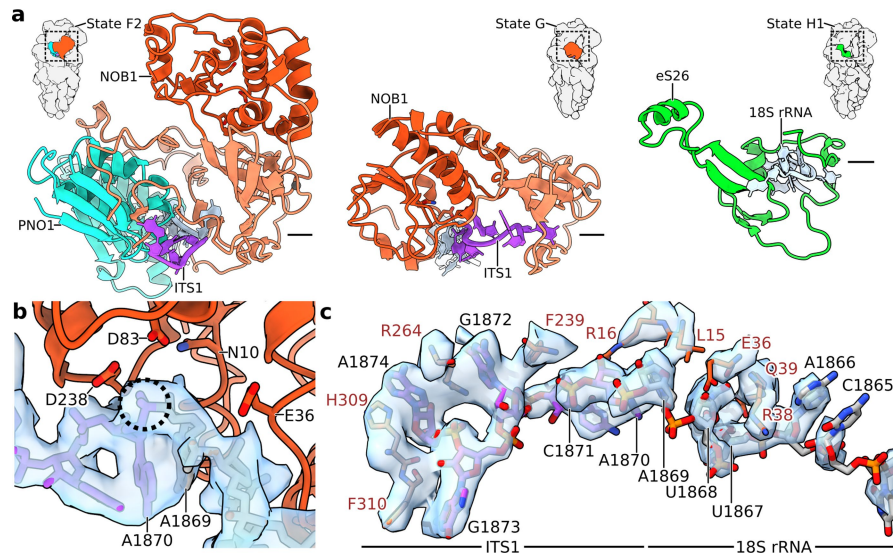
Extended Data Fig. 6 | Structural details of rRNA 3′ end maturation.
a, Cartoon representation of the 3′ end of 18S-E rRNA with PNO1, NOB1 and eS26 throughout the maturation process in states F2, G and H1. Panel of eS26 with 18S rRNA in state H1 has been shifted slightly upwards as indicated by the line to the right. **b**, Detailed view of NOB1(D10N) active site with its substrate. Dashed circle marks site 3 cleavage site. Electron density around ITS1 and the 3′end shown in blue. **c**, Stick representation of NOB1 residues that interact with the ITS1 and 3′ end with electron density shown in blue.

Article

Extended Data Table 1 | Data collection and refinement statistics

	State F1 (EMDB-11517) (PDB-6ZXD)	State F2 (EMDB-11518) (PDB-6ZXE)	State G (EMDB-11519) (PDB-6ZXF)	State H1 (EMDB-11520) (PDB-6ZXG)	State H2 (EMDB-11521) (PDB-6ZXH)
Data collection and processing					
Camera	Gatan K2 Summit	Gatan K2 Summit	Falcon II	Gatan K2 Summit	Gatan K2 Summit
Magnification	130,000	130,000	75,000/ 96,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e⁻Å⁻²)	48	48	25	48	48
Defocus range (μm)	0.5-2.5	0.5-2.5	1.0-3.0	0.5-2.5	0.5-2.5
Pixel size (Å)	1.059	1.059	1.084/0.846	1.059	1.059
Symmetry imposed	C1	C1	C1	C1	C1
Micrographs collected (no.)	7365	7365	14,397/14,548	7365	7365
Initial particle images (no.)	1,822,874	1,822,874	1,744,341/ 1,946,432	1,822,874	1,822,874
Final particle images (no.)	16,876	32,535	140,612	121,468	144,738
Map resolution (Å)	3.2	3.0	3.7	2.6	2.7
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.9 – 8.5	2.8 – 8.3	3.3 – 7.9	2.3 – 5.4	2.4 – 5.8
Refinement					
Initial model used (PDB code)	6G18/ 6G5H	6G18/ 6G5H	6G18/ 6G5H	6G5H	6G5H
Model resolution (Å)	3.2	3.1	3.7	2.7	2.7
FSC threshold	0.5	0.5	0.5	0.5	0.5
Model resolution range (Å)	2.9 – 8.5	2.8 – 8.3	3.3 – 7.9	2.3 – 5.4	2.4 – 5.8
Map sharpening B factor (Å²)	-38.1	-61.1	-158.6	-37.3	-37.7
Model composition					
Non-hydrogen atoms	80,922	80,500	80,134	79,866	76,980
Protein residues	5,944	5,903	5,758	5,625	5,195
Nucleotide residues	1,645	1,645	1,656	1,648	1,648
Ligands	2	2	6	148	138
R.m.s deviations					
Bond lengths (Å)	0.007	0.010	0.010	0.010	0.013
Bond angles (°)	1.101	1.213	1.127	1.082	1.504
Validation					
Molprobity score	1.78	1.87	1.96	1.40	1.45
Clash score	5.74	7.03	8.43	3.35	3.77
Poor rotamers (%)	0.21	0.54	0.65	0.23	0.45
Ramachandran plot					
Favored (%)	92.59	92.10	91.48	96.06	95.85
Allowed (%)	7.20	7.78	8.49	3.92	4.11
Disallowed (%)	0.21	0.12	0.04	0.02	0.04
Map vs. Model CC	0.79	0.76	0.82	0.83	0.83

nature research

Corresponding author(s):	Prof. Dr. Beckmann
Last updated by author(s):	Aug 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	EPU, EM-Tools
Data analysis	MotionCor2 v1.2.6 CTFFIND4 v4.1.13 Gctf v1.06 Relion 3.0 and 3.1 Gautomatch v0.56 Phenix 1.12 (includes MolProbity) Pymol 2.3.2 Coot 0.8.9 ChimeraX v1.0 Phyre 2 (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) cryoSPARC v2.15.0

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All cryo-EM density maps have been deposited in the EM Data Bank under the accession codes EMD-11517, EMD-11518, EMD-11519, EMD-11520, and EMD-11521. The atomic models have been deposited in the Protein Data Bank under PDB accession codes 6ZXD, 6ZXE, 6ZXF, 6ZXG, and 6ZXH.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were applied to pre-determine sample sizes. Number of images obtained for cryo-EM analysis of native pre-40S complexes were determined by the limited access to the microscope.
Data exclusions	No data were excluded from the analysis. During cryo-EM data processing, bad particle images which did not align well with the reference structures were removed from further consideration as described in the Method section and shown in Extended Data Fig. 1.
Replication	All immunofluorescence and FISH experiments were performed at least in biological triplicates and all attempts at replicating the purification of native 40S precursors were successful. Cryo-EM data collection and analysis has not been repeated, because each individual structure already represents the average of several thousand molecules.
Randomization	No randomization was required for the reported experiments, as all variables could be controlled.
Blinding	No blinding was required for the reported experiments and therefore not attempted.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies targeting PNO1 (1:2,000), NOB1 (1:5,000), ENP1 (1:15,000), RIOK1 (1:8,000) and TSR1 (1:1,000) have been previously described (see Ref. 10,26,44) Anti-eIF1AD (20528-1-AP; IF: 1:150; WB: 1:1,000) and anti-LRRC47 (23217-1-AP; 1:1,500) were purchased from Proteintech, anti-beta-actin from Sigma Aldrich (A1987; 1:20,000; Extended Data Fig. 5a) and from Santa Cruz (sc-47778; 1:1,000; Extended Data Fig. 5b). Goat anti-mouse Alexa Fluor 488 (A-11001; 1:300) secondary antibody for immunofluorescence, goat anti-mouse Alexa Fluor Plus 680 (A32729; 1:10,000) and goat anti-rabbit Alexa Fluor Plus 800 (A32735; 1:10,000) secondary antibodies for Western blotting were purchased from Invitrogen.
Validation	Antibodies used in this study are commercially available and have been validated by the manufacturers for the respective species and experiments applied to in this work. Additional information about validation of each antibody is provided by the manufacturer.

Eukaryotic cell lines

Cell line source(s)	HEK293 T-REx Flp-In: Thermo Fischer Scientific; HeLa Kyoto: D. Gerlich (IMBA, Vienna)
Authentication	Cells were not authenticated
Mycoplasma contamination	All cell lines used in this study have tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No cell lines used throughout the work are listed in the ICLAC database

Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2

Matthias Thoms[1]*, Robert Buschauer[1]*, Michael Ameismeier[1]*, Lennart Koepke[2], Timo Denk[1], Maximilian Hirschenberger[2], Hanna Kratzat[1], Manuel Hayn[2], Timur Mackens-Kiani[1], Jingdong Cheng[1], Jan H. Straub[2], Christina M. Stürzel[2], Thomas Fröhlich[3], Otto Berninghausen[1], Thomas Becker[1], Frank Kirchhoff[2], Konstantin M. J. Sparrer[2]†, Roland Beckmann[1]†

[1]Gene Center Munich, Department of Biochemistry, University of Munich, Munich, Germany. [2]Institute of Molecular Virology, Ulm University Medical Center, Ulm, Germany. [3]Laboratory of Functional Genome Analysis, University of Munich, Munich, Germany.

*These authors contributed equally to this work. †Corresponding author. Email: beckmann@genzentrum.lmu.de (R.B.); konstantin.sparrer@uni-ulm.de (K.M.J.S.)

SARS-CoV-2 is the causative agent of the current COVID-19 pandemic. A major virulence factor of SARS-CoVs is the nonstructural protein 1 (Nsp1) which suppresses host gene expression by ribosome association. Here, we show that Nsp1 from SARS-CoV-2 binds to the 40S ribosomal subunit, resulting in shutdown of mRNA translation both in vitro and in cells. Structural analysis by cryo-electron microscopy (cryo-EM) of in vitro reconstituted Nsp1-40S and various native Nsp1-40S and -80S complexes revealed that the Nsp1 C terminus binds to and obstructs the mRNA entry tunnel. Thereby, Nsp1 effectively blocks RIG-I-dependent innate immune responses that would otherwise facilitate clearance of the infection. Thus, the structural characterization of the inhibitory mechanism of Nsp1 may aid structure-based drug design against SARS-CoV-2.

Coronaviruses (CoVs) are enveloped, single-stranded viruses with a positive-sense RNA genome, which infect a large variety of vertebrate animal species. Currently, seven CoV species from two genera (alpha and beta) are known human pathogens, four of which usually cause only mild respiratory diseases like common colds (*1–5*). Over the last two decades, however, three *Betacoronaviruses* (beta-CoVs) – the severe acute respiratory syndrome-coronavirus (SARS-CoV), the Middle East respiratory syndrome-coronavirus (MERS-CoV) and the novel severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) – have emerged as the causative agents of epidemic and in the case of SARS-CoV-2 pandemic outbreaks of highly pathogenic respiratory diseases. COVID-19, the disease caused by SARS-CoV-2 has affected millions of people with a death toll amounting to hundreds of thousands worldwide (*6, 7*).

Coronavirus particles contain a single, 5′-capped and 3′-poly-adenylated RNA genome, which codes for two large overlapping open reading frames in gene 1 (ORF1a and ORF1b), as well as a variety of structural and nonstructural proteins at the 3′ end (*8, 9*). Following host infection, precursor proteins ORF1a and ORF1ab are translated and subsequently proteolytically cleaved into functional proteins, most of which play roles during viral replication (*10*). Amongst them is the N-terminal nonstructural protein 1 (Nsp1). Despite differences in protein size and mode of action, Nsp1 proteins from alpha- and beta-CoVs display a similar biological function in suppressing host gene expression (*11–14*). SARS-

CoV Nsp1 induces a near-complete shutdown of host protein translation by a two-pronged strategy: first, it binds the small ribosomal subunit and stalls canonical mRNA translation at various stages during initiation (*15, 16*). Second, Nsp1 binding to the ribosome leads to endonucleolytic cleavage and subsequent degradation of host mRNAs. Notably, interactions between Nsp1 and a conserved region in the 5′ untranslated region (UTR) of viral mRNA prevent shutdown of viral protein expression through an unknown mechanism (*17*). Taken together, Nsp1 inhibits all cellular anti-viral defense mechanisms that depend on the expression of host factors, including the interferon response. This shutdown of the key parts of the innate immune system may facilitate efficient viral replication (*13, 18*) and immune evasion. Its central role in weakening the anti-viral immune response makes SARS-CoV Nsp1 a potential therapeutic target (*19, 20*). Here, we set out to characterize the interaction of Nsp1 of SARS-CoV-2 with the human translation machinery.

Nsp1 of SARS-CoV-2 shows 84% amino acid sequence identity with SARS-CoV, suggesting similar properties and biological functions (Fig. 1A). The C-terminal residues K164 and H165 in SARS-CoV are conserved in beta-CoVs and essential for 40S interaction since mutations to alanine abolish 40S binding and relieve translational inhibition (*16*). To confirm an analogous function of Nsp1 from SARS-CoV-2, we expressed and purified recombinant Nsp1 and the K164A/H165A mutant (Nsp1-mt) of both SARS-CoV and SARS-CoV-2 from *E. coli*, and tested their binding efficiency

to purified human ribosomal subunits (Fig. 1B and fig. S1A). Nsp1 from both CoVs associated strongly with 40S subunits but not with 60S subunits, whereas both Nsp1-mt constructs showed no binding (Fig. 1B). Thus, ribosome binding to the 40S subunit is preserved and residues K164 and H165 of Nsp1 from both SARS-CoVs are important for this ribosome interaction. To further verify this, we expressed wildtype or mutant Nsp1 constructs in human HEK293T cells and analyzed ribosome association by sucrose-gradient centrifugation. Consistent with the behavior in vitro, Nsp1 of CoV and CoV-2 co-migrated with 40S ribosomal subunits and 80S ribosomes, but not with actively translating polyribosomes. In contrast, the mutant constructs barely penetrated the gradient, indicative of their loss of affinity for ribosomes (Fig. 1C). Notably, compared to the control the polysome profiles showed a shift from translating polyribosomes to 80S monosomes in the presence of Nsp1, indicating global inhibition of translation. This effect was less pronounced for the two Nsp1-mt constructs. Next, we performed in vitro translation assays of capped reporter mRNA in cell free translation extracts from human cells (HeLa S3) or rabbit reticulocytes in the presence of Nsp1 or Nsp1-mt. Probing for the translation products by Western blotting revealed a complete inhibition of translation by Nsp1 and only weak effects in the presence of Nsp1-mt constructs (Fig. 1D and fig. S1B). To test the inhibitory effect of Nsp1 on translation in cells, we expressed 3xFLAG-tagged Nsp1 of SARS-CoV-2 and SARS-CoV and their respective mutants in HEK293T cells and monitored translation of a co-transfected capped luciferase reporter mRNA. Consistent with the results of the in vitro assays, we observed a strong reduction of translation in presence of Nsp1 from SCoV-1 or -2, but not of the respective Nsp1-mt constructs (Fig. 1E). This phenotype was confirmed for differently tagged (V5) and codon-optimized versions of SCoV-2 Nsp1 (fig. S1, C and D). Nsp7, which is derived from the same polyprotein precursor as Nsp1 had no effect on translation (fig. S1C). In summary, Nsp1 from both, SARS-CoV and SARS-CoV-2 binds 40S and 80S ribosomes and disrupts cap-dependent translation. Moreover, the conserved KH motif close to the C terminus of Nsp1 is crucial for ribosome binding and translation inhibition.

To elucidate the molecular interaction of SARS-CoV-2 Nsp1 with human ribosomes, we reconstituted a complex from purified, recombinant Nsp1 and purified human 40S ribosomal subunits and determined its structure by cryo-EM at an average resolution of 2.6 Å (Fig. 2, A and B, and figs. S2 and S3). In addition to the 40S ribosomal subunit, we observed density corresponding to two α-helices inside the ribosomal mRNA entry channel, which could be unambiguously identified as the C-terminal part of Nsp1 from SARS-CoV-2 (Fig. 2C). In proximity to the helical density, we observed undefined globular density between rRNA helix h16 and

ribosomal proteins uS3 and uS10. The dimensions of this extra density roughly match the putative dimensions of the globular N-terminal domain of Nsp1 (Fig. 2, C and D), based on a structure of the highly similar N terminus of Nsp1 from SARS-CoV, previously determined by NMR (21). However, the resolution of this region in our cryo-EM density map was insufficient for unambiguous identification. The C terminus of Nsp1 is located close to the so called "latch" between rRNA helix h18 of the body and h34 of the head of the 40S subunit, which influences mRNA accommodation and movement during translation initiation (22, 23). When bound at this position, the Nsp1 C terminus blocks regular mRNA accommodation, thus providing an explanation for Nsp1 mediated host translation shutdown (Fig. 2D).

To characterize the ribosomal targets and the mode of interaction of Nsp1 in human cells, we expressed N-terminally 3xFLAG tagged Nsp1 in HEK293T cells and affinity purified associated native complexes for analysis by cryo-EM and mass spectrometry (Fig. 2E, figs. S2 and S3, and data S1). Structural analysis revealed 40S and 80S ribosomal complexes in nine compositionally different states (Fig. 2, F to N). Importantly, all of them displayed density for the Nsp1 C terminus in an identical position and conformation observed in the in vitro assembled complex, and all complexes lacked density corresponding to mRNA.

The Nsp1-bound 40S ribosomal complexes could be divided into three major populations. The first represents idle Nsp1-40S complexes (Fig. 2F), essentially resembling the in vitro reconstituted complex. The second population comprises unusual, pre-40S-like complexes (Fig. 2, G and H), in which the cytosolic ribosome biogenesis factor TSR1 is bound in two distinct conformations between the 40S head and body (24, 25). Notably, these complexes do not resemble any known on-pathway biogenesis intermediates. The third population represents eIF3-containing 43S pre-initiation complexes (PICs), and could be further divided into PICs with and without eIF1A, eIF1 and a fully assembled eIF2-tRNA$_i$-GTP complex (Fig. 2, I and J) (26–28). Both PICs adopt the previously observed open conformation (28). The stable association of Nsp1 in the cell with multiple different intermediates states of translation initiation besides empty 40S ribosomal complexes is in agreement with the proposed role of Nsp1 as an inhibitor of translation initiation (15).

The Nsp1-bound 80S complexes could be divided into two major populations of translationally inactive ribosomes. The first population (Fig. 2, K and L, and fig. S4, A to E) contained the protein coiled-coil domain containing short open reading frame 124 (CCDC124), a homolog of the ribosome protection and translation recovery factor Lso2 in *Saccharomyces cerevisiae* (29). A similar complex of inactive 80S ribosomes bound to CCDC124 was recently described (30). In addition to the known hibernation complex, a subpopulation of the

CCDC124-bound 80S contained also the ribosome recycling factor and ABC-type ATPase ABCE1 (*31–33*) and the class I translation termination factor eRF1 in an unusual conformation (fig. S4, C to E). The previously unresolved, flexible C-terminal part of CCDC124 was stably bound to the ribosomal A-site in this complex. This sub-population might represent a previously unidentified ribosome recycling-like state.

The second major population of Nsp1-bound 80S ribosomes (Fig. 2, M and N) lacked CCDC124, but contained the cell growth regulating nucleolar protein LYAR, which has been implicated in processing of pre-rRNA and in negative regulation of antiviral innate immune responses (*34, 35*). We found the C terminus of LYAR occupying the ribosomal A-site, similar to CCDC124 (Fig. 2M and fig. S4, F and G). Furthermore, we identified a subpopulation among the LYAR-bound inactive 80S ribosomes that contained a ternary eEF1A-GTP-tRNA complex (Fig. 2N and fig. S4, H to K). This ternary complex was in an unusual conformation, with the anticodon loop contacting an α-helix of the LYAR C terminus. Such a complex has not been previously described and its functional relevance is unknown.

Taken together, we found Nsp1 bound to the mRNA entry channel of a unique set of translationally inactive 80S ribosomes, among which were unusual complexes. It is unclear, whether these are a result of the presence of Nsp1, or whether they occur naturally and have an increased affinity for Nsp1 due to their distinct conformation or lack of mRNA.

All observed ribosomal complexes displayed the same binding mode of Nsp1 to the 40S subunit, in which the C-terminal domain of Nsp1 (Nsp1-C) is rigidly bound inside the mRNA entry channel. Here, it interacts with the rRNA helix h18, with the ribosomal protein uS5 of the 40S body and with uS3 of the 40S head. The local resolution of 2.6 Å (fig. S3) allowed for a detailed analysis of the molecular interactions of Nsp1 with the ribosome (Fig. 3A).

The shorter, first α-helix of Nsp1-C (α1; residues 154-160) interacts with uS3 and uS5. The helix is followed by a short loop, which contains the essential KH motif that interacts with h18. Notably, this part of h18 belongs to the so-called "530-loop", which actively participates in ribosomal decoding and has been reported to resemble a conserved structural motif in the 3'-UTR of beta-CoVs (*36*). The second, larger α-helix of Nsp1-C (α2; residues 166-179) also interacts with rRNA h18 and connects back to uS5 at its C-terminal end. The two helices stabilize each other through hydrophobic interactions. The electrostatic potential on the Nsp1-C surface displays three major patches (Fig. 3B). A negatively charged patch on α1 facing positively charged residues on uS3, a positively charged patch on α2 facing the phosphate backbone of h18, and a hydrophobic patch at the α1-α2 interface, which is exposed to hydrophobic residues on uS5. In addition to the matching surface charge, the shape of Nsp1-C matches the

shape of the mRNA channel and completely overlaps the regular mRNA path (Fig. 3, C and D). Together, this explains the strong inhibitory effect on translation observed in vitro and in vivo. A key interaction is established through the KH motif, which binds to a distinct site on rRNA helix h18 (Fig. 3, C and E); K164 of Nsp1 inserts into a negatively charged pocket, constituted mainly by the phosphate backbone of rRNA bases G625 and U630, whereas H165 stacks in between U607 and U630. The base U630 is stabilized in this position through interaction with the backbone of G168 of Nsp1. Further interactions involve R171 and R175 of Nsp1, which form salt bridges to the backbone phosphates of G601, C607, A605 and G606 of h18 (Fig. 3F). The interactions of Nsp1-C and uS3 are established through salt bridges and hydrogen bonds between D152, E155, E159 of Nsp1 and R116, R143 and M150 of uS3 (Fig. 3G). The interactions of Nsp1-C with uS5 occur within a hydrophobic surface of ~440 Å^2 involving residues Y154, F157, W161, T170, L173, M174, L177 of Nsp1 and residues V106, I109, P111, T122, F124, V147, I151 of uS5 (Fig. 3H). Taken together, specific molecular contacts (summarized in Fig. 3I) rigidly anchor Nsp1 and thereby obstruct the mRNA entry channel.

Type-I interferon induction and signaling represents one of the major innate anti-viral defense pathways, ultimately leading to the induction of several hundred anti-viral interferon-stimulated genes (ISGs) (*37*). Coronavirus infections are sensed by RIG-I which activates this defense system (*37, 38*). To assess the effects of SARS-CoV-2 Nsp1 on the interferon system, we stimulated HEK293T cells with Sendai Virus (SeV), a well-known trigger of RIG-I-dependent signaling (*39, 40*). Expression of Nsp1 completely abrogated the translation of Firefly Luciferase controlled by human interferon-beta (IFN-β) promoter, whereas the Nsp1-mt had no significant effect (Fig. 4A and fig. S5A), confirming the results of the in vitro translation assays. Rabies virus P protein (*41*) and SARS-CoV-2 Nsp7 were used as positive and negative control, respectively. After stimulation with SeV, the protein levels of endogenous IFN-β, IFN-λ1 and interleukin-8 (IL-8) (Fig. 4B and fig. S5, B and C) in the supernatant of Nsp1 expressing cells were drastically reduced, although transcription of the corresponding mRNAs was induced. Again, Nsp1-mt showed no inhibitory effect. Expression of Luciferase driven by the interferon stimulated response element (ISRE), which is part of the promoter of most ISGs, was effectively shut down by Nsp1 but not by Nsp1-mt in a dose-dependent manner (Fig. 4, C and D, and fig. S5D). SARS CoV-2 Nsp7 and Measles virus V protein (MeV V) (*40, 42*) served as negative and positive controls, respectively. In line with this, Nsp1 but not Nsp1-mt suppressed the induction of endogenous RIG-I and ISG15 upon IFN-β stimulation on the protein but not the mRNA level (Fig. 4E).

Notably, not all innate immune responses require active

translation for function. For example autophagy is barely affected by the expression of Nsp1 or its mutant (fig. S5E) even upon induction with Rapamycin (*43*). Tripartite Motif Protein 32 (TRIM32) was used as a positive control (*44*). Taken together, these results demonstrate that SARS-CoV-2 Nsp1 almost completely prevents translation not only of interferons and other pro-inflammatory cytokines but also of interferon-stimulated anti-viral ISGs.

Our data establish that one of the major immune evasion factors of SARS-CoV-2, Nsp1, efficiently interferes with the cellular translation machinery resulting in a shut-down of host protein production. Thus, major parts of the innate immune system, that depend on translation of antiviral defense factors such as IFN-β or RIG-I (*45*) are disarmed. Although SARS-CoV-2 encodes additional potential inhibitors of the innate immune defenses, a loss of Nsp1 function may render the virus vulnerable toward immune clearance. Thus, our data may provide a starting point for rational structure-based drug design, targeting the Nsp1-ribosome interaction.

However, important questions remain to be addressed. For example, how can the virus overcome the Nsp1-mediated block of translation for the production of its own viral proteins? Common structural features present in the 5′ untranslated region of all SARS-CoV mRNAs, may help to circumvent the ribosome blockage by Nsp1 (*46*).

REFERENCES AND NOTES

1. S. R. Weiss, S. Navas-Martin, Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol. Mol. Biol. Rev.* **69**, 635–664 (2005). doi:10.1128/MMBR.69.4.635-664.2005 Medline

2. P. C. Woo, S. K. P. Lau, C. S. F. Lam, C. C. Y. Lau, A. K. L. Tsang, J. H. N. Lau, R. Bai, J. L. L. Teng, C. C. C. Tsang, M. Wang, B.-J. Zheng, K.-H. Chan, K.-Y. Yuen, Discovery of seven novel mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. *J. Virol.* **86**, 3995–4008 (2012). doi:10.1128/ JVI.06540-11 Medline

3. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019). doi:10.1038/s41579-018-0118-9 Medline

4. L. van der Hoek, Human coronaviruses: What do they cause? *Antivir. Ther.* **12**, 651–658 (2007). Medline

5. L. van der Hoek, K. Pyrc, B. Berkhout, Human coronavirus NL63, a new respiratory virus. *FEMS Microbiol. Rev.* **30**, 760–773 (2006). doi:10.1111/j.1574-6976.2006.00032.x Medline

6. E. de Wit, N. van Doremalen, D. Falzarano, V. J. Munster, SARS and MERS: Recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016). doi:10.1038/nrmicro.2016.81 Medline

7. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020). doi:10.1001/jama.2020.1585 Medline

8. Y. X. Lim, Y. L. Ng, J. P. Tam, D. X. Liu, Human coronaviruses: A review of virus-host interactions. *Diseases* **4**, 26 (2016). doi:10.3390/diseases4030026 Medline

9. P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020). doi:10.1038/s41586-020-2012-7 Medline

10. P. S. Masters, The molecular biology of coronaviruses. *Adv. Virus Res.* **66**, 193–292 (2006). doi:10.1016/S0065-3527(06)66005-3 Medline

11. L. Lei, S. Ying, L. Baojun, Y. Yi, H. Xiang, S. Wenli, S. Zounan, G. Deyin, Z. Qingyu, L. Jingmei, C. Guohui, Attenuation of mouse hepatitis virus by deletion of the LLRKxGxKG region of Nsp1. *PLOS ONE* **8**, e61166 (2013). doi:10.1371/journal.pone.0061166 Medline

12. Y. Tohya, K. Narayanan, W. Kamitani, C. Huang, K. Lokugamage, S. Makino, Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. *J. Virol.* **83**, 5282–5288 (2009). doi:10.1128/ JVI.02485-08 Medline

13. K. Narayanan, C. Huang, K. Lokugamage, W. Kamitani, T. Ikegami, C.-T. K. Tseng, S. Makino, Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* **82**, 4471–4479 (2008). doi:10.1128/JVI.02472-07 Medline

14. C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, S. Makino, Alphacoronavirus transmissible gastroenteritis virus nsp1 protein suppresses protein translation in mammalian cells and in cell-free HeLa cell extracts but not in rabbit reticulocyte lysate. *J. Virol.* **85**, 638–643 (2011). doi:10.1128/ JVI.01806-10 Medline

15. K. G. Lokugamage, K. Narayanan, C. Huang, S. Makino, Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. *J. Virol.* **86**, 13598–13608 (2012). doi:10.1128/ JVI.01958-12 Medline

16. W. Kamitani, C. Huang, K. Narayanan, K. G. Lokugamage, S. Makino, A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* **16**, 1134–1140 (2009). doi:10.1038/nsmb.1680 Medline

17. C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, S. Makino, SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: Viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLOS Pathog.* **7**, e1002433 (2011). doi:10.1371/journal.ppat.1002433 Medline

18. M. G. Wathelet, M. Orr, M. B. Frieman, R. S. Baric, Severe acute respiratory syndrome coronavirus evades antiviral signaling: Role of nsp1 and rational design of an attenuated strain. *J. Virol.* **81**, 11620–11633 (2007). doi:10.1128/ JVI.00702-07 Medline

19. C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* **10**, 766–788 (2020). doi:10.1016/j.apsb.2020.02.008 Medline

20. A. R. Jauregui, D. Savalia, V. K. Lowry, C. M. Farrell, M. G. Wathelet, Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling. *PLOS ONE* **8**, e62416 (2013). doi:10.1371/journal.pone.0062416 Medline

21. M. S. Almeida, M. A. Johnson, T. Herrmann, M. Geralt, K. Wüthrich, Novel β-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J. Virol.* **81**, 3151–3161 (2007). doi:10.1128/ JVI.01939-06 Medline

22. F. Schluenzen, A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, A. Yonath, Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102**, 615–623 (2000). doi:10.1016/S0092-8674(00)00084-2 Medline

23. L. A. Passmore, T. M. Schmeing, D. Maag, D. J. Applefield, M. G. Acker, M. A. Algire, J. R. Lorsch, V. Ramakrishnan, The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol. Cell* **26**, 41–50 (2007). doi:10.1016/j.molcel.2007.03.018 Medline

24. M. Ameismeier, J. Cheng, O. Berninghausen, R. Beckmann, Visualizing late states of human 40S ribosomal subunit maturation. *Nature* **558**, 249–253 (2018). doi:10.1038/s41586-018-0193-0 Medline

25. A. Heuer, E. Thomson, C. Schmidt, O. Berninghausen, T. Becker, E. Hurt, R. Beckmann, Cryo-EM structure of a late pre-40S ribosomal subunit from *Saccharomyces cerevisiae*. *eLife* **6**, e30189 (2017). doi:10.7554/eLife.30189 Medline

26. A. des Georges, V. Dhote, L. Kuhn, C. U. T. Hellen, T. V. Pestova, J. Frank, Y. Hashem, Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525**, 491–495 (2015). doi:10.1038/nature14891 Medline

27. Y. Hashem, A. des Georges, V. Dhote, R. Langlois, H. Y. Liao, R. A. Grassucci, C. U.

T. Hellen, T. V. Pestova, J. Frank, Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119 (2013). doi:10.1016/j.cell.2013.04.036 Medline

28. J. L. Llácer, T. Hussain, L. Marler, C. E. Aitken, A. Thakur, J. R. Lorsch, A. G. Hinnebusch, V. Ramakrishnan, Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol. Cell* **59**, 399–412 (2015). doi:10.1016/j.molcel.2015.06.033 Medline

29. Y. J. Wang, P. P. Vaidyanathan, M. F. Rojas-Duran, N. D. Udeshi, K. M. Bartoli, S. A. Carr, W. V. Gilbert, Lso2 is a conserved ribosome-bound protein required for translational recovery in yeast. *PLOS Biol.* **16**, e2005903 (2018). doi:10.1371/journal.pbio.2005903 Medline

30. J. N. Wells, R. Buschauer, T. Mackens-Kiani, K. Best, H. Kratzat, O. Berninghausen, T. Becker, W. Gilbert, J. Cheng, R. Beckmann, Structure and function of yeast Lso2 and human CCDC124 bound to hibernating ribosomes. bioRxiv 944066 [Preprint]. 12 February 2020; https://doi.org/10.1101/2020.02.12.944066.

31. T. Becker, S. Franckenberg, S. Wickles, C. J. Shoemaker, A. M. Anger, J.-P. Armache, H. Sieber, C. Ungewickell, O. Berninghausen, I. Daberkow, A. Karcher, M. Thomm, K.-P. Hopfner, R. Green, R. Beckmann, Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature* **482**, 501–506 (2012). doi:10.1038/nature10829 Medline

32. A. Preis, A. Heuer, C. Barrio-Garcia, A. Hauser, D. E. Eyler, O. Berninghausen, R. Green, T. Becker, R. Beckmann, Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep.* **8**, 59–65 (2014). doi:10.1016/j.celrep.2014.04.058 Medline

33. A. Brown, S. Shao, J. Murray, R. S. Hegde, V. Ramakrishnan, Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493–496 (2015). doi:10.1038/nature14896 Medline

34. N. Miyazawa, H. Yoshikawa, S. Magae, H. Ishikawa, K. Izumikawa, G. Terukina, A. Suzuki, S. Nakamura-Fujiyama, Y. Miura, T. Hayano, W. Komatsu, T. Isobe, N. Takahashi, Human cell growth regulator Ly-1 antibody reactive homologue accelerates processing of preribosomal RNA. *Genes Cells* **19**, 273–286 (2014). doi:10.1111/gtc.12129 Medline

35. C. Yang, X. Liu, T. Cheng, R. Xiao, Q. Gao, F. Ming, M. Jin, H. Chen, H. Zhou, LYAR suppresses beta interferon induction by targeting phosphorylated interferon regulatory factor 3. *J. Virol.* **93**, e00769-19 (2019). doi:10.1128/JVI.00769-19 Medline

36. M. P. Robertson, H. Igel, R. Baertsch, D. Haussler, M. Ares Jr., W. G. Scott, The structure of a rigorously conserved RNA element within the SARS virus genome. *PLOS Biol.* **3**, e5 (2005). doi:10.1371/journal.pbio.0030005 Medline

37. K. M. Sparrer, M. U. Gack, Intracellular detection of viral nucleic acids. *Curr. Opin. Microbiol.* **26**, 1–9 (2015). doi:10.1016/j.mib.2015.03.001 Medline

38. Y. Hu, W. Li, T. Gao, Y. Cui, Y. Jin, P. Li, Q. Ma, X. Liu, C. Cao, The severe acute respiratory syndrome coronavirus nucleocapsid inhibits type I interferon production by interfering with TRIM25-mediated RIG-I ubiquitination. *J. Virol.* **91**, e02143-16 (2017). doi:10.1128/JVI.02143-16 Medline

39. L. Strähle, J.-B. Marq, A. Brini, S. Hausmann, D. Kolakofsky, D. Garcin, Activation of the beta interferon promoter by unnatural Sendai virus infection requires RIG-I and is inhibited by viral C proteins. *J. Virol.* **81**, 12227–12237 (2007). doi:10.1128/JVI.01300-07 Medline

40. K. M. Sparrer, C. K. Pfaller, K. K. Conzelmann, Measles virus C protein interferes with Beta interferon transcription in the nucleus. *J. Virol.* **86**, 796–805 (2012). doi:10.1128/JVI.05899-11 Medline

41. K. Brzózka, S. Finke, K. K. Conzelmann, Identification of the rabies virus alpha/beta interferon antagonist: Phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *J. Virol.* **79**, 7673–7681 (2005). doi:10.1128/JVI.79.12.7673-7681.2005 Medline

42. P. Devaux, V. von Messling, W. Songsungthong, C. Springfeld, R. Cattaneo, Tyrosine 110 in the measles virus phosphoprotein is required to block STAT1 phosphorylation. *Virology* **360**, 72–83 (2007). doi:10.1016/j.virol.2006.09.049 Medline

43. K. M. J. Sparrer, S. Gableske, M. A. Zurenski, Z. M. Parker, F. Full, G. J. Baumgart, J. Kato, G. Pacheco-Rodriguez, C. Liang, O. Pornillos, J. Moss, M. Vaughan, M. U. Gack, TRIM23 mediates virus-induced autophagy via activation of TBK1. *Nat. Microbiol.* **2**, 1543–1557 (2017). doi:10.1038/s41564-017-0017-2 Medline

44. M. Di Rienzo, M. Piacentini, G. M. Fimia, A TRIM32-AMBRA1-ULK1 complex initiates the autophagy response in atrophic muscle cells. *Autophagy* **15**, 1674–1676 (2019). doi:10.1080/15548627.2019.1635385 Medline

45. M. Z. Tay, C. M. Poh, L. Rénia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: Immunity, inflammation and intervention. *Nat. Rev. Immunol.* **20**, 363–374 (2020). doi:10.1038/s41577-020-0311-8 Medline

46. T. Tanaka, W. Kamitani, M. L. DeDiego, L. Enjuanes, Y. Matsuura, Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* **86**, 11128–11137 (2012). doi:10.1128/JVI.01700-12 Medline

47. V. Bhaskar, A. Graff-Meyer, A. D. Schenk, S. Cavadini, O. von Loeffelholz, S. K. Natchiar, C. G. Artus-Revel, H.-R. Hotz, G. Bretones, B. P. Klaholz, J. A. Chao, Dynamics of uS19 C-terminal tail during the translation elongation cycle in human ribosomes. *Cell Rep.* **31**, 107473 (2020). doi:10.1016/j.celrep.2020.03.037 Medline

48. S. Matheisl, O. Berninghausen, T. Becker, R. Beckmann, Structure of a human translation termination complex. *Nucleic Acids Res.* **43**, 8615–8626 (2015). doi:10.1093/nar/gkv909 Medline

49. C. Sidrauski, A. M. McGeachy, N. T. Ingolia, P. Walter, The small molecule ISRIB reverses the effects of eIF2α phosphorylation on translation and stress granule assembly. *eLife* **4**, e05033 (2015). doi:10.7554/eLife.05033 Medline

50. A. Sharma, M. Mariappan, S. Appathurai, R. S. Hegde, In vitro dissection of protein translocation into the mammalian endoplasmic reticulum. *Methods Mol. Biol.* **619**, 339–363 (2010). doi:10.1007/978-1-60327-412-8_20 Medline

51. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

52. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). doi:10.1016/j.jsb.2015.11.003 Medline

53. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). doi:10.1016/j.jsb.2015.08.008 Medline

54. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166–e42166 (2018). doi:10.7554/eLife.42166 Medline

55. J. Zivanov, T. Nakane, S. H. W. Scheres, Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in *RELION*-3.1. *IUCrJ* **7**, 253–267 (2020). doi:10.1107/S2052252520000081 Medline

56. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi:10.1038/nmeth.4169 Medline

57. T. Burnley, C. M. Palmer, M. Winn, Recent developments in the CCP-EM software suite. *Acta Crystallogr. D Struct. Biol.* **73**, 469–477 (2017). doi:10.1107/S2059798317007859 Medline

58. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi:10.1107/S0907444904019158 Medline

59. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). doi:10.1107/S0907444910007493 Medline

60. S. Sarkar, S. Witham, J. Zhang, M. Zhenirovskyy, W. Rocchia, E. Alexov, DelPhi Web Server: A comprehensive online suite for electrostatic calculations of biological macromolecules and their applications. *Commun. Comput. Phys.* **13**, 269–284 (2013). doi:10.4208/cicp.300611.201011s Medline

61. H. Kratzat, T. Mackens-Kiani, M. Ameismeier, J. Cheng, E. Dacheux, A. Namane, O. Berninghausen, M. Fromont-Racine, T. Becker, R. Beckmann, Structural inventory of native ribosomal ABCE1-43S pre-initiation complexes. bioRxiv 194902 [Preprint]. 9 July 2020; https://doi.org/10.1101/2020.07.09.194902.

62. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi:10.1002/pro.3235 Medline

SUPPLEMENTARY MATERIALS

science.sciencemag.org/cgi/content/full/science.abc8665/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
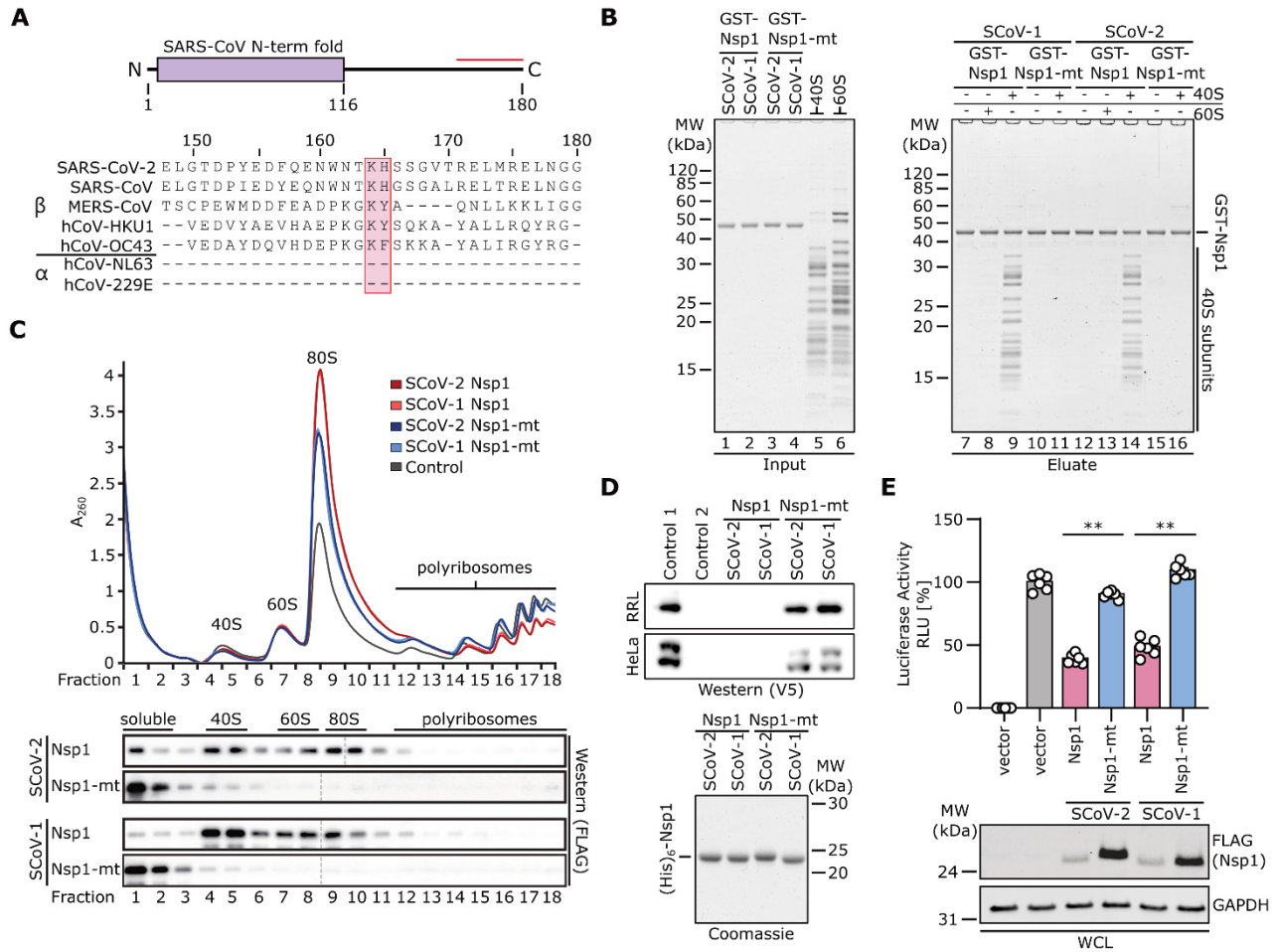References (*48–62*)
MDAR Reproducibility Checklist
Data S1

18 May 2020; accepted 13 July 2020
Published online 17 July 2020
10.1126/science.abc8665

Fig. 1. Nsp1 interacts with 40S ribosomal subunits and inhibits translation. (**A**) Domain organization of Nsp1 and sequence alignment of the C-terminal segment (red line) of Nsp1 from seven human CoVs. The KH motif is marked. (**B**) In vitro binding assay of GST-TEV (GST) tagged Nsp1 and Nsp1-mt from SCoV-1 and SCoV-2 with human 40S and 60S ribosomal subunits. Coomassie stained SDS-PAGE of inputs and eluates. (**C**) Polyribosome gradient analysis of HEK293T lysate (Control) and lysate from HEK293T cells transiently transfected with 3xFLAG tagged Nsp1 and Nsp1-mt constructs from SCoV-1 and SCoV-2 and Western blot analysis (anti-FLAG antibody, dashed lines: separate blots). (**D**) Western blot (top, anti-V5 antibody) and SDS-PAGE analysis (bottom) of cell-free in vitro translation of a capped reporter mRNA with rabbit reticulocytes (RRL) and HeLa S3 lysate. Controls 1 and 2, with and without capped reporter mRNA, respectively. Coomassie stained SDS-PAGE of the applied $(His)_6$-TEV (His_6) tagged Nsp1 constructs is shown below. (**E**) Quantification of luciferase in HEK293T cells transfected with indicated 3xFLAG-tagged proteins and in vitro transcribed firefly luciferase mRNA. Bars represent the mean of $n=6\pm$SEM. RLU, relative light units. Representative immunoblots of whole cell lysates (WCL) stained with anti-FLAG and anti-GAPDH. Unpaired student's t test (Welch correction), **, $p<0.001$.
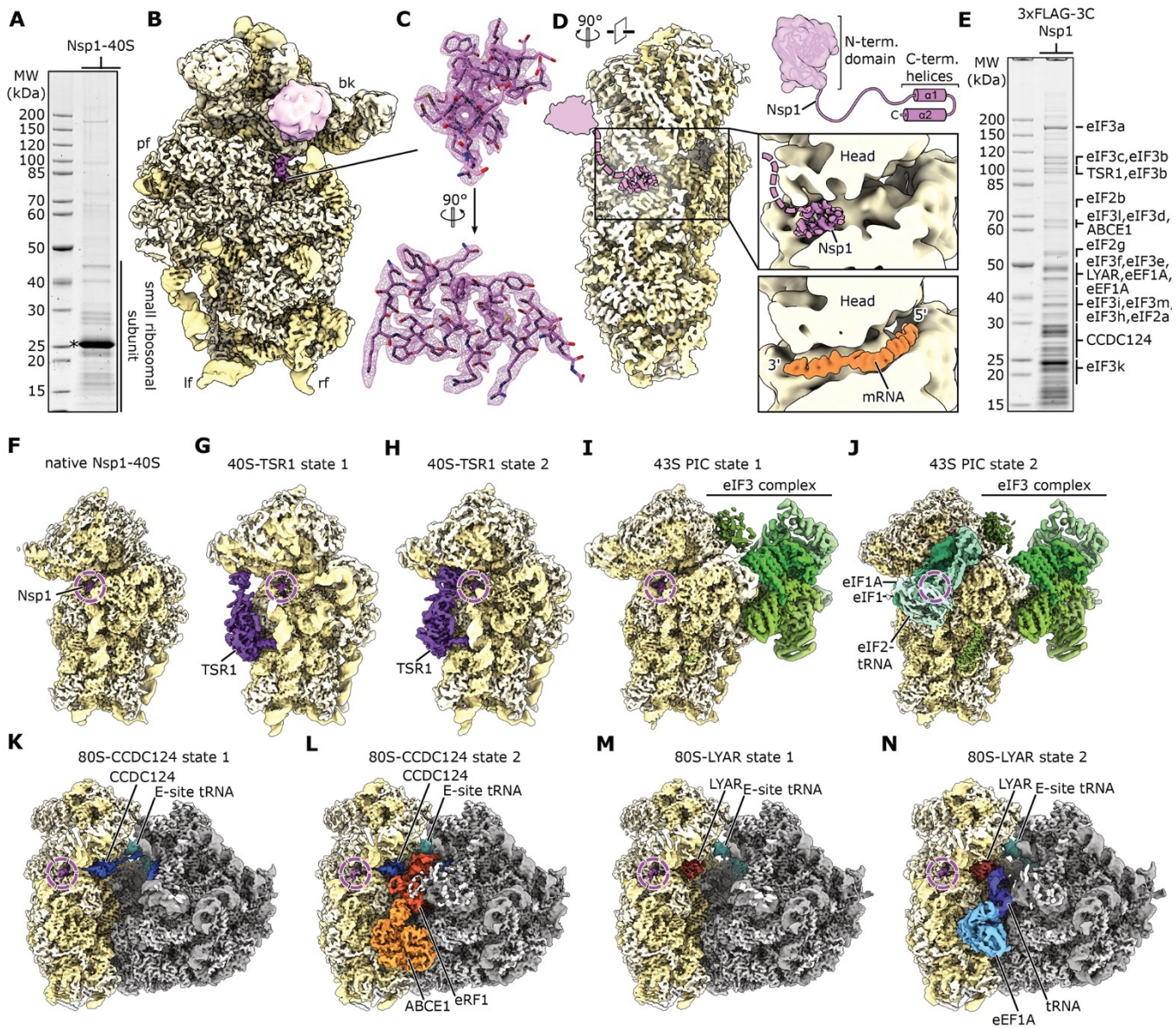
Fig. 2. Cryo-EM structures of Nsp1-bound ribosomal complexes. (**A**) SDS-PAGE analysis of reconstituted Nsp1-40S complexes, Nsp1 is labeled with an asterisk. (**B**) Reconstituted Nsp1-40S structure with Nsp1 in pink and ribosomal RNA and proteins in yellow. Additional density between uS3 and h16 assigned to the N-terminal fold of Nsp1. bk, beak; pf, platform; lf, left foot; rf, right foot. (**C**) C-terminal helix 1 and 2 of Nsp1 with corresponding density. (**D**) Cross-section of the 40S highlighting the central position of Nsp1 within the mRNA tunnel. The putative position of the N-terminal domain of Nsp1 is schematically indicated [models based on PDB-2HSX (*21*) and PDB-6Y0G (*47*)]. (**E**) SDS-PAGE analysis of Nsp1-ribosomal complexes affinity purified from HEK293T cells. Proteins identified in the cryo-EM structures were labeled according to mass spectrometry analysis (see data S1). (**F** to **N**) Cryo-EM maps of affinity purified Nsp1-ribosomal complexes. Additional factors are colored and labeled accordingly.
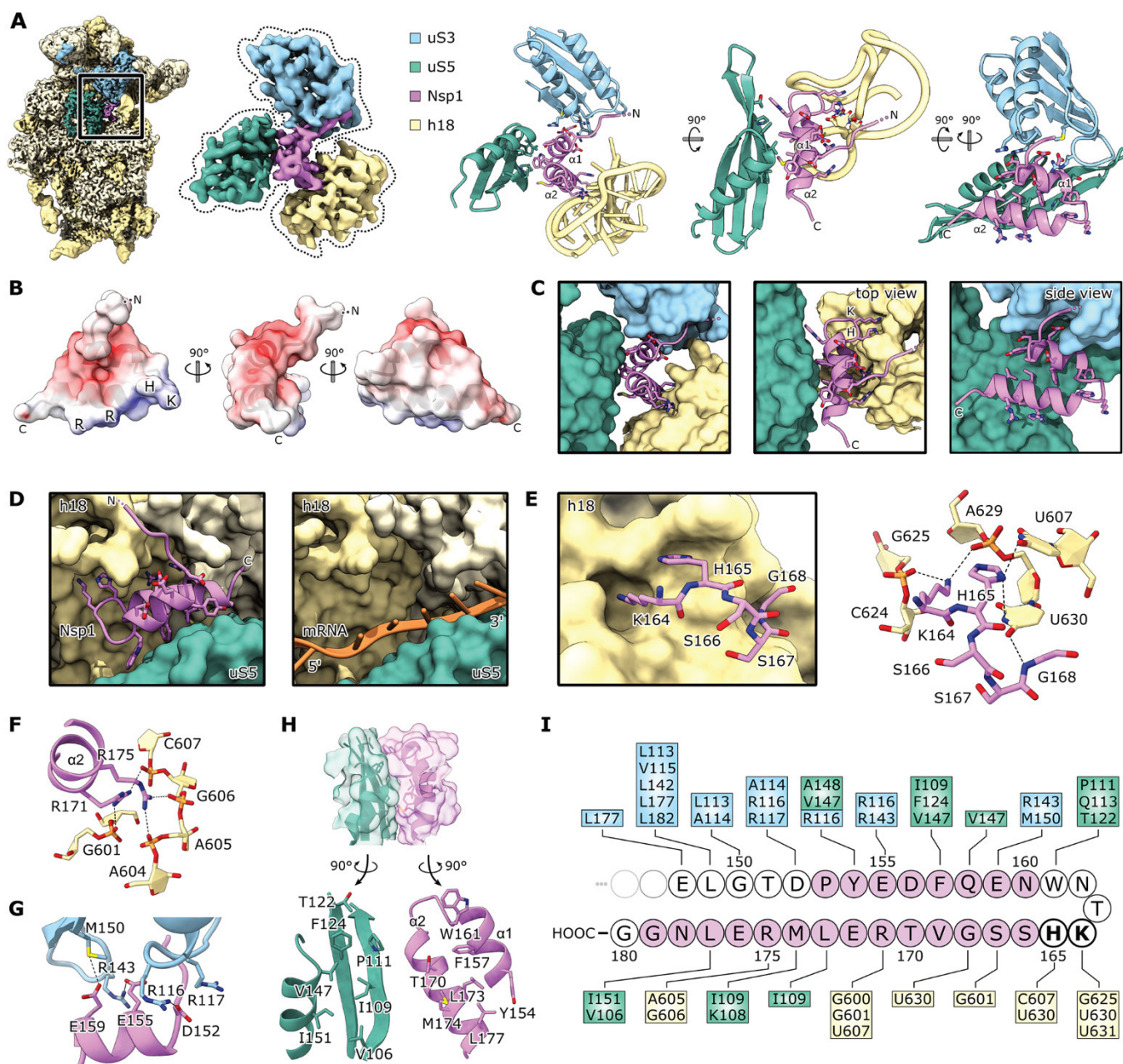
Fig. 3. Molecular basis of Nsp1 ribosome interaction and inhibition. (A) Cryo-EM map of in vitro reconstituted Nsp1-40S and segmented density of Nsp1-C, uS3 (97-153,168-189), uS5 (102-164) and rRNA helix h18 with the corresponding models; interacting residues are shown as sticks. (B) Nsp1-C surface, colored by electrostatic potential from -5 (red) to +5 (blue). (C) Model of Nsp1-C and surface representation of the models of uS3 (97-153,168-189), uS5 (102-164) and rRNA helix h18. Molecular interactions between Nsp1 and the ribosome. (D) mRNA entry channel, 40S head is removed. Nsp1-C occupies the mRNA path [mRNA based on PDB-6Y0G (47)]. (E) K164 and H165 of Nsp1 bind to a pocket on h18. (F) R171 and R175 of Nsp1 bind to the phosphate backbone of h18. (G) Negatively charged residues D152, E155 and E159 of α1 interact with uS3. (H) The hydrophobic interface of α1 and α2 binds to a hydrophobic patch on uS5. (I) Schematic summary of the interaction of Nsp1-C with uS3, uS5 and h18; residues belonging to α1 and α2 are colored in pink.
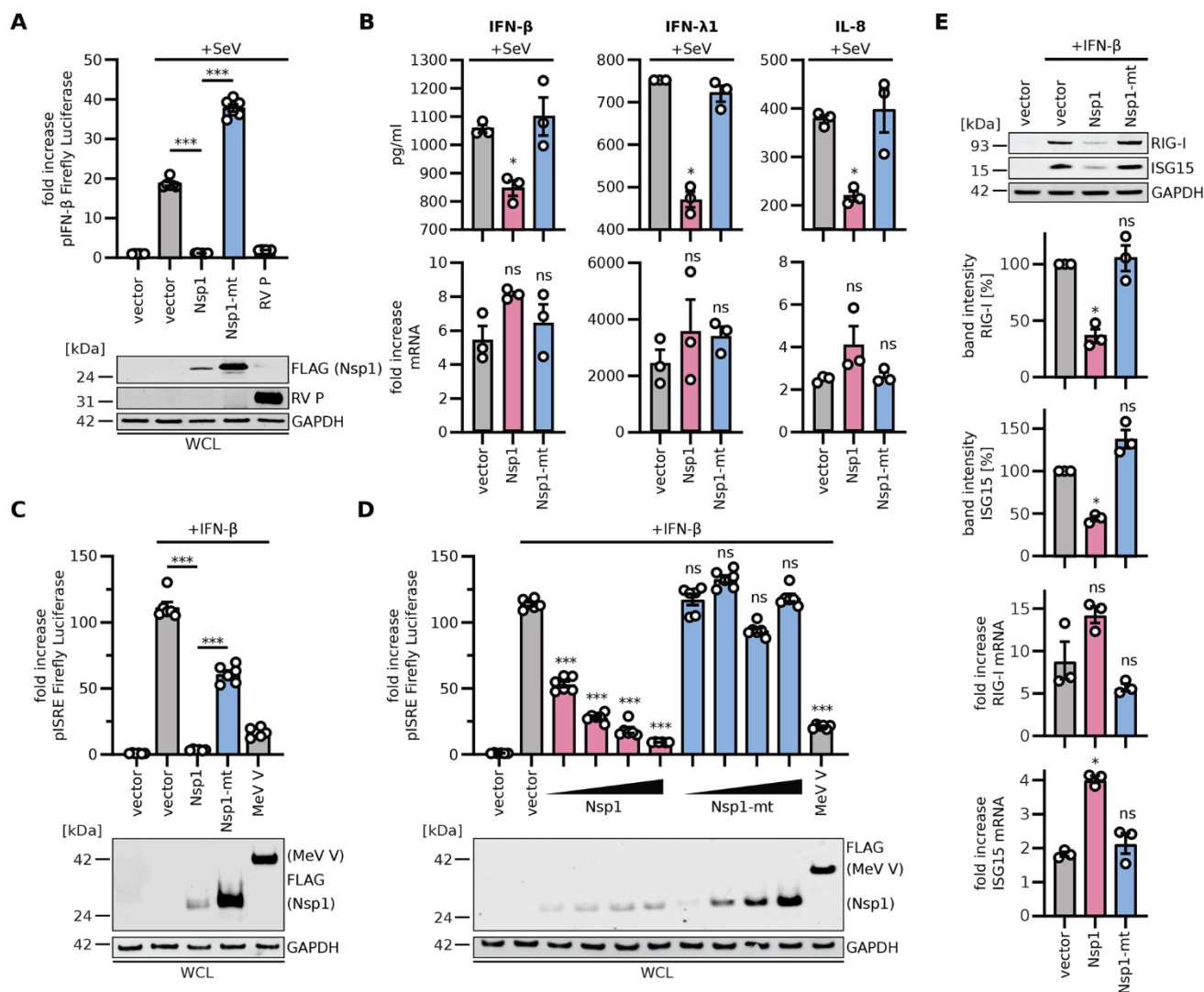
Fig. 4. Inhibition of the innate immune response by SARS-CoV-2 Nsp1. (A) Quantification of IFN-β-promotor controlled firefly luciferase activity in HEK293T cells transiently expressing 3xFLAG-tagged or non-tagged (RV P) proteins. Cells were infected with Sendai Virus (SeV) or left uninfected. Representative immunoblots of whole cell lysates (WCLs) stained with anti-RV P, anti-FLAG and anti-GAPDH (bottom panel). (B) ELISA of IFN-β, IFN-λ1 or IL-8 in the supernatant of HEK293T cells transiently expressing 3xFLAG-tagged proteins and infected with SeV (top panel) for 24 hours. qPCR of corresponding mRNAs (bottom panel). (C and D) Quantification of ISRE-promotor controlled firefly luciferase activity in HEK293T cells transiently expressing 3xFLAG-tagged proteins in single amounts (C) or increasing amounts (D) and treated with 1000 U/mL IFN-β as indicated. Representative immunoblots of WCLs stained with anti-FLAG and anti-GAPDH (bottom panel). (E) Representative immunoblots and quantification of WCLs of HEK293T cells stimulated with 200 U/mL IFN-β stained for endogenous RIG-I, ISG15 and GAPDH. qPCR of the corresponding mRNAs (bottom two panels). In (A), (C), and (D), bars represent the mean of n=6±SEM. In (B) and (E), bars represent the mean of n=3±SEM. Unpaired student's *t* test (Welch correction), ns, not significant; *, p<0.01, ***, p<0.0001.

Science

Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2

Matthias Thoms, Robert Buschauer, Michael Ameismeier, Lennart Koepke, Timo Denk, Maximilian Hirschenberger, Hanna Kratzat, Manuel Hayn, Timur Mackens-Kiani, Jingdong Cheng, Jan H. Straub, Christina M. Stürzel, Thomas Fröhlich, Otto Berninghausen, Thomas Becker, Frank Kirchhoff, Konstantin M. J. Sparrer and Roland Beckmann

ARTICLE TOOLS	http://science.sciencemag.org/content/early/2020/07/16/science.abc8665
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2020/07/16/science.abc8665.DC1
RELATED CONTENT	http://stm.sciencemag.org/content/scitransmed/12/555/eabc9396.full http://stm.sciencemag.org/content/scitransmed/12/550/eabc3539.full http://stm.sciencemag.org/content/scitransmed/12/554/eabc1126.full http://stm.sciencemag.org/content/scitransmed/12/549/eabb9401.full
REFERENCES	This article cites 62 articles, 16 of which you can access for free http://science.sciencemag.org/content/early/2020/07/16/science.abc8665#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Supplementary Materials for

Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2

Matthias Thoms*, Robert Buschauer*, Michael Ameismeier*, Lennart Koepke,
Timo Denk, Maximilian Hirschenberger, Hanna Kratzat, Manuel Hayn, Timur Mackens-Kiani,
Jingdong Cheng, Jan H. Straub, Christina M. Stürzel, Thomas Fröhlich, Otto Berninghausen,
Thomas Becker, Frank Kirchhoff, Konstantin M. J. Sparrer†, Roland Beckmann†

*These authors contributed equally to this work.
†Corresponding author. Email: beckmann@genzentrum.lmu.de (R.B.); konstantin.sparrer@uni-ulm.de (K.M.J.S.)

This PDF file includes:

Materials and Methods
Figs. S1 to S5
Tables S1 and S2
Caption for Data S1
References

Other Supporting Online Material for this manuscript includes the following:
(available at science.sciencemag.org/cgi/content/full/science.abc8665/DC1)

MDAR Reproducibility Checklist (.pdf)
Data S1 (.xlsx)

Materials and Methods

<u>Cell lines, plasmids, viruses and transfections</u>

HEK293T cells were purchased from American type culture collection (ATCC: #CRL-3216) and cultivated in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS, Gibco), 100 U/ml penicillin (PAN-Biotech), 100 µg/ml streptomycin (PAN-Biotech) or 1x penicillin/streptomycin (Gibco), and 2 mM L-glutamine (PAN-Biotech) or 1x GlutaMAX (Gibco) (hereafter called DMEM+3).

The open reading frames for SARS-CoV-2 (SCoV-2) Nsp1 and Nsp7 were ordered codon optimized and V5-tagged from Twist Bioscience and the SARS-CoV (SCoV-1) and SCoV-2 Nsp1 open reading frames were in addition synthesized by GenArts (ThermoFisher Scientific). Nsp1-V5 and Nsp7-V5 were amplified by PCR (using Primer 1 (Nsp1 rev): CGA CGC GTC TAG CCG CCA TTC AGC TCG CGC, primer 2 (Nsp7 rev): CGA CGC GTC TAT TGC AGC GTG GCA CG, primer 3 (XbaI fwd) CGT CTA GAG CCA CCATG) and the single restriction sites XbaI/MluI as well as the Kozak sequence GCCACC were introduced. Afterwards PCR-fragments were subcloned into pCG or pCG-IRES-GFP vectors to yield pCGSARS-CoV-2 Nsp1-V5 IRES eGFP, pCGSARS-CoV2 Nsp7-V5 IRES eGFP, pCGSARS-CoV-2 Nsp1-V5, pCGSARS-CoV-2 -Nsp7-V5. pCR3-MeV-V-FLAG, pIFNb-FFLuc, pISRE-FFLuc, pCMV-GLuc were kind gifts from Karl-Klaus Conzelmann and described previously (*40*). The SCoV-1 Nsp1 and SCoV-2 Nsp1 open reading frames were PCR amplified with Primer 4 (BamHI Nsp1 SCoV-1) CGC GGA TCC ATG GAG AGC CTT GTT CTT GGT GTC AAC G, Primer 5 (Nsp1 STOP SCoV-1 XhoI) CCG CTC GAG TCA ACC TCC ATT GAG CTC ACG AGT GAG TTC and Primer 6 (BamHI Nsp1 SCoV-2) CGC GGA TCC ATG GAG AGC CTT GTC CCT GGT TTC AAC G, Primer 7 (Nsp1 STOP SCoV-2 XhoI) CCG CTC GAG TCA CCC TCC GTT AAG CTC ACG CAT GAG TTC and cloned into a modified pcDNA5/FRT/TO-3xFLAG-3C expression vector for N-terminal tagging. In order to generate *E. coli* expression plasmids, Nsp1 from SCoV-1 and SCoV-2 were amplified with Primer 8 (NdeI Nsp1 SCoV-1) GGG AAT TCC ATA TGG AGA GCC TTG TTC TTG GTG TCA ACG, Primer 9 (Nsp1 STOP SCoV-1 BamHI) CGC GGA TCC TTA ACC TCC ATT GAG CTC ACG AGT GAG TTC and Primer 10 (NdeI Nsp1 SCoV-2) GGG AAT TCC ATA TGG AGA GCC TTG TCC CTG GTT TCA ACG, Primer 11 (Nsp1 STOP SCoV-2 BamHI) CGC GGA TCC TTA CCC TCC GTT AAG CTC ACG CAT GAG TTC and were cloned into the plasmid backbones pET-24d-(His)6-TEV and pET-24d-GST-TEV for N-terminal tagging. The K164A H165A mutations were introduced by overlap extension PCR using the following primer: Primer 12 (Nsp1 KH>AA SCoV-1_fwd) TGG AAC ACT GCC GCT GGC AGT GGT GCA CTC CGT GAA CTC A, primer 13 (Nsp1 KH>AA SCoV-1_rev) CTG CCA GCG GCA GTG TTC CAG TTT TGT TCA TAA TCT TCA ATG GG and Primer 14 (Nsp1 KH>AA SCoV-2_fwd) TGG AAC ACT GCA GCT AGC AGT GGT GTT ACC GT GAA CTC ATG, Primer 15 (Nsp1 KH>AA SCoV-2_rev) CTG CTA GCT GCA GTG TTC CAG TTT TCT TGA AAA TCT TCA TAA GG.

Firefly luciferase (Fluc) mRNA diluted in 25 µl Opti-MEM (Gibco) was mixed with 1 µl Lipofectamine 2000 (Invitrogen) diluted in 25 µl Opti-MEM, incubated for 5 min at room temperature and transferred on the cells. For all transfections of plasmid DNA, Polyethylenimine (PEI, 1 mg/ml in H2O, Sigma-Aldrich) or the TransIT-LT1 Transfection Reagent (Mirus) were used according to the manufacturers recommendations or as described previously (*47*).

Sendai virus (SeV, Cantell strain) was kindly provided by Georg Kochs (Freiburg University) and Daniel Sauter (Ulm University) and used to stimulate innate immune activation via RIG-I 16 h post transfection. Recombinant IFN beta was purchased from R&D Systems (8499-IF) and used at a concentration of 200 U/ml for stimulation either for 8 h (luciferase reporter gene assays) or for 24 h (ISG expression).

Whole-cell lysates

Whole-cell lysates were prepared by collecting cells in Phosphate-Buffered Saline (PBS, Gibco). The cell pellet (500 g, 4 °C, 5 min) was lysed in transmembrane lysis buffer [50 mM HEPES pH 7.4, 150 mM NaCl, 1% Triton X-100, 5 mM ethylenediaminetetraacetic acid (EDTA)] by vortexing at maximum speed for 30 s. Cell debris were pelleted by centrifugation (20,000 g, 4 °C, 20 min) and the cleared supernatants were stored for analysis at -20°C.

SDS-PAGE and immunoblotting

SDS-PAGE and immunoblotting was performed using standard techniques as previously described (*47*). The following antibodies were used throughout the study: anti-β-actin (1:10,000, AC-15, Invitrogen), anti-GAPDH (1:1,000, # 607902, BioLegend), anti-V5 (1:1,000, D3H8Q, Cell Signaling Technology), anti-FLAG (1:1,000 or 1:5,000, M2, Sigma-Aldrich), anti-ISG15 (1:1,000, F-9, Santa Cruz Biotechnology), anti-RIG-I (1:1,000, D14G6, Cell Signaling Technology), anti-RVP (1:10,000, this antibody was a kind gift by Karl-Klaus Conzelmann) and IRDye Secondary Antibodies, Li-Cor (1:20,000 in 0.05% (w/v) casein, Thermo Scientific).

Legendplex ELISA

The Legendplex ELISA (Anti-virus Panel, Biolegend) was performed according to the manufacturer's instructions. In brief, the supernatants were incubated for 2 h at room temperature with the antibody-coated beads, followed by washing and incubation with the detection antibodies. After incubation with the staining reagent, the beads were analysed in a high-throughput sampler via flow cytometry (FACSCanto II, BD Biosciences). Absolute quantification was performed using a standard and the Biolegend Legendplex v8.0 software.

Luciferase assay

16 h after mRNA transfection, 24 h after SeV infection or 8 h after cytokine stimulation transfection, the cells were lysed in 200 µl 1x Passive Lysis buffer (Promega). 25 µl of the lysate were transferred into a white 96-well plate, 25 µl of firefly substrate (Promega) were added and the luminescence was quantified as relative light units (Orion microplate luminator). To measure

Gaussia luciferase activity for normalisation, 25 µl of the supernatants were transferred into a white 96-well plate. Coelenterazine Substrate (pjk), diluted 1:120 in PBS, was added to the samples and luminescence quantified (Orion microplate luminator).

qRT-PCR

Total RNA was extracted from cells using the Quick-RNA Microprep Kit (Zymo Research) according to the manufacturer's instructions. Reverse transcription and qRT–PCR were performed in one step using 2 µl (~400 ng) of the purified RNA samples as templates (SuperScript III Platinum Kit, Invitrogen) on a StepOnePlus Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions. TaqMan probes for each individual gene were acquired as premixed TaqMan Gene Expression Assays (Thermo Fisher Scientific) and added to the reaction. Expression level for each target gene was calculated by normalizing against GAPDH using the $\Delta\Delta CT$ method and represented relative to the values for mock-transfected cells, which were set to 1.

In vitro transcription (IVT)

For IVT of the luciferase reporter, the HiScribe T7 Quick High Yield RNA Synthesis Kit und Cap analogue (both NEB) were used. Capped RNA synthesis, DNase treatment and LiCl precipitation were performed according to the supplier's protocol. As a template, the included linearized Firefly Luciferase template DNA was used. For IVT of the hCMV stalling mRNA (*in vitro* translation assay) the mMESAGE mMACHINE T7 Transcription Kit was used according to the manufacturer's protocol. Template for hCMV stalling mRNA was PCR amplified from plasmid pGEM4-CD4-CMV. The size of the transcripts was analysed by agarose gel electrophoresis, stained using ethidium bromide (AppliChem GmbH) and visualized in a Bio-Rad Gel Doc XR+.

Quantification of autophagy

Transiently transfected HEK293T cells stably expressing GFP-LC3B were harvested in PBS and treated for 20 min at 4 °C with PBS containing 0.05% Saponin to wash out non-membrane bound GFP-LC3B. Cells were subsequently fixed in 4% Paraformaldehyde (Santa Cruz) and fluorescence intensity quantified via flow cytometry (FACSCanto II, BD Biosciences). The GFP-LC3B mean fluorescence intensity of the control was subtracted.

Statistical calculations

Unpaired student's t-test (Welch correction) were used to calculate significances for luciferase assays, qPCRs and ELISA. Not significant values are indicated as ns; Significant samples denoted as *,$p<0.01$;**, $p<0.001$; ***, $p<0.0001$.

In vitro translation assay

HeLa S3 translation extract was prepared as described before (*48*), except that cells were treated with 200 nM integrated stress response inhibitor (ISRIB) 1 h prior to harvesting to ensure cap-dependent translation initiation (*49*). For the reaction, 12 µl HeLa translation reaction mix with 50% (v/v) extract was adjusted to 2.75 mM Mg(OAc)$_2$, 0.42 mM MgCl$_2$, 75 mM KOAc, 37.5 mM KCl, 42 mM NaCl, 2 mM DTT, 1.56 mM GTP, 0.25 mM ATP, 1.6 mM creatine phosphate, 0.45 mg/mL creatine kinase, 50 µg/mL yeast tRNA, 0.4 mM spermidine, 0.12 mM complete amino acid mixture (Promega) and 0.8 U/µl RNase inhibitor (Invitrogen).

Crude rabbit reticulocyte lysate (RRL, Green Hectares) was first treated with micrococcal nuclease, then supplemented with Hemin (*50*) and frozen in aliquots. The final 25 µl translation reaction mixture with 70% (v/v) RRL was adjusted to 50 mM NaCl, 0.5 mM MgCl$_2$, 0.3 mM complete amino acid mixture (Promega), 90 mM creatine phosphate, 0.1 mg/mL yeast tRNA and 7.3 U/µl SUPERase·In RNase inhibitor (Invitrogen).

HeLa and RRL translation reactions were preincubated with purified N-terminal (His)$_6$-TEV tagged Nsp1 or Nsp1-mt from SARS-CoV-1 and -2 for 30 min on ice with final concentrations of 6.25 µM (HeLa) and 7.5 µM (RRL). Subsequently, the reactions were initiated by addition of 1 µg of mRNA encoding the gp48 uORF2 peptide that leads to stalling of eukaryotic ribosomes (fig. S1B). Both *in vitro* translation reactions were incubated for 20 min at 30°C. Subsequently, ribosomes were isolated from the RRL reaction by pelleting through a sucrose cushion [50 mM HEPES-KOH pH 7.5, 10 mM Mg(OAc)$_2$, 200 mM KOAc, 1 mM DTT, 0.5 M sucrose] in a TLA100 rotor for 1 h at 434,513 g and 4°C. The ribosome pellets were resuspended in reducing sample buffer [50 mM Tris-HCl pH 6.8, 2% SDS, 0.005% bromophenol blue, 10% glycerol, 100 mM DTT] and separated using SDS-PAGE. Samples from HeLa translation reactions were loaded directly, without pelleting. Peptidyl-tRNA was detected by immunoblotting with primary antibodies against the V5-tag and secondary HRP-coupled antibodies.

Polyribosome gradient analysis

HEK293T were cultured to 40% confluency and transiently transfected with the pcDNA/FRT/TO-3xFLAG-3C-Nsp1 plasmids using PEI. After 22 h cells from one 15 cm culture dish (~1.8x10^7) were treated with 100 µg/ml cycloheximide (CHX) for 15 min, then washed with 5 ml ice-cold PBS containing 100 µg/ml CHX and scraped off the culture dish. Cells were harvested by centrifugation at 120 x g for 10 min. The cells were washed once with 1 ml PBS/CHX and pelleted by centrifugation at 300 x g for 5 min. The cell pellet was resuspended in 500 ml lysis buffer [5 mM Tris-HCl, pH 7.5, 1.5 mM KCl, 2.5 mM MgCl$_2$, complete EDTA-free protease inhibitor cocktail (Roche), 100 µg/ml CHX, 2 mM DTT, 200 U/ml SUPERase·In RNase Inhibitor (Invitrogen), 0.5% v/w Triton X-100, and 0.5% v/w sodium deoxycholate], incubated for 1 minute and cleared by centrifugation at 15000 x g for 5 min. The nucleic acid concentration in the cleared samples was determined by measurement of the absorption at 260 nm and an equivalent amount of each sample was separated on a 10%-50% sucrose gradient [20 mM HEPES-KOH, pH 7.5, 100 mM KOAc, 5 mM MgCl$_2$, 1 mM DTT, 10 U/ml RNase inhibitor, protease inhibitor cocktail] at

202048 x g for 2.5 h using a SW 40 Ti rotor (Beckman Coulter). Gradients were fractionated using a Biocomp piston gradient fractionator and A_{260} was observed using a Biocomp Triax flow cell. Proteins in the collected fractions were precipitated using trichloroacetic acid and separated on a 15% SDS-PAGE gel, then blotted onto a PVDF membrane (0.45 µm pore size, Immobilon-P, Merck). Detection of 3xFLAG-3C-Nsp1 on the membranes was performed using an M2 anti-FLAG HRP antibody (Sigma, A8592) at 1/1000 working dilution according to the manufacturer's protocol. Fractions from one gradient were blotted on two separate membranes, decorated with AB at the same time, and developed using the same imaging parameters. Images were then combined, and band intensities equalized based on samples loaded on both membranes.

Expression and purification from *E. coli*

The constructs were transformed and expressed in *E. coli* BL21(DE3) cells. Cells were grown at 37°C in LB medium supplemented with kanamycin (50 µg/ml). When the cultures reached an OD_{600} of about 0.5-0.7, cells were shifted to 18°C and expression was induced after 30 min by the addition of IPTG to a final concentration of 0.25 mM. Cells were grown for 16 h, harvested by centrifugation and stored at -80°C. For purification of the GST-TEV tagged Nsp1 constructs cell pellets were resuspended in lysis buffer [20 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$] and lysed with a M-110L Microfluidizer (Microfluidics). The lysate was cleared by centrifugation (43,200 x g, 30 min, 4°C) and the supernatant was incubated with Glutathione Sepharose 4 Fast Flow beads (GE Healthcare) for 1 h at 4°C. The supernatant was removed and beads were washed two times with 40 ml lysis buffer. Samples were eluted with lysis buffer supplemented with 25 mM reduced L-gluthathione (Sigma-Aldrich). Cells expressing the different $(His)_6$-TEV-Nsp1 and the Nsp1-AviTag-$(His)_6$ construct were resuspended and lysed in lysis buffer 2 [20 mM HEPES-NaOH pH 8.0, 500 mM NaCl, 5 mM $MgCl_2$, 10 mM imidazole] as described above. Ni-NTA agarose (Qiagen) was added to the cleared lysate and incubated for 1 h at 4°C. Beads were washed once with 40 ml lysis buffer 2 and once with 40 ml lysis buffer 2 containing 20 mM imidazole. Samples were eluted with buffer containing: 20 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$ and 300 mM imidazole. All samples were further purified by size-exclusion chromatography in buffer containing: 20 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 5 mM $MgCl_2$. The $(His)_6$-TEV-Nsp1 and the Nsp1-AviTag-$(His)_6$ proteins were purified through a Superdex 75 10/300 GL column (GE Healthcare) and for the GST-TEV-Nsp1 samples a Superdex 200 Increase 10/300 GL (GE Healthcare) column was used. Fractions containing Nsp1 were pooled, concentrated with a Amicon Ultra centrifugal filter (Millipore), flash frozen in liquid nitrogen and stored at -80°C.

In vitro binding assay

Human 40S and 60S ribosomal subunits were prepared as previously described (*24*). The binding assay was performed in binding buffer [20 mM HEPES-KOH pH 7.6, 150 mM KOAc, 2.5 mM $MgCl_2$, 0.01% NP-40 and 2 mM DTT]. The GST-TEV tagged Nsp1 bait proteins (60 pmol) were incubated with purified 40S or 60S ribosomal subunits (see fig. S1A) in a total volume of

300 µl for 45 min at 4°C. The samples were transferred to 1 ml Mobicols (MoBiTec) containing 30 µl Glutathione Sepharose 4 Fast Flow (GE Healthcare) and incubated for an hour at 4°C. The unbound material was removed by centrifugation in a pre-cooled table top centrifuge and the beads were washed 3x times with binding buffer (1x 800 µl, 2x 500 µl). The beads were incubated for 1 h at 4°C with binding buffer supplemented with 25 mM reduced L-glutathione (Sigma-Aldrich) and the samples were eluted by centrifugation. Samples were analyzed on a 12% polyacrylamide gel (NuPAGE, Invitrogen) and stained with Coomassie.

Affinity purification of Nsp1 from HEK293T cells

HEK293T cells transiently transfected with the pcDNA5/FRT/TO-3xFLAG-3C-Nsp1 (SCoV-2) construct were lysed in purification buffer [20 mM Hepes-KOH pH 7.5, 150 mM KOAc, 5 mM $MgCl_2$, 1 mM DTT, 0.5 mM NaF, 0.1 mM $Na_2V_3O_4$] supplemented with 5% glycerol and 0.5% NP-40. The lysate was sonicated 4 times for 10 s with 30 s on ice in between (Branson Sonifier 250). The lysate was cleared by centrifugation for 15 min at 2,960 g and 25 min at 36,500 g and the supernatant was incubated with ANTI-FLAG M2 agarose beads (Sigma-Aldrich) on a rotating wheel for 120 min at 4°C. The FLAG beads were washed twice with 10 ml purification buffer supplemented with 0.01% NP-40 and once with purification buffer containing 0.05% octaethylene glycol monododecyl ether. The FLAG beads were transferred to a 1 ml Mobicol (MoBiTec) and washed with 5 ml buffer + 0.05% octaethylene glycol monododecyl ether. The beads were incubated with buffer containing 20 mM HEPES-KOH pH 7.5, 150 mM KOAc, 5 mM $MgCl_2$, 1 mM DTT, 0.05% octaethylene glycol monododecyl ether and 40 µg 3C protease for 1 h at 4°C and the eluted samples were collected by centrifugation. Eluates were used for Cryo-EM analysis and analyzed on a 4-12% Bis-Tris polyacrylamide gel (NuPAGE, Invitrogen).

Mass spectrometry

The gel bands were de-stained using 50% acetonitrile in 50 mM NH_4HCO_3. For protein reduction, 45 mM dithioerythritol (DTE) in 50 mM NH_4HCO_3 was added to the gel pieces and incubated for 30 min at 55°C. Carbamidomethylation of cysteines was done twice by incubation at RT in the dark for 15 min in 100 mM iodoacetamide/50 mM NH_4HCO_3. Gel slices were minced and digestion was performed overnight at 37°C using 70 ng porcine trypsin (Promega, Fitchburg, WI, USA). Tryptic peptides were extracted using 70% ACN. Prior to liquid chromatography, the samples were dried using a SpeedVac vacuum concentrator. Chromatography was done with an Ultimate 3000 nano-LC system (Thermo Fisher Scientific) using an Acclaim PepMap 100 trap column (nanoViper C18, 2 cm length, 100 µM ID, Thermo Scientific) and an EasySpray separation column (PepMap RSLC C18, 50 cm length, 75 µm ID, Thermo Fisher Scientific). For peptide separation a flow rate of 250 nl/min and 0.1% formic acid as solvent A was used. The method consisted on gradients from 3% to 25% solvent B (0.1% formic acid in acetonitrile) in 30 min and from 25% to 40% B in 5 min. Data dependent mass spectrometry was performed on a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) using cycles of one full scan (350 to 1600 m/z) at 60k resolution and up to 12 data-dependent MS/MS scans at 15k resolution. Spectra were

searched using MASCOT V2.4 (Matrix Science Ltd, London, UK) and the human subset of the UniProt database. Common contaminants like keratins were removed and the results were filtered for an FDR < 1%. Exclusively protein identifications with at least two individual peptides were considered.

Electron microscopy and image processing

3.5 µl of sample solution were applied to holey carbon support grids (R3/3 with 2 nm continuous carbon support, Quantifoil), which had been glow discharged at 2.1×10^{-1} mbar for 20 s. Grids were incubated for 45 s at 4°C and subsequently plunge frozen in liquid ethane using a Vitrobot Mark IV (FEI Company). 11,270 (dataset 1, reconstituted complex) and 6,610 (dataset 2, native complex) movies were recorded on a Titan Krios at 300 kV using a K2 Summit direct electron detector with a nominal pixel size of 1.059 Å and a defocus range from 0.5 – 3.0 µm at low-dose conditions. For each movie, 40 frames with approximately 1.12 e$^-$ Å^{-2} exposure were gain corrected and aligned using MotionCor2 (*51*). Contrast-transfer function (CTF) parameters of the summed micrographs were estimated with Gctf (*52*) and CTFFIND4 (*53*), before micrographs were manually screened for quality. 1,690,969 (dataset 1) and 701,423 (dataset 2) particles from 11,270 (dataset 1) and 5,799 (dataset 2) micrographs were then picked using Gautomatch (dataset 1), or the Relion 3.1 AutoPick function (dataset 2) (*54, 55*). The particles from dataset 1 displayed severe orientation bias and were further processed using cryoSPARC (*56*). The particles from dataset 2 were further processed in Relion 3.1. All particles were subjected to extensive 2D and 3D classification in cryoSPARC (dataset 1) and Relion 3.1 (dataset 2) (fig. S2, C and D). Particles were then subjected to CTF parameter refinement in cryoSPARC (dataset 1) and Relion 3.1 (dataset 2), before final reconstructions were prepared. In case of the reconstituted Nsp1-40S complex (dataset 1), local refinements were performed in cryoSPARC using masks on the body and head of the 40S. The focused refined maps were filtered according to local resolution and sharpened using a B-factor of -80, before they were combined to a composite map using Phenix (*57*). Furthermore, both 43S PIC volumes (dataset 2) were subjected to multi-body refinement in Relion 3.1 using a masks on the 40S body, head and eIF3, as well as eIF2-tRNA in case of 43S PIC state 2. Finally, the local resolution of each reconstituted volume was estimated using cryoSPARC or Relion 3.1.

Model building and refinement.

The molecular model of Nsp1-40S was manually built in Coot (*58, 59*) using PDB-6G5H as an initial model of the small ribosomal subunit. All proteins and ribosomal RNA were checked and refined in Coot, before the C-terminal helices of Nsp1 were built *de novo*. The model was subsequently real-space refined in Phenix and the surface potential of Nsp1-C determined using DelPhi (*60*). Models of all Nsp1 initiation complexes were based on previously described structures of native PICs (61). Models of all Nsp1-80S ribosomal complexes were based on the structure of CCDC124-bound 80S (*30*). C-terminal domains of CCDC124 and LYAR were built de novo in Coot. Models for ABCE1 and eRF1 were based on PDB-5LZV and rigid-body fit into

the density before manually adjusting domain 3 of eRF1 in Coot. Models for the ternary eEF1A-tRNA-GTP complex were based on PDB-5LZS and rigid-body fitted into density. The sequence of pre-accommodated tRNA was manually changed to leucyl-tRNA and the extended variable loop built. All models were then real-space refined in Phenix.

Cryo-EM densities and molecular models were visualized using ChimeraX (*62*).
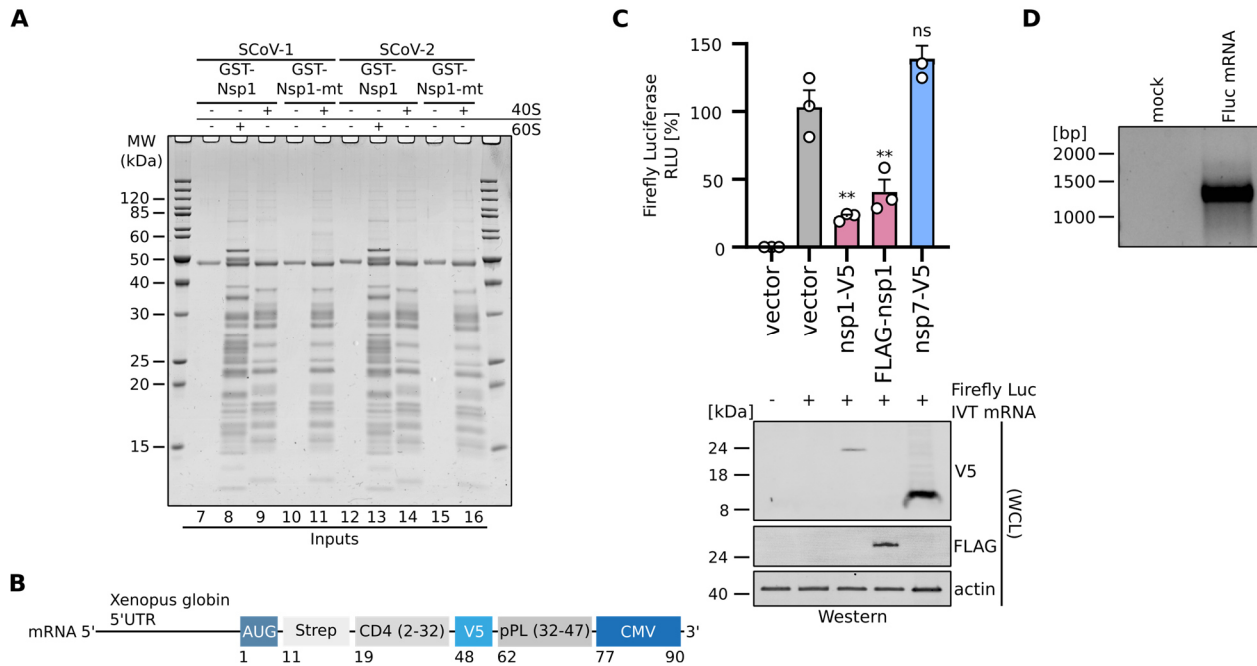
Fig. S1. Characterization of Nsp1 from SARS-CoV and SARS-CoV-2.
(**A**) Input samples of the *in vitro* binding assay described in Fig. 2B. The samples were analyzed by SDS-PAGE and Coomassie staining. SCoV-1, SARS-CoV. SCoV-2, SARS-CoV-2. (**B**) Organization of the reporter mRNA construct used in the *in vitro* translation assay in Fig. 1D: mRNA coding for an N-terminal Strep-tag, amino acids 2-32 of CD4, a V5-tag, amino acids 32-47 of preprolactin (pPL) and a sequence coding for the arrest peptide of gp48 uORF2 from cytomegalovirus (CMV). (**C**) Quantification of luciferase in HEK293T cells transiently transfected with indicated codon-optimized (V5) and non-codon-optimized (3xFLAG) plasmids (empty vector, Nsp1-V5, FLAG-Nsp1-FLAG, Nsp7-V5) and *in vitro* transcribed Firefly luciferase mRNA. Bars represent the mean of n=3±SEM. RLU, relative light units. Immunoblot of whole cell lysates of the reporter gene assay stained with anti-V5, anti-FLAG and anti-actin antibodies (bottom panel). (**D**) Agarose gel of the *in vitro* transcribed Firefly luciferase mRNA used in Fig. 1E and (**C**).
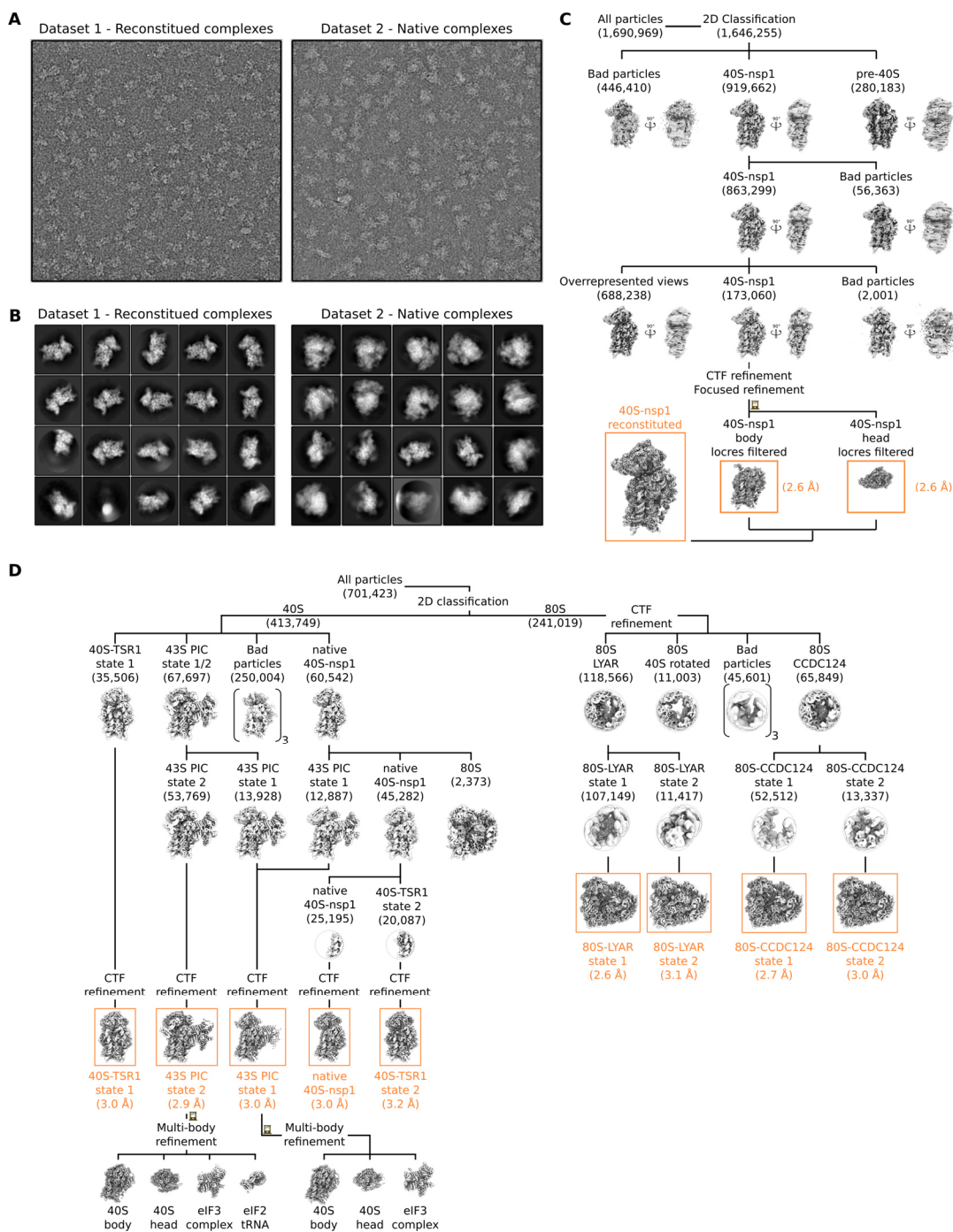
Fig. S2. Cryo-electron microscopy analysis.
(**A**) Representative electron micrographs from both reconstituted and native datasets, respectively, displayed with inverted contrast and after low-pass filtering at 20 Å. (**B**) Selected image averages after particle image 2D classification. (**C** and **D**) Cryo-EM data processing scheme for the reconstituted Nsp1-40S dataset (**C**) and affinity purified ribosomal complexes (**D**). Respective image numbers are provided in brackets and final volumes and their resolution marked in orange.
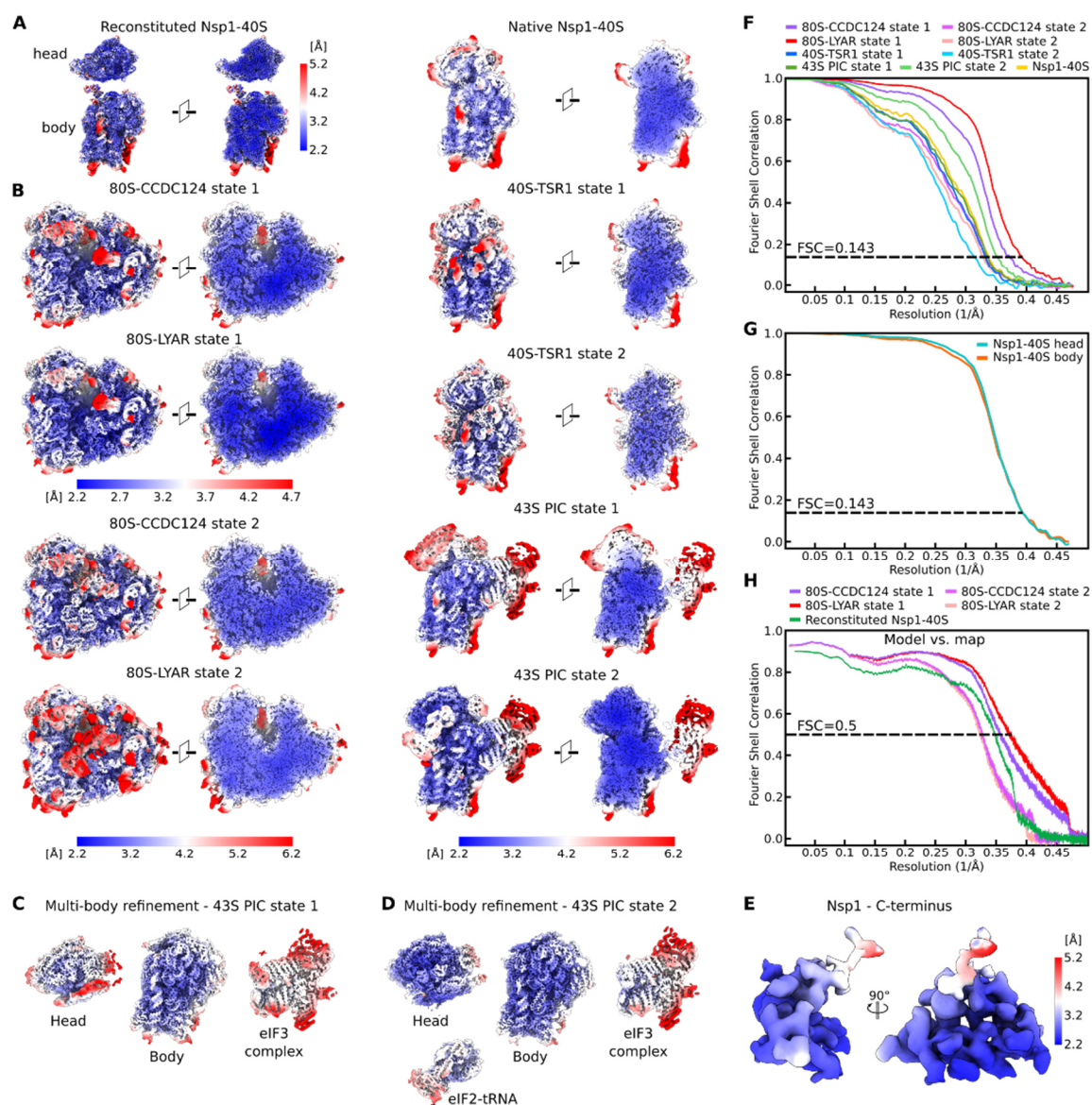
Fig. S3. Local resolution and model statistics.
(**A** to **E**) Final cryo-EM maps filtered and colored according to local resolution as estimated by cryoSPARC (**A** and **E**) and Relion (**B** to **D**). Reconstituted Nsp1-40S (**A**) and both 43S PIC states (**C** and **D**) were locally or multi-body refined in cryoSPARC and Relion, respectively, and their respective subvolumes are shown. Colors refer to local resolution averages as shown by the nearest color key to the right (**A** and **C** to **E**) or below each image (**B**). (**F** and **G**) Fourier shell correlation (FSC) curves of all native Nsp1-ribosomal complexes (**F**) and locally refined maps of reconstituted Nsp1-40S (**G**) as estimated by Relion. Threshold for final resolution estimation according to the 'gold-standard' set at FSC=0.143. (**H**) FSC curve of the final models against their respective cryo-EM maps as provided by Phenix.
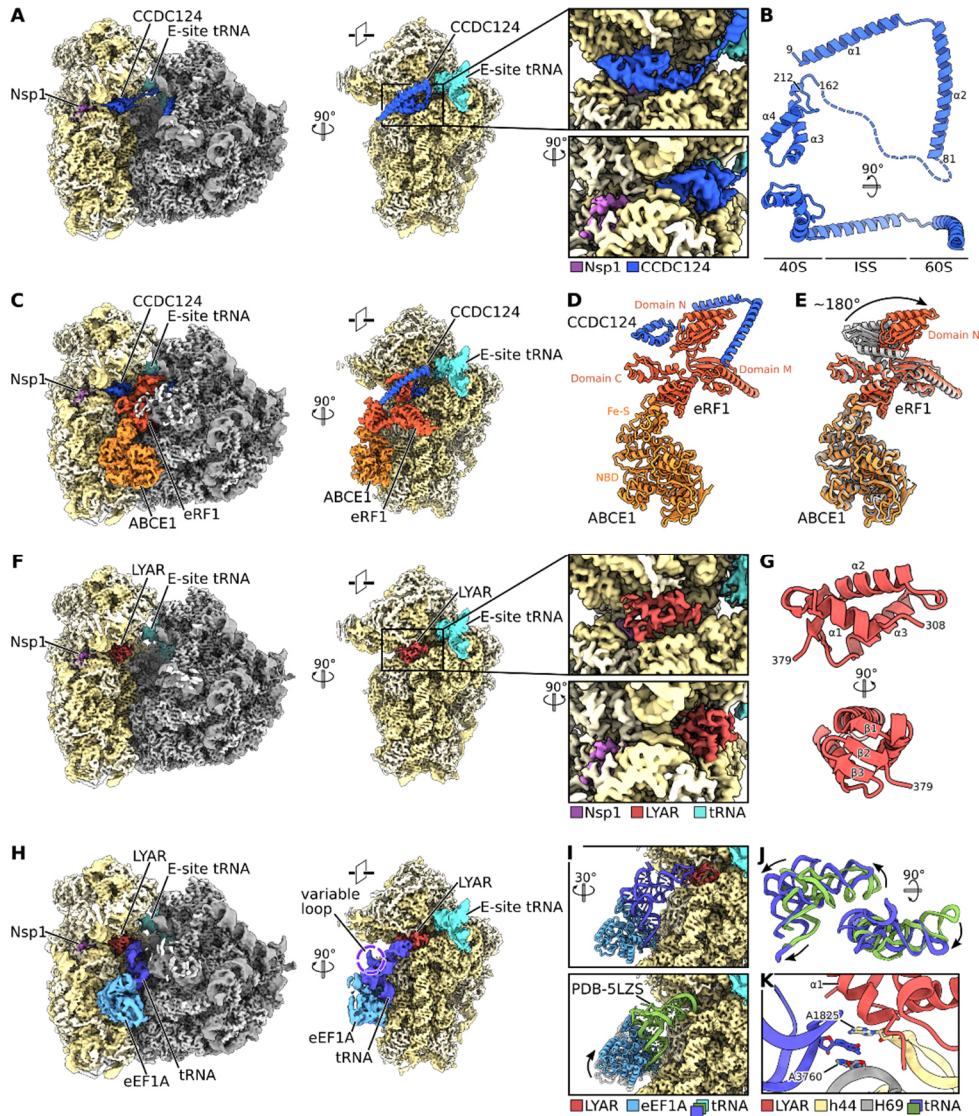
Fig. S4. Nsp1-bound 80S ribosomal complexes display unusual factor compositions.
(**A**) Cryo-EM map of Nsp1-80S bound by CCDC124 with its C-terminal domain positioned at the A-site. (**B**) Molecular model of CCDC124. CCDC124 forms two long alpha-helices (α1 and α2), which span the intersubunit side (ISS) and bind to the 60S as previously described (*30*). The newly observed C-terminal domain comprises two alpha-helices (α3 and α4) and is linked to α2 by a long flexible linker (dashed line). (**C**) Cryo-EM volume of CCDC124 bound Nsp1-80S with splitting and translation termination factors ABCE1 and eRF1. (**D**) Molecular models of ABCE1, eRF1 and CCDC124 with their domains labeled show the central position of eRF1 domain N encompassed by CCDC124. (**E**) Superposition of ABCE1-eRF1 from the Nsp1-80S volume (red and orange) and PDB-3JAH (gray) highlighting the rotated position of eRF1 domain N. Domain N, which is usually positioned within the A-site decoding the stop codon, is displaced by the C-terminal fold of CCDC124. (**F**) Cryo-EM volume of a Nsp1-80S complex, containing LYAR. LYAR occupies the mRNA channel at the A-site with its C-terminal fold. (**G**) Three alpha-helices (α1 – α3) and a

three-stranded beta-sheet (β1 – β3) form the observed part of LYAR in a α1-β1-α2-α3-β2-β3 sequence. (**H**) Cryo-EM volume of a Nsp1-80S ribosome with LYAR bound by the ternary eEF1A-tRNA-GTP complex. The large variable loop of the pre-accommodated type II tRNA is marked. (**I**) Molecular model of eEF1A, tRNA and LYAR bound to the 40S subunit volume shows the anticodon loop displaced by LYAR (top). Model of a canonical elongation complex (PDB-5LZS) docked into the cryo-EM map shows a different position of eEF1A and tRNA (bottom). (**J**) Superposition of the tRNA in the Nsp1-80S state and PDB-5LZS highlights the repositioned tRNA with a stretched decoding loop. (**K**) Molecular interactions between LYAR, tRNA and rRNA helices h44 and H69. Base stacking occurs between A3760 of the 28S rRNA, A1825 of the 18S rRNA and a base of the anticodon loop of the bound tRNA. A model of a leucyl-tRNA was used, however, local resolution prevented unambiguous identification of the bound tRNA.
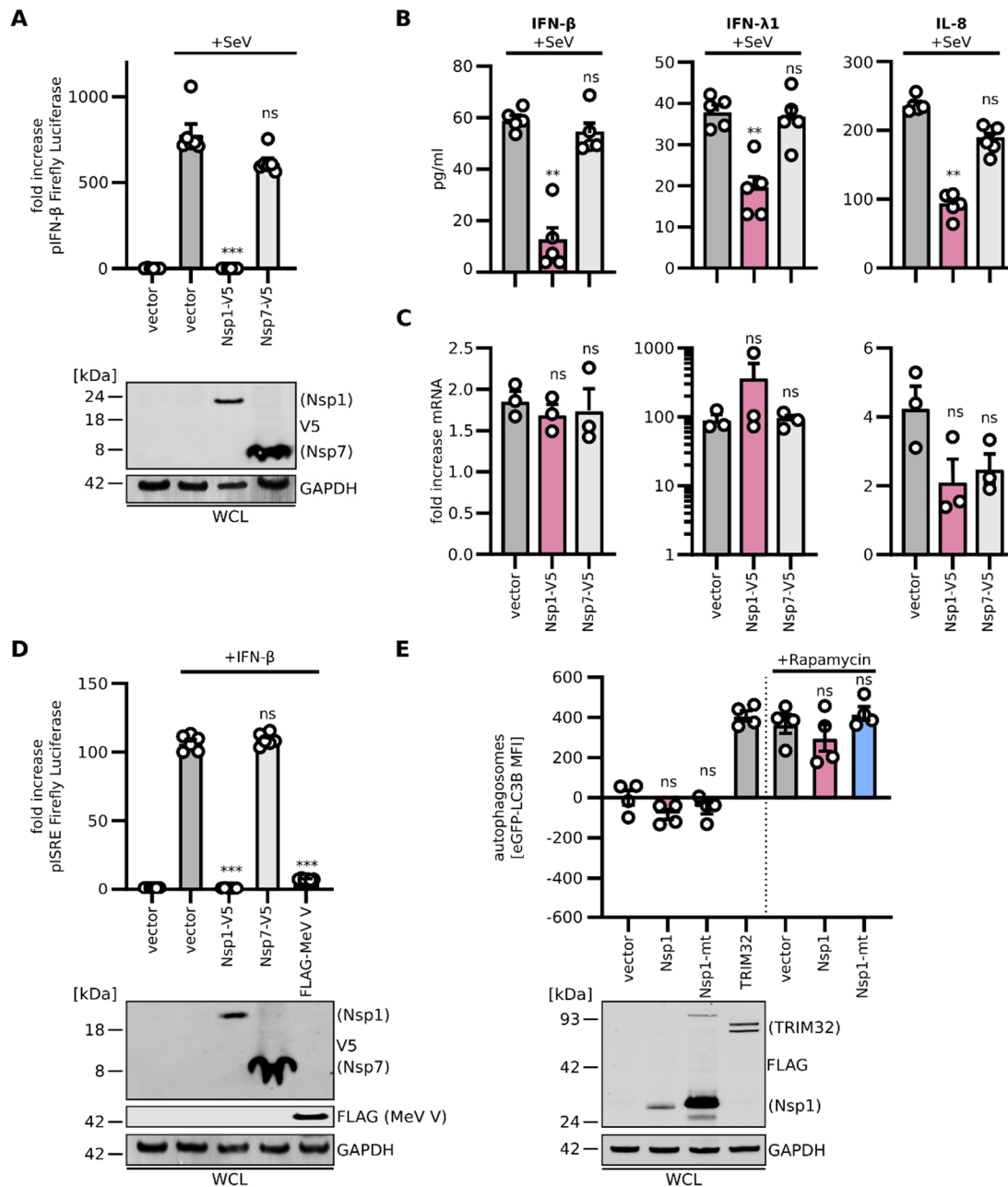
Fig. S5. Inhibition of the innate immune response by SARS-CoV-2 Nsp1. (A) Quantification of IFN-β-promotor controlled firefly luciferase activity in HEK293T cells transiently expressing indicated proteins. Cells were infected with Sendai Virus (SeV) for 24 h or left uninfected. Bars represent the mean (fold induction) of n=6 SEM. Representative immunoblots of whole cell lysates (WCLs) stained with anti-V5, anti-FLAG and anti-GAPDH (bottom panel). **(B)** ELISA of IFN-β, IFN-λ1 or IL-8 released by HEK293T cells transiently expressing indicated proteins and infected with SeV (top panel) for 24 h. Bars represent the mean (fold induction) of n=5±SEM. **(C)** qPCR of the corresponding mRNAs (bottom panel). Bars represent the mean (fold induction) of n=3±SEM. **(D)** Quantification of Interferon Stimulated Response Element (ISRE)-promotor

controlled firefly luciferase activity in HEK293T cells transiently expressing indicated proteins and treated with 1000 U/mL IFN-β for 8 h or left untreated. Bars represent the mean (fold induction) of n=6±SEM. Immunoblot of WCLs stained with anti-V5, anti-FLAG and anti-GAPDH (bottom panel). (**E**) Autophagy levels in HEK293T cells stably expressing GFP-LC3B and transiently transfected with the indicated plasmids. Autophagosomes were quantified by flow cytometry as mean fluorescence intensity of GFP-LC3B-positive vesicles in saponin-permeabilised cells. Bars represent the mean of n=3±SEM. The MFI of the vector control was set to 0. Representative immunoblots of WCLs in parallel stained with anti-FLAG and anti-GAPDH (bottom panel). Unpaired student's t-test (Welch correction), ns, not significant; **, $p < 0.001$, ***, $p < 0.0001$.

Table S1. Data collection, refinement and model statistics

	Nsp1-40S reconstituted (EMDB-11276) (PDB-6ZLW)	80S-CCDC124 state 1 (EMDB-11288) (PDB-6ZM7)	80S-CCDC124 state 2 (EMDB-11289) (PDB-6ZME)	80S-LYAR state 1 (EMDB-11292) (PDB-6ZMI)	80S-LYAR state 2 (EMDB-11299) (PDB-6ZMO)
Data collection and processing					
Camera	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit
Magnification	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e-Å^{-2})	44.8	44.8	44.8	44.8	44.8
Defocus range (μm)	0.5 - 3.0	0.5-3.0	0.5-3.0	0.5-3.0	0.5-3.0
Pixel size (Å)	1.059	1.059	1.059	1.059	1.059
Symmetry imposed	C1	C1	C1	C1	C1
Micrographs collected (no.)	11,270	6610	6610	6610	6610
Initial particle images (no.)	1,690,969	701,423	701,423	701,423	701,423
Final particle images (no.)	173,060	52,512	13,337	107,149	11,417
Map resolution (Å)	2.6	2.7	3.0	2.6	3.1
FSC threshold	0.143	0.143	0.143	0.143	0.143
Refinement					
Model resolution (Å)	2.7	2.7	3.0	2.6	3.1
FSC threshold	0.5	0.5	0.5	0.5	0.5
Map sharpening Bfactor (Å2)	-80	-20	-20	-20	-20
Model composition					
Non-hydrogen atoms	74,562	225,532	233,386	225,120	230,350
Protein residues	4,887	12,547	13,537	12,499	12,939
Nucleotide residues	1,665	5,864	5,864	5,864	5,951
Ligands	3	264	267	264	264
R.m.s deviations					
Bond lengths (Å)	0.008	0.007	0.006	0.007	0.009
Bond angles (°)	1.083	0.926	0.923	0.958	1.040
Validation					
Molprobity score	1.44	1.63	1.71	1.63	1.74
Clash score	3.70	3.66	4.39	3.59	4.67
Poor rotamers (%)	0.40	0.48	0.58	0.55	0.83
Ramachandran plot					
Favored (%)	95.05	92.20	91.79	92.22	91.50
Allowed (%)	3.82	7.57	8.00	7.48	8.25
Disallowed (%)	0.23	0.23	0.21	0.30	0.25
Map vs. Model CC (mask)	0.82	0.85	0.82	0.85	0.74
EMRinger score (Nsp1[148-180])	5.44				

Table S2. Data collection statistics.

	Native 40S-Nsp1 (EMDB-11276)	40S-TSR1 state 1 (EMDB-11301) (PDB-6ZMT)	40S-TSR1 state 2 (EMDB-11310) (PDB-6ZN5)	43S PIC state 1 (EMDB-11325) (PDB-6ZON)	43S PIC state 2 (EMDB-11335) (PDB-6ZP4)
Data collection and processing					
Camera	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit	Gatan K2 Summit
Magnification	130,000	130,000	130,000	130,000	130,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e⁻Å⁻²)	44.8	44.8	44.8	44.8	44.8
Defocus range (μm)	0.5-3.0	0.5-3.0	0.5-3.0	0.5-3.0	0.5-3.0
Pixel size (Å)	1.059	1.059	1.059	1.059	1.059
Symmetry imposed	C1	C1	C1	C1	C1
Micrographs collected (no.)	6610	6610	6610	6610	6610
Initial particle images (no.)	701,423	701,423	701,423	701,423	701,423
Final particle images (no.)	25,195	35,506	20,087	13,928	53,769
Map resolution (Å)	3.0	3.0	3.2	3.0	2.9
FSC threshold	0.143	0.143	0.143	0.143	0.143
Refinement					
Model resolution (Å)		3.0	3.1	3.1	2.9
FSC threshold		0.5	0.5	0.5	0.5
Map sharpening Bfactor (Å2)	-20	-20	-20	-20	-20
Model composition					
Non-hydrogen atoms		76,211	79,377	109,242	117,860
Protein residues		5,357	5,478	9,349	10,413
Nucleotide residues		1,569	1,671	1,721	1,796
Ligands		2	3	3	6
R.m.s deviations					
Bond lengths (Å)		0.008	0.011	0.010	0.006
Bond angles (°)		0.989	1.153	1.040	0.907
Validation					
Molprobity score		1.66	1.72	1.83	1.62
Clash score		5.44	5.08	7.70	4.87
Poor rotamers (%)		0.46	0.53	0.51	0.33
Ramachandran plot					
Favored (%)		94.64	92.93	93.81	94.69
Allowed (%)		5.17	6.94	6.17	5.25
Disallowed (%)		0.19	0.13	0.01	0.06
Map vs. Model CC (mask)		0.81	0.80	0.54	0.57

Data S1. Mass spectrometry of SARS-CoV-2 Nsp1 purified from HEK293T
Related to Fig. 2E

References and Notes

1. S. R. Weiss, S. Navas-Martin, Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus. *Microbiol. Mol. Biol. Rev.* **69**, 635–664 (2005). doi:10.1128/MMBR.69.4.635-664.2005 Medline

2. P. C. Woo, S. K. P. Lau, C. S. F. Lam, C. C. Y. Lau, A. K. L. Tsang, J. H. N. Lau, R. Bai, J. L. L. Teng, C. C. C. Tsang, M. Wang, B.-J. Zheng, K.-H. Chan, K.-Y. Yuen, Discovery of seven novel mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. *J. Virol.* **86**, 3995–4008 (2012). doi:10.1128/JVI.06540-11 Medline

3. J. Cui, F. Li, Z. L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019). doi:10.1038/s41579-018-0118-9 Medline

4. L. van der Hoek, Human coronaviruses: What do they cause? *Antivir. Ther.* **12**, 651–658 (2007). Medline

5. L. van der Hoek, K. Pyrc, B. Berkhout, Human coronavirus NL63, a new respiratory virus. *FEMS Microbiol. Rev.* **30**, 760–773 (2006). doi:10.1111/j.1574-6976.2006.00032.x Medline

6. E. de Wit, N. van Doremalen, D. Falzarano, V. J. Munster, SARS and MERS: Recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016). doi:10.1038/nrmicro.2016.81 Medline

7. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020). doi:10.1001/jama.2020.1585 Medline

8. Y. X. Lim, Y. L. Ng, J. P. Tam, D. X. Liu, Human coronaviruses: A review of virus-host interactions. *Diseases* **4**, 26 (2016). doi:10.3390/diseases4030026 Medline

9. P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020). doi:10.1038/s41586-020-2012-7 Medline

10. P. S. Masters, The molecular biology of coronaviruses. *Adv. Virus Res.* **66**, 193–292 (2006). doi:10.1016/S0065-3527(06)66005-3 Medline

11. L. Lei, S. Ying, L. Baojun, Y. Yi, H. Xiang, S. Wenli, S. Zounan, G. Deyin, Z. Qingyu, L. Jingmei, C. Guohui, Attenuation of mouse hepatitis virus by deletion of the LLRKxGxKG region of Nsp1. *PLOS ONE* **8**, e61166 (2013). doi:10.1371/journal.pone.0061166 Medline

12. Y. Tohya, K. Narayanan, W. Kamitani, C. Huang, K. Lokugamage, S. Makino, Suppression of host gene expression by nsp1 proteins of group 2 bat coronaviruses. *J. Virol.* **83**, 5282–5288 (2009). doi:10.1128/JVI.02485-08 Medline

13. K. Narayanan, C. Huang, K. Lokugamage, W. Kamitani, T. Ikegami, C.-T. K. Tseng, S. Makino, Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. *J. Virol.* **82**, 4471–4479 (2008). [doi:10.1128/JVI.02472-07](doi:10.1128/JVI.02472-07) Medline

14. C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, S. Makino, Alphacoronavirus transmissible gastroenteritis virus nsp1 protein suppresses protein translation in mammalian cells and in cell-free HeLa cell extracts but not in rabbit reticulocyte lysate. *J. Virol.* **85**, 638–643 (2011). [doi:10.1128/JVI.01806-10](doi:10.1128/JVI.01806-10) Medline

15. K. G. Lokugamage, K. Narayanan, C. Huang, S. Makino, Severe acute respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation. *J. Virol.* **86**, 13598–13608 (2012). [doi:10.1128/JVI.01958-12](doi:10.1128/JVI.01958-12) Medline

16. W. Kamitani, C. Huang, K. Narayanan, K. G. Lokugamage, S. Makino, A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat. Struct. Mol. Biol.* **16**, 1134–1140 (2009). doi:10.1038/nsmb.1680 Medline

17. C. Huang, K. G. Lokugamage, J. M. Rozovics, K. Narayanan, B. L. Semler, S. Makino, SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: Viral mRNAs are resistant to nsp1-induced RNA cleavage. *PLOS Pathog.* **7**, e1002433 (2011). doi:10.1371/journal.ppat.1002433 Medline

18. M. G. Wathelet, M. Orr, M. B. Frieman, R. S. Baric, Severe acute respiratory syndrome coronavirus evades antiviral signaling: Role of nsp1 and rational design of an attenuated strain. *J. Virol.* **81**, 11620–11633 (2007). [doi:10.1128/JVI.00702-07](doi:10.1128/JVI.00702-07) Medline

19. C. Wu, Y. Liu, Y. Yang, P. Zhang, W. Zhong, Y. Wang, Q. Wang, Y. Xu, M. Li, X. Li, M. Zheng, L. Chen, H. Li, Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm. Sin. B* **10**, 766–788 (2020). doi:10.1016/j.apsb.2020.02.008 Medline

20. A. R. Jauregui, D. Savalia, V. K. Lowry, C. M. Farrell, M. G. Wathelet, Identification of residues of SARS-CoV nsp1 that differentially affect inhibition of gene expression and antiviral signaling. *PLOS ONE* **8**, e62416 (2013). doi:10.1371/journal.pone.0062416 Medline

21. M. S. Almeida, M. A. Johnson, T. Herrmann, M. Geralt, K. Wüthrich, Novel β-barrel fold in the nuclear magnetic resonance structure of the replicase nonstructural protein 1 from the severe acute respiratory syndrome coronavirus. *J. Virol.* **81**, 3151–3161 (2007). [doi:10.1128/JVI.01939-06](doi:10.1128/JVI.01939-06) Medline

22. F. Schluenzen, A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, H. Bartels, I. Agmon, F. Franceschi, A. Yonath, Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102**, 615–623 (2000). [doi:10.1016/S0092-8674(00)00084-2](doi:10.1016/S0092-8674(00)00084-2) Medline

23. L. A. Passmore, T. M. Schmeing, D. Maag, D. J. Applefield, M. G. Acker, M. A. Algire, J. R. Lorsch, V. Ramakrishnan, The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol. Cell* **26**, 41–50 (2007). doi:10.1016/j.molcel.2007.03.018 Medline

24. M. Ameismeier, J. Cheng, O. Berninghausen, R. Beckmann, Visualizing late states of human 40S ribosomal subunit maturation. *Nature* **558**, 249–253 (2018). doi:10.1038/s41586-018-0193-0 Medline

25. A. Heuer, E. Thomson, C. Schmidt, O. Berninghausen, T. Becker, E. Hurt, R. Beckmann, Cryo-EM structure of a late pre-40S ribosomal subunit from *Saccharomyces cerevisiae*. *eLife* **6**, e30189 (2017). doi:10.7554/eLife.30189 Medline

26. A. des Georges, V. Dhote, L. Kuhn, C. U. T. Hellen, T. V. Pestova, J. Frank, Y. Hashem, Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525**, 491–495 (2015). doi:10.1038/nature14891 Medline

27. Y. Hashem, A. des Georges, V. Dhote, R. Langlois, H. Y. Liao, R. A. Grassucci, C. U. T. Hellen, T. V. Pestova, J. Frank, Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119 (2013). doi:10.1016/j.cell.2013.04.036 Medline

28. J. L. Llácer, T. Hussain, L. Marler, C. E. Aitken, A. Thakur, J. R. Lorsch, A. G. Hinnebusch, V. Ramakrishnan, Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol. Cell* **59**, 399–412 (2015). doi:10.1016/j.molcel.2015.06.033 Medline

29. Y. J. Wang, P. P. Vaidyanathan, M. F. Rojas-Duran, N. D. Udeshi, K. M. Bartoli, S. A. Carr, W. V. Gilbert, Lso2 is a conserved ribosome-bound protein required for translational recovery in yeast. *PLOS Biol.* **16**, e2005903 (2018). doi:10.1371/journal.pbio.2005903 Medline

30. J. N. Wells, R. Buschauer, T. Mackens-Kiani, K. Best, H. Kratzat, O. Berninghausen, T. Becker, W. Gilbert, J. Cheng, R. Beckmann, Structure and function of yeast Lso2 and human CCDC124 bound to hibernating ribosomes. bioRxiv 944066 [Preprint]. 12 February 2020; https://doi.org/10.1101/2020.02.12.944066.

31. T. Becker, S. Franckenberg, S. Wickles, C. J. Shoemaker, A. M. Anger, J.-P. Armache, H. Sieber, C. Ungewickell, O. Berninghausen, I. Daberkow, A. Karcher, M. Thomm, K.-P. Hopfner, R. Green, R. Beckmann, Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature* **482**, 501–506 (2012). doi:10.1038/nature10829 Medline

32. A. Preis, A. Heuer, C. Barrio-Garcia, A. Hauser, D. E. Eyler, O. Berninghausen, R. Green, T. Becker, R. Beckmann, Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep.* **8**, 59–65 (2014). doi:10.1016/j.celrep.2014.04.058 Medline

33. A. Brown, S. Shao, J. Murray, R. S. Hegde, V. Ramakrishnan, Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493–496 (2015). doi:10.1038/nature14896 Medline

34. N. Miyazawa, H. Yoshikawa, S. Magae, H. Ishikawa, K. Izumikawa, G. Terukina, A. Suzuki, S. Nakamura-Fujiyama, Y. Miura, T. Hayano, W. Komatsu, T. Isobe, N. Takahashi, Human cell growth regulator Ly-1 antibody reactive homologue accelerates processing of preribosomal RNA. *Genes Cells* **19**, 273–286 (2014). doi:10.1111/gtc.12129 Medline

35. C. Yang, X. Liu, T. Cheng, R. Xiao, Q. Gao, F. Ming, M. Jin, H. Chen, H. Zhou, LYAR suppresses beta interferon induction by targeting phosphorylated interferon regulatory factor 3. *J. Virol.* **93**, e00769-19 (2019). doi:10.1128/JVI.00769-19 Medline

36. M. P. Robertson, H. Igel, R. Baertsch, D. Haussler, M. Ares Jr., W. G. Scott, The structure of a rigorously conserved RNA element within the SARS virus genome. *PLOS Biol.* **3**, e5 (2005). doi:10.1371/journal.pbio.0030005 Medline

37. K. M. Sparrer, M. U. Gack, Intracellular detection of viral nucleic acids. *Curr. Opin. Microbiol.* **26**, 1–9 (2015). doi:10.1016/j.mib.2015.03.001 Medline

38. Y. Hu, W. Li, T. Gao, Y. Cui, Y. Jin, P. Li, Q. Ma, X. Liu, C. Cao, The severe acute respiratory syndrome coronavirus nucleocapsid inhibits type I interferon production by interfering with TRIM25-mediated RIG-I ubiquitination. *J. Virol.* **91**, e02143-16 (2017). doi:10.1128/JVI.02143-16 Medline

39. L. Strähle, J.-B. Marq, A. Brini, S. Hausmann, D. Kolakofsky, D. Garcin, Activation of the beta interferon promoter by unnatural Sendai virus infection requires RIG-I and is inhibited by viral C proteins. *J. Virol.* **81**, 12227–12237 (2007). doi:10.1128/JVI.01300-07 Medline

40. K. M. Sparrer, C. K. Pfaller, K. K. Conzelmann, Measles virus C protein interferes with Beta interferon transcription in the nucleus. *J. Virol.* **86**, 796–805 (2012). doi:10.1128/JVI.05899-11 Medline

41. K. Brzózka, S. Finke, K. K. Conzelmann, Identification of the rabies virus alpha/beta interferon antagonist: Phosphoprotein P interferes with phosphorylation of interferon regulatory factor 3. *J. Virol.* **79**, 7673–7681 (2005). doi:10.1128/JVI.79.12.7673-7681.2005 Medline

42. P. Devaux, V. von Messling, W. Songsungthong, C. Springfeld, R. Cattaneo, Tyrosine 110 in the measles virus phosphoprotein is required to block STAT1 phosphorylation. *Virology* **360**, 72–83 (2007). doi:10.1016/j.virol.2006.09.049 Medline

43. K. M. J. Sparrer, S. Gableske, M. A. Zurenski, Z. M. Parker, F. Full, G. J. Baumgart, J. Kato, G. Pacheco-Rodriguez, C. Liang, O. Pornillos, J. Moss, M. Vaughan, M. U. Gack, TRIM23 mediates virus-induced autophagy via activation of TBK1. *Nat. Microbiol.* **2**, 1543–1557 (2017). doi:10.1038/s41564-017-0017-2 Medline

44. M. Di Rienzo, M. Piacentini, G. M. Fimia, A TRIM32-AMBRA1-ULK1 complex initiates the autophagy response in atrophic muscle cells. *Autophagy* **15**, 1674–1676 (2019). doi:10.1080/15548627.2019.1635385 Medline

45. M. Z. Tay, C. M. Poh, L. Rénia, P. A. MacAry, L. F. P. Ng, The trinity of COVID-19: Immunity, inflammation and intervention. *Nat. Rev. Immunol.* **20**, 363–374 (2020). doi:10.1038/s41577-020-0311-8 Medline

46. T. Tanaka, W. Kamitani, M. L. DeDiego, L. Enjuanes, Y. Matsuura, Severe acute respiratory syndrome coronavirus nsp1 facilitates efficient propagation in cells through a specific translational shutoff of host mRNA. *J. Virol.* **86**, 11128–11137 (2012). doi:10.1128/JVI.01700-12 Medline

47. V. Bhaskar, A. Graff-Meyer, A. D. Schenk, S. Cavadini, O. von Loeffelholz, S. K. Natchiar, C. G. Artus-Revel, H.-R. Hotz, G. Bretones, B. P. Klaholz, J. A. Chao, Dynamics of uS19 C-terminal tail during the translation elongation cycle in human ribosomes. *Cell Rep.* **31**, 107473 (2020). doi:10.1016/j.celrep.2020.03.037 Medline

48. S. Matheisl, O. Berninghausen, T. Becker, R. Beckmann, Structure of a human translation termination complex. *Nucleic Acids Res.* **43**, 8615–8626 (2015). [doi:10.1093/nar/gkv909](doi:10.1093/nar/gkv909) Medline

49. C. Sidrauski, A. M. McGeachy, N. T. Ingolia, P. Walter, The small molecule ISRIB reverses the effects of eIF2α phosphorylation on translation and stress granule assembly. *eLife* **4**, e05033 (2015). doi:10.7554/eLife.05033 Medline

50. A. Sharma, M. Mariappan, S. Appathurai, R. S. Hegde, In vitro dissection of protein translocation into the mammalian endoplasmic reticulum. *Methods Mol. Biol.* **619**, 339–363 (2010). [doi:10.1007/978-1-60327-412-8_20](doi:10.1007/978-1-60327-412-8_20) Medline

51. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

52. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). [doi:10.1016/j.jsb.2015.11.003](doi:10.1016/j.jsb.2015.11.003) Medline

53. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015). [doi:10.1016/j.jsb.2015.08.008](doi:10.1016/j.jsb.2015.08.008) Medline

54. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166–e42166 (2018). doi:10.7554/eLife.42166 Medline

55. J. Zivanov, T. Nakane, S. H. W. Scheres, Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in *RELION*-3.1. *IUCrJ* **7**, 253–267 (2020). [doi:10.1107/S2052252520000081](doi:10.1107/S2052252520000081) Medline

56. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi:10.1038/nmeth.4169 Medline

57. T. Burnley, C. M. Palmer, M. Winn, Recent developments in the CCP-EM software suite. *Acta Crystallogr. D Struct. Biol.* **73**, 469–477 (2017). [doi:10.1107/S2059798317007859](doi:10.1107/S2059798317007859) Medline

58. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). [doi:10.1107/S0907444904019158](doi:10.1107/S0907444904019158) Medline

59. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010). [doi:10.1107/S0907444910007493](doi:10.1107/S0907444910007493) Medline

60. S. Sarkar, S. Witham, J. Zhang, M. Zhenirovskyy, W. Rocchia, E. Alexov, DelPhi Web Server: A comprehensive online suite for electrostatic calculations of biological

macromolecules and their complexes. *Commun. Comput. Phys.* **13**, 269–284 (2013). doi:10.4208/cicp.300611.201011s Medline

61. H. Kratzat, T. Mackens-Kiani, M. Ameismeier, J. Cheng, E. Dacheux, A. Namane, O. Berninghausen, M. Fromont-Racine, T. Becker, R. Beckmann, Structural inventory of native ribosomal ABCE1-43S pre-initiation complexes. bioRxiv 194902 [Preprint]. 9 July 2020; https://doi.org/10.1101/2020.07.09.194902.

62. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi:10.1002/pro.3235 Medline

A structural inventory of native ribosomal ABCE1-43S pre-initiation complexes

Hanna Kratzat[1,†], Timur Mackens-Kiani[1,†], Michael Ameismeier[1] (iD), Mia Potocnjak[1],
Jingdong Cheng[1] (iD), Estelle Dacheux[2], Abdelkader Namane[2] (iD), Otto Berninghausen[1], Franz Herzog[1],
Micheline Fromont-Racine[2] (iD), Thomas Becker[1,*] (iD) & Roland Beckmann[1,**] (iD)

Abstract

In eukaryotic translation, termination and ribosome recycling phases are linked to subsequent initiation of a new round of translation by persistence of several factors at ribosomal sub-complexes. These comprise/include the large eIF3 complex, eIF3j (Hcr1 in yeast) and the ATP-binding cassette protein ABCE1 (Rli1 in yeast). The ATPase is mainly active as a recycling factor, but it can remain bound to the dissociated 40S subunit until formation of the next 43S pre-initiation complexes. However, its functional role and native architectural context remains largely enigmatic. Here, we present an architectural inventory of native yeast and human ABCE1-containing pre-initiation complexes by cryo-EM. We found that ABCE1 was mostly associated with early 43S, but also with later 48S phases of initiation. It adopted a novel hybrid conformation of its nucleotide-binding domains, while interacting with the N-terminus of eIF3j. Further, eIF3j occupied the mRNA entry channel via its ultimate C-terminus providing a structural explanation for its antagonistic role with respect to mRNA binding. Overall, the native human samples provide a near-complete molecular picture of the architecture and sophisticated interaction network of the 43S-bound eIF3 complex and the eIF2 ternary complex containing the initiator tRNA.

Introduction

Translation of an mRNA into a polypeptide sequence is a central cellular process, which is highly regulated and linked to other cellular processes like ribosome biogenesis, mRNA turnover, and ribosome quality control. Most decisive for translational efficiency and regulation is the initiation phase; however, in eukaryotes the

individual phases of translation were found to be coupled, especially termination with ribosome recycling and a new round of initiation. Two prominent examples are the conserved multisubunit complex eIF3, which has been described as a factor functioning across the translation cycle (Valasek *et al*, 2017), as well as the ATP-binding cassette (ABC) ATPase ABCE1 (Rli1 in *Saccharomyces cerevisiae*), which was shown to enhance termination activity of the eRF1 release factor and which represents the key enzyme for ATP-dependent ribosome recycling (Pisarev *et al*, 2010; Shoemaker & Green, 2011). Moreover, ABCE1 was found associated with initiation factors (Chen *et al*, 2006; Dong *et al*, 2004) and as a part of eIF3-containing 43S or 48S pre-initiation complexes (Andersen & Leevers, 2007; Preis *et al*, 2014; Mancera-Martinez *et al*, 2017).

The ABCE1 ATPase consists of two nucleotide-binding domains (NBDs) that are forming two nucleotide-binding sites (NBSs) at their interface, as well as an essential iron–sulfur cluster domain (FeSD) at its N-terminus (Barthelme *et al*, 2007; Hopfner, 2016). ABCE1 binds the 80S ribosome during canonical stop codon-dependent termination or during rescue of stalled ribosomes and splits the 80S ribosomes into 40S and 60S small (SSU) and large (LSU) subunits, respectively. This recycling reaction requires an A site factor in the ribosome, either release factor eRF1 (after termination) or its homologue Pelota (Dom34 in *S.c.*; for ribosome rescue), in order to form part of the interaction network for ABCE1 (Becker *et al*, 2012; Brown *et al*, 2015; Preis *et al*, 2014). ABCE1 binds these pre-splitting complexes in a semi-open state with respect to its NBSs. Splitting requires binding of ATP and site-occlusion to both NBS (Barthelme *et al*, 2011; Gouridis *et al*, 2019; Nurenberg-Goloub *et al*, 2018). According to current models, the conformational change occurring during site-occlusion would be transmitted *via* the FeSD of ABCE1 to the bound A site factor (eRF1 or Dom34), whereby the FeSD exerts a force on the A site factor which ultimately leads to ribosome splitting (Becker *et al*, 2012; Heuer *et al*, 2017; Nürenberg-Goloub *et al*, 2020). The splitting reaction can be recapitulated *in vitro* (Becker *et al*, 2012; Nurenberg-Goloub & Tampe, 2019; Pisareva *et al*, 2011; Shao *et al*, 2015; Shoemaker & Green, 2011), where ABCE1 was observed to remain bound to the 40S small subunit to form a post-splitting complex (PSC), in

1 Gene Center and Center for Integrated Protein Science Munich, Department of Biochemistry, University of Munich, Munich, Germany
2 Génétique des Interactions Macromoléculaires, UMR3525 CNRS, Institut Pasteur, Paris, France
   *Corresponding author. Tel: +49 89 2180 76915; E-mail: becker@genzentrum.lmu.de
   **Corresponding author. Tel: +49 89 2180 76900; E-mail: beckmann@genzentrum.lmu.de
   †These authors contributed equally to this work

which the two NBDs are present in a closed, nucleotide-occluding state (Heuer et al, 2017; Kiosze-Becker et al, 2016; Nürenberg-Goloub et al, 2020). Therefore, it was assumed that in vivo as well, ABCE1 may remain bound to the 40S for a defined time span (Gerovac & Tampe, 2019) to prevent re-association of the LSU (Heuer et al, 2017) or to coordinate assembly of initiation factors on the 40S subunit. However, a direct physical involvement of ABCE1 in the translation initiation process has not been shown to date.

In eukaryotes, the start of translation initiation requires the assembly of the 43S pre-initiation complex (PIC). It consists of the 40S subunit, eIF3, eIF1, eIF1A, eIF5, and the ternary complex (TC) formed by the trimeric eIF2-$\alpha\beta\gamma$, initiator methionyl tRNA (tRNA$_i$), and GTP. After 43S PIC assembly, the mRNA—in collaboration with the eIF4F complex (the cap-binding protein eIF4E, the helicase eIF4A, and the scaffolding protein eIF4G)—can be recruited to the 43S PIC, forming the 48S initiation complex (IC). This event is coordinated by interactions between eIF3 and eIF4F as well as eIF4B, a single-stranded RNA-binding protein that attaches to the 40S subunit (Walker et al, 2013) and stimulates the helicase activity of eIF4A. The 48S complex then scans the mRNA for the first cognate AUG codon. After start-codon recognition, inorganic phosphate (P$_i$) is released from the eIF2 complex, which is stimulated by eIF5 acting as a GTPase-activating protein, likely via an arginine-finger mechanism (Algire et al, 2005; Das et al, 2001; Paulin et al, 2001). Subsequently, initiation factors apart from eIF1A and eIF3 dissociate (Mohammad et al, 2017; Sha et al, 2009) and subunit joining with the 60S LSU is then mediated by the GTPase eIF5B (Acker et al, 2006; Acker et al, 2009; Lee et al, 2002; Pestova et al, 2000).

An important regulatory and scaffolding role in these processes is taken on by the multisubunit complex eIF3 (Cate, 2017; Hinnebusch, 2006), which can be structurally divided into the so-called PCI-MPN core and the peripheral subunits. In yeast, the PCI-MPN core consists of the two subunits eIF3a (Rpg1/Tif32) and eIF3c (Nip1), whereas in mammals, it is formed by an octamer of eIFs 3a, 3c, 3e, 3f, 3h, 3i, 3k, and 3l (Valasek et al, 2017). The peripheral subunits consist of the so-called yeast-like core (YLC) module, containing eIF3b (Prt1), eIF3g (Tif35), and eIF3i (Tif34), as well as the C-terminus of eIF3a, the N-terminal domain of eIF3c that interacts with eIF1 and eIF5 (Valasek et al, 2003; Valasek et al, 2004; Yamamoto et al, 2005; Zeman et al, 2019), and in mammals eIF3d. In addition, eIF3j is associated with eIF3 but does not belong to its core, and plays a special role (Block et al, 1998; Valasek et al, 1999). It was shown that eIF3j participates during termination by recycling eRF3 (Beznoskova et al, 2013) and during ribosome recycling by assisting ABCE1 in subunit splitting (Young & Guydosh, 2019). Furthermore, it is involved in dissociation of mRNA from the 40S subunit (Pisarev et al, 2007; Pisarev et al, 2010). In the context of initiation, eIF3j is believed to participate in the recruitment of eIF3 to the 40S (Elantak et al, 2010; Fraser et al, 2004; Nielsen et al, 2006), to antagonize premature mRNA recruitment (Fraser et al, 2007), and to regulate start-site selection (Elantak et al, 2010).

For a better mechanistic understanding of this complicated interplay, a number of cryo-EM structures of 43S PICs and partial 48S ICs gave first insights into the architectural variety of initiation complexes (Aylett et al, 2015; des Georges et al, 2015; Eliseev et al, 2018; Erzberger et al, 2014; Hashem et al, 2013; Hussain et al, 2014; Llacer et al, 2015; Llacer et al, 2018; Mancera-Martinez et al, 2017).

During 43S assembly, the 40S subunit gets prepared to thread the mRNA into the mRNA-binding channel between the 40S body and the head. The main constriction for mRNA is at the so-called "latch", a structural element formed between ribosomal RNA (rRNA) helix h18 and ribosomal protein (r-protein) uS12 on the 40S body, and h34 and uS3 on the head (Schluenzen et al, 2000). Empty or only ABCE1-bound 40S usually does not adopt a defined head conformation, and the latch is rather closed (Heuer et al, 2017; Passmore et al, 2007). Binding of eIF1 and especially eIF1A, which bridges the body with the head, seems to prime and confine the 40S by inducing a small rotation of the 40S head (Llacer et al, 2015; Passmore et al, 2007), but the latch still remains in a closed position (Llacer et al, 2015). Latch opening was only observed in in vitro reconstituted partial 48S ICs containing mRNA and both eIF3 and the eIF2 TC in addition to eIF1 and eIF1A (Llacer et al, 2015; Llacer et al, 2018). Here, two conformations of the 48S IC can be distinguished: the open P$_{OUT}$ and the closed P$_{IN}$ conformation, which differ in the orientation of the 40S head and the TC. Compared to the empty and eIF1/1A-bound structures, the head is moved upwards away from the body in the P$_{OUT}$ conformation. This leads to widening of the latch and the P site tRNA$_i$ in the TC is only bound via the anticodon loop (AL) to the 40S head but not the body. In the P$_{IN}$ conformation, the AL moves down and engages in stable codon–anticodon interactions with the cognate start codon in the P site, accompanied by a downward movement of the 40S head.

In all eIF3-containing structures, the PCI-MPN core was located on the back of the 40S subunit, from where peripheral subunits stretch out. In 43S PICs, the YLC was found close to the mRNA entry site of the 40S (Aylett et al, 2015; des Georges et al, 2015; Eliseev et al, 2018; Erzberger et al, 2014), however only at low resolution. Moreover, the YLC module has been shown to relocate to the intersubunit space (ISS), as observed in in vitro reconstituted partial 48S complexes (Llacer et al, 2015), thereby occupying the position of ABCE1. The other peripheral subunits eIF3d and the eIF3c N-terminal domain have been localized near the mRNA exit site (eIF3d: Eliseev et al, 2018) and in the ISS (eIF3c-NTD: Llacer et al, 2015; Obayashi et al, 2017). Interestingly, two structures of partial native 43S/48S complexes exist in which ABCE1 could be visualized in substantial quantities (Simonetti et al, 2016, re-interpreted in Mancera-Martinez et al, 2017; Heuer et al, 2017). Notably, both samples were obtained after adding non-hydrolyzable AMP-PNP and/or GMP-PNP to either yeast (Heuer et al, 2017) or rabbit reticulocyte (Simonetti et al, 2016) lysates and subsequent isolation of the 43S peak from a sucrose gradient. This may have led to non-physiological locking of ABCE1 on the 40S subunit, thereby limiting any conclusions about a putative role of ABCE1 during the phase connecting recycling with initiation. Furthermore, apart from a low-resolution cryo-EM map (Aylett et al, 2015) no structural data exist on eIF3j in the context of the native 43S PIC. Therefore, the native structural landscape enabling the transition from translation termination via recycling to initiation is not yet well-understood.

Results

In this work, we set out to provide a structural inventory of ABCE1-containing 43S or 48S initiation complexes from native small ribosomal subunits (SSU). We first asked if substantial amounts of ABCE1

are associated with initiation factor-bound 40S under native conditions. To that end, lysates from a yeast strain (*S.c.*) containing TAP-tagged ABCE1 (Rli1) were subjected to density gradient centrifugation followed by Western blotting of fractions (Fig EV1A). In agreement with previous studies (Andersen & Leevers, 2007; Pisarev *et al*, 2010; Pisareva *et al*, 2011), we observed that ABCE1 was especially enriched on 40S and 80S ribosomes. We further performed affinity purification from the lysates under varying buffer conditions but without any stabilizing non-hydrolyzable ATP or GTP analogs, and analyzed the elution fractions by quantitative mass spectrometry (LC-MS/MS) (Figs 1A and EV1B and C). We found that the expected SSU proteins but also eIF3 core components and especially eIF3j (Hcr1) were enriched by ABCE1 affinity purification, indicating that both proteins were indeed integral components of native pre-initiation complexes. Because of this finding and since eIF3j was implicated in ABCE1-dependent ribosome splitting *in vivo* (Young & Guydosh, 2019), we tested if eIF3j together with ABCE1 had a direct impact on ribosome splitting in a reconstituted system. To this end, we performed *in vitro* splitting assays in yeast and tested if eIF3j can play a stimulatory role. Purified 80S ribosomes were incubated with the purified splitting factors Dom34, Hbs1, Rli1 (ABCE1), eIF6 to prevent re-association of ribosomal subunits, ATP and GTP as well as different amounts of eIF3j. Splitting efficiency was assessed from sucrose density gradient UV profiles by monitoring 80S versus ribosomal subunit amounts (Figs 1B and C, and EV1D). Indeed, we observed that an addition of eIF3j in molar excess increased the ratio of split subunits to 80S when compared to a reaction containing the splitting factors only (Fig 1C). Increasing amounts of eIF3j resulted in higher splitting activity. However, eIF3j alone did not exhibit any activity (Fig EV1E). In addition, we found that eIF3j and substoichiometric amounts of ABCE1 remained bound to the 40S after splitting (Fig EV1F). To further confirm that eIF3j can still be associated with the 40S-ABCE1 complex after splitting, we employed the "facilitated splitting" assay as described before (Heuer *et al*, 2017). In this assay, ribosomes are allowed to dissociate under splitting-promoting conditions (low Mg^{2+} and high salt) and in the presence of putative subunit-binding factors (see Materials and Methods). Indeed, in this assay we observed that eIF3j remained on the 40S SSU together with ABCE1, confirming that the two factors remain together on the 40S for downstream events such as initiation after collaborating during splitting (Fig EV1G and H).

To gain further insights into the composition of native small subunits in yeast and human cells, we adopted a shotgun cryo-EM approach. Yeast SSU complexes were obtained after harvesting the crude 43S/48S peak from a preparative sucrose density gradient of yeast cell lysate that was not further treated or stabilized with a non-hydrolyzable nucleotide analog. Similarly, human native 40S was obtained from untreated lysates of HEK Flp-In 293 T-Rex cells after serendipitous non-specific enrichment on sepharose material during unrelated affinity pullouts (see Materials and Methods). Of these samples, large enough cryo-EM data sets were collected in order to analyze their complex composition by extensive 3D classification (Appendix Figs S1 and S2).

In the yeast data set, as expected, the selected particles contained pre-initiation complexes, which could be further classified into defined states varying in composition and conformation of eIF-associated 40S subunits. The majority of these complexes (62%) contained ABCE1, and the most interesting classes consisted of 43S

particles containing ABCE1, eIF3, eIF1, eIF1A, and eIF3j on the 40S (Aylett *et al*, 2015; Heuer *et al*, 2017). The mRNA path (latch) was in the closed conformation (Passmore *et al*, 2007), and at the mRNA entry, we found a density for a typical RNA recognition motif (RRM) (see below). Importantly, in these classes we observed an interaction between the FeSD of ABCE1 and eIF3j (Fig 1D). Moreover, we found one class of particles with mRNA bound, apparently representing a partial 48S IC complex. It contained eIF3, eIF1, tRNA_i in the P_{IN} conformation, as well as the N-terminal domain (NTD) of eIF5 as observed before (Llacer *et al*, 2018), and, to our surprise, also ABCE1 (Fig 1E). The classes representing 43S PIC and 48S IC were refined to a resolution of 5.3 and 6.2 Å, respectively, allowing us to fit molecular models of existing structures as rigid bodies (Fig 1D and E, Appendix Fig S3, Appendix Table S1).

In the human sample, we also found 40S subunits associated with initiation factors, similar to the yeast sample. After classification, four major stable eIF3-containing classes could be obtained (Fig 2A). The 40S in State I resembled the state of an empty 40S subunit with a closed latch (Heuer *et al*, 2017; Passmore *et al*, 2007), and only the core eIF3 subunits and weakly bound eIF1 were found. State II had a similar conformation, and we found extra densities in the ISS for eIF1, eIF3j, and ABCE1. State III additionally contained eIF1A and the ternary eIF2-GTP-tRNA_i complex (TC) in the open P_{OUT} conformation (Llacer *et al*, 2015), whereas State IV was similar to State III but lacked ABCE1. Notably, in contrast to the yeast sample, we did not find any 48S classes containing mRNA. Thus, our human sample mainly represented 43S post-splitting or pre-initiation complexes prior to mRNA recruitment.

Independent focused classification and multi-body refinements focusing on individual sub-complexes (Fig EV2 and Appendix Fig S2) enabled us to obtain molecular resolution for large parts of the human 43S sub-complexes. Therefore, we were able to build models for the octameric eIF3 PCI-MPN core at the backside of the 40S, parts of the YLC at the mRNA entry site and most factors located in the ISS, including ABCE1, eIF3j, eIF1 (including the N-terminal tail), eIF1A, the full eIF2 TC, and the eIF3c N-terminal domain, thus resulting in a near-complete molecular model of the human 43S particle bound to ABCE1 (Fig 2B and C, Appendix Table S2).

Conformation of ABCE1-bound 40S-initiation complexes

Strikingly, we observed ABCE1 associated with 40S subunits during all stages of 43S PIC assembly in humans and even with 48S IC complexes in the yeast sample. In all complexes, the FeSD of ABCE1 was in the extended conformation packed against h44, and the ATPase body occupied the universal translation factor binding site on the 40S, which is highly similar to previous observations of non-native complexes (h8-h14 junction; h5-h15 junction) (Heuer *et al*, 2017; Mancera-Martinez *et al*, 2017; Nürenberg-Goloub *et al*, 2020) (Fig 3A). Here, the 40S subunit is engaged in a very similar way as in the archaeal 30S-ABCE1 structure (Nürenberg-Goloub *et al*, 2020) *via* the ABCE1-specific helix-loop-helix (HLH) domain and the open conformation with respect to the composite hinge regions (h1 and h2). Surprisingly, however, in all structures we observed the ATPase in a novel state that has not yet been described for ABC-type ATPases (Figs 3B, C and D, and EV3A): Compared to the closed
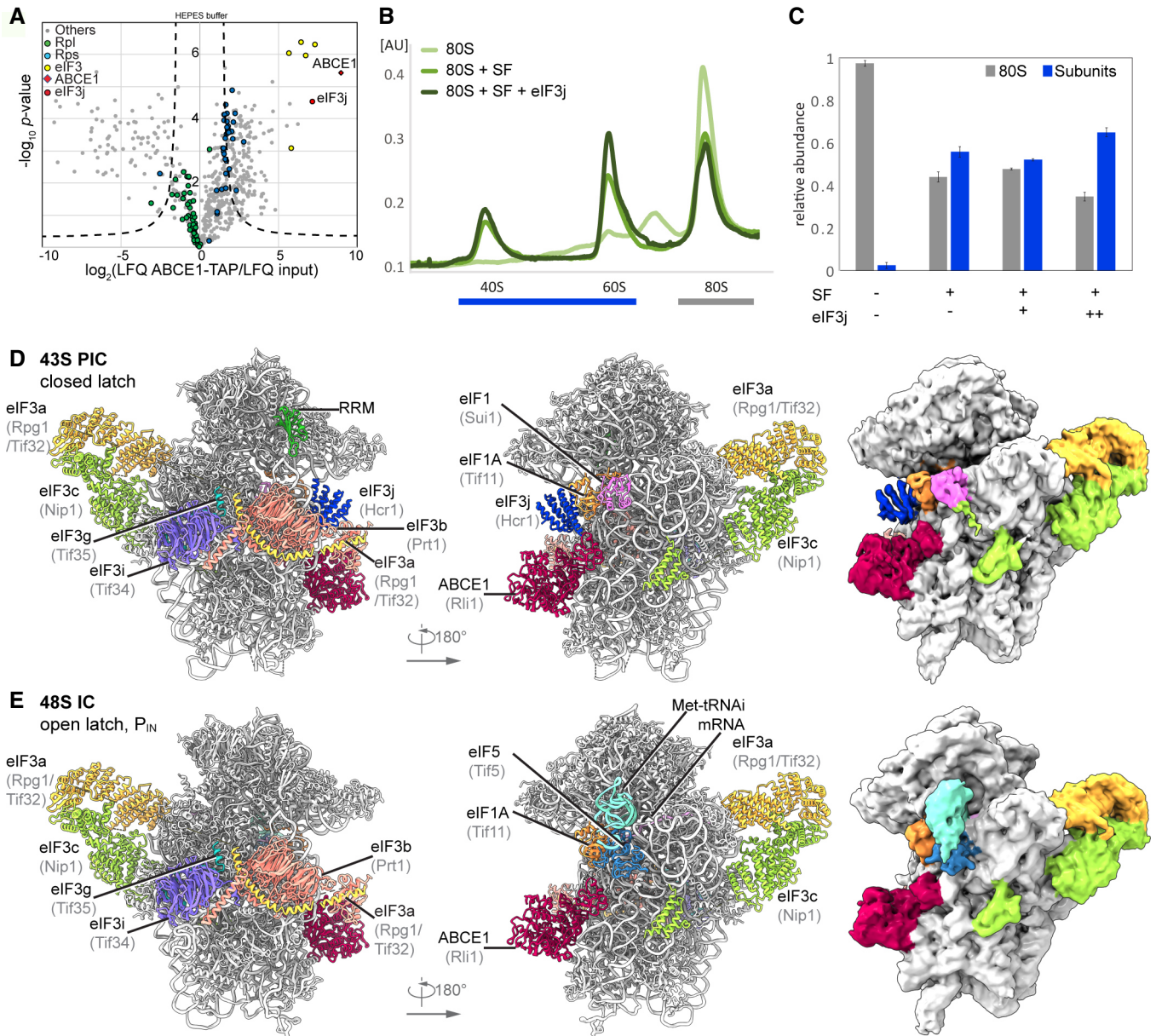
Figure 1. Biochemical analysis and cryo-EM structures of yeast ABCE1-containing initiation complexes.

A   Volcano plot representing the statistical analysis of the fold enrichment of proteins after affinity purification in HEPES buffer of ABCE1-TAP followed by label-free quantification (LFQ) using liquid chromatography–tandem mass spectrometry (LC-MS/MS). Proteins above the curved lines show a statistically significant enrichment according to the t-test value.

B, C   Sucrose density gradient UV profile after *in vitro* splitting assays (B) and relative abundance of 80S and subunits as calculated from triplicates and displayed as mean ± SD. (C); SF = splitting factors including Dom34, Hbs1, ABCE1, and eIF6; (+) = 4-fold molar excess of eIF3j; (++) = 20-fold molar excess of eIF3j.

D, E   Cryo-EM maps low-pass filtered at 6 Å and models of the yeast subclasses representing an ABCE1- and eIF3j-containing 43S PIC (D) and an ABCE1- and eIF5-containing partial 48S IC (E).

Source data are available online for this figure.

conformation as observed in *in vitro* reconstituted 30S and 40S PSCs (Heuer *et al*, 2017; Nürenberg-Goloub *et al*, 2020), we found that only NBSII is closed whereas NBSI adopts a half-open conformation comparable to the one observed in several 80S pre-splitting complexes (Fig 3B) (Becker *et al*, 2012; Brown *et al*, 2015; Preis *et al*, 2014). When analyzing our best-resolved human map, which

was obtained after focused classification on ABCE1, we unambiguously identified an Mg^{2+}-ATP (Fig 3E) occluded in NBSII, similar to the archaeal 30S-ABCE1 structure with Mg^{2+}-AMP-PNP (Nürenberg-Goloub *et al*, 2020). In the human structure, residues of the typical conserved motifs of ABC-type ATPases are involved: Lys386 of the Walker A, Gly220 of the NBD1-Signature loop, and His521 of H-loop

contact the γ-phosphate, and the Mg^{2+} ion is coordinated by Thr387 (Walker A) and Gln415 (Q-loop). In contrast, for NBSI we observed Mg^{2+}-ADP bound exactly as observed in the crystal structures of open archaeal ABCE1 (Barthelme *et al*, 2011; Karcher *et al*, 2008): Y87 of the A-loop stacks on the adenine base, F92 on the ribose, the Walker A-loop (Asn112-Ser117) binds the α- and β-phosphates, and the Mg^{2+} ion is coordinated by the β-phosphate, Ser117, Gln171 (Q-loop) and Asp241, Glu242 (Walker B). Importantly, the signature loop of NBD2 (Leu463-Glu467), which occludes ATP in the catalytically active closed state, is moved by 3.5 Å away from NBD1. In conclusion, our data suggest that—in contrast to the nucleotide-occluded state observed *in vitro*—in native SSU-ABCE1 complexes, ATP hydrolysis in NBSI has already occurred, whereas NBSII is still inhibited.

As an additional difference to previous structures, we observed a rod-like extra density (ED) after low-pass filtering in all native 43S PIC structures, protruding from h17 of the 40S body via the HLH motif into the cleft between NBD1 and NBD2 of ABCE1 (Fig 3F). However, local resolution in both human and yeast samples was too low to identify this factor. To stabilize this assembly, we generated a chemically crosslinked yeast initiation complex sample derived from a strain harboring TAP-tagged eIF3c (Nip1) and performed a cryo-EM analysis focused on the ABCE1 and the adjacent eIF3j (Appendix Figs S1B and S4). Indeed, in this reconstruction, we clearly observed an extra density protruding from eIF3j into the composite NBSs of ABCE1. At a resolution of 3.0 Å, we built the model for yeast eIF3j (Fig 3G and 4, Appendix Fig S4) based on the human eIF3j dimer (unpublished; PDB 3BPJ; lacking 137 residues at the N-terminus and 28 residues at the C-terminus). In brief, this dimer folds into a stable entangled 6-helix bundle that is arranged such that the N-termini are in close vicinity. Yet, the C-termini face into opposite directions, whereby the C-terminal tail of one protomer reaches into the mRNA entry channel (see below). On this basis, we could assign the extra density in ABCE1 as a part of the eIF3j N-terminus. This assignment was further confirmed by protein crosslinking coupled with mass spectrometry (XL-MS) using a lysine-specific BS2G crosslinker (Appendix Fig S5, Appendix Table S3). Two crosslinks between the Lys118 of eIF3j with Lys121 and Lys181 of ABCE1, both located near the ATP-binding site of NBD1, were identified (Fig 3G). In this position, the eIF3j N-terminus may easily modulate the ATPase activity of ABCE1 by restricting further movements of the HLH or the two NBDs with respect to each other. Interestingly, the position of the eIF3j-NTD on ABCE1 is similar to the one observed in a recent structure of archaeal ABCE1 co-crystallized with an 18-mer fragment from the C-terminus of the archaeal 50S stalk protein aP1 (Imai *et al*, 2018). This suggests that ABCE1 possesses a multivalent interaction patch in this region, which would allow for regulation of its ATPase activity. The observed stabilization of ABCE1 in the half-open conformation with one ADP still bound in NBS1 may indicate an inhibition of ADP release, which would explain its rather stable association with the 40S subunit.

Conformation of eIF3j in human and yeast 40S-initiation complexes

As described above, we found yeast and human 43S PIC sub-populations concomitantly bound to ABCE1 and eIF3j. The eIF3j subunit was positioned on the intersubunit side, roughly resembling the

location previously described in low-resolution maps (Aylett *et al*, 2015) (Fig 4). The main difference between the maps was the absence (human) or presence (yeast) of eIF1A. However, apart from a small rotation around the neck (approx. 3°), we did not observe significant conformational changes in the 40S when comparing the two structures.

In the low-pass-filtered human State II, which lacks eIF1A, we identified the eIF3j 6-helix bundle located above the ABCE1 ATPase body and in close vicinity to NBD1 (Fig 4A and B), but no direct contacts were formed with ABCE1. On the 40S, eIF3j contacted the N-terminal tail of eS30 (protomer 1) and the C-terminus of uS12 (protomer 2). The C-terminal helix of protomer 2 further projects toward the three-way junction formed by h32, h33, and h34 at the 40S head, whereas in protomer 1 it points toward h17 and the HLH of ABCE1 (Figs 4B and EV3E). In this position, the N-termini of eIF3j are located above the ABCE1 ATPase body close to the NBD1-NBD2 cleft.

In the yeast 43S PIC, in which eIF1A was present, we found eIF3j in a similar position, but different conformation compared to the human structure (Fig 4C and D). Here, the 6-helix bundle is stably anchored between the 40S beak at rRNA h33 on one side and the 40S body near the ABCE1 FeSD and eIF1A on the other side. The two sides of the anchor are formed again by the C-terminal helices of eIF3j: protomer 2 contacts eS30 at a similar site as in the human structure but now the entire helix bundle was rotated by approximately 100 degrees (Fig EV3B, C, and D). Consequently, the tip of the protomer 2 C-terminal helix now pointed toward the 40S head, whereas the C-terminal helix of protomer 1 projected toward the ABCE1-FeSD, thereby passing along eIF1A (Figs 4D and EV3F). Molecular details of the eIF3j-40S interaction were derived from the high-resolution structure of the crosslinked 43S-PIC (Fig 4E). In brief, the 6-helix bundle accommodates between the 40S body and head *via* interactions of both protomers. The body is contacted by the first and third helix of protomer 2 (to the h17-h18 junction and eS30) mainly by basic residues. The third helix projects toward the beak to contact the phosphate backbone of h33 (G1264). Following this helix, the ultimate eIF3j C-terminus forms a loop inside a pocket formed by h33, h34, and eS10 and from there runs along h18 and uS3, parallel to the latch, to position the ultimate C-terminal tail inside the mRNA entry channel (Figs 4F, EV3G, and H; for a detailed description of molecular contacts see Appendix Text 1). In this position, eIF3j directly overlaps with the mRNA path and would possibly interfere with mRNA loading during 48S-IC formation (Fig 4G).

Taken together, our structural data explain how eIF3j could exert its functions during key steps of translation initiation in conjunction with eIF1A.

Molecular architecture of the PCI-MPN core and its interactions with 40S

State I of the human sample represented a stable class with mainly eIF3 and weak density for eIF1 bound to the 40S SSU. This appears plausible when considering that eIF3 activity during termination and ribosome recycling has been proposed (Beznoskova *et al*, 2013; Pisarev *et al*, 2007; Valasek *et al*, 2017), which further indicates that eIF3 can already bind the 40S before eIF1A comes into play. The lack of ABCE1 in this complex may be a result of fast dissociation after splitting or of an alternative splitting mechanism. In any case,
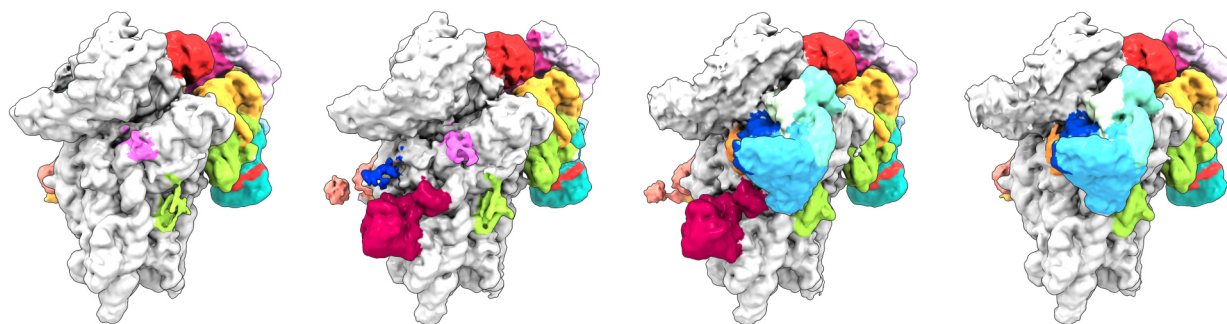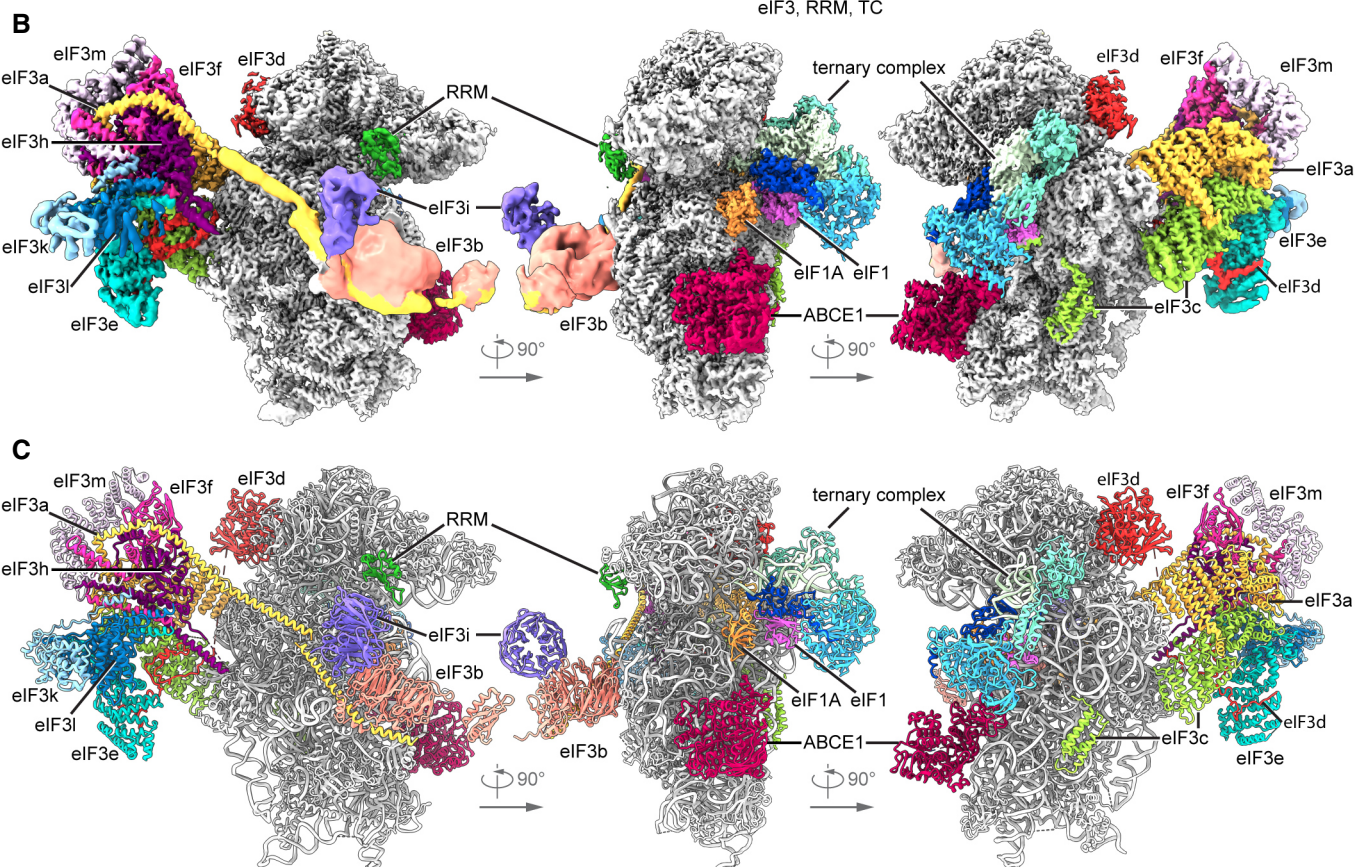
Figure 2. Cryo-EM structures of human 43S PICs in different assembly states.

A Overview of four selected compositional states present in the human 43S PIC data set.

B Composite map of the complete human 43S PIC after focused and multi-body refinements on individual sub-complexes, filtered at local resolution.

C Composite model of the complete human 43S PIC, as represented by state III.

after accommodation of eIF1 and eIF1A, the eIF2 TC binds to the 43S to induce the P_{OUT} conformation (State III-IV). Here, the improved resolution allowed us to describe the interaction network of these factors at unprecedented molecular detail.

The PCI-MPN core is located at the backside of the 40S as observed before (des Georges *et al*, 2015; Hashem *et al*, 2013;

Srivastava *et al*, 1992), and high resolution of the core was obtained by multi-body refinement of State I and State II particles. The structure assembles into β-sheets with the shape of an arc formed by PCI domains of eIF3 subunits a, c, e, l, k, and m. The arc wraps around a seven-helix bundle formed by the C-terminal helices of subunits c, e, f, h, k, and l (Figs 5A and EV4A), resulting in the typical five-
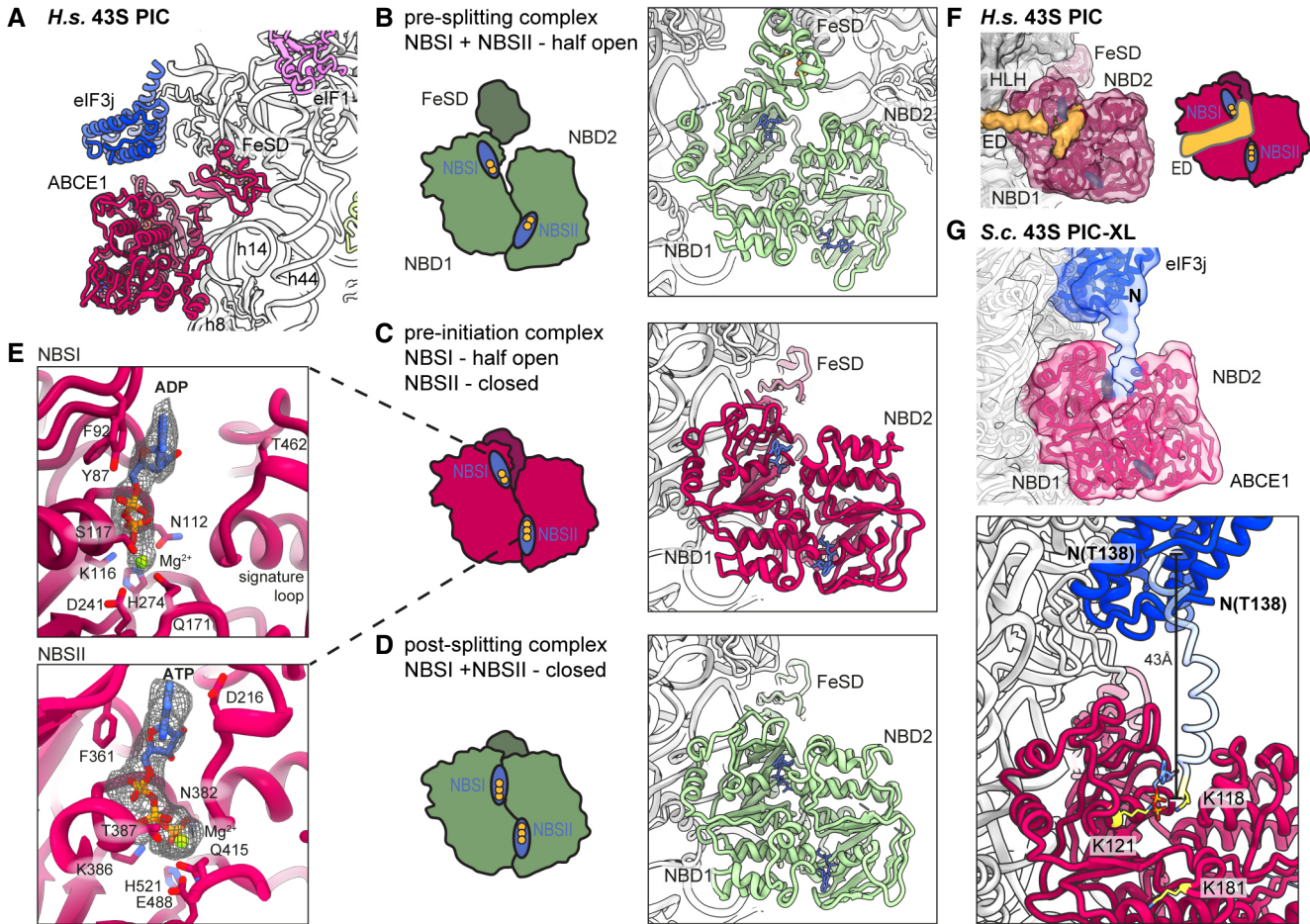
Figure 3. Conformation of ABCE1 in native 40S initiation complexes.

A   Overall position of ABCE1 in 40S initiation complexes, here representatively shown for the human State II with eIF3j.

B–D   Schematic representation and structure of semi-open ABCE1 as in 80S-pre-splitting complexes (Brown *et al*, 2015, PDB 3JAH) (B), hybrid semi-open/closed ABCE1 as in native 40S initiation complexes (C) and fully closed ABCE1 as in *in vitro* reconstituted post-splitting complexes (Nürenberg-Goloub *et al*, 2020, PDB 6TMF) (D). Nucleotide-binding sites colored in blue and bound nucleotide indicated with yellow circles (one circle per phosphate group).

E   Zoom into NBSI and NBSII showing bound Mg^{2+}-ADP (in NBSI) and Mg^{2+}-ATP (in NBSII) fit in density as obtained after focused classification on ABCE1 and refinement.

F   View focusing on the NBDs and the unassigned extra density (ED) reaching from the 40S *via* the HLH into NBSI. The ABCE1 map was low-pass filtered at 6 Å. Schematic representation highlighting the position of the ED with respect to the NBSs.

G   Upper panel: Position of eIF3j and ABCE1 in the crosslinked yeast 43S-PIC (43S-PIC-XL) sample. View focusing on the ABCE1-eIF3j interaction (same view as (F)): An extra density attributing to the eIF3j N-terminal region is connecting the eIF3j 6-helix bundle with NBSI of ABCE1. The map was low-pass filtered at 8 Å. Lower panel: N-terminally extended model of eIF3j (transparent blue) highlighting the position of K118, which was found crosslinked to K121 and K181 of ABCE1 (atoms colored in yellow).

lobed structure (left and right arm, left and right leg and head), which was visualized at a local resolution of 3.4 Å (left arm, head and right arm) and 3.8 Å (left and right leg) (Fig EV2C). This allowed for an almost complete molecular interpretation (Fig EV4A, Appendix Table S4), thus refining previous low-resolution models (des Georges *et al*, 2015; Eliseev *et al*, 2018; Erzberger *et al*, 2014), for example, by correcting the register of helices and extending molecular models (Appendix Fig S6).

The main anchor of the eIF3 PCI-MPN core to the 40S is provided by the eIF3a and eIF3c subunits, which form the "head" and the "right arm" of the PCI-MPN core, respectively. eIF3a contacts eS1 *via* its N-terminal PCI helix H1 and the loop between H1 and H2. Here, Arg14

forms salt bridges to Glu78 and Asp77 of eS1 (Fig 5B and C, see Appendix Table S4 for an inventory of observed molecular interactions). A second contact site was established between Glu17, Phe18, and Val21 of eIF3a and the eS1 Pro190 as well as adjacent residues. The loop H1-H2 of eIF3c (residues 340-345) interacts with rRNA h22 (G929, C930) and multiple sites at the Zn-knuckle domain of eS27 (Figs 5C and EV4B). Furthermore, the β-sheet insert between PCI helices 4 and 5 (residues 417-441) of eIF3c forms interactions with uS15, and basic residues in the PCI loops of both eIF3a and eIF3c are positioned to interact with the flexible tip of rRNA ES7 (Fig 5B).

An additional anchor of the eIF3 PCI-MPN to the 40S is provided by the N-terminus of eIF3d (from A2 to D84) (Figs 5C and D, and
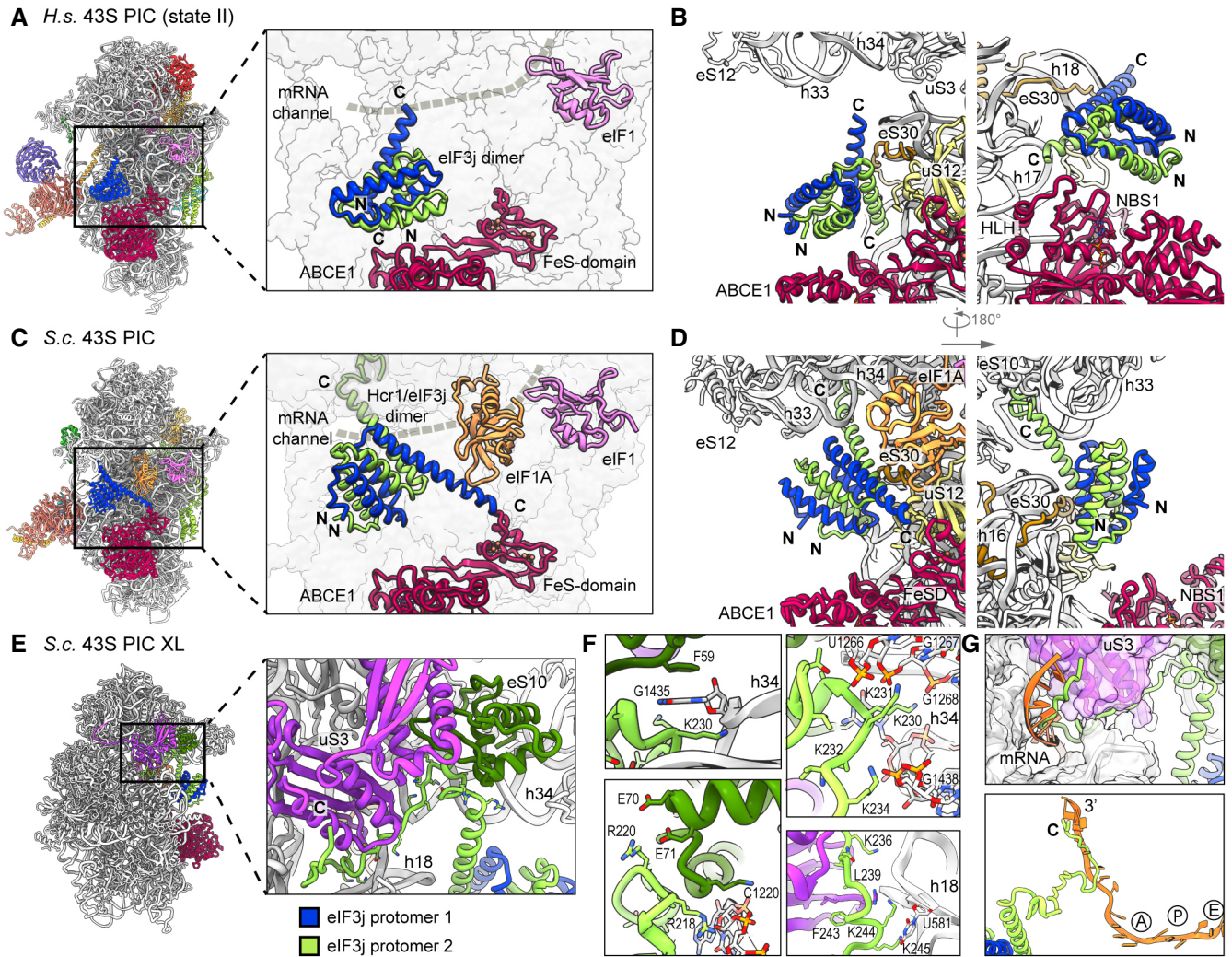
Figure 4. Two conformations of eIF3j in human and yeast 43S PICs.

A Overview and zoomed view on the model of human 43S PIC II (lacking eIF1A), focusing on the two protomers of the dimeric eIF3j 6-helix bundle in the ISS. eIF3j is in close vicinity to NBD of ABCE1 but only forms contacts to the 40S. The mRNA channel is indicated by a dashed gray line.

B Two different views showing the interaction of the two *Homo sapiens (H.s.)* eIF3j protomers with the 40S.

C Same views as in (A) on the model of the yeast 43S PIC. Here, eIF3j (Hcr1) is turned approximately 100 degrees around a pivot formed by the C-terminal helices contacting eS30 and uS12. Protomer 1 thereby contacts eIF1A and the FeSD of ABCE1 and protomer 2 contacts h33.

D Two different views showing the interaction of the two *S.c.* eIF3j protomers with the 40S and ABCE1.

E Overview and zoomed view highlighting the position of the eIF3j C-terminus in the yeast 43S-PIC-XL structure.

F Zoomed views focusing on interactions of the eIF3j C-terminus with the 40S. The loop following the third helix of eIF3j protomer 2 is in a pocket formed by the 40S h33, h34, and eS10. Lys230 of eIF3j C-terminus (protomer 2) and Phe59 of eS10 are sandwiching the flipped-out G1435 base of h34 (upper left); Lys231 and Lys234 interact with h33 (U1266 and G1267) and h34 (G1438) (upper right); salt bridges between Arg220 and Glu70-Glu71 of eS10 further stabilize the loop (lower left). Following the loop, the eIF3j C-terminus bridges the 40S body and head in the latch and contacts are formed with h18 (*via* Lys236) and *via* hydrophobic interactions with uS3 (lower right). See Appendix Text 1 for more molecular details.

G Position of the ultimate eIF3j C-terminus in the mRNA entry tunnel (upper panel) and steric clash with mRNA as positioned in an 80S ribosome stalled during translation (PDB 5MC6); for clarity, in the lower panel only eIF3j and mRNA are shown, A/P/E, respectively indicate the positions of aminoacyl, peptidyl, and exit site in the 80S ribosome.

EV4C). Interestingly, we found that it meanders along the PCI helices 1 to 3, 7, 9, 10, and 12 of eIF3e (left arm) and bridges eIF3e with eIF3c (head) by interacting with PCI helices 12, 14, and 16 (eIF3e) and PCI helix 11 (eIF3c). Another specific contact between eIF3c and eIF3e is formed by stacking of Y286 (eIF3e) to Y583 (eIF3c). Moreover, eIF3d also interacts with PCI helices 10, 13, and 14 of eIF3c by forming a large loop, which is anchored by the conserved Trp45

(interactions to Pro603, Ile607, and Glu666 of eIF3c). The interaction to eS27 is established *via* its Zn knuckle, where Phe80 of eIF3d is sandwiched between the side chains R80 and K36 of eS27.

Taken together, the PCI-MPN core of eIF3 establishes a multi-modal molecular interaction pattern with the 40S involving the eIF3a, c, and d subunits, which display an unexpected degree of inter-connectivity.

Structure and location of the peripheral subunits

The peripheral subunits, which consist of the YLC, the eIF3c-NTD, and in humans the eIF3d cap-binding protein domain, are connected to the PCI-MPN scaffold *via* flexible linkers. While eIF3a connects *via* its CTD to the YLC module located close to the mRNA entry site, the N-terminus of eIF3c protrudes from the mRNA exit toward the ISS, where it interacts directly with eIF1. While the N-terminus of eIF3d as an integral part of the PCI-MPN core is anchored to the 40S body, the cap-binding protein domain of eIF3d is located on the 40S head close to the mRNA exit site as observed before (Eliseev *et al*, 2018). Here, it contacts the 40S SSU *via* its highly conserved helix α10 (Lee *et al*, 2016) that packs upon eS28 *via* Gln416, Thr423, and Lys426 and reaches into the interface between eS28 and uS7, where Gln416 stacks on Arg51 (eS28), which in turn stacks on Phe61 (uS7). The eIF3d helix α12 lies on top of uS7 and forms contacts *via* Lys472, Glu475, Ser478, and Gln479. Notably, since eIF3d is bridging the 40S head with the eIF3 PCI-MPN core anchored to the 40S body, it could serve to relay conformational rearrangements of the 40S head—as occurring during the assembly of 43S and 48S complexes—to the PCI-MPN core or, vice versa, allow the eIF3 complex to directly control the conformational state of the 40S head (Figs 5D and EV4C and D).

For eIF3c, only a part of its NTD could be located on the ISS of the 40S so far, where it forms a helix bundle (Llacer *et al*, 2015). We found a particularly stable arrangement of the eIF3c NTD in classes containing the eIF2 TC and, after multi-body refinement, local resolution of 3 to 4 Å (Figs EV2B, EV5A and B) allowed us to determine the register of the four eIF3c-NTD helices (Val47 to Y149) (Fig 6). A stretch preceding the first helix (47-51) contacts h24 and h27 *via* R47 to the backbone phosphate of C1039 and the 2'-OH of A1181. The peptide bond of Val49 of eIF3c stacks on base C1180, which is also contacted by the first helix (52-74) of the bundle. Here, the two charged residues K55 and R56 interact with the backbone of rRNA G1179 and C1180. Backbone–phosphate interactions were also formed by the second helix (76-92) to rRNA h11 (A364) and h27 (U1178), by the fourth helix (136-143) to rRNA h11 *via* K136 (to U367), and finally by the peptide bond of Thr140 stacking upon the U367 base, as well as Gln143 hydrogen bonding to U367. Additional but less rigid contacts were established by the K-rich loop between helix 3 and helix 4 of eIF3c (Figs 6E and F, and EV5A and B, Appendix Table S4).

Notably, when low-pass filtered, a rod-like extra density for the eIF3c N-terminus became apparent, bridging the 4-helix bundle with eIF1 near rRNA h23 and h24. This density was neither present in our nor in other (Llacer *et al*, 2015) yeast 43S/48S reconstructions, where the four-helix bundle was directly connected to the eIF3c core moiety, and a site N-terminal of this region interacted with eIF1 (Fig 6A and B). Sequence alignments of the yeast and human eIF3c N-termini revealed an insertion on the C-terminal side of the conserved four-helix bundle in humans (Figs 6C and D, EV5C). This insertion from residue 165 to 213 displays 32.0% sequence identity and 56.0% sequence similarity with a stretch at the N-terminus of yeast (42-92), which was previously shown to be involved in the interaction of eIF3c with eIF1 by NMR studies (Obayashi *et al*, 2017). Here, chemical shift perturbation after eIF1 binding is observed for Glu51, Ala67, and a stretch between Lys68 and Lys77. Moreover, in our human complex one stretch of well-resolved density for the eIF3c-NTD was present at the eIF1 loop between helix α 1 and helix α 2 (Asp53-Lys58) as well as Ile100 and Gly101 of α 2 (Fig EV5D). This observation is highly consistent with the NMR study, in which the same interacting region on eIF1 is identified for the eIF3c-NTD of yeast. Together, these observations lead us to the conclusion that the density observed near eIF1 in the human structure corresponds to this insertion C-terminal of the helix bundle, fulfilling an analogous role to the previously characterized N-terminal stretch of eIF3c in yeast.

From local classification, we also obtained one class with strong density for the YLC module including the eIF3a-linker that connects it to the PCI-MPN core (Appendix Fig S7). In brief, the YLC module contains two β-propellers: the 7-bladed WD40 repeat of eIF3i and the 9-bladed WD40 repeat near the C-terminus of eIF3b. The two propellers are held together by the C-terminal helical domain of eIF3b, which is formed by 3 α-helices: the most C-terminal one binds to eIF3i, while the two preceding α-helices are bracketing the eIF3a C-terminus against the eIF3b β-propeller (des Georges *et al*, 2015; Herrmannova *et al*, 2012). N-terminal of its β-propeller, eIF3b contains a noncanonical RNA recognition motif (RRM) (ElAntak *et al*, 2007) that can form further interactions with the eIF3a-CTD (Dong *et al*, 2013; Khoshnevis *et al*, 2014; Valasek *et al*, 2002; Valasek *et al*, 2001) as well as the N-terminus of eIF3j (Elantak *et al*, 2010; Valasek *et al*, 2001).

For the CTD of eIF3a, we could build a long α-helix (residues 602-743) into the elongated rod-like density protruding from the PCI-MPN core to contact uS2 and eS21 (Appendix Fig S7A). This helix extends further toward the YLC where it forms a hinge-like structure and then connects to the stretch of the eIF3a helix that is bound to the eIF3b β-propeller. It thereby contacts the tip of the otherwise flexible rRNA expansion segment ES6C, which in turn contacts the loop between the first two helices of the eIF3b helical domain. In this arrangement, the eIF3b WD40 is rigidly confined between rRNA h16 and uS4 on one side, and ES6C on the other side, and is thus well resolved in the proximity of the 40S (Appendix Fig S7B, Appendix Table S4). The eIF3i-eIF3g complex and the eIF3b-RRM, however, remained rather flexible as observed before (Erzberger *et al*, 2014). Nonetheless, we observed a stabilization of the eIF3b-RRM in ABCE1- and eIF3j-containing classes, possibly due to an interaction of the eIF3b-RRM with the eIF3j N-terminus (Elantak *et al*, 2010; Valasek *et al*, 2001).

In yeast, the positioning of the YLC module at the mRNA exit was the same, because here it was also held in place by ES6C (Appendix Fig S7C). However, in the majority of particles in the yeast dataset (approximately 85%), we could observe a conformational change in the eIF3i-eIF3g module relative to the ES6 anchor point. Especially in the eIF3j-containing 43S class, the eIF3i-eIF3g entity rotates by approximately 120 degrees away from the mRNA entry toward ES6C and ES6B. The loop preceding the eIF3i-contacting helix of eIF3b (Thr697-Asp701) appears to serve as a hinge for this rotation (Appendix Fig S7D).

Apart from the YLC, we observed an additional density near the mRNA entry at the tip of h16 in all of our 43S structures, which was previously assigned to the RRM of eIF4B (Eliseev *et al*, 2018) (Appendix Fig S8). This density is especially prominent in subclasses of the human dataset lacking the TC, in which we could unambiguously identify the typical RRM fold at a local resolution around 4 Å (Appendix Figs S8C and D). Notably, besides eIF4B, the
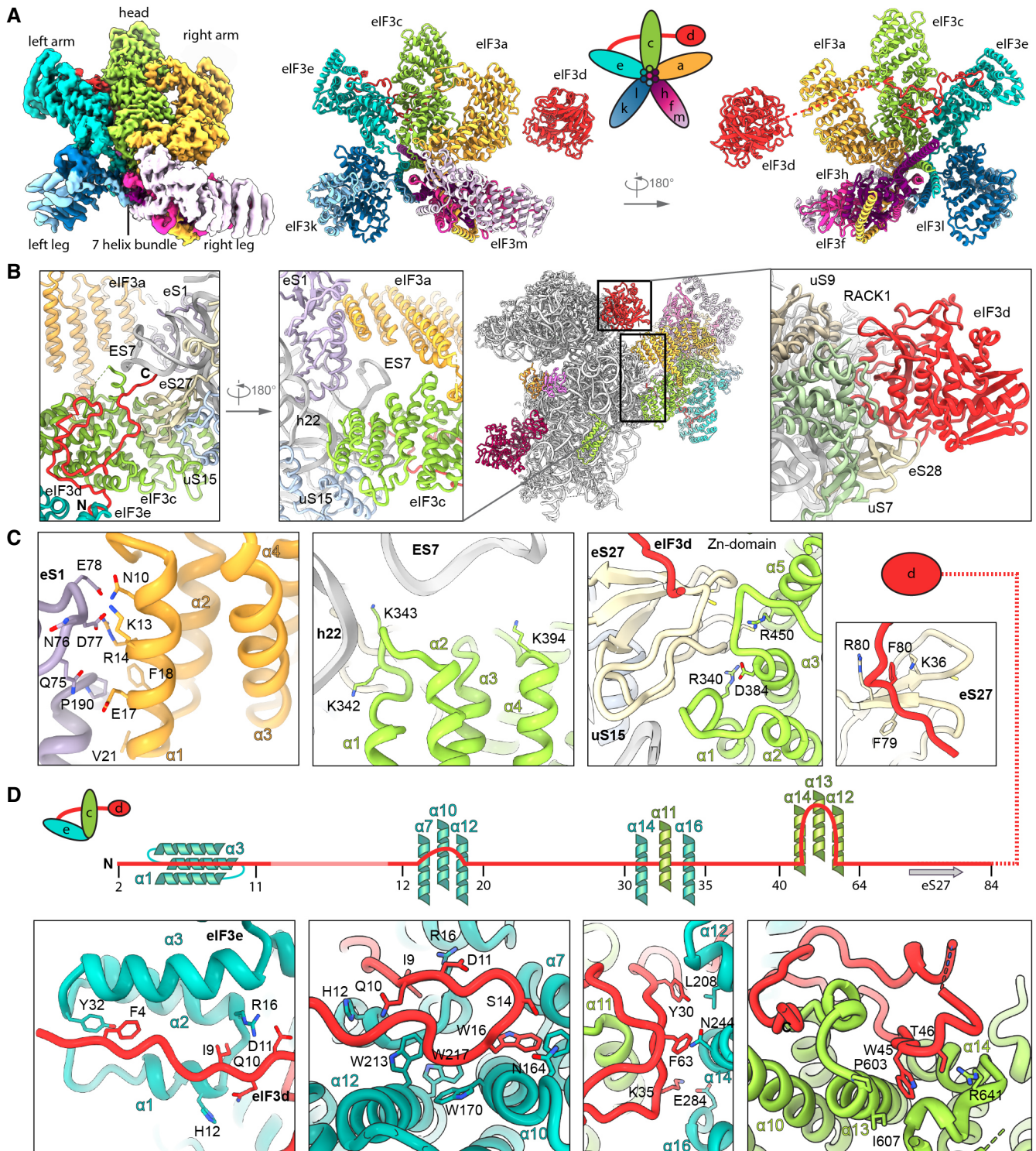
Figure 5. Molecular interactions of the human PCI-MPN core of eIF3 in the 43S PIC.

A   Isolated map and molecular model of the eIF3 PCI-MPN core color coded as in Fig 2. Structural hallmarks are indicated, and a scheme shows the composition of the lobes.

B, C   Interactions of eIF3a, eIF3c, and eIF3d with the ribosome: (B) shows an overview of the structure and zoomed views highlighting the interactions of eIF3a, eIF3c, the eIF3d N-terminal tail and the eIF3d cap-binding domain with the 40S, (C) shows molecular details of eIF3a interacting with eS1; eIF3c interacting with rRNA h22 and eIF3c and the N-terminal tail of eIF3d with the Zn-knuckle domain of eS27.

D   Interactions of the eIF3d N-terminal tail with the PCI-MPN core.

largely flexible eIF3g subunit is a potential candidate for this density because it also contains an RRM, which shares very high structural and sequence similarity (50.0%) to eIF4B (Appendix Fig S8D and E), and it was crosslinked to the nearby proteins uS10 and uS3 (Cuchalova *et al*, 2010). Unfortunately, at the current resolution we cannot unambiguously distinguish these two RRMs in our maps and it is possible that both compete for the same binding site. Next to this domain, we observed density reaching from the RRM into the mRNA channel in all human early 43S PIC structures with a closed latch (Appendix Fig S8A and B). Close to the RRM, this density forms a loop that shows multiple contacts to uS3 before winding along uS3 toward the mRNA channel. Within the channel, one side chain can clearly be identified as a tryptophan facing toward uS3 (contacting Lys148 and Met150) and further interacting with uS3 Leu142 and Val115. The stretch also contacts 18S rRNA G626, A628, and U630 of h18 as well as C1698 of h28, C1331, and A1489 of h34 (all in the A site). Thereby, this peptide stretch blocks the entire mRNA channel down to the P site where it contacts the flipped-out base C1701 at the tip of h44. Unfortunately, local resolution in this region is insufficient to provide further molecular detail and clearly identify this entity, yet considerable candidates may be further missing parts of eIF3g, eIF4B, the C-terminus of eIF3j as

observed in yeast maps, the CTD of eIF3a, or the ribosome hibernation factor SERBP1 (Stm1 in yeast) (Anger *et al*, 2013; Ben-Shem *et al*, 2011; Brown *et al*, 2018). In any case, it is apparent that accommodation of mRNA in the 48S IC complex would require its relocation, which may allow for allosteric communication between the different eIFs.

Conformation of the ternary complex

After analyzing the eIF3 complex, we also gained molecular information on the human eIF2 TC by focused classification. The TC as well as eIF1 and eIF1A were observed on the intersubunit side in a similar overall position and conformation as described before for other ICs in P_{OUT} conformation at low resolution (PDB 6GSM, PDB 3JAQ (Llacer *et al*, 2015)) (Appendix Fig S2). Briefly, eIF2 consists of three subunits, α, β, and γ. The eIF2γ subunit shares structural homology to EF-Tu-like translational GTPases (e.g., Schmitt *et al*, 2002) and consists of a G-domain (domain I), including the regulatory switch loops (swI and swII), followed by two β-barrel domains. eIF2α consists of an N-terminal OB-fold domain, a central helical domain, and a C-terminal α-β domain. The eIF2β subunit has an unstructured N-terminal domain, followed by a central helix-turn-
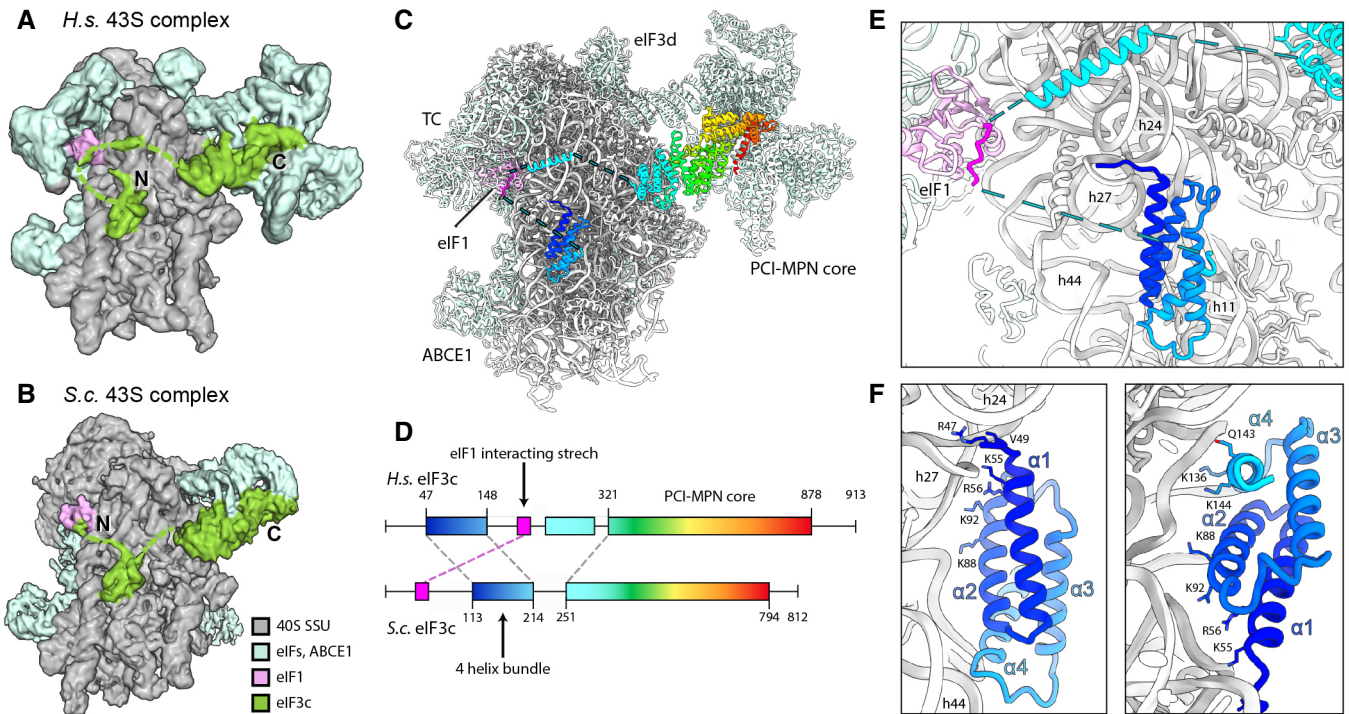
Figure 6. Arrangement of the eIF3c-NTD in human and yeast 43S PICs.

A   Cryo-EM map obtained after focused sorting of the human 43S PIC on the TC: when low-pass filtered at 6 Å, it shows the density of almost complete eIF3c-NTD in the ISS.

B   Cryo-EM map of the yeast 43S PIC low-pass filtered at 6 Å.

C   Model for human eIF3c in the TC-containing 43S colored in rainbow (C) and scheme of the alignment between human and yeast eIF3c sequences, colored accordingly (D). The eIF1-interacting stretch present in the N-terminus of *S.c.* eIF3c shows 32.0/56.0% sequence identity/similarity with an insert C-terminal of the conserved 4-helix bundle conserved in mammals.

D   Zoomed view highlighting the position of the eIF3c NTD and eIF1 in the 40S ISS.

E   Molecular model for the 4-helix bundle interacting with 40S rRNA and r-proteins.

helix (HTH) domain and C-terminal zinc binding domain (ZBD). In solution, tRNA$_i$ was shown to be bound to the TC in a distinct way different to canonical tRNA-bound EF-Tu/eEF1A by employing additional composite interactions with both eIF2α and eIF2γ (Schmitt et al, 2012). The eIF2β subunit, however, has never been sufficiently resolved to elucidate its molecular contribution to tRNAi binding and 43S PIC formation.

In our structure, we found the tRNAi embraced by all three eIF2 subunits (Fig 7A and B). Similar to the 5 Å resolution crystal structure (3V11 Schmitt et al, 2012), the methionylated CCA-end is sandwiched between the GTPase domain and domain II of eIF2γ. The terminal adenine base A76 is accommodated in a pocket formed by the β-sheets of the eIF2γ domain II including Val278, Phe322, Gly340, and Arg260 (Fig 7C, Appendix Fig S9C). The 2'-OH group of the ribose moiety interacts with the carbonyl group of Ala323 and the methionyl side chain stacks on Tyr83 of eIF2γ G-domain. The CCA-end is further stabilized by contacts including a cation-π stack of Lys266 on tRNA$_i$ C75 and Asn71 of the eIF2γ swI loop with tRNA$_i$ C74. Moreover, Arg296 of the eIF2β ZBD intercalates into the major groove of the acceptor-stem helix (G70; supported by Lys293 contacting the phosphate backbone of U69) (Fig 7D, Appendix Fig S9C). eIF2α contacts the T- and D-loops mainly via its central helical domain whereas the N-terminal OB-fold domain intercalates between anticodon stem and uS7 in the E site on the head of the 40S. The central eIF2β HTH domain contacts the anticodon from the A site and thereby forms multiple contacts to eIF1, also involving residues of the newly built C-terminus (I314-R329), which stretches below the tRNA$_i$ anticodon stem toward the E site and contacts C1057 of rRNA h24 (via N327).

Notably, in the GTP binding pocket of eIF2γ we clearly identified a Mg^{2+}-GTP (Fig 7D). Ser55 of the conserved P-loop and Thr78 of swI coordinate the Mg^{2+}-ion, whereas Asp134 and Pro135 of swII likely contact the γ-phosphate. Compared to the crystal structure of the archaeal TC (Schmitt et al, 2012), th, 2012 so that this citation matches the Reference List. Please confirm that this is correct."-->e guanine base is rotated by 90° and accommodated in a pocket between Asn190 and Ala226 of eIF2γ and Cys305 of the eIF2β ZBD, which is tightly packed upon the nucleotide-binding pocket.

Interestingly, both switch loops were embedded in a tight interaction network involving interactions with tRNA$_i$, eIF2β, and the eIF1 N-terminal tail, which we built de novo. The N-terminal tail of eIF1 protrudes from the 5-stranded β-sheet and binds to Arg446 of eIF2γ domain III, where it forms a loop and projects toward Arg75 of eIF2γ swI, forming a cation-π stack with Phe13 (Fig 7D, Appendix Fig S9C and D). Furthermore, the conformation of the swI loop was stabilized by the tRNA$_i$ via Asn71 (see above) and an interaction between conserved Ser310 of the ZBD of eIF2β with Glu74.

In close vicinity to the guanosine binding pocket, we find eIF2β Ser307, the equivalent of yeast eIF2β Ser264. In yeast, a Ser264Tyr mutation causes the Sui$^-$ (suppressor of initiation codon) phenotype, leading to increased utilization of UUG start codons (Huang et al, 1997). This mutation was shown to increase GTP hydrolysis rates and stabilize the closed P$_{IN}$ conformation of the 43S PIC (Martin-Marcos et al, 2014). In the observed position, the tyrosine mutation of Ser307 could easily interfere with the bound nucleotide, for example, by stacking on the guanine base, and thus alter the geometry of the nucleotide-binding pocket.

Taken together, we found the TC in a stable state within the 43S PIC, in an open conformation in the absence of mRNA. An intricate interaction framework is established by the 40S and eIF1 to accommodate the GTP-bound eIF2-tRNA$_i$ in a rigid position. The switch loops are kept in a rigid conformation stabilized by tRNA$_i$, eIF2β, and the eIF1 N-terminal tail, and the GTPase pocket of eIF2γ is closed by eIF2β. This may prevent premature release of the bound nucleotide and, at the same time, may restrict access for eIF5-NTD to avoid premature GAP activity.

Following TC assembly on 43S PIC and opening of the latch, mRNA can be threaded into the mRNA binding site, followed by scanning for the first AUG codon by the 48S particle. While we do not find scanning intermediates in either yeast or human datasets, in our yeast native 40S population we find one state containing eIF1A, tRNA$_i$ in the P$_{IN}$ state, and the eIF5-NTD instead of eIF1 (yeast 43S PIC). Apart from weaker density for eIF2, this state is similar to one observed before (Llacer et al, 2018), where it was interpreted as a late state after start-codon recognition. However, to our surprise we still find ABCE1 in this complex. This suggests that ABCE1 may play further roles even in later stages of initiation, or that its dissociation is not required at this stage.

Discussion

While the role of highly conserved ABCE1 during ribosome recycling has been studied in mechanistic details (Becker et al, 2012; Nurenberg-Goloub et al, 2018; Nürenberg-Goloub et al, 2020), its role after 60S dissociation remained largely elusive. However, when first characterized biochemically, ABCE1 was found associated with 43S/48S pre-initiation complexes in yeast, humans, and Drosophila (Andersen & Leevers, 2007; Chen et al, 2006; Dong et al, 2004). Since then, it is a long-standing question what the function of ABCE1 in these complexes is. Our extensive single particle analysis of native small subunits from yeast and human cells captured a variety of states throughout the assembly of the 43S PIC prior to mRNA loading, in which ABCE1 can stay associated with the 40S. Surprisingly, in yeast we even find ABCE1-48S complexes beyond the stage of mRNA engagement and start-codon recognition as indicated by the presence of the eIF5-NTD (Fig 8).

We further observe that in all ABCE1-containing 43S structures its NBDs are in an unusual hybrid conformation, where NBS2 is closed and NBS1 is semi-open. This is contrary to previous in vitro studies showing SSU-associated ABCE1 in the ATP-occluded fully closed state. Notably, the two NBSs in ABCE1 were shown to be highly asymmetric and NBSII has a low ATP-turnover rate compared to NBSI (Gouridis et al, 2019; Nurenberg-Goloub et al, 2018). Consistent with this behavior, we find Mg^{2+}-ATP still bound in the closed NBSII, whereas Mg^{2+}-ADP is present in NBSI. This is in agreement with the most recent model for the ABCE1 ATPase cycle, in which closure of the NBSII was discussed to be the decisive step for disassembly of 80S pre-splitting complexes, a process that is then triggered by subsequent closure and ATP hydrolysis in NBS1. Subsequently, re-opening of NBSI would be expected on the small subunit. But if ATP hydrolysis is prevented either by usage of a non-hydrolyzable ATP analog or by hydrolysis-deficient Walker B mutants, ABCE1 can be trapped in the fully closed state on the small subunit under facilitated splitting conditions (Heuer et al, 2017; Kiosze-Becker et al, 2016; Nürenberg-Goloub et al, 2020). In native ABCE1-associated complexes, however, NBSI is already in a more

open conformation and additionally obstructed by a part of the eIF3j N-terminal domain, which intercalates between the two NBDs close to NBSI. Thus, eIF3j may keep NBSI from closing (after putative binding of another ATP), or alternatively, prevents further opening into a state as observed in free ABCE1. This brings up the question of why ATP hydrolysis in NBSII, which would lead to dissociation from the 40S SSU, is inhibited. We find NBSII in a very similar conformation as in the fully closed archaeal structure (Nürenberg-Goloub et al, 2020), and the structure reveals no clues to explain why ATP hydrolysis is slowed down. Thus, we speculate that a further and likely only small-scale allosteric signal into NBSII may be necessary for its activation. This may occur after dissociation of the eIF3j N-terminus upon further opening of NBSI and be accompanied by changes in the ABCE1-specific HLH and hinge regions.

The observation that ABCE1 dissociation can apparently be actively prevented points toward a direct role in 43S PIC and even 48S IC assembly, most likely in concert with eIF3j. We could corroborate the finding that eIF3j assists in ABCE1-dependent splitting by *in vitro* dissociation assays, and furthermore, we established that eIF3j remains bound to the 40S together with ABCE1 after the splitting cycle. A high-resolution structure of a crosslinked yeast 43S-PIC revealed that dimeric eIF3j is highly stabilized in the presence of ABCE1, positioning the ultimate C-terminus of one protomer in the mRNA channel near the entry site. This position explains, how eIF3j could exert its roles as an antagonist of mRNA binding, for example

by recycling of mRNA from the 40S subunit (Pisarev et al, 2007; Pisarev et al, 2010), or during initiation by preventing premature mRNA recruitment (Fraser et al, 2007). Notably, its position close to eIF1A and thus near the A site may also explain its suggested role in regulating start-site selection (Elantak et al, 2010). Moreover, the comparison of yeast with the human structures of early 43S PICs suggests that eIF3j and ABCE1 may be beneficial for binding of eIF1A. In the yeast conformation, eIF3j appears like a molecular ruler reading out the exact distance between the post-splitting-specific FeSD conformation of ABCE1 and the 40S head and beak conformation as adopted after eIF1A binding. Thus, it is tempting to speculate that the observed conformational change in eIF3j may play a role in priming the 40S for eIF1A binding and/or stabilizing the early closed-latch conformation of the 43S PIC when eIF1A is bound. Notably, eIF1A is the only factor that was not found to be pre-assembled in a 40S-free multi-factor complex (MFC) consisting of eIF1, eIF2-tRNAi-GTP, eIF3, and eIF5 in yeast (Asano et al, 2000; Zeman et al, 2019), plants, and mammals (Sokabe et al, 2012). While eIF1A is capable of binding 40S SSU independently and adopting a similar conformation as within the context of initiation (Yu et al, 2009), it is possible that after binding of the MFC eIF3j binding between the 40S head and body in concert with rigidifying the latch structure may be constructive for its productive integration into the 43S complex.

Concluding our cryo-EM analysis of native initiation complexes, we can deduce a putative order of events during 43S PIC and 48S IC
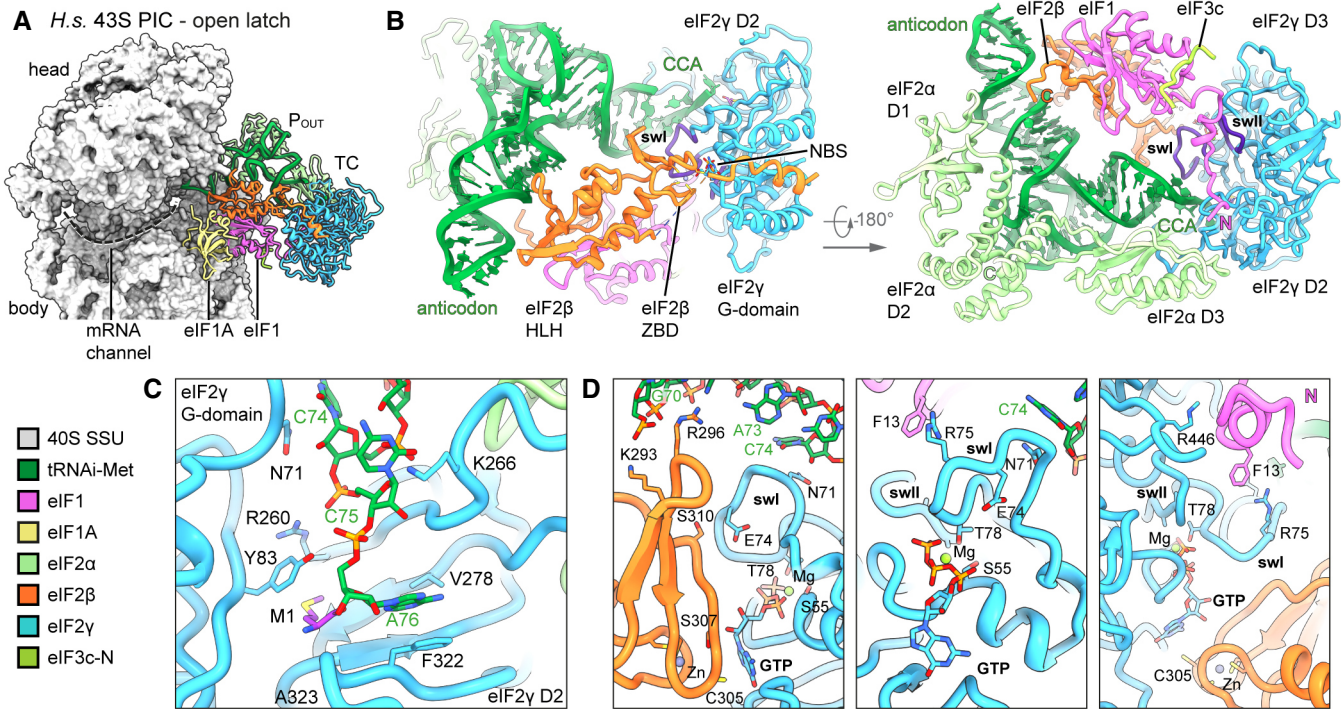
Figure 7. Conformation of the TC in the complete human 43S PIC.

A Overview highlighting the positions of TC, eIF1, and eIF1A in the complete human 43S PIC.

B Interactions of eIF2 subunits and domains and eIF1 with methionylated tRNAi; switch loops (sw) of eIF2γ are labeled and colored in purple; nucleotide-binding site (NBS) with Mg-GTP bound; the *de novo* built N-terminal tail of eIF1; and the C-terminus of eIF2α and eIF2β are labeled with N and C, respectively.

C Molecular interactions of the methioninylated CCA-end of tRNAi and eIF2γ.

D Molecular interactions within the nucleotide-binding pocket and conformation of sw loops stabilized by the eIF1 N-terminal tail, the eIF2β ZBD, and tRNAi.
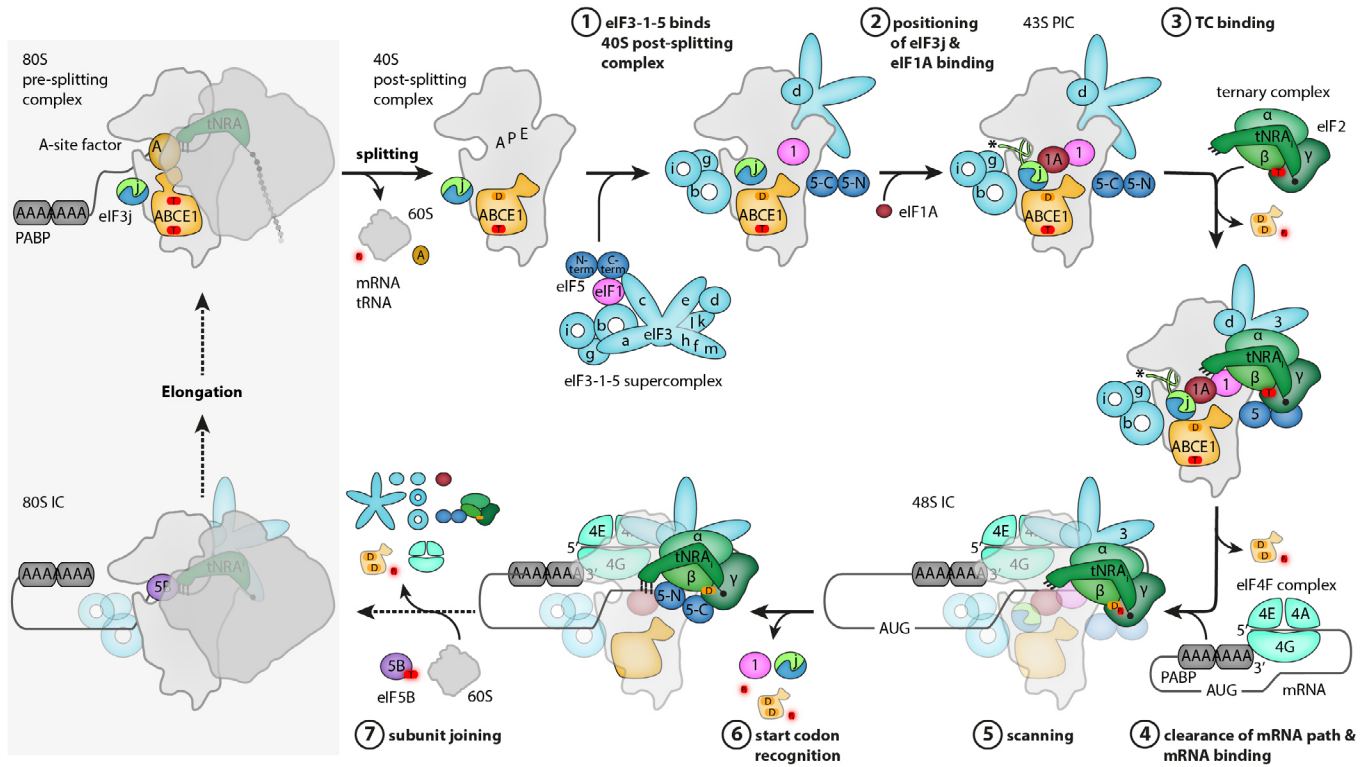
Figure 8.   Role of ABCE1 in eukaryotic translation initiation.

Schematic representation of eukaryotic translation initiation. ABCE1 is shown in orange, nucleotide states are symbolized with "T" for ATP bound and "D" for ADP bound; the two protomers of eIF3j are shown in blue and green, the eIF3j C-terminus of protomer 2 in mRNA entry channel is marked with an asterisk. For details, see main text (discussion).

assembly by formation of several structural hallmarks (Fig 8). 80S ribosomes are recycled by ABCE1 after canonical or noncanonical termination. eIF3j may assist the recycling by actively aiding ABCE1 during splitting or by destabilizing mRNA while inserting with its C-terminus into the mRNA channel. As a first step during initiation, the MFC binds to the recycled 40S as indicated by the highly populated eIF3-eIF1 bound classes. While the PCI-MPN core is stably anchored at the solvent side of the 40S, the eIF3c-NTD locates into the ISS *via* the 4-helix bundle, positioning eIF1 in the process. The YLC module is guided to the mRNA entry by stable positioning of the eIF3b β-propeller between h16 and of rRNA expansion segment ES6c. Here, the eIF3i-eIF3g complex can adopt variable positions that may be important for the role of eIF3g-eIF3i during scanning (Cuchalova *et al*, 2010). Concomitantly, the RRM of either eIF3g or eIF4B accommodates on the mRNA entry, and in the human 43S complexes, the mRNA entry channel is blocked by a yet unidentified density. After eIF1A accommodation, the TC can be stably integrated to form the complete mRNA-free P_{OUT} state 43S. This opens up the latch and leads to clearance of the mRNA path, since in P_{OUT} complexes no density in the mRNA path is visible.

With respect to a fully accommodated TC, our structure reveals for the first time a network of interactions between the tRNA$_i$ and all subunits of eIF2 as well as eIF1 at molecular resolution. The eIF2γ switch loops are highly confined, and the GTPase pocket is closed by the ZBD of eIF2β, thus restricting the access for the eIF5-

NTD to exert its GAP activity. Notably, GTP hydrolysis in eIF2γ may already occur during scanning. This would require that the eIF5 N-terminal tail could reach into the eIF2γ GTPase pocket and, thus, result in a rearrangement of the eIF2β ZBD. A structure of a scanning 48S, however, is still lacking. Yet, large structural rearrangements have been observed after start-codon recognition, during which the 48S IC adopts the closed P_{IN} state. Here, the entire TC rearranges, and especially, eIF2β alters its location on the 40S head and relative to eIF1 and eIF1A. It is likely that this conformational switch could already partially occur during scanning and that this would also affect the position of the eIF2β ZBD, which was too flexible to be resolved in all previous cryo-EM structures (Llacer *et al*, 2015; Llacer *et al*, 2018; Simonetti *et al*, 2016; Eliseev *et al*, 2018). After eIF5-dependent GTP hydrolysis, release of inorganic phosphate (P_i) would still be inhibited until start-codon recognition. During or after this process, the eIF5 NTD replaces the gatekeeper eIF1 and leads to a further stabilization and compaction of the P_{IN} state, which may be a prerequisite for the following step of eIF5B-mediated subunit joining (Llacer *et al*, 2018).

Our analysis shows that ABCE1 can still be associated with initiating 40S. Yet, which role might ABCE1 play during formation of the full 43S and—as observed in yeast—even in context of the eIF5-accommodated partial 48S? Currently, ABCE1 is assumed to act as an anti-association factor, ensuring that premature 60S interaction is prevented after termination and ribosome splitting. However, in this

function it would likely become redundant after the formation of the 43S PIC, failing to explain its presence in later stages of initiation. Another possibility is that its observed interplay with eIF3j as early as during the splitting reaction supports the timely recruitment of the remaining eIFs to the vacant 40S. Furthermore, we speculate that the inhibiting peptide close to NBSI would need to be ejected to facilitate ATP-hydrolysis in NBSII. Here, it is possible that dynamics of the rather flexible YLC module could play a role. In fact, this module is able to relocate into the ISS to occupy the position of ABCE1 (Llacer *et al*, 2015). With this steric competition in mind, it would be plausible that it contributes to ABCE1 dissociation, although it is not entirely clear at which stage this relocation happens. In addition, eIF3j, which is still present at least as fuzzy density in the fully assembled 43S, may also contribute in coordinating such events, for example, *via* its known interaction with eIF1A and the eIF3b-RRM (Elantak *et al*, 2010). Finally, since ABCE1 is even present on 48S IC complexes after start-codon recognition, events during subunit joining may be the final trigger for ABCE1 dissociation. In this context, the P proteins of the 60S subunit may not only play a role during ribosome splitting as suggested before (Imai *et al*, 2018), but also for ABCE1 removal after initiation. Yet to reveal exact timing of these events and the mechanistic interplay of these factors, future work will be needed.

Materials and Methods

Yeast strains

Saccharomyces cerevisiae ribosomes for biochemical assays were purified from a wild-type BY4741 strain, which was grown on YPD medium.

Samples for LC-MS/MS analyses were purified from a BY4741 (*MATa, ura3Δ0, his3Δ1, leu2Δ0, met15Δ0*), Rli1-TAP:HIS3MX6 strain (Ghaemmaghami *et al*, 2003).

For the preparation of native yeast 40S initiation complexes, a BY4741 strain containing genomic TAP-tagged SKI3 and a plasmid overexpressing SKA1 (pCM190) (Zhang *et al*, 2019) were used; the crosslinked yeast 43S pre-initiation complex was derived from a *S.c.* W303 strain (*MATa, ade2Δ1, trp1Δ1, can1Δ100, leo2Δ3,112, his3Δ11, ura3, GAL*) expressing genomically TAP-tagged Nip1 (eIF3c).

ABCE1-TAP polysome profile and sucrose density gradient fractionation

Yeast (*Saccharomyces cerevisiae*; *S.c.*) cells from the BY4741 strain expressing C-terminally TAP-tagged ABCE1 (Rli1) were grown in 200 mL YPD to an OD_{600} of 0.8. The cells were then treated with 50 μg ml^{-1} cycloheximide on ice for 5 min. and collected by centrifugation. The cells were lysed in lysis buffer (20 mM Tris–HCl, pH 7.4, 50 mM KCl, 10 mM MgCl$_2$, 50 μg ml^{-1} cyclohex-imide, and EDTA-free protease inhibitors (Roche)) by vortexing them with glass beads (12 cycles of 30 sec. vortex/30 sec. on ice). The lysate was cleared by centrifugation for 10 min. at 16,000 *g*, 4 °C and stored at −80 °C. Ten A_{260} units were loaded on a 10–50% sucrose gradient and centrifuged at 187,813 *g* for 2.75 h at 4 °C in a SW41Ti rotor (Beckman Coulter). The fractions of the gradient were collected, and proteins were

precipitated with trichloroacetic acid and separated on a 10% acrylamide gel. The proteins were detected with antibodies after Western blotting: ABCE1-TAP with peroxidase anti-peroxidase (PAP) complex (Sigma-Aldrich) at 1:2,000, and Nog1 with a rabbit anti-Nog1 antibody at 1:5,000 dilution.

ABCE1-TAP tandem affinity purifications

Cells expressing C-terminally TAP-tagged ABCE1 (Rli1) were cultivated in rich medium (YPD) until OD_{600} of 2, and cultures were centrifuged at 4 °C, rinsed in cold water, and frozen at −80 °C. Cells were thawed on ice, resuspended in lysis buffer (50 mM Tris–HCl pH 8.0, 100 mM NaCl, 10 mM MgCl$_2$, complete EDTA-free protease inhibitor mix or: 20 mM HEPES/KOAc pH 7.4, 100 mM KOAc, 10 mM MgCl$_2$, complete EDTA-free protease inhibitor mix), and lysed with glass beads using a Magnalyser. The lysates were clarified by centrifugation at 16,000 *g* for 10 min. at 4 °C. Supernatants were collected, and triton (0.5% final) or NP-40 (0.1% final) was added to the lysate. Binding to magnetic beads coupled with IgG was performed on a wheel at 4 °C overnight. Beads were collected on a magnet, flow-through was discarded, and beads were washed in lysis buffer. Elution was performed by resuspension in 2% SDS, 1× Tris-EDTA buffer and incubation at 65 °C for 10 min. Eluted beads were discarded on a magnet, and eluate was purified on HiPPR Detergent Removal Resin (Thermo Scientific, 88305). Purified proteins were eluted in PBS. The rest of the eluates was precipitated by the methanol/chloroform technique (Wessel & Flugge, 1984) and analyzed by mass spectrometry.

To control the quality of the affinity purification, a sample of eluates (3%) was separated on acrylamide NuPAGE Novex 4–12% Bis-Tris gels (Life Technologies) and analyzed by silver staining.

Mass Spectrometry: data acquisition and analysis

After reduction and alkylation, protein samples were treated with endoprotease Lys-C (Wako) and trypsin (Trypsin Gold Mass Spec Grade; Promega). Peptide samples were desalted by OMIX C18 pipette tips (Agilent Technologies) and then analyzed by LC-MS/MS on an LTQ-Orbitrap velos instrument (Thermo Fisher Scientific) connected online to an EASY-nLC system (Thermo Fisher Scientific). Raw mass spectrometry (MS) data from the LTQ-Orbitrap were analyzed using MaxQuant software (Cox & Mann, 2008) version 1.6.10.43, which uses Andromeda search engine (Cox *et al*, 2011). Bioinformatic analysis of the MaxQuant/Andromeda workflow output and the analysis of the abundances of the identified proteins were performed with the Perseus module (Tyanova *et al*, 2016) version 1.6.10.43. Only protein identifications based on a minimum of two peptides were selected for further quantitative studies. After data processing, label-free quantification (LFQ) values from the "proteinGroups.txt" output file of MaxQuant were further analyzed. To distinguish specifically enriched proteins from the background, protein abundances were compared between sample and control groups using Student's t-test statistic, and results were visualized as volcano plots (Hubner & Mann, 2011).

Preparation of puromycin-treated 80S ribosomes from yeast

S.c. BY4741 wild-type cells were grown in YP medium with 2% glucose to an OD_{600} of 2.5, then harvested by spinning at 4,400 *g* for

10 min. Cells were washed first with water, then 1% KCl, then resuspended in 30 ml lysis buffer (20 mM HEPES/KOH pH 7.4, 100 mM KOAc, 7.5 mM Mg(OAc)$_2$, 1 mM DTT, 0.5 mM PMSF, complete EDTA-free protease inhibitor mix). Lysis was performed using a Microfluidics M-110L microfluidizer at 15k psi.

The lysate was cleared by centrifugation first at 26,892 g for 15 min., then at 140,531 g for 30 min. 15 ml of cleared lysate was loaded on a layered sucrose cushion consisting of 4 ml 2 M sucrose and 4 ml 1.5 M sucrose (buffer: 20 mM HEPES/KOH pH 7.4, 500 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT, 0.5 mM PMSF) and centrifuged at 246,468 g for 21 h and 15 min.

The pellet containing ribosomal components was resuspended in water and mixed with 2× puromycin buffer (40 mM HEPES pH 7.5, 1 M KOAc, 25 mM Mg(OAc)$_2$, 2 mM puromycin, 2 mM DTT, 1 U/ml SUPERase-In RNase Inhibitor (Invitrogen)). The mixture was incubated for 30 min at room temperature and then loaded on 10-40% sucrose density gradients (20 mM HEPES/KOH pH 7.4, 500 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT, 0.5 mM PMSF). Gradients were centrifuged at 20,755 g in an SW 32 Ti rotor (Beckman Coulter) for 20 h. 80S ribosomal fractions were identified using a Biocomp Gradient station ip and a Triax Flow cell and were manually collected. Fractions were then pelleted in a TLA110 rotor at 417,200 g for 45 min and resuspended in storage buffer (20 mM HEPES/KOH pH 7.5, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT). Aliquots were frozen in liquid nitrogen and stored at -80 °C until use.

Protein expression and purification

eIF3j (Hcr1) purification

Escherichia coli (*E. coli*) BL21(DE3) cells were transformed with the pTYB2 plasmid containing full-length *HCR1* and selected on LB plates containing ampicillin. Cells from a pre-culture were inoculated into 1.5 L of LB medium with ampicillin, and cell growth was monitored at 37 °C. At an OD$_{600}$ of 0.6, the cultures were transferred to an ice-water bath and incubated for 20 min. 0.1 mM IPTG was added to induce protein expression, and cells were incubated for 15 h at 16 °C while shaking. Cells were harvested by centrifugation at 3,500 g for 10 min and washed with 1% KCl, then resuspended in lysis buffer (20 mM HEPES pH 7.5, 500 mM NaCl). Cells were then pelleted again at 2,600 g, frozen in liquid nitrogen, and stored at -80 °C until further use.

Frozen cell pellets were thawed, resuspended in lysis buffer, and lysed using a Microfluidics M-110L microfluidizer at 15k psi. The lysate was cleared by centrifugation at 20,000 g for 30 min. Clear lysate fraction was added to 1.5 ml magnetic chitin beads (NEB E8036S) equilibrated in lysis buffer. Binding was performed for 1.5 h at 4 °C on a wheel. Beads were harvested on a magnet and washed once using 5 ml lysis buffer, twice using washing buffer (20 mM HEPES pH 7.4, 1 M NaCl, 1 mM EDTA) and once again using lysis buffer. The protein was then eluted from the beads using 5 ml elution buffer (20 mM HEPES pH 7.4, 500 mM KCl, 50 mM DTT) by incubating on a wheel at 4 °C overnight. A second elution step was performed using the same buffer for one hour after removal of the first elution fraction. Both elution volumes were combined and concentrated using an Amicon Ultra 10k MWCO concentrator. Aliquots of pure eIF3j were flash-frozen in liquid nitrogen and stored at -80°C.

ABCE1 (Rli1) purification

ABCE1 (Rli1) was overexpressed in *S. cerivisiae* strain WCGα using the pYes2-ABCE1-His$_6$ plasmid (kindly provided by R. Green, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine) (Shoemaker & Green, 2011). Cells were grown in YP medium lacking uracil and containing 2% galactose, 1% raffinose at 30 °C to mid-log phase and were harvested at a final OD$_{600}$ of 1.0 by centrifugation at 3,500 g for 10 min. Cells were washed once with 1% KCl, pelleted again, and resuspended in lysis buffer (75 mM HEPES pH 8.0, 300 mM NaCl, 5mM beta-mercaptoethanol (β-ME), 1% Tween, 20 mM imidazole, 2 mM MgCl$_2$, 10% glycerol). Excess buffer was removed by centrifugation at 2,600 g, and the cells were frozen in liquid nitrogen. Frozen cells were ground using a Spex SamplePrep Freezer Mill and the powder stored at -80 °C until further use. The cell powder was thawed and resuspended in lysis buffer. Cell debris was removed by centrifugation at 47,807.6 g for 30 min and filtered using a 1.6-μm membrane.

ABCE1 was purified first by metal affinity chromatography. Cleared lysate was applied to a HisTrap HP column (GE 5 mL column). The column was washed with 15 column volumes (CV) wash buffer (50 mM HEPES pH 8.0, 500 mM NaCl, 5 mM β-ME, 20 mM imidazole, 2 mM MgCl$_2$, 10% glycerol), and the protein was eluted with 4 CV over a gradient from 20 mM to 300 mM imidazole. Fractions containing ABCE1 were combined and dialyzed against Buffer A (20 mM HEPES pH 7.6, 100 mM KCl, 5 mM β-ME, 0.1 mM EDTA, 10% glycerol, 0.1 mM PMSF) overnight. The sample was diluted to 50 mL and loaded onto a cation exchange column (HiTrap SP 5mL, GE). The column was washed with 6 CV Buffer A, and ABCE1 was eluted over gradient from 100 mM to 1 M KCl over 8 CV. ABCE1-containing fractions were concentrated using Amicon® 50k MWCO concentrator before loading onto a gel filtration column (Superdex200) for size-exclusion chromatography. The fractions containing ABCE1 were concentrated, and aliquots of pure ABCE1 in 20 mM HEPES pH 7.5, 200 mM KCl, 1.5 mM MgCl$_2$, 2 mM β-ME, and 5% glycerol were flash-frozen and stored at -80 °C.

eIF6 purification

E. coli BL21 (DE3) cells were transformed with a p7XC3GH plasmid expressing eIF6 fused to 3C protease cleavage site, GFP, and 10-His. Cells were grown on LB medium to mid-log phase (OD$_{600}$ = 0.7-0.8) at 37 °C and induced with 1 mM IPTG at 16 °C for 20 h. Cells were harvested by centrifugation at 4,400 g and 4 °C for 8 min, washed with PBS, and resuspended in lysis buffer (20 mM Tris–HCl pH 8.0, 300 mM NaCl, 2 mM β-ME) with 10% glycerol. Resuspended cells were flash-frozen in liquid nitrogen and stored at -80 °C until further use. For purification, frozen cells were thawed and resuspended in lysis buffer without glycerol. Lysis was performed using a Microfluidics M-100L microfluidizer at 15k psi. Crude lysate was cleared by centrifugation at 30,596 g for 20 min. TALON metal affinity resin was equilibrated in lysis buffer and added to the cleared lysate, then incubated at 4 °C for 40 min on a wheel. After collection of the flow-through, the column was washed using lysis buffer with 10 mM imidazole. Elution was performed by incubating the resin with lysis buffer with 10 mM imidazole and 0.25 mg ml^{-1} 3C protease for 30 min at 4 °C on a wheel. The elution fraction was concentrated using an Amicon 10k MWCO concentrator and loaded onto a Superdex200 column for size-exclusion chromatography using storage buffer (50 mM HEPES/KOH pH 7.5, 500 mM KCl, 2 mM MgCl$_2$, 2 mM β-ME).

The purified protein in storage buffer was flash-frozen in liquid nitrogen and stored at -80 °C.

Dom34 and Hbs1 were purified as described before (Lee *et al*, 2007).

Splitting assays

In vitro splitting assays

Ribosome splitting assays were carried out to test the influence of eIF3j (Hcr1) on the canonical splitting reaction mediated by Dom34, Hbs1, and ABCE1 in yeast. For each reaction, 5 pmol of yeast 80S ribosomes (see above) was mixed with fivefold molar excess of splitting factors Dom34, Hbs1, and ABCE1 as well as the anti-association factor eIF6 under physiological buffer conditions (20 mM HEPES/KOH pH 7.5, 100 mM KOAc, 4 mM Mg(OAc)$_2$, 5 mM β-ME, 1 mM ATP, 1 mM GTP). Varying amounts of eIF3j were added to the reactions, ranging from twofold to twentyfold molar excess over the 80S ribosomes.

The samples were incubated on ice for 30 min and then loaded on 10-50% sucrose density gradients (20 mM HEPES/KOH pH 7.5, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT, 10-50% (w/v) sucrose). Gradients were spun in an SW 40 Ti rotor (Beckman Coulter) at 202,048 *g* for 4 h and fractionated at a BioComp Gradient Station *ip* using a Triax Flow Cell for UV measurement.

Ribosomal peak fractions were collected manually, and from each fraction, proteins were precipitated using 0.015% sodium deoxycholate and 7.2% trichloroacetic acid at 4 °C.

Proteins were separated on a 15% SDS-PAGE gel and visualized using SimplyBlue staining reagent.

"*Facilitated*" *splitting assays*

"Facilitated" splitting assays were performed to test the association of yeast ABCE1 and eIF3j to ribosomal particles under non-physiological high-salt conditions and in the presence of ATP or the non-hydrolyzable ATP analog AMP-PNP. To induce splitting, purified 80S ribosomes were mixed with tenfold molar excess of ABCE1 in splitting facilitating buffer (20 mM HEPES/KOH pH 7.4, 500 mM KCl, 1.5 mM MgCl$_2$, 1mM DTT). Depending on the experiment, 0.5 mM AMP-PNP or ATP and 10-fold molar excess of eIF3j were added. For the experiments described here, approx. 50 pmol ribosomes in a total reaction volume of 250 μL were used. The samples were incubated for 20 min at 25°C and then cooled down to 4 °C on ice and loaded on 10-50% sucrose density gradients (20 mM HEPES/KOH pH 7.5, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT, 10-50% (w/v) sucrose). All following procedures were carried out as described above for splitting assays.

Cryo-EM sample preparation

Preparation of native yeast 40S complexes

A BY4741 strain containing genomic TAP-tagged SKI3 and a plasmid overexpressing SKA1 (pCM190) (Zhang *et al*, 2019) was used for generation of the cryo-EM sample.

Yeast cells were grown in synthetic medium lacking uracil (SL-Ura) with 2% glucose at 30 °C to an OD$_{600}$ of 3.0, whereupon the cultures were chilled in ice water. The cells were harvested by centrifugation at 4,422 *g* for 10 min in a Sorvall SLC-6000 rotor, washed with water, and resuspended in lysis buffer (20 mM HEPES/KOH pH 7.4, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT,

0.5 mM PMSF, complete EDTA-free protease inhibitor mix). Cells were frozen in liquid nitrogen and ground using a Spex SamplePrep Freezer/Mill.

Frozen cell powder was resuspended in lysis buffer (1:3 w/v), and the lysate was cleared by centrifugation in an SS-34 rotor (Thermo Scientific) at 26891.8 *g* for 15 min.

Approximately 150 A$_{260}$ absorption units were loaded on a 10-50% sucrose density gradient (buffer composition identical to lysis buffer). Gradients were spun in an SW40 Ti rotor (Beckman Coulter) at 202,048 *g* for 3 h, and the 40S peak was harvested manually using a Triax Flow Cell.

Total A$_{260}$ of the collected 40S fraction from yeast lysate was measured, and the buffer was exchanged to cryo-EM grid buffer (20 mM HEPES/KOH pH 7.4, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT, 0.5 mM PMSF, complete EDTA-free protease inhibitor mix, 0.05% Nikkol) by three successive rounds of concentration and dilution by a factor of approx. 1:5 using an Amicon Ultra Centrifugal Filter (MWCO 100k) (total dilution factor approx. 1:125). The sample was then concentrated again. The A$_{260}$ was measured as A$_{260}$/ ml = 6.3.

Freshly prepared sample was diluted to approx. 1.25 A$_{260}$ / ml, and 3.5 μL was applied to 2 nm pre-coated Quantifoil R3/3 holey carbon support grids and vitrified in liquid ethane using a Vitrobot mark IV (FEI Company, Netherlands). (wait time 45 s, blotting time 2 s).

Preparation of crosslinked yeast 43S pre-initiation complexes

S.c. W303 cells expressing genomically TAP-tagged Nip1 (eIF3c) were grown in YP medium with 2% glucose at 30 °C to an OD$_{600}$ of 2.0 and harvested by centrifugation at 4,422 *g* for 10 min in a Sorvall SLC-6000 rotor, washed with water, and resuspended in lysis buffer (50 mM HEPES/KOH pH 7.4, 100 mM KOAc, 1.5 mM Mg(OAc)$_2$, 1 mM DTT, 0.5 mM PMSF, complete EDTA-free protease inhibitor mix). Cells were frozen in liquid nitrogen and ground using a Spex SamplePrep Freezer/Mill.

Frozen powder was resuspended in lysis buffer with 0.15% (v/v) NP-40, and the lysate was cleared by centrifugation first at 20,000 *g* for 30 min in an SS-34 rotor and then at 26891.8 *g* for 15 min in the same rotor.

The cleared lysate was applied to IgG Sepharose (GE17-0969-01) equilibrated in lysis buffer and incubated for 1 h at 4 °C on a rotating wheel. After binding, the supernatant was removed by centrifugation at 1,383 *g* for 3 min and the sepharose was transferred to a Bio-Rad Micro Bio-Spin Chromatography Column. The sepharose was sequentially washed with 10 column volumes each of lysis buffer containing 0.15% NP-40 and elution buffer (50 mM HEPES/KOH pH 7.4, 100 mM KOAc, 1.5 mM Mg(OAc)$_2$, 1 mM DTT).

Elution was performed by incubating the sepharose with elution buffer and approx. 2.4 U/μL AcTEV (Invitrogen) for 90 min at 20 °C. The eluted sample was harvested by spinning the columns at 380 *g* for 2 min.

The eluted sample was transferred onto a 10-50% sucrose density gradient (50 mM HEPES/KOH pH 7.4, 100 mM KOAc, 5 mM Mg(OAc)$_2$, 1 mM DTT). Gradients were spun in an SW40 Ti rotor (Beckman Coulter) at 202,048 *g* for 3 h, and the 40S peak was harvested manually using a Triax Flow Cell for UV absorption measurement.

The 40S fraction was concentrated and the buffer exchanged to elution buffer using an Amicon Ultra Centrifugal Filter (MWCO

100k). The sample was crosslinked by shaking with 0.5 mM BS3 at 10 °C, 1,200 rpm for 10 min and then further incubated at 4 °C for approx. 10 min. The reaction was quenched by addition of 40 mM Tris–HCl. To the sample, 0.05% ß-OG was added and cryogenic freezing was performed.

Preparation of native human 40S complexes

Human 40S initiation complexes were found as byproducts in an affinity purification using internally tagged RIOK1 and mutant RIOK1-D324A as bait. In brief, HEK Flp-In 293 T-Rex (Invitrogen) was grown in a 10-cm cell-culture dish to approximately 70% confluency and transfected with 0.5 μg of a pcDNA5/FRT/TO vector containing RIOK1 or RIOK1-D324A and 4.5 μg pOG44 (Invitrogen), using 20 μg polyethylenimine (PEI). Cells were selected using 150 μg ml-1 hygromycin B (Thermo Scientific) and maintained in DMEM (Thermo Scientific) containing 10% fetal calf serum, 100 μg ml-1 hygromycin B, 10 μg ml-1 blasticidin and 1x penicillin/streptomycin and GlutaMAX (Thermo Scientific). Stable cell lines were subsequently grown in multiple 15-cm cell-culture dishes, protein expression induced with 1.6 μg ml-1 tetracycline and harvested in 0.025% trypsin/EDTA (Thermo Scientific) after 24 h. Cells were washed one in 1x phosphate-buffered saline (PBS) and subsequently pelleted at 1,600 g at 4 °C. Cells were then resuspended in lysis buffer (20 mM HEPES pH 7.6, 150 mM potassium acetate, 5 mM $MgCl_2$, 1 mM DTT, 0.5 mM NaF, 0.1 mM Na_3VO_4, 1x protease inhibitor (Sigma-Aldrich), 0.5% NP-40 substitute) and incubated for 30 min in an over-head rotator at 4 °C, before centrifugation at 4,000 g for 15 min at 4 °C. The cleared lysate was then added to 100 μl of anti-Flag affinity beads (Sigma-Aldrich) and rotated for 2 h at 4 °C. Beads were harvested and 4 times washed with 1 ml wash buffer (20 mM HEPES pH 7.6, 150 mM potassium acetate, 5 mM $MgCl_2$, 1 mM DTT, 0.5 mM NaF, 0.1 mM Na_3VO_4, 1x protease inhibitor (Sigma-Aldrich)), before bound complexes were eluted 6 times with 100 μl of 20 mM HEPES pH 7.6, 150 mM potassium acetate, 5 mM $MgCl_2$, 1mM DTT, 0.05% Nikkol, and 0.2 mg ml-1 3x Flag peptide (Sigma-Aldrich). All eluate fractions were combined and concentrated on 300 kDa molecular mass cut-off filters (Sartorius).

3.5 μl of the concentrated sample was applied to glow discharged copper grids with holey carbon support and a 2 nm continuous carbon layer (R3/3, Quantifoil). Grids were blotted in a Vitrobot Mark IV (FEI Company) for 2 s after incubation for 45 s at 4°C and frozen in liquid ethane.

Cryo-EM data collection and processing

Data collection and processing of the yeast 40S complex sample

Cryo-EM data were collected on a Titan Krios TEM, using a Falcon II DED at 300 kV, with an electron dose of approx. 2.5 e⁻/Å^2 per frame for 10 frames (defocus range of 1.1 to 2.3 μm). The magnified pixel size was 1.084 Å/pixel.

Micrograph stacks collected at the TEM were summed and corrected using MotionCor2 (Zheng *et al*, 2017). Micrograph quality was assessed individually, and CTF parameters were estimated using GCTF (Zhang, 2016). Particle picking was performed using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/), and all further processing was performed using RELION 3.0 (Scheres, 2012; Zivanov *et al*, 2018).

Data collection and processing of the crosslinked yeast 43S PIC sample

For the crosslinked yeast sample, 5126 micrograph movies were collected at a Titan Krios at 300 kV, at a nominal pixel size of 1.059 Å, and a defocus range from 0.5 to 2.5 μm. Movies were recorded on a K2 Summit direct electron detector using low-dose conditions with 40 frames at approximately 1.12 e⁻/Å^2. each. All frames were gain corrected and subsequently aligned and summed using MotionCor2 (Zheng *et al*, 2017), and CTF parameters were determined using CTFFIND (Rohou & Grigorieff, 2015) and Gctf (Zhang, 2016). Particles were picked using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/), and particle images were extracted in RELION 3.1 (Zivanov *et al*, 2018). 2D classification was performed using a previously generated cryo-EM map of an idle 40S subunit as reference, and all particles with recognizable features of the small subunit were selected and subjected to an initial 3D refinement using the same reference, followed by 3D classification. All classes with density corresponding to ABCE1 were selected, grouped, and classified again using an ellipsoid mask around ABCE1. Particles with ABCE1 in the semi-open conformation as observed in the native yeast sample were selected, refined, and subjected to one round of CTF refinement and Bayesian polishing. Thereupon, these particles were further sub-classified using an ellipsoid mask around eIF3j, and all particles containing eIF3j were selected as the final class containing both semi-open ABCE1 and eIF3j. This class represented 4.8% of the total dataset. Particles in this class were subjected to one more round of CTF refinement and Bayesian polishing, before 3D refinement and post-processing as well as local resolution estimation (all within RELION 3.1) yielded the final reconstructions at 3.0 Å average resolution, as shown in (Appendix Fig S1).

Data collection and processing of the human 40S complex sample

Data collection was performed on a Titan Krios at 300 kV, where 7,365 and 4,499 movies were collected for RIOK1-D324A and RIOK1, respectively, at a nominal pixel size of 1.059 Å and at a defocus range from 0.5 to 2.5 μm. Movies were recorded on a K2 Summit direct electron detector using low-dose conditions with 48 frames at approximately 1 e⁻/Å^2. All frames were gain corrected and subsequently aligned and summed using MotionCor2 (Zheng *et al*, 2017), and CTF parameters were determined using CTFFIND (Rohou & Grigorieff, 2015) and Gctf (Zhang, 2016). Particles were then picked using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/). Particle images were extracted in RELION 3.0 (Zivanov *et al*, 2018) and subjected to reference-free 2D classification. Good particles were selected, 3D refined, and classified. Besides the expected pre-40S classes (unpublished), one class containing the initiation complex was obtained in both datasets, comprising approximately 2% (RIOK1-D324A data set) and 8.7% (RIOK1 data set) of the total particle number. The two datasets were subsequently subjected to Bayesian polishing and CTF refinement, combined and further classified extensively as shown in Appendix Fig S2. Final reconstructions were then B-factor sharpened with RELION and the local resolution estimated. Where indicated (Appendix Fig S2), local or multi-body refinement was performed.

Model building and refinement

For rigid body fits and figures, Chimera version 1.13.1 (Pettersen *et al*, 2004) and ChimeraX version 0.91 (Goddard *et al*, 2018) were

used. Homology models were created using SWISS-MODEL Repository (Bienert *et al*, 2017; Waterhouse *et al*, 2018).

Yeast 43S PIC and 48S IC model

The atomic models PDB 5NDG (Prokhorova *et al*, 2017), 6FYY, 6FYX (Llacer *et al*, 2018), and 6TB3 (Buschauer *et al*, 2020) containing the models for *S.c.* 40S rRNA, r-proteins, and eIFs were fitted as rigid bodies into the cryo-EM maps of the *S.c.* 43S PIC and 48S IC. For the 43S PIC, the 40S rRNA and ribosomal proteins were fitted from PDB 5NDG and eIFs were fitted from PDB 6FYY. For the 48S IC, the 40S rRNA and ribosomal proteins were fitted from PDB 6TB3 and eIFs were fitted from PDB 6FYX. For ABCE1, the hybrid semi-open/closed model derived from the human 43S PIC (see below) was fitted into the density. For Hcr1, a homology model was created based on the structure of the human eIF3j dimer (PDB 3BPJ). The C-terminus of protomer 1 was extended by 3 amino acids, and the C-terminus of protomer 2 was extended by 19 amino acids based on a comparison with the data from the yeast 43S PIC-XL sample with the density observed for the native sample.

Models for the "mRNA entry position" of the YLC were obtained by fitting the crystal structure of eIF3i/g (PDB 4U1E, Erzberger *et al*, 2014) to the observed density as a rigid body and matching it to the structure of eIF3b CTD from PDB 6FYY; to obtain the "ES6 position", the eIF3i-eIF3g moiety bound to the C-terminal helix of eIF3b was rotated by 120 degrees around the Thr697-Asp701 hinge in the CTD of eIF3b as a rigid body.

Yeast 43S PIC-XL model

The atomic model 6TB3 (Buschauer *et al*, 2020) containing the models for *S.c.* 40S rRNA and r-proteins was split into 40S head and body and fitted as rigid bodies into the best-resolved cryo-EM map of the complex using ChimeraX version 1.0 (Goddard *et al*, 2018). The homology models for eIF3j and ABCE1 previously generated for the native 43S complex were also fitted, and the entire model was adjusted in WinCoot 0.8.9.2 (Emsley & Cowtan, 2004). The C-terminus of eIF3j protomer 2 was built *de novo*. A focused refined cryo-EM map for the NBD2 of ABCE1 was consulted to improve accuracy in the model for this domain during initial model building. The model was real space refined using Phenix 1.18 (Afonine *et al*, 2018; Liebschner *et al*, 2019).

Human 43S PIC

To obtain the atomic model, the best-resolved maps as obtained after local focused refinement or multi-body refinement (Fig EV2, Appendix Fig S2) were used to build the different parts of the *H.s.* 43S PIC. The 40S subunit was fit into maps of 40S body and 40S head obtained from multi-body refinement III (Appendix Fig S2) starting with the 40S model (PDB 6G5H, Ameismeier *et al*, 2018). After rigid body fitting, side chains of ribosomal proteins and rRNA were adjusted using Coot (version 0.8.9.2) (Emsley & Cowtan, 2004). Further, the 60S ribosomal protein eL41 was added to the model using PDB 6EK0 (chain h, Natchiar *et al*, 2017). For eIF1A, the homology model based on PDB 3J81 (Hussain *et al*, 2014) was fitted and adjusted using the 40S body map. The N-terminal helix bundle of eIF3c (47-149) was built *de novo* into the same map.

The homology model of the crystal structure of the C-terminal part of eIF3d (162-527; PDB 5K4B, Lee *et al*, 2016) was fitted into

the map for the 40S head obtained from multi-body refinement III (Appendix Fig S2). The atomic model was only modified in the regions interacting with the 40S head. Similarly, the model for eIF3b (PDB 5K1H, Simonetti *et al*, 2016) was only adjusted in blades 5 and 6, which contact the 40S body. Here, the best-resolved cryo-EM map, obtained by focused classification on the YLC, could be used (Appendix Fig S2). Also, the homology model of eIF3i (PDB 5K0Y, Simonetti *et al*, 2016) and an α-helix corresponding to the C-terminal part of eIF3a were fitted into this map.

The eIF3-PCI-MPN core (including eIF3a, c, e, f, h, k, l, m) was modeled into the two maps of multi-body refinement II (Appendix Fig S2) using the human homology model based on PDB 5A5T (des Georges *et al*, 2015) as starting model. eIF3d-N (2-84) was built *de novo* into the density.

For eIF3j, the unpublished crystal structure of the human eiF3j dimer (PDB 3BPJ) was fitted as rigid body into the density of 43S PIC state II.

Classification of the entire 43S dataset focusing on ABCE1 followed by focused refinement yielded a well-resolved map, which could be used for model building. A homology model based on the closed-state yeast ABCE1 bound to the 40S (PDB 5LL6, Heuer et al, 2017) was used as starting model. ATP and ADP were added to the NBSs.

One class obtained by focused classification on the YLC represents a very stable 43S complex in P_{OUT} state and yielded a well-resolved map of the TC after focused refinement. The models of tRNA$_i$ (PDB 6FEC, Eliseev *et al*, 2018), eIF2α, and eIF2γ (PDB 6O85, Kenner *et al*, 2019) and the homology models of eIF2β and eIF1 (based on PDB 6GSM) were fitted into the map and adjusted using Coot. Further, a stretch of 8 amino acids was modeled into the density adjacent to eIF1, which corresponds to eIF3c. For the unassigned RRM on top of 18S rRNA h16, we generated a poly-alanine model.

All models were real space refined using Phenix (version 1.17 and 1.18) (Afonine *et al*, 2018; Liebschner *et al*, 2019). In order to model the interfaces between the different parts of the structure, maps before and after multi-body refinement were used. Furthermore, neighboring parts were included in the real space refinement using focused cryo-EM maps. The final composite model was subjected to a final refinement using the overall cryo-EM map of state II and state III (Appendix Fig S2, Appendix Table S2). In regions with local resolution lower than 4 Å, side chains were not modeled.

Chemical crosslinking and mass spectrometry

For the crosslinking analysis followed by mass spectrometry of the ABCE1 43S pre-initiation complex, the sample was prepared as described above (see section Preparation of crosslinked yeast 43S pre-initiation complexes). After the sucrose density gradient, the buffer was exchanged to 50 mM HEPES/KOH pH 7.4, 50 mM KCl, 5 mM MgCl$_2$, 1 mM DTT, and crosslinking was performed using an equimolar mixture of isotopically labeled BS2G (H6/D6) (bis(sulfosuccinimidyl) 2,2,4,4-glutarate, Creative molecules) for 30 min at 1,200 rpm and 10°C. The reaction was quenched with 100 mM ammonium bicarbonate for 10 min. Digestion and peptide clean-up were performed using the EasyPep Mini MS Sample Prep Kit (Thermo Scientific), according to the manufacturer's protocol. Crosslinks were further enriched on a Superdex

Peptide PC 3.2/30 column (300 × 3.2 mm), and the fractions were analyzed by liquid chromatography coupled to tandem mass spectrometry with an LTQ-Orbitrap Elite (Thermo Scientific) instrument (Herzog *et al*, 2012). Identification of the crosslinked peptides was done using xQuest (Walzthoeni *et al*, 2012). The results were filtered with an MS1 tolerance window of −4 to 4 ppm and score ≥ 22 followed by manual validation.

Sequence alignments

In order to quantify the conservation of protein sequences between human and yeast proteins of interest, pairwise alignments were conducted using the T-Coffee implementation at https://toolkit.tuebingen.mpg.de (Notredame *et al*, 2000; Zimmermann *et al*, 2018) and visualized using JalView (Waterhouse *et al*, 2009). Multiple sequence alignments of the conserved elements of the eIF3c N-terminus were created using MAFFT (Katoh *et al*, 2019).

Data availability

Cryo-EM density map of the yeast 43S PIC: Electron Microscopy Data Bank 11160 (https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11160).

Atomic model of the yeast 43S PIC: Protein Data Bank 6ZCE (http://www.rcsb.org/structure/6ZCE).

Cryo-EM density map of the yeast 48S IC: Electron Microscopy Data Bank 11439 (https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11439).

Atomic model of the yeast 48S IC: Protein Data Bank 6ZU9 (http://www.rcsb.org/structure/6ZU9).

Cryo-EM density map of the yeast 43S PIC-XL: Electron Microscopy Data Bank 11608 (https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11608).

Atomic model of the yeast 43S PIC: Protein Data Bank 7A1G (http://www.rcsb.org/structure/7A1G).

Cryo-EM density map of the human 43S PIC—state II: Electron Microscopy Data Bank 11458 (https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11458).

Atomic model of the human 43S PIC—state II: Protein Data Bank 6ZVJ (http://www.rcsb.org/structure/6ZVJ).

Cryo-EM density map of the human 43S PIC—state III: Electron Microscopy Data Bank 11602 (https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-11602).

Atomic model of the human 43S PIC—state III: Protein Data Bank 7A09 (http://www.rcsb.org/structure/7A09).

Mass spectrometry proteomics data of the yeast 43S PIC: ProteomeXchange Consortium PXD020849 (http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD020849).

The cryo-EM density maps of the yeast 43S PIC, the yeast 48S IC, the yeast 43S PIC-XL, the human 43S-PIC state II, and the human 43S-PIC state III have been deposited in the Electron Microscopy Data Bank under accession numbers EMD-11160, EMD-11429, EMD-11608, EMD-11458, and EMD-11602, respectively (https://www.emdataresource.org/). Atomic coordinates for the atomic models have been deposited in the Protein Data Bank under accession number PDB ID 6ZCE, 6ZU9, 7A1G, 6ZVJ, and 7A09, respectively. (https://www.wwpdb.org/) (see Appendix Table S1 and S2). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE (Perez-Riverol *et al*, 2019) partner repository with the dataset identifier PXD020849.

Correspondence and requests for materials should be addressed to T.B. (becker@genzentrum.lmu.de) or R.B. (beckmann@genzentrum.lmu.de).

Expanded View for this article is available online.

Author contributions

HK, T M-K, TB, and RB designed the study. MA prepared the sample for the human and TM-K for the yeast native 40S complexes. TM-K and HK prepared components for *in vitro* splitting assays and TM-K performed splitting assays. MA, TM-K, and OB collected and MA and TM-K processed the cryo-EM data. HK, JC, and TM-K built and refined the model. HK, TM-K, TB, and RB analyzed and interpreted the structures. ED, AN, and MF-R performed ABCE1-TAP purification and quantitative label-free MS. MP and FH performed crosslinking mass spectrometry. TB, HK, TM-K, and RB wrote the manuscript with contributions from all authors.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Algire MA, Maag D, Lorsch JR (2005) Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* 20: 251–262

Acker MG, Shin BS, Dever TE, Lorsch JR (2006) Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. *J Biol Chem* 281: 8469–8475

Acker MG, Shin BS, Nanda JS, Saini AK, Dever TE, Lorsch JR (2009) Kinetic analysis of late steps of eukaryotic translation initiation. *J Mol Biol* 385: 491–506

Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* 74: 531–544

Ameismeier M, Cheng J, Berninghausen O, Beckmann R (2018) Visualizing late states of human 40S ribosomal subunit maturation. *Nature* 558: 249–253

Andersen DS, Leevers SJ (2007) The essential Drosophila ATP-binding cassette domain protein, pixie, binds the 40 S ribosome in an ATP-dependent manner and is required for translation initiation. *J Biol Chem* 282: 14752–14760

Anger AM, Armache JP, Berninghausen O, Habeck M, Subklewe M, Wilson DN, Beckmann R (2013) Structures of the human and Drosophila 80S ribosome. *Nature* 497: 80–85

Asano K, Clayton J, Shalev A, Hinnebusch AG (2000) A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. *Genes Dev* 14: 2534–2546

Aylett CH, Boehringer D, Erzberger JP, Schaefer T, Ban N (2015) Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nat Struct Mol Biol* 22: 269–271

Barthelme D, Dinkelaker S, Albers SV, Londei P, Ermler U, Tampe R (2011) Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. *Proc Natl Acad Sci USA* 108: 3228–3233

Barthelme D, Scheele U, Dinkelaker S, Janoschka A, Macmillan F, Albers SV, Driessen AJ, Stagni MS, Bill E, Meyer-Klaucke W, *et al* (2007) Structural organization of essential iron-sulfur clusters in the evolutionarily highly conserved ATP-binding cassette protein ABCE1. *J Biol Chem* 282: 14598–14607

Becker T, Franckenberg S, Wickles S, Shoemaker CJ, Anger AM, Armache J-P, Sieber H, Ungewickell C, Berninghausen O, Daberkow I, *et al* (2012) Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature* 482: 501–506

Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. *Science* 334: 1524–1529

Beznoskova P, Cuchalova L, Wagner S, Shoemaker CJ, Gunisova S, von der Haar T, Valasek LS (2013) Translation initiation factors eIF3 and HCR1 control translation termination and stop codon read-through in yeast cells. *PLoS Genet* 9: e1003962

Bienert S, Waterhouse A, de Beer TA, Tauriello G, Studer G, Bordoli L, Schwede T (2017) The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* 45: D313–D319

Block KL, Vornlocher HP, Hershey JW (1998) Characterization of cDNAs encoding the p44 and p35 subunits of human translation initiation factor eIF3. *J Biol Chem* 273: 31901–31908

Brown A, Baird MR, Yip MC, Murray J, Shao S (2018) Structures of translationally inactive mammalian ribosomes. *Elife* 7

Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V (2015) Structural basis for stop codon recognition in eukaryotes. *Nature* 524: 493–496

Cate JH (2017) Human eIF3: from 'blobology' to biological insight. *Philos Trans R Soc Lond B Biol Sci* 372

Buschauer R, Matsuo Y, Sugiyama T, Chen YH, Alhusaini N, Sweet T, Ikeuchi K, Cheng J, Matsuki Y, Nobuta R, *et al* (2020) The Ccr4-Not complex monitors the translating ribosome for codon optimality. *Science* 368

Chen ZQ, Dong J, Ishimura A, Daar I, Hinnebusch AG, Dean M (2006) The essential vertebrate ABCE1 protein interacts with eukaryotic initiation factors. *J Biol Chem* 281: 7452–7457

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26: 1367–1372

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10: 1794–1805

Cuchalova L, Kouba T, Herrmannova A, Danyi I, Chiu WL, Valasek L (2010) The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. *Mol Cell Biol* 30: 4671–4686

Das S, Ghosh R, Maitra U (2001) Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *J Biol Chem* 276: 6720–6726

des Georges A, Dhote V, Kuhn L, Hellen CUT, Pestova TV, Frank J, Hashem Y (2015) Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* 525: 491–495

Dong J, Lai R, Nielsen K, Fekete CA, Qiu H, Hinnebusch AG (2004) The essential ATP-binding cassette protein RLI1 functions in translation by promoting preinitiation complex assembly. *J Biol Chem* 279: 42157–42168

Dong Z, Qi J, Peng H, Liu J, Zhang JT (2013) Spectrin domain of eukaryotic initiation factor 3a is the docking site for formation of the a:b:i:g subcomplex. *J Biol Chem* 288: 27951–27959

ElAntak L, Tzakos AG, Locker N, Lukavsky PJ (2007) Structure of eIF3b RNA recognition motif and its interaction with eIF3j: structural insights into the recruitment of eIF3b to the 40 S ribosomal subunit. *J Biol Chem* 282: 8165–8174

Elantak L, Wagner S, Herrmannova A, Karaskova M, Rutkai E, Lukavsky PJ, Valasek L (2010) The indispensable N-terminal half of eIF3j/HCR1 cooperates with its structurally conserved binding partner eIF3b/PRT1-RRM and with eIF1A in stringent AUG selection. *J Mol Biol* 396: 1097–1116

Eliseev B, Yeramala L, Leitner A, Karuppasamy M, Raimondeau E, Huard K, Alkalaeva E, Aebersold R, Schaffitzel C (2018) Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res* 46: 2678–2689

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

Erzberger JP, Stengel F, Pellarin R, Zhang S, Schaefer T, Aylett CHS, Cimermancic P, Boehringer D, Sali A, Aebersold R, *et al* (2014) Molecular Architecture of the 40SeIF1eIF3 Translation Initiation Complex. *Cell* 159: 1227–1228

Fraser CS, Berry KE, Hershey JW, Doudna JA (2007) eIF3j is located in the decoding center of the human 40S ribosomal subunit. *Mol Cell* 26: 811–819

Fraser CS, Lee JY, Mayeur GL, Bushell M, Doudna JA, Hershey JW (2004) The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40 S ribosomal subunits in vitro. *J Biol Chem* 279: 8946–8956

Gerovac M, Tampe R (2019) Control of mRNA Translation by Versatile ATP-Driven Machines. *Trends Biochem Sci* 44: 167–180

Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. *Nature* 425: 737–741

Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* 27: 14–25

Gouridis G, Hetzert B, Kiosze-Becker K, de Boer M, Heinemann H, Nurenberg-Goloub E, Cordes T, Tampe R (2019) ABCE1 Controls Ribosome Recycling by an Asymmetric Dynamic Conformational Equilibrium. *Cell Rep* 28 (723–734): e726

Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, Hellen CUT, Pestova TV, Frank J (2013) Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* 153: 1108–1119

Herrmannova A, Daujotyte D, Yang JC, Cuchalova L, Gorrec F, Wagner S, Danyi I, Lukavsky PJ, Valasek LS (2012) Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. *Nucleic Acids Res* 40: 2294–2311

Herzog F, Kahraman A, Boehringer D, Mak R, Bracher A, Walzthoeni T, Leitner A, Beck M, Hartl FU, Ban N, *et al* (2012) Structural probing of a protein

phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* 337: 1348–1352

Heuer A, Gerovac M, Schmidt C, Trowitzsch S, Preis A, Kötter P, Berninghausen O, Becker T, Beckmann R, Tampé R (2017) Structure of the 40S–ABCE1 post-splitting complex in ribosome recycling and translation initiation. *Nature Structural & Molecular Biology* 24: 453–460

Hinnebusch AG (2006) eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem Sci* 31: 553–562

Hopfner KP (2016) Invited review: Architectures and mechanisms of ATP binding cassette proteins. *Biopolymers* 105: 492–504

Huang HK, Yoon H, Hannig EM, Donahue TF (1997) GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in Saccharomyces cerevisiae. *Genes Dev* 11: 2396–2413

Hubner NC, Mann M (2011) Extracting gene function from protein-protein interactions using Quantitative BAC InteraCtomics (QUBIC). *Methods* 53: 453–459

Hussain T, Llacer JL, Fernandez IS, Munoz A, Martin-Marcos P, Savva CG, Lorsch JR, Hinnebusch AG, Ramakrishnan V (2014) Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* 159: 597–607

Imai H, Abe T, Miyoshi T, Nishikawa SI, Ito K, Uchiumi T (2018) The ribosomal stalk protein is crucial for the action of the conserved ATPase ABCE1. *Nucleic Acids Res* 46: 7820–7830

Karcher A, Schele A, Hopfner KP (2008) X-ray structure of the complete ABC enzyme ABCE1 from Pyrococcus abyssi. *J Biol Chem* 283: 7962–7971

Katoh K, Rozewicki J, Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20: 1160–1166

Kenner LR, Anand AA, Nguyen HC, Myasnikov AG, Klose CJ, McGeever LA, Tsai JC, Miller-Vedam LE, Walter P, Frost A (2019) eIF2B-catalyzed nucleotide exchange and phosphoregulation by the integrated stress response. *Science* 364: 491–495

Khoshnevis S, Gunisova S, Vlckova V, Kouba T, Neumann P, Beznoskova P, Ficner R, Valasek LS (2014) Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes. *Nucleic Acids Res* 42: 4123–4139

Kiosze-Becker K, Ori A, Gerovac M, Heuer A, Nürenberg-Goloub E, Rashid UJ, Becker T, Beckmann R, Beck M, Tampé R (2016) Structure of the ribosome post-recycling complex probed by chemical cross-linking and mass spectrometry. *Nature Communications* 7: 13248

Lee AS, Kranzusch PJ, Doudna JA, Cate JH (2016) eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* 536: 96–99

Lee HH, Kim YS, Kim KH, Heo I, Kim SK, Kim O, Kim HK, Yoon JY, Kim HS, Kim DJ, *et al* (2007) Structural and functional insights into Dom34, a key component of no-go mRNA decay. *Mol Cell* 27: 938–950

Lee JH, Pestova TV, Shin BS, Cao C, Choi SK, Dever TE (2002) Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proc Natl Acad Sci USA* 99: 16689–16694

Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ, *et al* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* 75: 861–877

Llacer JL, Hussain T, Marler L, Aitken CE, Thakur A, Lorsch JR, Hinnebusch AG, Ramakrishnan V (2015) Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol Cell* 59: 399–412

Llacer JL, Hussain T, Saini AK, Nanda JS, Kaur S, Gordiyenko Y, Kumar R, Hinnebusch AG, Lorsch JR, Ramakrishnan V (2018) Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *Elife* 7

Mancera-Martinez E, Brito Querido J, Valasek LS, Simonetti A, Hashem Y (2017) ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. *RNA Biol* 14: 1279–1285

Martin-Marcos P, Nanda JS, Luna RE, Zhang F, Saini AK, Cherkasova VA, Wagner G, Lorsch JR, Hinnebusch AG (2014) Enhanced eIF1 binding to the 40S ribosome impedes conformational rearrangements of the preinitiation complex and elevates initiation accuracy. *RNA* 20: 150–167

Mohammad MP, Munzarova Pondelickova V, Zeman J, Gunisova S, Valasek LS (2017) In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. *Nucleic Acids Res* 45: 2658–2674

Natchiar SK, Myasnikov AG, Kratzat H, Hazemann I, Klaholz BP (2017) Visualization of chemical modifications in the human 80S ribosome structure. *Nature* 551: 472–477

Nielsen KH, Valasek L, Sykes C, Jivotovskaya A, Hinnebusch AG (2006) Interaction of the RNP1 motif in PRT1 with HCR1 promotes 40S binding of eukaryotic initiation factor 3 in yeast. *Mol Cell Biol* 26: 2984–2998

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302: 205–217

Nürenberg-Goloub E, Heinemann H, Gerovac M, Tampé R (2018) Ribosome recycling is coordinated by processive events in two asymmetric ATP sites of ABCE1. *Life Science Alliance* 1(3): e201800095

Nürenberg-Goloub E, Kratzat H, Heinemann H, Heuer A, Kötter P, Berninghausen O, Becker T, Tampé R, Beckmann R (2020) Molecular analysis of the ribosome recycling factor ABCE1 bound to the 30S post-splitting complex. *The EMBO Journal* n/a e103788

Nurenberg-Goloub E, Tampe R (2019) Ribosome recycling in mRNA translation, quality control, and homeostasis. *Biol Chem* 401: 47–61

Obayashi E, Luna R, Nagata T, Martin-Marcos P, Hiraishi H, Singh C, Erzberger J, Zhang F, Arthanari H, Morris J, *et al* (2017) Molecular Landscape of the Ribosome Pre-initiation Complex during mRNA Scanning: Structural Role for eIF3c and Its Control by eIF5. *Cell Reports* 18: 2651–2663

Passmore LA, Schmeing TM, Maag D, Applefield DJ, Acker MG, Algire MA, Lorsch JR, Ramakrishnan V (2007) The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* 26: 41–50

Paulin FE, Campbell LE, O'Brien K, Loughlin J, Proud CG (2001) Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Curr Biol* 11: 55–59

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

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, *et al* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47: D442–D450

Pestova TV, Lomakin IB, Lee JH, Choi SK, Dever TE, Hellen CU (2000) The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403: 332–335

Pisarev AV, Hellen CU, Pestova TV (2007) Recycling of eukaryotic posttermination ribosomal complexes. *Cell* 131: 286–299

Pisarev AV, Skabkin MA, Pisareva VP, Skabkina OV, Rakotondrafara AM, Hentze MW, Hellen CU, Pestova TV (2010) The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell* 37: 196–210

Pisareva VP, Skabkin MA, Hellen CU, Pestova TV, Pisarev AV (2011) Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J* 30: 1804–1817

Preis A, Heuer A, Barrio-Garcia C, Hauser A, Eyler DE, Berninghausen O, Green R, Becker T, Beckmann R (2014) Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. *Cell Rep* 8: 59–65

Prokhorova I, Altman RB, Djumagulov M, Shrestha JP, Urzhumtsev A, Ferguson A, Chang CT, Yusupov M, Blanchard SC, Yusupova G (2017) Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc Natl Acad Sci USA* 114: E10899–E10908

Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192: 216–221

Scheres SH (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180: 519–530

Schluenzen F, Tocilj A, Zarivach R, Harms J, Gluehmann M, Janell D, Bashan A, Bartels H, Agmon I, Franceschi F, *et al* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* 102: 615–623

Schmitt E, Blanquet S, Mechulam Y (2002) The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. *EMBO J* 21: 1821–1832

Schmitt E, Panvert M, Lazennec-Schurdevin C, Coureux PD, Perez J, Thompson A, Mechulam Y (2012) Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat Struct Mol Biol* 19: 450–454

Sha Z, Brill LM, Cabrera R, Kleifeld O, Scheliga JS, Glickman MH, Chang EC, Wolf DA (2009) The eIF3 interactome reveals the translasome, a supercomplex linking protein synthesis and degradation machineries. *Mol Cell* 36: 141–152

Shao S, Brown A, Santhanam B, Hegde RS (2015) Structure and assembly pathway of the ribosome quality control complex. *Mol Cell* 57: 433–444

Shoemaker CJ, Green R (2011) Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc Natl Acad Sci USA* 108: E1392–1398

Simonetti A, Brito Querido J, Myasnikov AG, Mancera-Martinez E, Renaud A, Kuhn L, Hashem Y (2016) eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. *Mol Cell* 63: 206–217

Sokabe M, Fraser CS, Hershey JW (2012) The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. *Nucleic Acids Res* 40: 905–913

Srivastava S, Verschoor A, Frank J (1992) Eukaryotic initiation factor 3 does not prevent association through physical blockage of the ribosomal subunit-subunit interface. *J Mol Biol* 226: 301–304

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13: 731–740

Valasek L, Hasek J, Trachsel H, Imre EM, Ruis H (1999) The Saccharomyces cerevisiae HCR1 gene encoding a homologue of the p35 subunit of human translation initiation factor 3 (eIF3) is a high copy suppressor of a temperature-sensitive mutation in the Rpg1p subunit of yeast eIF3. *J Biol Chem* 274: 27567–27572

Valasek L, Mathew AA, Shin BS, Nielsen KH, Szamecz B, Hinnebusch AG (2003) The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome in vivo. *Genes Dev* 17: 786–799

Valasek L, Nielsen KH, Hinnebusch AG (2002) Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. *EMBO J* 21: 5886–5898

Valasek L, Nielsen KH, Zhang F, Fekete CA, Hinnebusch AG (2004) Interactions of eukaryotic translation initiation factor 3 (eIF3) subunit NIP1/c with eIF1 and eIF5 promote preinitiation complex assembly and regulate start codon selection. *Mol Cell Biol* 24: 9437–9455

Valasek L, Phan L, Schoenfeld LW, Valaskova V, Hinnebusch AG (2001) Related eIF3 subunits TIF32 and HCR1 interact with an RNA recognition motif in PRT1 required for eIF3 integrity and ribosome binding. *EMBO J* 20: 891–904

Valasek LS, Zeman J, Wagner S, Beznoskova P, Pavlikova Z, Mohammad MP, Hronova V, Herrmannova A, Hashem Y, Gunisova S (2017) Embraced by eIF3: structural and functional insights into the roles of eIF3 across the translation cycle. *Nucleic Acids Res* 45: 10948–10968

Walker SE, Zhou F, Mitchell SF, Larson VS, Valasek L, Hinnebusch AG, Lorsch JR (2013) Yeast eIF4B binds to the head of the 40S ribosomal subunit and promotes mRNA recruitment through its N-terminal and internal repeat domains. *RNA* 19: 191–207

Walzthoeni T, Claassen M, Leitner A, Herzog F, Bohn S, Forster F, Beck M, Aebersold R (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat Methods* 9: 901–903

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, *et al* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46: W296–W303

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25: 1189–1191

Wessel D, Flugge UI (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* 138: 141–143

Yamamoto Y, Singh CR, Marintchev A, Hall NS, Hannig EM, Wagner G, Asano K (2005) The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G. *Proc Natl Acad Sci USA* 102: 16164–16169

Young DJ, Guydosh NR (2019) Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. *Cell Rep* 28(39–50): e34

Yu Y, Marintchev A, Kolupaeva VG, Unbehaun A, Veryasova T, Lai SC, Hong P, Wagner G, Hellen CU, Pestova TV (2009) Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing. *Nucleic Acids Res* 37: 5167–5182

Zeman J, Itoh Y, Kukacka Z, Rosulek M, Kavan D, Kouba T, Jansen ME, Mohammad MP, Novak P, Valasek LS (2019) Binding of eIF3 in complex with eIF5 and eIF1 to the 40S ribosomal subunit is accompanied by dramatic structural changes. *Nucleic Acids Res* 47: 8282–8300

Zhang E, Khanna V, Dacheux E, Namane A, Doyen A, Gomard M, Turcotte B, Jacquier A, Fromont-Racine M (2019) A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. *EMBO J* 38: e100640

Zhang K (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol* 193: 1–12

Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y, Agard DA (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* 14: 331–332

Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding J, Lupas AN, Alva V (2018) A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* 430: 2237–2243

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7: e42166

STRUCTURAL BIOLOGY

90S pre-ribosome transformation into the primordial 40S subunit

Jingdong Cheng[1]*, Benjamin Lau[2]*, Giuseppe La Venuta[2], Michael Ameismeier[1], Otto Berninghausen[1], Ed Hurt[2]†, Roland Beckmann[1]†

Production of small ribosomal subunits initially requires the formation of a 90S precursor followed by an enigmatic process of restructuring into the primordial pre-40S subunit. We elucidate this process by biochemical and cryo–electron microscopy analysis of intermediates along this pathway in yeast. First, the remodeling RNA helicase Dhr1 engages the 90S pre-ribosome, followed by Utp24 endonuclease–driven RNA cleavage at site A₁, thereby separating the 5′-external transcribed spacer (ETS) from 18S ribosomal RNA. Next, the 5′-ETS and 90S assembly factors become dislodged, but this occurs sequentially, not en bloc. Eventually, the primordial pre-40S emerges, still retaining some 90S factors including Dhr1, now ready to unwind the final small nucleolar U3–18S RNA hybrid. Our data shed light on the elusive 90S to pre-40S transition and clarify the principles of assembly and remodeling of large ribonucleoproteins.

The formation of eukaryotic ribosomes requires transcription, processing, and modification of ribosomal RNA (rRNA) and the integration of ~80 ribosomal proteins. This highly complex process starts in the nucleolus with the transcription of a large rRNA precursor (35S in yeast, 47S in human), which contains the 18S rRNA of the 40S small subunit and the 5.8S and 25S rRNAs of the 60S large subunit, separated by the internal transcribed spacers ITS1 and ITS2 and flanked by the external transcribed spacers 5′-ETS and 3′-ETS. During transcription of the 35S pre-rRNA by RNA polymerase I, the earliest stable assembly intermediate, called the 90S pre-ribosome or small subunit processome, forms. This comprises more than 50 assembly factors, the U3 small nucleolar ribonucleoprotein (U3 snoRNP), the nascent pre-rRNA, and dozens of the small subunit ribosomal proteins (1, 2). In the subsequent maturation steps, coordinated endonucleolytic cleavage at site A₁ within the 5′-ETS and at A₂ within the ITS1 is key to the separation of the 18S rRNA and 5.8S/25S rRNA (Fig. 1A), which then follow independent biogenesis pathways [reviewed in (3–5)]. The remaining 5′-ETS is degraded by RNA nucleases, including the nuclear exosome (6, 7).

Recently, several structures of the 90S pre-ribosome in stages prior to A₁ cleavage were solved by cryo–electron microscopy (cryo-EM) single-particle analysis from *Chaetomium thermophilum* and *Saccharomyces cerevisiae* (8–12). The four subdomains of the 18S rRNA (5′, central, 3′ major, and 3′ minor) were found to fold independently and to associate with their set of early ribosomal proteins of the small subunit (S-proteins) cotranscriptionally, before integrating into the 90S pre-ribosome (13, 14). This integration happens in a reverse order, in which the 3′ major and 3′ minor domains incorporate first, followed by the central and 5′ domain (15). Together with its

[1]Gene Center, Department of Biochemistry, University of Munich, 81377 Munich, Germany. [2]Biochemistry Center (BZH), University of Heidelberg, 69120 Heidelberg, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: beckmann@genzentrum.lmu.de (R.B.); ed.hurt@bzh.uni-heidelberg.de (E.H.)
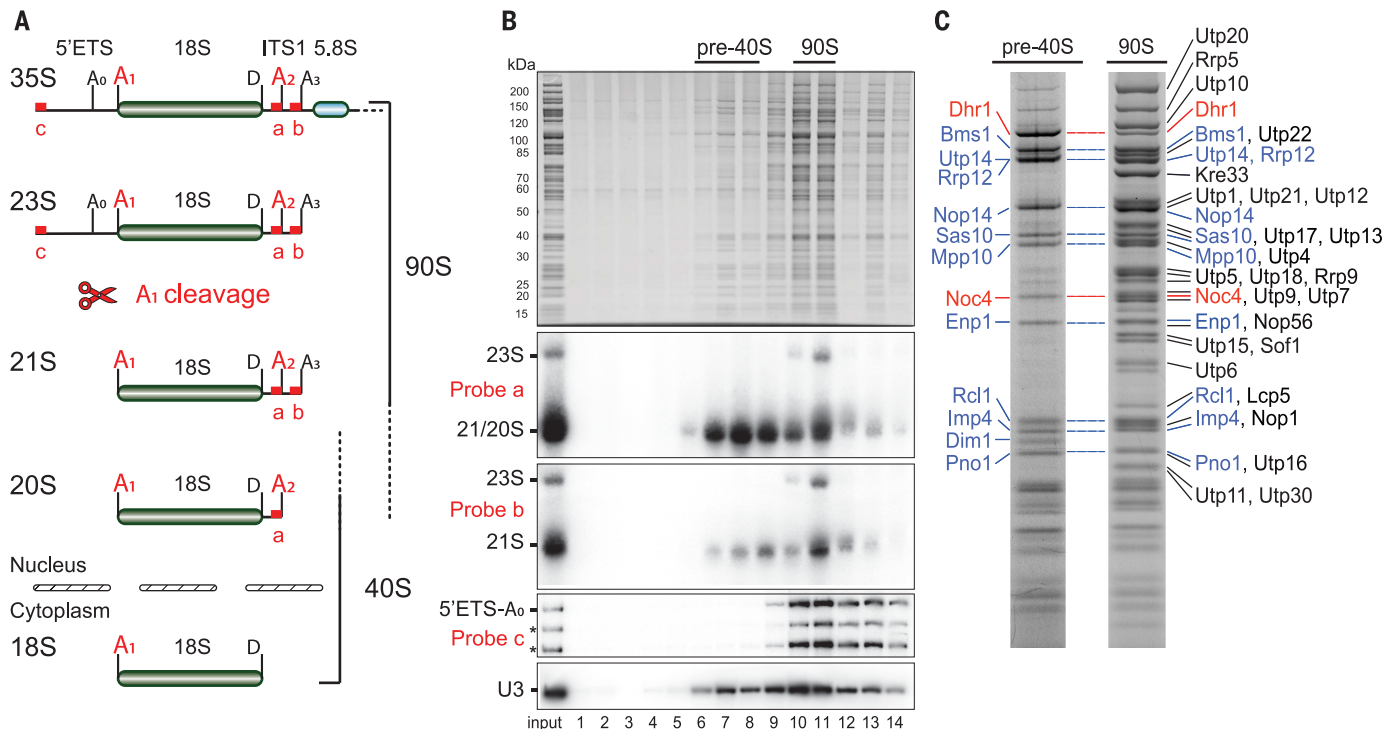
Fig. 1. Pre-rRNA processing and purification of 90S to pre-40S transition intermediates. (**A**) Sequence of rRNA processing events during 40S biogenesis in yeast with A₀, A₁, A₂, A₃, and D as cleavage sites on the pre-rRNA. (**B**) Analysis of split-tag (Noc4-TAP–Dhr1-Flag) affinity-purified assembly intermediates by sucrose gradient centrifugation, followed by SDS–polyacrylamide gel electrophoresis (PAGE) and Coomassie staining or Northern blotting, using the RNA probes a, b, and c to detect the various indicated pre-rRNA forms (23S, 21S, 20S, 5′-ETS-A₀). (**C**) SDS-PAGE of sucrose gradient fractions containing the pre-40S and 90S intermediates, with ribosome assembly factors identified by mass spectrometry indicated (bait proteins in red, factors shared by pre-40S and 90S in blue). See fig. S1B for the whole sucrose gradient.

numerous associated 90S assembly factors, which predominantly are organized in modules (e.g., UTP-A, UTP-B, UTP-C, U3 snoRNP, Mpp10 complex, Bms1-Rcl1 complex, Kre33 module, Noc4 module, Krr1-Faf1 complex), the 90S pre-ribosome is fully assembled and primed for subsequent A_1 cleavage (15–20).

After this cleavage, the 90S pre-ribosome is transformed into a primordial pre-40S intermediate (a process hereafter referred to as 90S transition) by a still-elusive mechanism; only later pre-40S intermediates have been described so far (21–30). The early 90S transition was initially uncovered by biochemical studies of Enp1, a ribosome assembly factor that is present in both early 90S and late pre-40S intermediates (21). However, many additional proteins have since been suggested to be intimately involved in this transition. Candidate factors include the methyltransferase Dim1 (31), the KH domain proteins Krr1 and Pno1 (previously called Dim2) (32), RNA helicase Dhr1 (33, 34), and the nuclear RNA exosome system (6). To decipher the mechanisms of this transition, we aimed at a structural analysis of 90S transition intermediates by cryo-EM.

Isolation of ribosome assembly intermediates marking the 90S to pre-40S transition

Late 90S and early 40S pre-ribosomal intermediates were purified from *S. cerevisiae* using 90S assembly factors Noc4 and Dhr1 as baits in split-tag affinity purifications. Dhr1 was chosen because this RNA helicase regulates A_1 cleavage and eventually dissociates distinct heteroduplexes between U3 and pre-18S rRNA (35, 36). Subsequent separation of these preparations on a sucrose gradient yielded 90S and smaller unusual pre-40S intermediates, which contained not only classical pre-40S factors (e.g., Rrp12, Enp1, Dim1, Pno1) (21), but also Dhr1 and factors so far only assigned to 90S complexes (e.g., Bms1, Rcl1, Utp14, Sas10, Mpp10, and Imp4; Fig. 1, B and C, fig. S1, and data S1). After finding Dim1 enriched in this unusual pre-40S fraction (data S1; see also Fig. 1C), we used it together with Dhr1 in another split-tag affinity purification to enrich further intermediates of the 90S transition (fig. S1C).

Northern blotting of the Noc4-Dhr1 purification revealed that the 90S fraction contains the 5′-ETS–A_0 fragment, which is also typical for other 90S preparations [e.g., Enp1-FTpA,

Utp22-FTpA, or Pwp2-FTpA (8)]. It also contains 23S pre-rRNA (cleavage fragment 5′ end–A_3) and 21S pre-rRNA (A_1-A_3 fragment), which together are indicative of both pre-A_1 and post-A_1 cleavage states. In contrast, the unusual pre-40S fraction exhibited the strong presence of 20S pre-rRNA (A_1-A_2 fragment), which suggests that this intermediate completed both A_1 and A_2 processing. Both the 90S and 40S fractions contained the U3 snoRNA (Fig. 1, A and B).

## Cryo-EM structures revealing the 90S to pre-40S transition

After 3D classification, cryo-EM of the Noc4-Dhr1 and Dhr1-Dim1 samples revealed seven well-defined and distinct classes of pre-ribosomal particles, which were put in a temporal order covering the 90S transition (Fig. 2). Four of these states (called B2, Pre-A_1, Post-A_1, and Dis-C) were resolved at average resolution between 3.5 and 3.8 Å, allowing us to build and refine complete models. The other three states (A, Dis-A, and Dis-B) displayed average resolution between 4.4 and 7.1 Å, permitting rigid-body docking of molecular models (Fig. 2, figs. S2 to S6, and table S1).
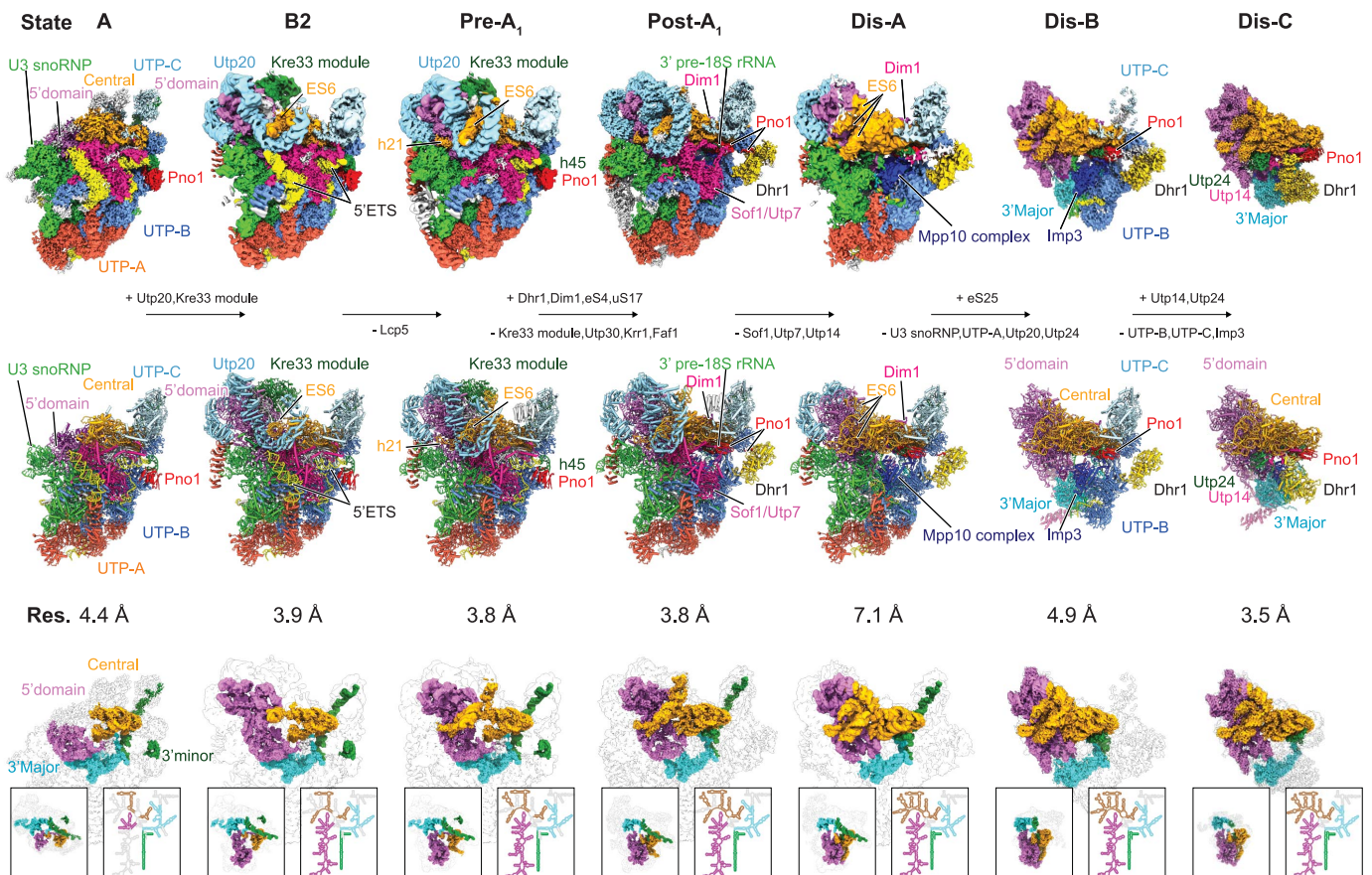
Fig. 2. Cryo-EM analysis of 90S to pre-40S transition intermediates. Cryo-EM maps (top) and molecular models (middle) of distinct states of yeast 90S to pre-40S transition observed after split-tag affinity purification using Noc4-Dhr1 (all states) and Dhr1-Dim1 (states B2, Post-A_1, Dis-A, Dis-B, Dis-C) are shown. Assembly factors and modules are labeled and compositional changes indicated. Bottom: Depiction of pre-18S rRNA density using 90S views with corresponding rRNA secondary structures; 40S views are shown in boxes (color code: magenta, 5′ domain; orange, central domain; cyan, 3′ major domain; green, 3′ minor domain).
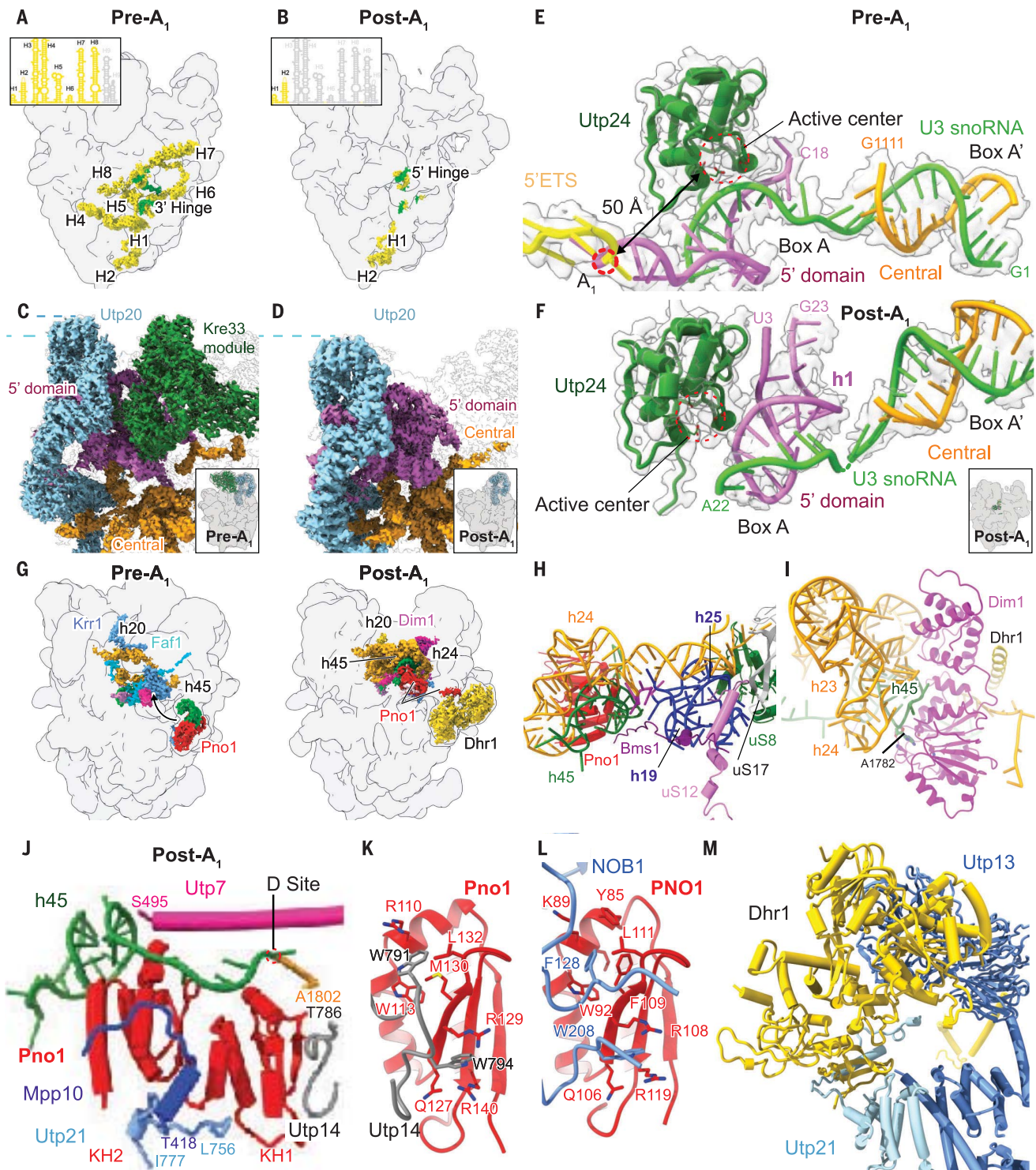
Fig. 3. Conformational, positional, and compositional changes upon A1 cleavage. (**A** and **B**) Dismantling of 5′-ETS RNA helices (yellow) from Pre-A₁ (A) and Post-A₁ (B) is shown with interacting U3 snoRNA (green) in transparent 90S density. The secondary structure of the 5′-ETS RNA is shown with observed helices in yellow and the dislodged helices in gray. (**C** and **D**) Volume representation of Utp20, the Kre33 module, and the 5′ and central domains of the 18S pre-RNA highlights the compaction from state Pre-A₁ (C) to Post-A₁ (D). Dashed lines indicate the movement of Utp20. (**E** and **F**) Model of the h1 region of the 18S pre-rRNA aligned at Utp24 shown before (E) and after (F) A₁ cleavage. The catalytic center of Utp24 and the A₁ cleavage site are indicated by dashed circles. (**G**) Differences in the 18S central domain region are highlighted in states Pre-A₁ and Post-A₁. (**H**) Ribbon representation of the platform region in state Post-A₁. Nucleotides 1021 to 1025 of the 18S rRNA are shown in magenta. (**I**) Model of Dim1 bound to its substrate adenosine 1782 of the pre-18S in state Post-A₁. (**J**) Model of Pno1 in state Post-A₁ with the 3′ end of the 18S pre-rRNA and interacting segments of Mpp10, Utp7, Utp14, and Utp21. (**K** and **L**) Models of yeast and human Pno1 bound to Utp14 (K) and NOB1 [(L), PDB ID 6G18], respectively, highlight overlapping interaction sites. Key residues are shown as sticks. (**M**) Ribbon representation showing the interaction of Dhr1 with Utp21 and Utp13 in state Post-A₁. Amino acid abbreviations: F, Phe; I, Ile; K, Lys; L, Leu; M, Met; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr.
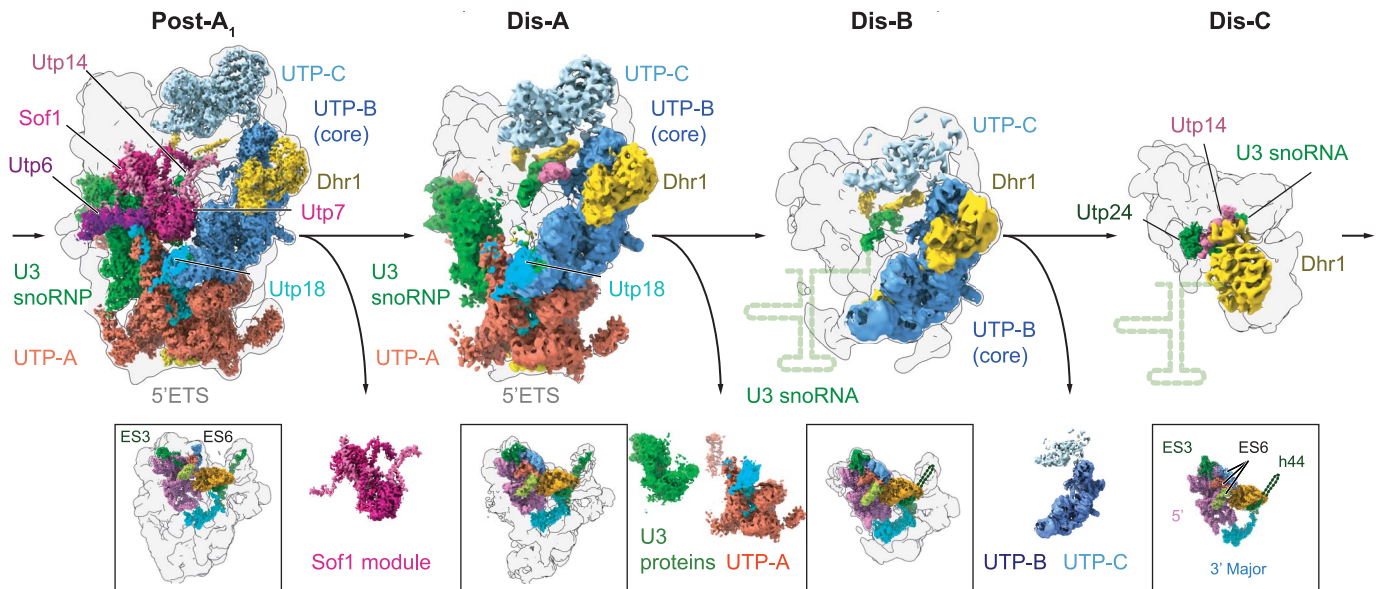
Fig. 4. Successive shedding of factors during transition from 90*S* to primordial pre-40*S*. EM densities reveal the shedding process of major 90*S* assembly factor modules during transition from state Post-A$_1$ to state Dis-C. Proteins, modules, and the U3 snoRNP are colored and labeled accordingly. The last remaining helices of the 5′-ETS (H1, H2) and Dhr1 are shown in yellow. A local resolution-filtered map of Dhr1 was used in state Dis-C. The color-coded 18*S* pre-rRNA and the dismantled modules are shown below.

The first two particles in this series, states A and B2, closely resemble previously described 90*S* assembly intermediates, with a noncleaved A$_1$ site as a characteristic feature (Fig. 2 and fig. S6) (*8*, *11*, *12*, *15*). In the next state, Pre-A$_1$, which is the last state preceding A$_1$ processing, we observed the integration of helix 21 (h21) of the pre-18*S* rRNA at its mature position (Fig. 2 and fig. S6). Subsequently and concomitant with cleavage at site A$_1$, major structural changes lead to the Post-A$_1$ intermediate. Specifically, Pno1, together with h45 of the 18*S* rRNA, replaces the Krr1-Faf1 complex, bringing h45 from a peripheral to an integrated mature position. Furthermore, the remodeling helicase Dhr1 is observed for the first time in this post-A$_1$ state intermediate, positioned at the site vacated by Pno1 (Fig. 2 and fig. S6).

The succeeding group of intermediates comprises states Dis-A, Dis-B, and Dis-C (where Dis stands for dissociation), which represent the actual 90*S* to pre-40*S* transition intermediate states. Unexpectedly, the canonical pre-40*S* is not released as a whole entity after A$_1$ cleavage by leaving a presumed 5′-ETS particle behind. Instead, several prominent 90*S* assembly factor modules dissociate one after the other, leading to a progressive simplification of these complexes (Fig. 2 and fig. S6). The last state of this series, Dis-C, after ultimately losing the conspicuous UTP-B module, resembles a 40*S* shape for the first time, clearly displaying the 5′, central (platform), and flexible and immature 3′ major and 3′ minor domains. Here, the characteristic h44 of the 3′ minor domain, which in preceding states is held in a defined immature position by UTP-B and UTP-C, is mostly

delocalized. This is because the immature conformation is no longer stabilized by the 90*S* modules, and the mature position is still masked by the Bms1-Rcl1 complex (Fig. 2 and fig. S6).

Conformational changes and mechanism of A$_1$ cleavage

In contrast to the expected en bloc release of a 5′-ETS particle upon A$_1$ cleavage (*5*, *7*, *8*, *37*), we observed that the 5′-ETS RNA becomes increasingly disordered during a stepwise 90*S* transition. Helix H9 at the 3′ end of the 5′-ETS is the first to become detached from Utp20 (fig. S7, A to F). Next, after A$_1$ cleavage, other prominent 5′-ETS helices (H3 to H8) disappear, leaving behind empty cavities in this part of the 90*S* intermediate (Fig. 3A and fig. S7, A to C). Only the first two helices, H1 and H2, and two distinct internal 5′-ETS segments that form short heteroduplexes with the 3′ hinge and 5′ hinge of U3 snoRNA remain in position (Fig. 3B and figs. S4 and S7C). The cause for the increased structural freedom or partial degradation of the 5′-ETS is not clear. Notably, despite the dismantling of the 5′-ETS RNA, the overall shape and composition of this 90*S* intermediate remained largely unchanged at first.

Nonetheless, during this 90*S* transition, a clear structural compaction was observed. Initially, in state B2, the C terminus of Utp20 exhibits a superhelical HEAT repeat structure that wraps around the rRNA expansion segment ES6c (fig. S7D). Then, after transition to the Pre-A$_1$ state, the Utp20 superhelix adopts a more closed conformation (changing from a

"C" to an "O" shape), caused by the movement of the immature 5′ domain toward the central domain. As a consequence, binding of Lcp5 and H9 to Utp20 is disrupted, allowing h21 (ES6d) to enter this previously occupied site (fig. S7, E and F). Similar to Utp20, the prominent Kre33 module also moves with the 5′ domain during this transition; however, the Kre33 module finally dissociates when the intermediate reaches the Post-A$_1$ state (Fig. 3, C and D, and fig. S7, G to I). This brings the 5′ domain even closer to the central domain (fig. S7H), and when state Post-A$_1$ is reached, the 5′ and central domains eventually interact and adopt a mature-like conformation (Fig. 3D and fig. S7I).

This apparent domain compaction—together with the 5′-ETS RNA remodeling, which takes place during the first steps of 90*S* transition—results in a conformational state sufficiently dynamic to facilitate A$_1$ processing: The state Pre-A$_1$ shows the noncleaved A$_1$ site ~50 Å away from the catalytic center of the PIN domain endonuclease Utp24, which was suggested to carry out this cut (Fig. 3E) (*38*–*41*). The Post-A$_1$ state reveals the formation of a short RNA stem-loop (h1) at the 5′ end of the mature 18*S* rRNA. The initial formation of this helix may have triggered endonucleolytic cleavage, because it results in repositioning the A$_1$ site close to the catalytic center of the Utp24 endonuclease (Fig. 3F). Consistent with this interpretation, at the base of the newly formed h1 we can retrace the RNA sequence until the third nucleotide (uridine 3) of the cleaved and mature 5′ end of 18*S* rRNA. Such a sequence-independent but h1 stem-loop–dependent mechanism for

A_1 cleavage, requiring an arbitrary three-nucleotide spacer relative to the h1 base, has been predicted by genetic studies from the Tollervey lab (*42*). Unexpectedly, A_1 cleavage does not require unwinding of the Box A and Box A′ heteroduplexes formed by pre-18S rRNA and U3 snoRNA, the latter two being retained in the Post-A_1 state (Fig. 3F). Together, the data show that A_1 cleavage is initiated by remodeling and relocation of the pre-18S rRNA substrate rather than by movement of the endonuclease Utp24 toward the A_1 cleavage site.

A_1 cleavage coupled to Pno1 and 90S remodeling

A series of additional remodeling events can be observed when comparing the Pre-A_1 and Post-A_1 states, which includes the release of the Krr1-Faf1 complex and its replacement by the relocated Pno1 and h45 (Fig. 3, G and H, and fig. S8). With the translocation of Pno1 and the associated h45, the RNA platform region adopts a mature-like conformation, typically seen in both yeast and human later cytoplasmic pre-40S intermediates (*26, 28, 29*). This results in the formation of rRNA helices h19 and h25 (Fig. 3H) and, as a consequence,

a new interface is formed and stabilized by the newly recruited methyltransferase Dim1 (Fig. 3I, fig. S9, and supplementary text), thereby rendering these rearrangements irreversible.

The binding mode of Pno1 to h45 differs between the Pre-A_1 and Post-A_1 states (Fig. 3J and fig. S8B). Relocation of h45 relative to Pno1 is stabilized by the long C-terminal α helix of Utp7, which previously was bound to H7 of the 5′-ETS before its dismantling (Fig. 3J and fig. S8C). The interaction of Utp14 with Pno1 prevents premature recruitment of the endonuclease Nob1, which later catalyzes the last processing step of the pre-18S rRNA at site D (Fig. 3, K and L) (*28*). Finally, by relocating to the central platform, Pno1 and its clamped h45 dissociate from Utp21 and Utp13 of the UTP-B module (fig. S8B), allowing for the first time the recruitment of the RNA helicase Dhr1 to the 90S via a two-site contact to Utp21 and Utp13 (see below and Fig. 3M).

Successive factor shedding during 90S transition to primordial pre-40S

Following the formation of state Post-A_1, the dismantled 5′-ETS allows for a cascade of structural and compositional changes that re-

sult in a stepwise reduction of the 90S complex (Fig. 4). First, the Sof1 module (Utp7, Utp14, and Sof1) is released together with Utp6 from its binding site close to the former helix H9 at the 3′ region of the 5′-ETS, resulting in the state Dis-A. Then, in the second step, the UTP-A complex dissociates together with Utp18 and all protein components of the U3 snoRNP, leading to state Dis-B. In the last transition step from Dis-B to Dis-C, the core of the UTP-B complex and the remains of UTP-C are released. Also, the 5′-ETS, which served as RNA scaffold for all binding modules (UTP-A, UTP-B, UTP-C, and Imp3), is no longer detectable in this final intermediate. Of the U3 snoRNP, only a short segment of its RNA remains localized in states Dis-B and Dis-C, in particular the heteroduplexes with 18S rRNA; the rest becomes detached from these intermediates (Figs. 1B and 4). Interestingly, the domains of the pre-40S undergo rather limited maturation rearrangements during this shedding phase. Only rRNA expansion segments ES3a, ES3b, and ES6a to ES6d on the solvent side of the subunit adopt mature conformations, whereas the rest of the domains remain essentially unchanged (see boxes in Fig. 4 and fig. S10A).
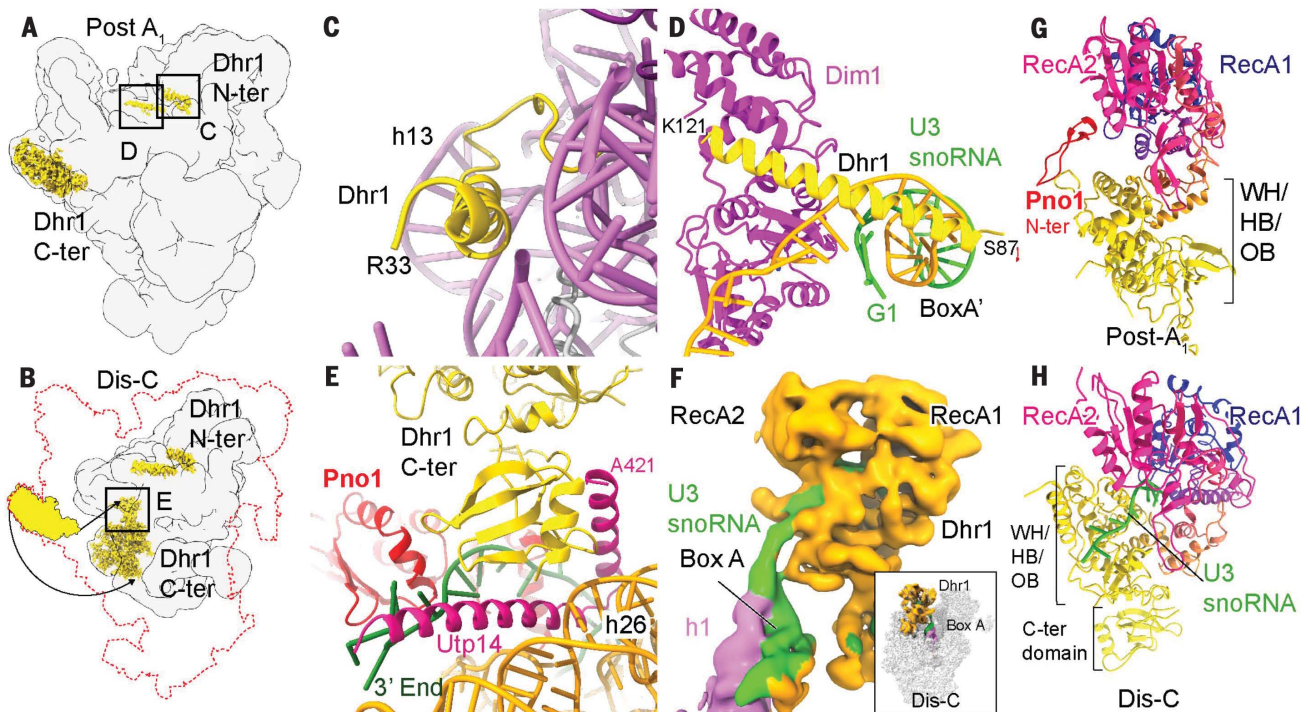
Fig. 5. Differential interaction of Dhr1 during 90S transition. (**A** and **B**) Locations of N and C termini (N-ter and C-ter) of Dhr1 in states Post-A_1 (A) and Dis-C (B). (**C** and **D**) Ribbon representation of the N-terminal helices of Dhr1. The first helix (yellow) binds to the h13 region of the 5′ domain (C), and the second one interacts with Box A′ duplex and Dim1 in all states after Pre-A_1 (D). (**E**) Model of the C-terminal domain of Dhr1 binding to the interface formed by Utp14, Pno1, h26, and the 3′ end of the 18S rRNA in state Dis-C. (**F**) Density revealing that near the Box A duplex, U3 snoRNA directly connects to Dhr1. Dhr1 density is shown in orange; U3 snoRNA density is shown in green. (**G** and **H**) Model of Dhr1 in states Post-A_1 (G) and Dis-C (H), showing two different conformations. Dhr1 is colored from blue to red and yellow. Pno1 and the U3 snoRNA are highlighted in red and green, respectively.
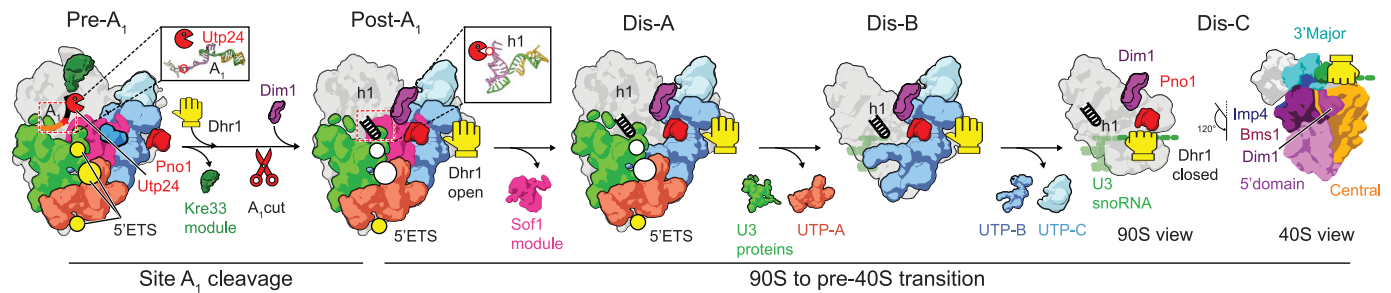
Fig. 6. Scheme of the 90S to pre-40S transition upon A1 cleavage. Cartoons depict the shedding-like transition from a 90S to a primordial pre-40S ribosome. Assembly factor modules and selected proteins are colored and labeled accordingly. The helicase Dhr1 is shown as a grabbing hand, representing the open and closed conformations.

Although this successive shedding of large protein modules is the distinct hallmark of the 90S transition, several structural changes occur in addition. First, Dhr1 moves from its previous UTP-B binding site (Fig. 3M) toward Box A in state Dis-C, thereby bringing the helicase into an ideal position to dismantle the last remaining U3::18S rRNA hybrid on the intermediate (fig. S10, B and C). Concomitantly, the Utp24 endonuclease becomes detectable again in this state, where it occupies the binding site for the uS5 ribosomal protein. Moreover, Utp14, previously bound to Utp6 and helix H9 of the 5′ETS, relocates to a new position in state Dis-C, where it binds Pno1 and Dhr1 via four α helices (fig. S10, D to F). Notably, two arginine residues of Utp14 inhibit the endonuclease activity of Utp24 by coordinating active-site residues, thus allowing Utp24 to serve as an endonuclease-inactive placeholder (fig. S10, G to I).

Sequential Dhr1 helicase function during the 90S to pre-40S transition

Finally, the Dis-C intermediate sheds light on the diverse functioning of the Dhr1 helicase during the 90S transition by showing how it is primed to remove the U3 snoRNA from its last contact point to the 18S rRNA, Box A. From state Post-A1 to state Dis-C, Dhr1 is tethered to the assembly intermediates via two invariant N-terminal helices: one that interacts with rRNA helix h13 at a site that is later occupied by Tsr1 (Fig. 5, A to C, and fig. S11, A and B) (*26, 28, 29*), and another that binds to the methyltransferase Dim1 (Fig. 5D and fig. S9B). The catalytic C-terminal domain, however, binds distantly from its N-terminal anchor region (Fig. 5, A and B, and fig. S11A). From state Post-A1 to state Dis-B, this globular part is bound to Utp13 and Utp21, which is mediated by a β barrel–like domain of Dhr1 (fig. S11C). In state Dis-C, however, after dissociation of Utp21, Dhr1 relocates to the Box A and h1 region, close to its U3 snoRNA substrate. Here, its β barrel–like domain is bound to the interface formed by Utp14, Pno1, and the 3′ region of the pre-18S rRNA (Fig. 5, E and F). The observed inter-

play between Utp14 and Dhr1 highlights the importance of Utp14 for the recruitment and relocation of Dhr1 during the 90S transition (*36, 43–45*).

In our ensemble of 90S transition structures, Dhr1 was observed in two distinct conformations. First, between states Post-A1 and Dis-B, Dhr1 was in a conformation with an open RNA binding tunnel. Here, the N terminus of Pno1 prevents substrate binding and closing of the catalytic domain (Fig. 5G). This suggests that Dhr1 is bound to the 90S pre-ribosome in an adenosine diphosphate–bound, open conformation after A1 cleavage (fig. S11D). After the transition to state Dis-C and relocation of Dhr1 toward the Box A duplex, however, the helicase domain engages a segment of the U3 snoRNA as substrate (Fig. 5F). Furthermore, Dhr1 appears now in a closed conformation, resembling the adenosine triphosphate (ATP)–free form of previously reported RNA-bound DHX37, the human homolog of Dhr1 (Fig. 5H and fig. S11E) (*18*). We conclude from these data that after the transition from state Dis-B to state Dis-C, Dhr1 is in a closed apo-state, awaiting ATP binding, and through successive cycles of ATP hydrolysis could exert its pulling force to completely expel the U3 snoRNA from this primordial pre-40S intermediate.

Discussion

Our work shows how the huge 90S pre-ribosome transforms into its next major biogenesis intermediate, the primordial pre-40S, by large structural rearrangements, including the successive shedding of assembly factor modules. Several findings support the notion that the observed transition states Dis-A to Dis-C are of physiological relevance and are not merely the products of random disintegration. First, successive maturation of rRNA expansion segments occurs between the states: In the post-A1 state, the expansion segments ES3a, ES3b, ES6a, ES6b, ES6c, and ES6d are still in an immature conformation (even with immature base pairing in ES6). In the Dis-A state, ES3b and ES6 adopt their mature conformation, even including the kissing loop formed between ES3b and ES6a.

When further progressing to Dis-B and Dis-C, ES3a was also observed to adopt the mature conformation. Second, from Dis-A to Dis-B, the 90S pre-ribosome incorporated ribosomal protein in its mature position. Third, in all the states until Dis-B, Dhr1 was present in an inactive conformation. Only after reaching the Dis-C state and UTP-B dissociation was Dhr1 captured in an active state, interacting with the U3 snoRNA. Fourth, Utp14 had already engaged in the pre-A1 stage but was also present in the Dis-C state, yet adopted different and highly defined positions. Taken together, all these findings are difficult to reconcile with the notion of randomly disintegrating particles.

Our structures may also provide insight into the coordination of pre-RNA processing at sites A1 and A2, which is assumed to be performed by the same Utp24 endonuclease (*38–40*). After site A1 cleavage, Utp24 is dislodged in state Dis-B while still tethered to the pre-ribosome. During this phase, Utp24 could access the second cleavage site A2 in the flexible ITS1 to finally generate the 20S pre-rRNA, the predominant species in Dis-C.

Our findings do not support the current view of a sudden release of a 5′-ETS particle and its associated principal modules UTP-A, UTP-B, and U3 snoRNP during the 90S to pre-40S transition. It was previously shown that 5′-ETS particles can, in principle, exist on their own—for example, by inhibiting the exosome in combination with a Utp18 mutation (*8*) or during arrest of 90S assembly using 3′-truncated pre-rRNA (*13, 14, 37*). Here we show, however, that contrary to the implications of earlier research, the 5′-ETS and its associated principal modules are not released simultaneously upon A1 cleavage, but rather in a sequential and stepwise shedding process. This mechanism for the 90S transition appears to be the more plausible, because it is difficult to imagine how segregation of an entire 5′-ETS particle can be induced by a single A1 cut, in light of the highly interconnected nature of the 90S structure.

The initial trigger for the 90S to pre-40S transition is still unknown. We suggest that the decisive step in dismantling the 90S

intermediate depends on an advanced state of the pre-40S domain maturation as reflected by the degree of its compaction. Sufficient compaction may trigger the gradual deconstruction of the 5′-ETS scaffold, as previously observed in several states preceding A$_1$ cleavage (15). This assumption is in agreement with the earlier observation that A$_1$ cleavage is dependent on the presence of the helicase Mtr4 as a potential 5′-ETS remodeler (12). Successive dislocation of the RNA helices starting at the 3′ region of the 5′-ETS then could allow for the relocation of Pno1 and h45 and, concomitantly, the recruitment and engagement of the helicase Dhr1. Dhr1 in turn may facilitate the partial formation of rRNA helix h1, which results in the cleavage at the A$_1$ site. This coincides with the dissociation of several factors and further destabilization of the 90S intermediate by 5′-ETS dismantling, eventually causing the sequential release of modules until the primordial pre-40S has emerged (Fig. 6).

Defects in ribosome biogenesis can have drastic consequences for human health as ribosomopathies (46). Therefore, gaining more mechanistic insights into this elaborate maturation process and its integration into cellular signaling pathways is desirable.

REFERENCES AND NOTES

1. F. Dragon et al., Nature 417, 967–970 (2002).
2. P. Grandi et al., Mol. Cell 10, 105–115 (2002).
3. J. L. Woolford Jr., S. J. Baserga, Genetics 195, 643–681 (2013).
4. J. Baßler, E. Hurt, Annu. Rev. Biochem. 88, 281–306 (2019).
5. S. Klinge, J. L. Woolford Jr., Nat. Rev. Mol. Cell Biol. 20, 116–131 (2019).
6. J. de la Cruz, D. Kressler, D. Tollervey, P. Linder, EMBO J. 17, 1128–1140 (1998).
7. M. Thoms et al., Cell 162, 1029–1038 (2015).
8. M. Kornprobst et al., Cell 166, 380–393 (2016).
9. J. Barandun et al., Nat. Struct. Mol. Biol. 24, 944–953 (2017).
10. M. Chaker-Margot, J. Barandun, M. Hunziker, S. Klinge, Science 355, eaal1880 (2017).
11. J. Cheng, N. Kellner, O. Berninghausen, E. Hurt, R. Beckmann, Nat. Struct. Mol. Biol. 24, 954–964 (2017).
12. Q. Sun et al., eLife 6, e22086 (2017).
13. M. Chaker-Margot, M. Hunziker, J. Barandun, B. D. Dill, S. Klinge, Nat. Struct. Mol. Biol. 22, 920–923 (2015).
14. L. Zhang, C. Wu, G. Cai, S. Chen, K. Ye, Genes Dev. 30, 718–732 (2016).
15. J. Cheng et al., Mol. Cell 75, 1256–1269.e7 (2019).
16. T. Wegierski, E. Billy, F. Nasr, W. Filipowicz, RNA 7, 1254–1267 (2001).
17. S. Granneman et al., Nucleic Acids Res. 31, 1877–1887 (2003).
18. N. J. Krogan et al., Mol. Cell 13, 225–239 (2004).
19. J. Pérez-Fernández, A. Román, J. De Las Rivas, X. R. Bustelo, M. Dosil, Mol. Cell. Biol. 27, 5414–5429 (2007).
20. J. Pérez-Fernández, P. Martín-Marcos, M. Dosil, Nucleic Acids Res. 39, 8105–8121 (2011).
21. T. Schäfer, D. Strauss, E. Petfalski, D. Tollervey, E. Hurt, EMBO J. 22, 1370–1380 (2003).
22. T. Schäfer et al., Nature 441, 651–655 (2006).
23. B. S. Strunk et al., Science 333, 1449–1453 (2011).
24. E. Wyler et al., RNA 17, 189–200 (2011).
25. N. Larburu et al., Nucleic Acids Res. 44, 8465–8478 (2016).
26. A. Heuer et al., eLife 6, e30189 (2017).
27. M. C. Johnson, H. Ghalei, K. A. Doxtader, K. Karbstein, M. E. Stroupe, Structure 25, 329–340 (2017).
28. M. Ameismeier, J. Cheng, O. Berninghausen, R. Beckmann, Nature 558, 249–253 (2018).
29. A. Scaiola et al., EMBO J. 37, e98499 (2018).
30. V. Mitterer et al., Nat. Commun. 10, 2754 (2019).
31. D. Lafontaine, J. Vandenhaute, D. Tollervey, Genes Dev. 9, 2470–2481 (1995).
32. M. Sturm, J. Cheng, J. Baßler, R. Beckmann, E. Hurt, Nat. Commun. 8, 2213 (2017).
33. A. Colley, J. D. Beggs, D. Tollervey, D. L. Lafontaine, Mol. Cell. Biol. 20, 7238–7246 (2000).
34. R. Sardana, J. Zhu, M. Gill, A. W. Johnson, Mol. Cell. Biol. 34, 2208–2220 (2014).
35. R. Sardana et al., PLOS Biol. 13, e1002083 (2015).
36. J. Zhu, X. Liu, M. Anjos, C. C. Correll, A. W. Johnson, Mol. Cell. Biol. 36, 965–978 (2016).
37. M. Hunziker et al., eLife 8, e45185 (2019).
38. F. Bleichert, S. Granneman, Y. N. Osheim, A. L. Beyer, S. J. Baserga, Proc. Natl. Acad. Sci. U.S.A. 103, 9464–9469 (2006).
39. R. Tomecki, A. Labno, K. Drazkowska, D. Cysewski, A. Dziembowski, RNA Biol. 12, 1010–1029 (2015).
40. G. R. Wells et al., Nucleic Acids Res. 44, 5399–5409 (2016).
41. W. An, Y. Du, K. Ye, PLOS ONE 13, e0195723 (2018).
42. K. Sharma, D. Tollervey, Mol. Cell. Biol. 19, 6012–6019 (1999).
43. J. J. Black, Z. Wang, L. M. Goering, A. W. Johnson, RNA 24, 1214–1228 (2018).
44. F. M. Boneberg et al., RNA 25, 685–701 (2019).
45. A. Roychowdhury et al., Nucleic Acids Res. 47, 7548–7563 (2019).
46. A. Narla, B. L. Ebert, Blood 115, 3196–3205 (2010).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6510/1470/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Table S1
References (47–66)
Data S1
Movie S1
MDAR Reproducibility Checklist

View/request a protocol for this paper from Bio-protocol.

Science

90*S* pre-ribosome transformation into the primordial 40*S* subunit

Jingdong Cheng, Benjamin Lau, Giuseppe La Venuta, Michael Ameismeier, Otto Berninghausen, Ed Hurt and Roland Beckmann

How ribosomes are made

The formation of eukaryotic ribosomes is a complex process that starts with transcription of a large precursor RNA that assembles into a large 90S preribosome, which matures to finally give the 40S small subunit of the ribosome. Cheng *et al.* and Du *et al.* give insight into this process, using cryo-electron microscopy to look at intermediates along the pathway. Together, these studies reveal how a cast of molecular players act to coordinate the compositional and structural changes that transform the 90S preribosome into a pre-40S subunit.
Science, this issue p. 1470, p. 1477

ARTICLE TOOLS	http://science.sciencemag.org/content/369/6510/1470
SUPPLEMENTARY MATERIALS	http://science.sciencemag.org/content/suppl/2020/09/16/369.6510.1470.DC1
REFERENCES	This article cites 66 articles, 19 of which you can access for free http://science.sciencemag.org/content/369/6510/1470#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service

Supplementary Materials for

90*S* pre-ribosome transformation into the primordial 40*S* subunit

Jingdong Cheng, Benjamin Lau, Giuseppe La Venuta, Michael Ameismeier, Otto Berninghausen, Ed Hurt*, Roland Beckmann*

*Corresponding author. Email: beckmann@genzentrum.lmu.de (R.B.); ed.hurt@bzh.uni-heidelberg.de (E.H.)

This PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S11
Table S1
References
Caption for data S1
Caption for movie S1

Other supplementary material for this manuscript includes:

Data S1 (Excel format)
Movie S1
MDAR Reproducibility Checklist (PDF)

Materials and Methods

Yeast strains and genetic methods

The *S. cerevisiae* strains w303 (genotype: *ade2-1, trp1-1, leu2-3,112, his3-11,15, ura3-1, can1-100*), Noc4-TAP/Dhr1-Flag (genotype: *NOC4-TAP::HIS3, DHR1-Flag::natNT2, W303*) and Dhr1-TAP/Dim1-Flag (genotype: *DHR1-TAP::HIS3, DIM1-Flag::TRP1, W303*) were used in this study.

Split-tag tandem affinity-purification from yeast cells

Yeast strains expressing C-terminus-tagged bait proteins, used for split-tag tandem affinity purification, were harvested during the logarithmic growth phase. Cells were mechanically disrupted in a cryogenic cell mill (Retsch MM400) and lysed in buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1.5 mM $MgCl_2$, 5% glycerol, 0.1% NP-40, and 1 mM DTT. The lysate was cleared by centrifugation followed by transferring the supernatant onto immunoglobulin G Sepharose 6 Fast Flow resin (GE Healthcare) for 2 h at 4 °C. Bound proteins were washed and TEV cleavage was performed for 45 min at 16 °C. In a second affinity purification step, the eluate was loaded onto Flag-agarose beads (Anti-Flag M2 Affinity Gel, Sigma–Aldrich), which were incubated for 1 h at 4 °C. Beads were washed and eluted with buffer containing Flag peptide. For the cryo-EM analysis, the elution buffer used contained 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5 mM $MgCl_2$, 2% glycerol, 0.01% NP-40, and 1 mM DTT. Eluates were analyzed by SDS-PAGE and staining with colloidal Coomassie (Roti-Blue, Roth), or further fractionated by sucrose gradient centrifugation, and analyzed by SDS-PAGE.

Sucrose gradient centrifugation

Eluates from the split-tag tandem affinity purifications were transferred to a linear 15%–40% (w/v) sucrose gradient containing a buffer of 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1.5

mM MgCl$_2$, 0.003% NP-40, and 1 mM DTT, and centrifuged for 16 h at 27,000 rpm at 4 °C. The sucrose gradient was fractionated and fractions were split for RNA analysis (see below) or precipitated with TCA (10%) for proteins, which were resuspended in sample buffer and analyzed by 4%–12% gradient PAGE (NUPAGE, Invitrogen) followed by staining with colloidal Coomassie Blue (Roti-Blue, Roth).

Mass spectrometry

The sucrose gradient fractions #8 (pre-40S pool) and #12 (90S pool) from the Noc4–Dhr1 split-tag sample were analyzed by semi-quantitative mass spectrometry at FingerPrints proteomics (University of Dundee, UK). Co-precipitating proteins were identified by 1D nLC–ESI-MS/MS using MaxQuant software (47). IBAQ values of label-free quantification are shown in Data Table S1.

RNA extraction and Northern analysis

RNA present in the sucrose gradient fractions (Noc4–Dhr1 split-tag sample) was extracted and precipitated with ethanol as previously described (48). 5'-ETS rRNA and U3 snoRNA were separated on a 6% polyacrylamide gel containing 8 M urea after denaturation with formaldehyde. Larger pre-rRNA species (23S, 21S, 20S) were resolved on 1.2% agarose gel after denaturation with glyoxal (8). After transferring to positively charged nylon membranes and UV crosslinking, the following 5'-^{32}P-labeled oligonucleotide probes were used for Northern analysis: 5'-CGGTTTTAATTGTCCTA-3' (OMK002; probe "a" for 23S, 21S, 20S) (8), 5'-GCAAAGATATGAAAACTCCAC-3' (OMK800; probe "b" for 23S, 21S) (49), 5'-GTCTTCAACTGCTTTCGCA-3' (OMK1455; probe "c" right at the beginning of 5'-ETS) (50), 5'-TATTCCCTCTTGCTAGAAG-3' (OMK1036; probe "d" right downstream A$_0$ for 5'-ETS-A$_1$), 5'-TACTTATTGAGTTTGGAAACAG-3' (probe "e" right upstream A$_0$ for 5'-ETS-A$_1$ and 5'-ETS-

A0), 5'-GGTTATGGGACTCATCA-3' (probe for U3) (8).

To produce different 5'-ETS fragments used as marker for the Northern analysis (Fig. S1D), five 5'-ETS rDNA sequences, all starting at the authentic 5' end of the 35S pre-rRNA and varying in length at the 3' end, were amplified by PCR from yeast genomic DNA and inserted 3' downstream of the T7 promoter in pBluescript II KS (–) plasmid. The corresponding 5'-ETS RNA fragments were in vitro transcribed, following the manufacturer's instructions (HiScribe® T7 High Yield RNA Synthesis Kit, New England Biolabs), extracted and ethanol precipitated, yielding 5'-ETS-A1 (700 nucleotides), 5'-ETS-A0 (609 nucleotides), and 5'-ETS-shorter-than-A1 fragments (500, 400 and 300 nucleotides). These five 5'-ETS RNA fragments carry additional 17 nucleotides (16 at the 5' and one at the 3'-end), which were transcribed from the pBluescript II KS (–) backbone. These fragments together with Noc4-Dhr1 extracted RNA and a commercially available RNA ladder standard were analyzed by polyacrylamide denaturing gel electrophoresis and Northern blotting, using the specific 5'-ETS probes as indicated above and in the corresponding figure legends. RiboRuler High Range RNA Ladder (Thermo Scientific, #SM1821) was used as RNA standard.

Electron microscopy and image processing

Purified Dhr1–Dim1 or Noc4–Dhr1 samples (3.5 µl) were directly applied onto pre-coated (2 nm) R3/3 holey-carbon-supported copper grids (Quantifoil), blotted for 2–3 s at 4 °C and plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV. Cryo-EM data was acquired on an FEI Titan Krios transmission electron microscope at 300 kV under low-dose conditions (10 frames at about 2.5 e− Å^{-2}). Both samples were collected twice: once on a Falcon II direct electron detector and once on a K2 Summit direct electron detector, at nominal defocus values ranging from −1.0 to −2.5 µm and from -0.8 to -2.5 µm, respectively. Falcon II and K2 data were collected with

a nominal pixel size of 1.084 and 1.059 Å per pixel on the object scale using the software EM-TOOLS (TVIPS) and EPU (Thermo Fisher), respectively. Original image stacks were dose weighted, aligned, summed and drift-corrected using MotionCor2 (*50*). Contrast-transfer function (CTF) parameters and resolutions were estimated for each micrograph using CTFFIND4 and GCTF, respectively (*51, 52*). Micrographs with an estimated resolution better than 5 Å and an astigmatism below 5% were manually screened for contamination or carbon rupture.

For the Dhr1–Dim1 sample, a total of 11,884 micrographs were selected from the Falcon II dataset and submitted to automated particle picking using Gautomatch with the 90S map as reference (11), resulting in 1,141,220 picked particles. A total 9,984 micrographs were selected from the K2 dataset and 707,080 particles were initially picked. For the Noc4–Dhr1 sample, a total of 10,508 micrographs were selected from the Falcon II dataset, resulting in 1,242,390 picked particles. A total of 9,224 micrographs were selected from K2 dataset and 753,312 particles were initially picked.

Reference-free 2D classification was used to sort out all the contaminated particles, following by 3D refinement and 3D classification using Relion V3.0 (*53*). For the dataset 1 (Noc4-Dhr1) initially, we used our published low-resolution filtered C.t. 90S as a reference(*11*) for particle picking and also for the initial 3D refinement, since we did not expect to find pre-40S like particles. This resulted in an initial average reconstruction essentially representing a 90S particle, probably at least in part caused by reference bias. After noticing that there were many 40S-like particles in this class, we took a low resolution 5' domain as a reference for a second step of classification. This resulted in the generation of clean 40S-like Dis-B and Dis-C state classes. After reaching conservative classes, beam tilt and per-particle CTF correction were done to get the final reconstruction. Particles belonging to state Post-A1 from all four datasets were combined and

downscaled to pixel size 1.084 Å before 3D refinement. Particles representing states Dis-B and Dis-C from both K2 datasets were combined to obtain the final reconstructions. The other datasets were classified as described in the scheme (fig. S2). In order to facilitate model building, multibody refinement was carried out for states B2, Pre-A_1, Post-A_1 and Dis-C. In states B2, Pre-A_1 and Post-A_1, the 90S was divided into four bodies: the head region which includes the 5' domain of the 18S pre-rRNA, Utp20, C terminus of Utp10 and Kre33 module if present; the body region which includes UTP-B, U3 snoRNP, Bms1-Rcl1, Mpp10 complex and one half of the 3' domain of the 18S pre-rRNA; the platform region which includes UTP-C, Krr1-Faf1/Pno1, Utp13, central domain of the 18S pre-rRNA and Dhr1 if present; the base region which includs UTP-A and part of the Noc4 module. For state Dis-C, pre-40S was divided into three bodies: the body and shoulder region which includes 5' domain and central domain of the 18S pre-rRNA, Box A/A' duplexes, Dim1 and Bms1-Rcl1 complex; the 3' domain region which includes 3' domain of the 18S pre-rRNA and the Noc4 module; the Dhr1 which only includes the C terminus helicase domain of Dhr1. All the masks were automatically generated using Relion. All the final maps were post-processed and local resolution filtered using Relion.

Model building and refinement

In general, the *C. thermophilum* 90S pre-ribosome structure (PDB ID: 6RXU or 6RXV) was used as an initial reference to generate homology models (*15*), followed by manual adjustment in Coot (*54*). The *S. cerevisiae* 90S pre-ribosome structure (PDB ID: 5WLC) was also rigid-body fitted in Chimera and Coot (*9, 54, 55*). Because certain sub-regions were not well resolved, mixed full models (poly-alanine based and sidechain based) were nevertheless prepared for states B2, Pre-A_1, Post-A_1 and Dis-C. In general, whenever a protein or parts of it are represented as poly-alanine the resolution of the corresponding map did not allow for building a *bona fide* molecular

model. Since the states B2, Pre-A$_1$ and Post-A$_1$ share most of the assembly factors, whenever appropriate all three maps were considered together to generate the final models. Due to flexibility and resulting low resolution of the corresponding region, Enp1 in all the models was represented as poly-alanine. In addition, rigid-body-fitted poly-alanine models were also generated for states A, Dis-A and Dis-B due to the lower resolution.

For state B2 and Pre-A$_1$, the homology models were generated on the SWISS-MODEL server using state B2 of the *C. thermophilum* 90S pre-ribosome (PDB ID 6RXV) as a reference (*56*). Together with the *S. cerevisiae* 90S pre-ribosome structure (PDB ID: 5WLC)(*9*), after rigid-body fitting the models, manual adjustment was carried out in Coot for the high-resolution region. The missing parts of the biogenesis factors were *de novo* built based on the available density (from overall and multibody refinement). The expansion segment 6, h21 (Pre-A$_1$) and most of h44 were adjusted from PDB ID 5NDG (*57*). The region from nucleotides 137 to 191 of the U3 snoRNA was generated using the RNA composer server (*58*); the same was done for U3 snoRNA in the state Pre-A$_1$. Due to the flexibility of Utp20, of the C-terminus of Utp10 (1348-1769) and of Kre33 (chain JB) in state B2, poly-alanine models of these factors were used to rigid body fit into the density. In state Pre-A$_1$ map, only the C-terminus of Utp20 (2010-2421) and Kre33 (chain JB) were not well resolved, thus poly-alanine models were used for these two regions. After rigid body fitting, the C-terminus of Utp20 was adjusted to fit the open conformation. Similarly, the middle part of Utp10 (488-806, in state B2) and the WD40 domain of Utp8 were presented as poly-alanine. The overall structure of the state B2 may be refined further in the future, but already reveals the key positions of the proteins required to gain mechanistic insights.

For state Post-A$_1$, the preliminary model was generated as state B2. In addition, Dhr1 was homology modeled and rigid-body docked, based on the crystal structures of yeast and mammalian

Dhr1 (PDB IDs: 6H57, 6O16) and Prp43 (PDB ID: 5D0U)(*44, 45, 59*). After docking, manual adjustment was carried out if necessary, whereas the N terminus was de novo built. Dim1 was first homology modeled based on human DIM1L (PDB ID: 1ZQ9), and then manually adjusted in Coot according to the density map. The newly identified regions of Utp14, Mpp10, Utp20, Utp21, and Bms1 were *de novo* built. h1, h19, h24, h25, h41, and h42 regions of the 18S rRNA were adopted from 5NDG(*57*). RNA regions and helices of the 5'-ETS and U3 snoRNA that were not visible in the map were simply removed from the model, as was the case for the Kre33 module.

For states A, Dis-A, and Dis-B, models from states B2 and Post-A1 were used for rigid body fitting into the densities. The missing factors or regions when compared with states B2 and Post-A$_1$ were simply removed from the final model. In Dis-A and Dis-B, the expansion 3 and expansion 6 regions of the 18S rRNA were adopted from 5NDG(*57*).

For state Dis-C, missing factors and regions were further removed from Dis-B. Assembly factors, Bms1, Rcl1, Pno1, Utp24, Utp14, Mpp10 (295-351), N-terminus (2-121) and C-terminus (1175-1263) of Dhr1 and Dim1 were built based on high resolution density. In this state, the helicase domain of Dhr1 was homology modeled and rigid-body docked based on both the yeast (PDB ID: 6H57)(*45*) and human (PDB ID: 6O16)(*44*) crystal structures. However, due to lower local resolution, only a poly-alanine model was provided for the helicase region (376-1174). The re-bound Utp24 was taken from state B2, but the Utp14 was *de novo* built into density in Coot. Due to lower local resolution, a poly-U RNA chain was used as a place holder for the U3 snoRNA region, which binds within the helicase domain of Dhr1. Due to the high flexibility, two halves of the 3' domain (including uS7, uS9, eS19, eS25 and eS28) together with the associated Noc4 module (Noc4, Nop14/Utp2, Enp1 and Emg1 dimer), as well as Imp4 and Mpp10 (352-386), displayed lower resolution, therefore also here only a poly-alanine model was provided for this

region.

The final models for states B2, Pre-A1, Post-A1, and Dis-C were real-space refined with secondary structure restraints using the PHENIX suite (*60*). Final model evaluation was performed with MolProbity (*61*). Maps and models were visualized and figures created with ChimeraX (*62*).

<u>Ordering of the structural states</u>

The ordering was done on the following basis: (i) We start with the transition from state A to state B2. Essentially the same states have been observed before from *C. thermophilum* 90S particles, which suggested to us that state A must precede state B2 because it has not yet acquired the Kre33 module. (ii) The Pre-A1 state must follow state B2 because rRNA helix 21 is matured in Pre-A1 whereas Lcp5 as well as helix 9 and helix 9' of the 5' ETS are dismantled. Moreover, the 5' and central domains of pre-18S moved closer together in Pre-A1. (iii) The transition from Pre-A1 to Post-A1 is clearest and discussed in detail in the main text, including why it is expected from functional analysis. (iv) State Post-A1 is followed by Dis-A because this new state is clearly in a Post-A1 conformation, yet, starting with Dis-A the Sof1 module and Utp6 have left the assembly. Moreover, as described above, in the Dis-A state the rRNA expansion segment 3b (ES3b) and ES6 region adopt the mature conformation, even including the kissing loop formed between ES3b and ES6a. (v) When further progressing to Dis-B (and Dis-C), also ES3a was observed to adopt the mature conformation. Most prominently, from Dis-A to Dis-B, the UTP-A complex and the U3 snoRNP as well as the H1-2, the 3' hinge and 5' hinge heteroduplexes of the 5'ETS disappeared. In addition, from Dis-A to Dis-B, the Dis-B intermediate gained the ribosomal protein eS25, which was bound in its mature position. (vi) Dis-C has to follow Dis-B because in this state the major modules UTP-B, UTP-C and Imp3 were dissociated. Moreover, it is well established that Dhr1 is the driving force behind dismantling the U3 snoRNA from the pre-

ribosome. However, in all the states before Dis-C, Dhr1 is kept in an inactive conformation with respect to U3 interaction by the UTP-B module. Only after reaching the Dis-C state and UTP-B dissociation, Dhr1 was captured interacting with the U3 snoRNA and primed to extract it from the pre-40S ribosome. Interesting in this case is Utp14, which already engages the Pre-A$_1$ 90S intermediate, but after the transition from Post-A$_1$ to the Dis-C state it adopts a different and highly defined position. Strikingly, Utp14 finally could exert a new function in its final position on the Dis-C particle when bound to the Dhr1 (see main text).

Taken all arguments together, these different 90S/pre-40S states can be unambiguously put together into the suggested order of transitions. This is based in each case on observations of a specific combination of progressing rRNA maturation, incorporation of ribosomal proteins, successive dismantling of the 5'-ETS and U3snoRNP, and sequential loss of 90S factors and modules.

Quantification and statistical analysis

The semiquantitative mass spectra were analyzed with the software MaxQuant (*47*), according to the user manual. The normalization of the mass spectrometry data is given in Data S1. Details of the cryo-EM analysis is described in the Method Details.

Supplementary Text

Dim1 methylates A1782 within the pre-18S rRNA around Pre/Post-A$_1$ states

At the Post-A$_1$ state, we observed association of the methyltransferase Dim1 with the 90S pre-ribosome, which remains bound during subsequent transition phases. Dim1 is a highly conserved KsgA-family methyltransferase that methylates two adenosines, A1781 and A1782, that are part of a loop in h45 (*31, 63, 64*). In the Post-A$_1$ state, Dim1 is stably bound next to h24, where it is sandwiched between the central domain (h23, h24, and h45) and the Box A' heteroduplex (fig.

9

S9, A and B). Notably, the incorporation of Dim1 into the 90S pre-ribosome can only occur around the A1 cleavage when the necessary docking sites (Pno1-h45, h24, Box A') develop.
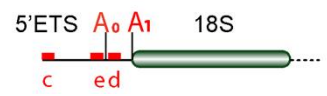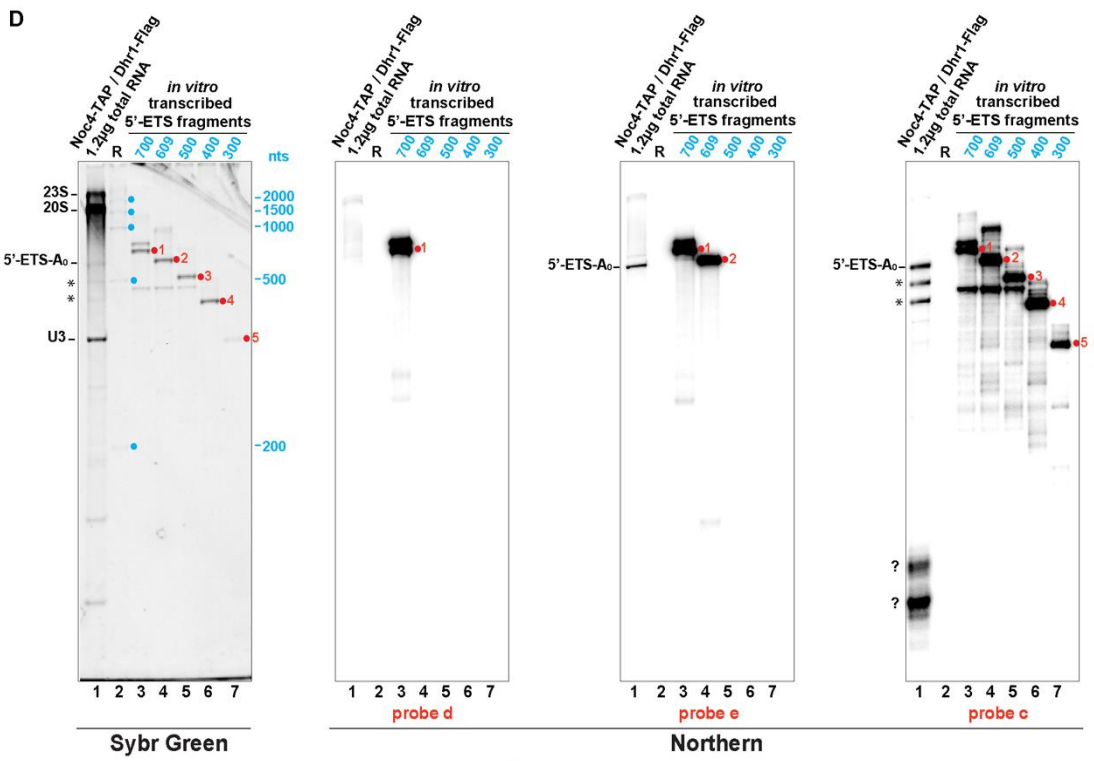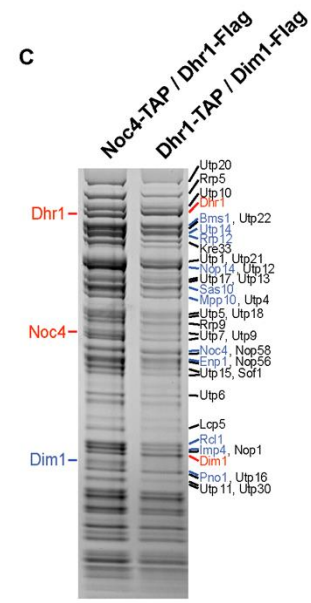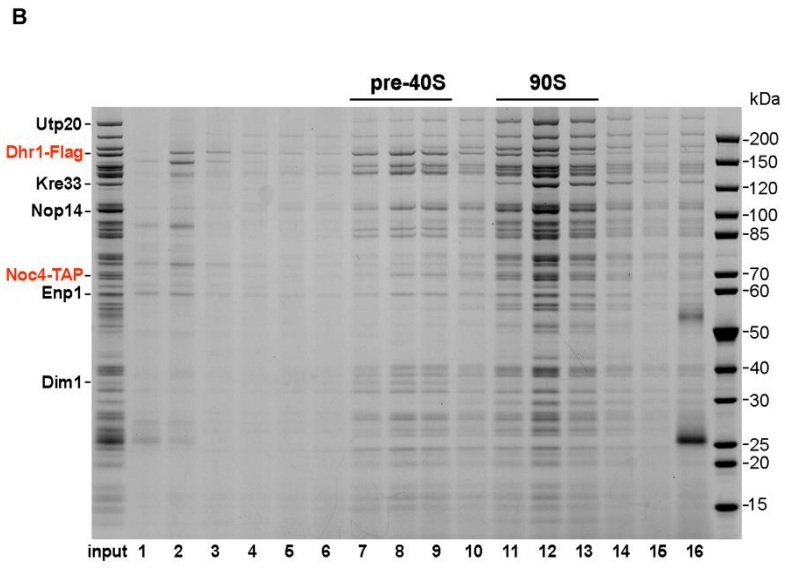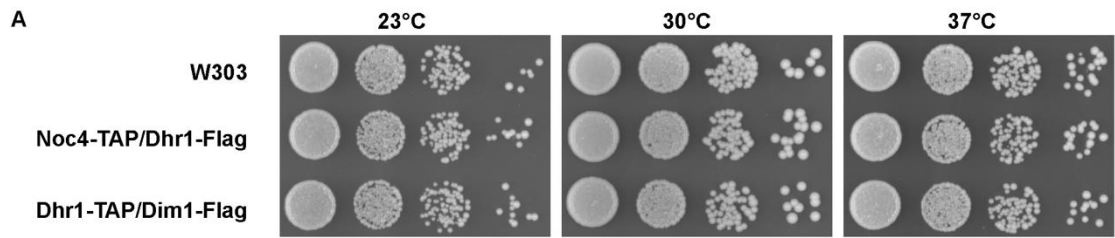
The cryo-EM structure of the Post-A$_1$ intermediate reveals the molecular details of the Dim1 interaction, such as how the N-terminal methyltransferase domain deeply penetrates the 90S particle, thereby reaching the critical loop of h45 with its A1782 substrate nucleotide, and the C-terminal helix bundle domain directly interacting with h24. At this stage, A1782 is sandwiched by residues H37 and Y131 of Dim1, which are close to N128, the catalytic residue that transfers the methyl group from its donor S-adenosyl methionine (SAM) (fig. S9C). Altogether, this data explains how A1782 can be specifically methylated upon A1 cleavage during the 90S to pre-40S transition.

Two different active modes of Utp24

The endonuclease Utp24 is dissociated from its rigid binding site on h1 in state Dis-B, however, it rebinds to the very same position in state Dis-C, where it occupies the binding site for the uS5 ribosomal protein, likely serving as a placeholder (fig. S10, G and H). Moreover, the enzymatic activity of Utp24 is inhibited by interactions between the catalytic PIN domain and specific residues of Utp14, which occupy Utp24's endonuclease site. Specifically, M440 of Utp14 inserts into a hydrophobic pocket and R439 directly reaches into the catalytic center of Utp24, substituting a Mg^{2+} (alternatively Mn^{2+}) ion, which is essential for the endonuclease activity (fig. S10I). However, it is conceivable that during a final quality check on pre-40S particles, Utp24's endonuclease activity is reactivated by the withdrawal of the arginine finger of Utp14, in order to perform an abnormal endonucleolytic cleavage. This was observed at an internal site Q1 (quality

control site 1) within the 18S rRNA, which is induced upon inhibition of 90S biogenesis and attracts the exosome-TRAMP system for subsequent 18S rRNA turnover (*65*).

Usage of an inhibitory arginine finger is reminiscent of how toxin–antitoxin systems regulate PIN domain endonucleases in bacteria (*66*) (fig. S10I). In the context of 90S maturation, it might serve to prevent unwanted cleavage of nearby 18S rRNA sequences.

A

23°C 30°C 37°C

W303

Noc4-TAP/Dhr1-Flag

Dhr1-TAP/Dim1-Flag

B

pre-40S 90S

Utp20
Dhr1-Flag
Kre33
Nop14

Noc4-TAP
Enp1

Dim1

kDa
–200
–150
–120
–100
–85
–70
–60
–40
–30
–25
–20
–15

input 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

C

Noc4-TAP / Dhr1-Flag
Dhr1-TAP / Dim1-Flag

Utp20
Rrp5
Utp10
Dhr1
Dhr1 — Bms1, Utp22
Utp14
Rrp12
Kre33
Utp1, Utp21
Nop14, Utp12
Utp17, Utp13
Sas10
Mpp10, Utp4
Utp5, Utp18
Rrp9
Utp7, Utp9
Noc4 — Noc4, Nop58
Enp1, Nop56
Utp15, Sof1
Utp6
Lcp5
Rcl1
Imp4, Nop1
Dim1 — Dim1
Pno1, Utp16
Utp 11, Utp30

D

Sybr Green

Northern

probe d probe e probe c

5'ETS A₀ A₁ 18S

c e d

Fig. S1. Analysis of yeast pre-ribosome intermediates at the 90S to pe-40S transition.

(**A**) Growth analysis of yeast *Saccharomyces cerevisiae* strains with chromosomally integrated Noc4-TAP/Dhr1-Flag and Dhr1-TAP/Dim1-Flag in comparison to the W303 wild-type strain. Strains were grown on YPD plates at the indicated temperatures for 2 days. (**B**) Analysis of split-tag (Noc4-TAP/Dhr1-Flag) affinity-purified assembly intermediates by 15%–40% sucrose gradient centrifugation (left panel). The final eluate (input) and fractions 1–16 from the sucrose gradient were analyzed by SDS-PAGE and Coomassie staining. Fractions containing the pre-40S and 90S particles are indicated. Fractions #8 and #12 are separately shown and the major bands are labeled in Fig. 1C. (**C**) Comparison of split-tag affinity-purified Noc4-TAP/Dhr1-Flag and Dhr1-TAP/Dim1-Flag preparations, analyzed by SDS-PAGE and Coomassie staining. The bands identified by mass spectrometry are labeled on the right. The bait proteins are shown in red, common ribosome assembly factors found in 90S and pre-40S particles are in blue. (D) RNA analysis of split-tag (Noc4-TAP/Dhr1-Flag) affinity-purified 90S/pre-40S particles with the focus on the 5'-ETS RNA. RNA present in the final Flag eluate (1.2 mg loaded on the gel, lane 1), a commercial RNA ladder (R, lane 2), and the *in vitro* transcribed 5'-ETS fragments of various length (please note that the five *in vitro* transcribed fragments carry additional 17 nucleotides derived from the plasmid-backbone) as indicated (lanes 3-7) were analyzed by denaturing polyacrylamide gel electrophoresis. The gel was first stained with the Sybr Green fluorescent dye (Sigma-Aldrich, #S9305) to visualize all the RNA present in the Noc4-TAP/Dhr1-Flag preparation (left panel; 5'-ETS-A_0 and U3 snoRNA were labeled, but the two prominent high molecular weight RNA bands indicated by lines were not assigned, but may include the 20S/21S pre-rRNA, which is also present in the preparation; see Fig. 1B), before RNA was blotted on a nylon membrane for Northern analysis. It was sequentially hybridized with the probes 'd', 'e' , and 'c',

which anneal in the 5'-ETS like indicated in the scheme below the Northern blots. Probe 'e' in principle detects both 5'-ETS-A_0 and 5'-ETS-A_1, while probe 'd' detects exclusively the 5'-ETS-A_1. Probe 'c' recognizes all the 5'-ETS forms with an intact 5'end, including two bands below the 200 nts marker (marked with '?'), which apparently could be 3'-shortened 5'-ETS-A_0 fragments.
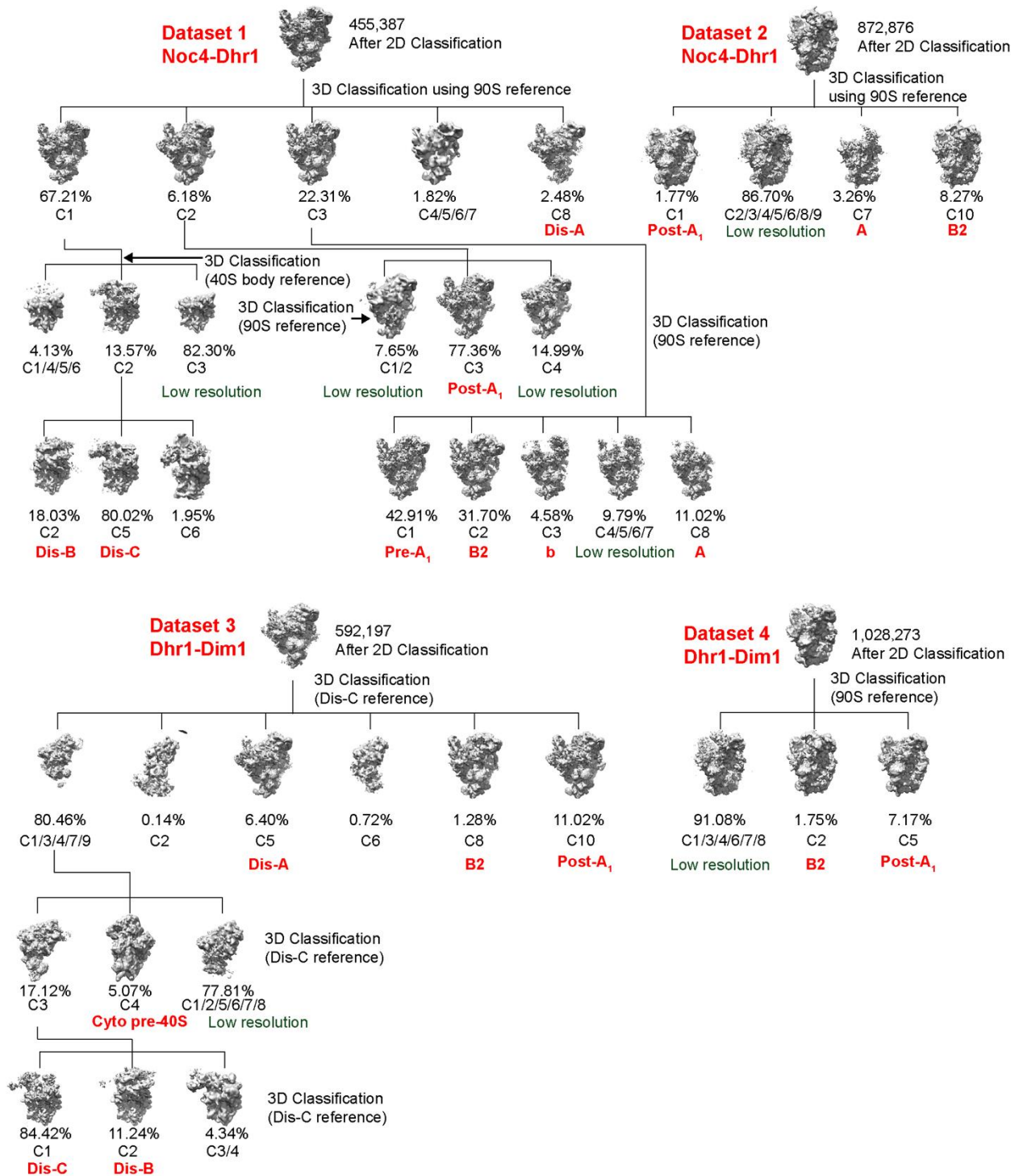
Fig. S2. Classification of the Noc4–Dhr1 and Dhr1–Dim1 split-Tag samples.

Classification scheme of the four datasets used in this study. Datasets 1 and 2 were collected using split-tag Noc4–Dhr1 pull-out with K2 and Falcon II detector, respectively. Datasets 3 and 4 were

collected using the split-tag Dhr1–Dim1 also with K2 and Falcon II detector, respectively. All

classes after 3D classification resembling one of the states in this study are labeled accordingly.

For the final reconstructions, all Post-A$_1$, Dis-B and Dis-C classes from the four datasets were

combined. Notably, a class resembling cytoplasmic pre-40S (state Cyto pre-40S) and a class

resembling state b of 90S (related to the *C. thermophilum* state b) have been described previously

and are not further discussed here. Classes that are not labeled could either not be refined further

to higher resolution or could not be classified further despite the presence of large highly flexible

domains.
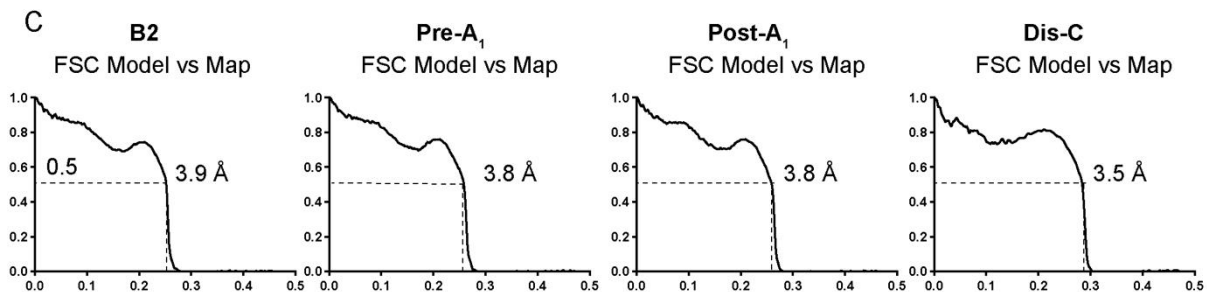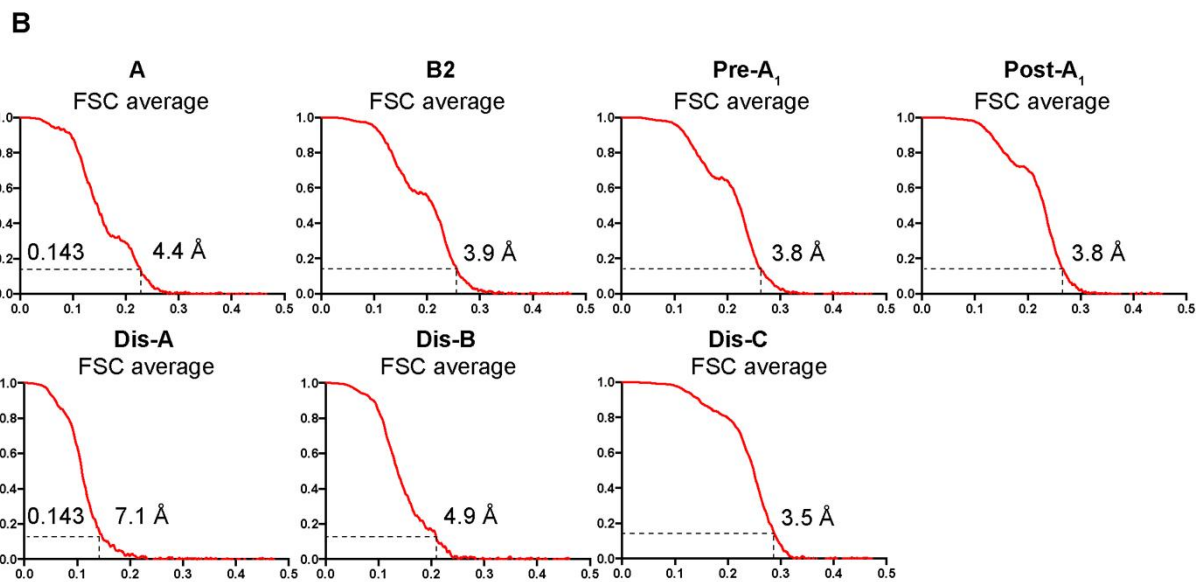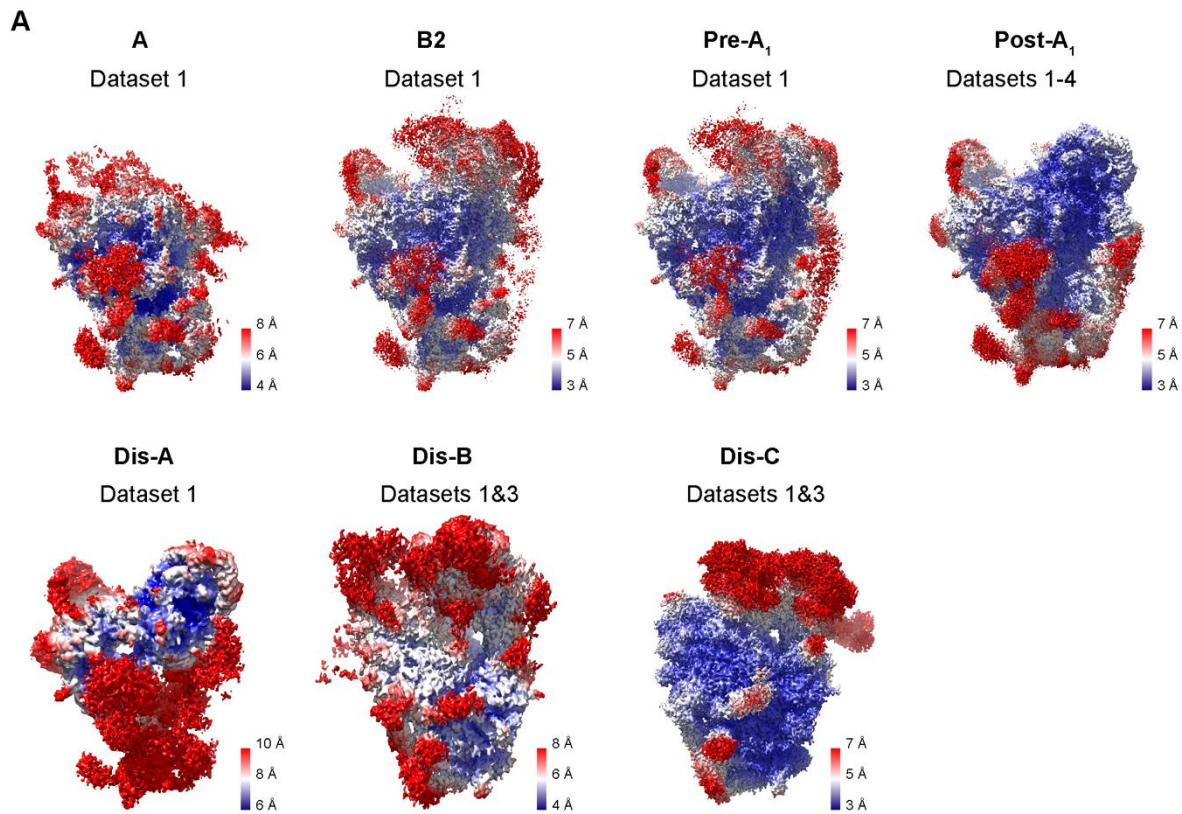
Fig. S3. Local resolution and structural data validation.

(**A**) The local-resolution distributions of the seven intermediates, as estimated by Relion and indicated by a blue-to-red color scale (see individual scale bars). All states are shown in the classical 90S view, whereas states Dis-B and Dis-C are shown in a 40S view. (**B**) The gold-standard Fourier shell correlation (FSC) curves corresponding to the reconstructions of the seven states are shown as estimated by Relion. (**C**) The FSC plots of the molecular models against the cryo-EM maps are shown for the four intermediates for which full models were built.
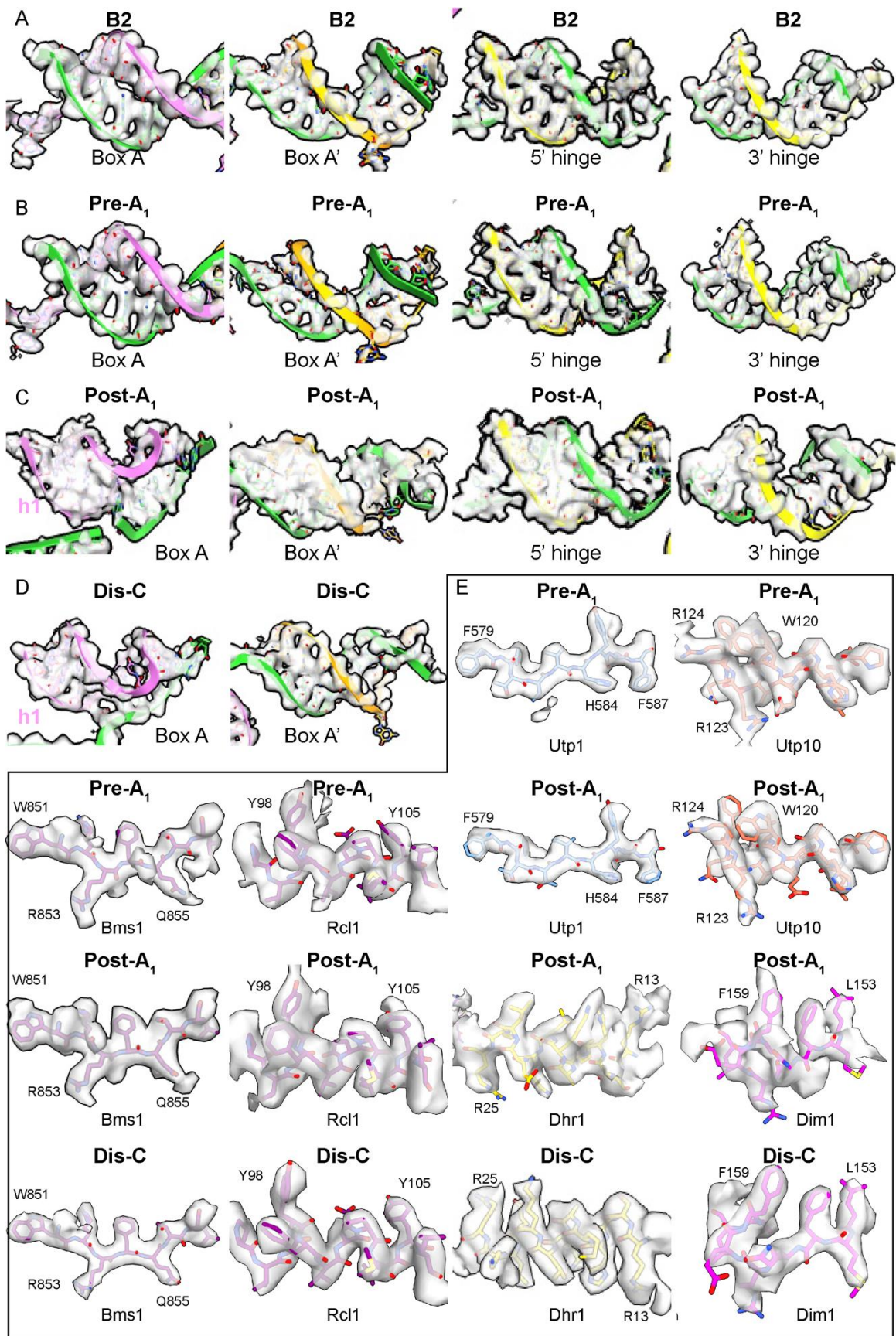
Fig. S4. Model validation for all the heteroduplexes and representative protein regions.

(**A-D**) All four heteroduplexes Box A (U3 snoRNA with 5' domain of the 18S rRNA), Box A' (U3 snoRNA with central domain of the 18S rRNA), 5' hinge (U3 snoRNA with 5' ETS RNA) and 3' hinge (U3 snoRNA with 5' ETS RNA) in state B2 (**A**), Pre-A$_1$ (**B**), Post-A$_1$ (**C**) and Dis-C (**D**) are shown with base paring. Cryo-EM densities representing the heteroduplexes are shown together with models. The h1 of the 18S rRNA is also shown with cryo-EM density. (**E**) A gallery collection of assembly factors shown together with their respective density in states Pre-A1, Post-A1 and Dis-C.
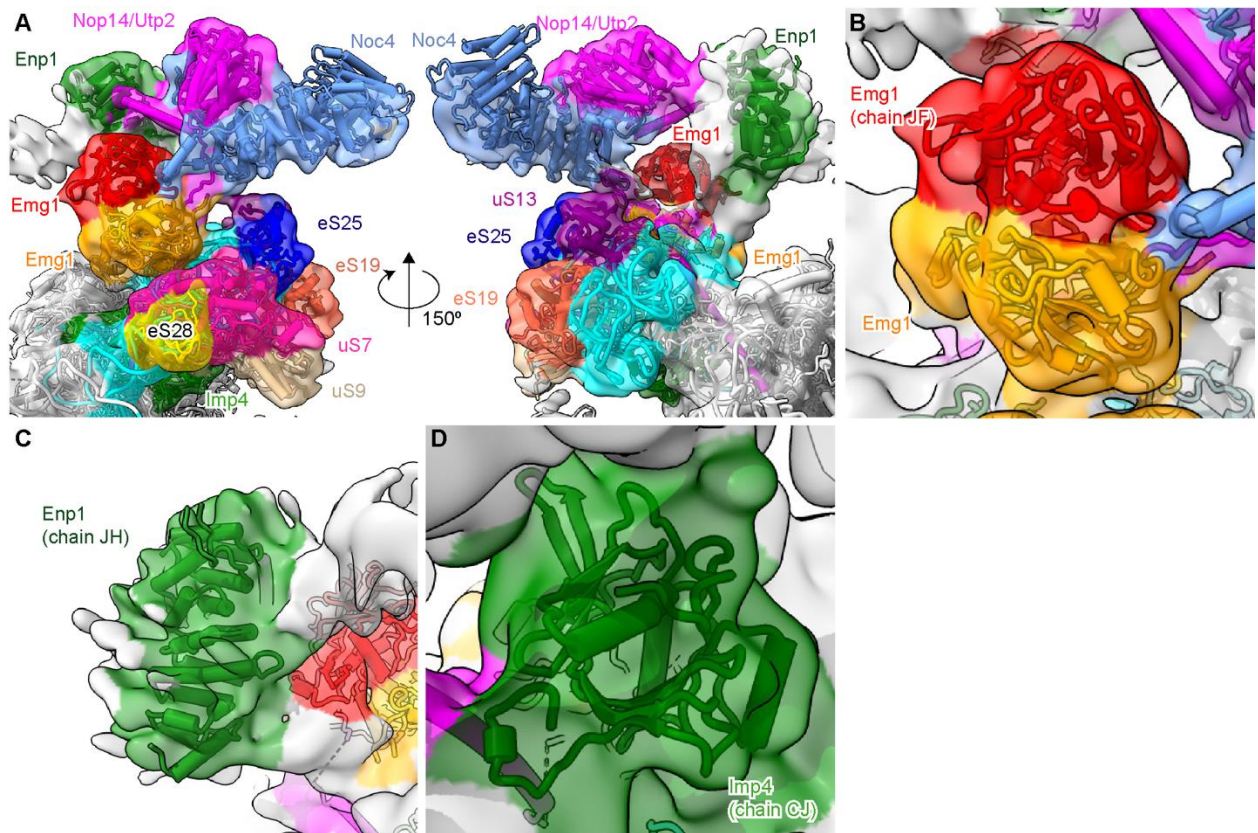
Fig. S5. Rigid body fit of the Noc4 module and Imp4 in state Dis-C.

(**A**) Two different views of the color-coded Noc4 module (Noc4, Nop14/Utp2, Enp1 and Emg1 dimer) fitted into the 10 Å low-pass filtered map of state Dis-C. The map used is a result of multibody refinement. Ribosomal proteins uS7, uS9, eS19, eS25 and eS28 are shown. (**B-D**) Focused views on the assembly factors Emg1 dimer (**B**), Enp1 (**C**) and Imp4 (**D**) rigid-body docked into the same low-pass filtered map of the state Dis-C.
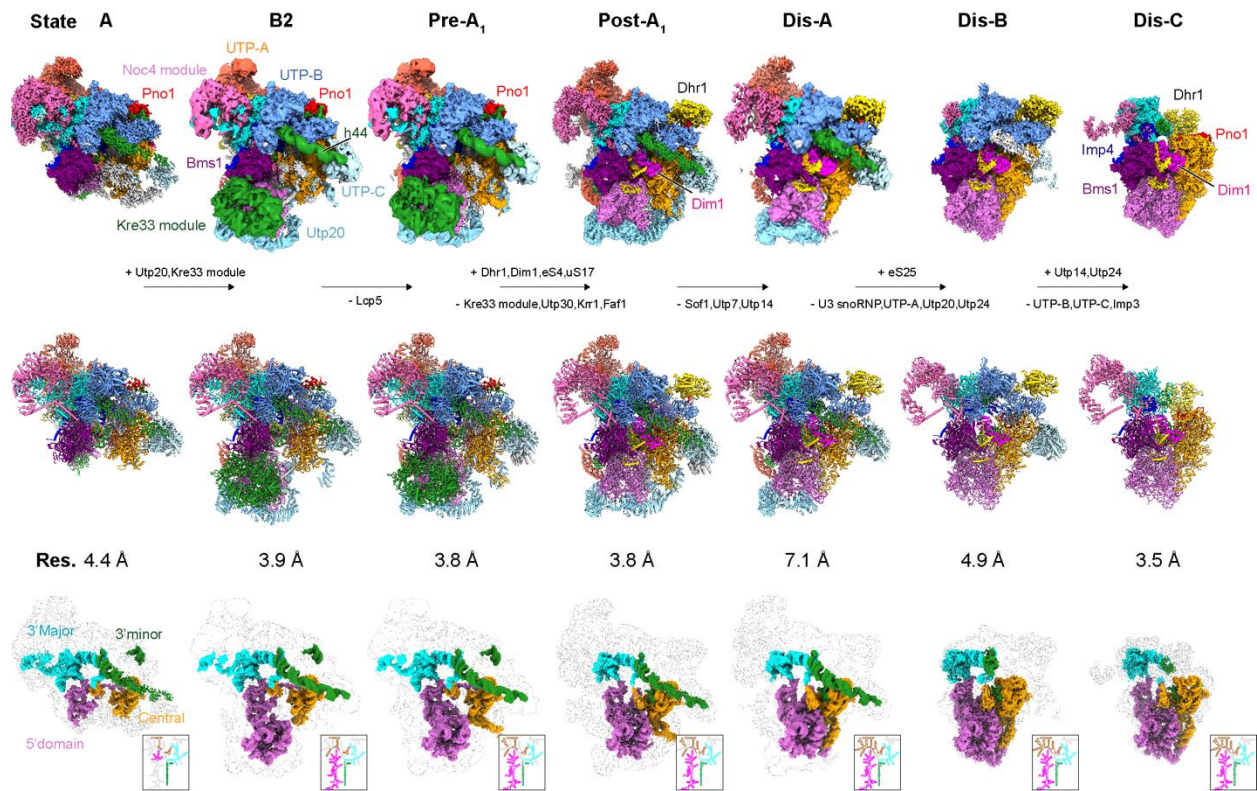
Fig. S6. Cryo-EM analysis of 90S to pre-40S transition intermediates shown in 40S view.

Cryo-EM maps (top) and molecular models (middle) of distinct states of yeast 90S to pre-40S transition observed after split-tag affinity-purification using Noc4–Dhr1 (all states) and Dhr1–Dim1 (states B2, Post-A$_1$, Dis-A, Dis-B, Dis-C). Assembly factors and modules are labeled and compositional changes indicated. Depiction of pre-18S rRNA density using 40S views with corresponding rRNA secondary structures (bottom; color code: 5' domain, magenta; central domain, orange; 3' major domain, cyan; 3' minor domain, green).
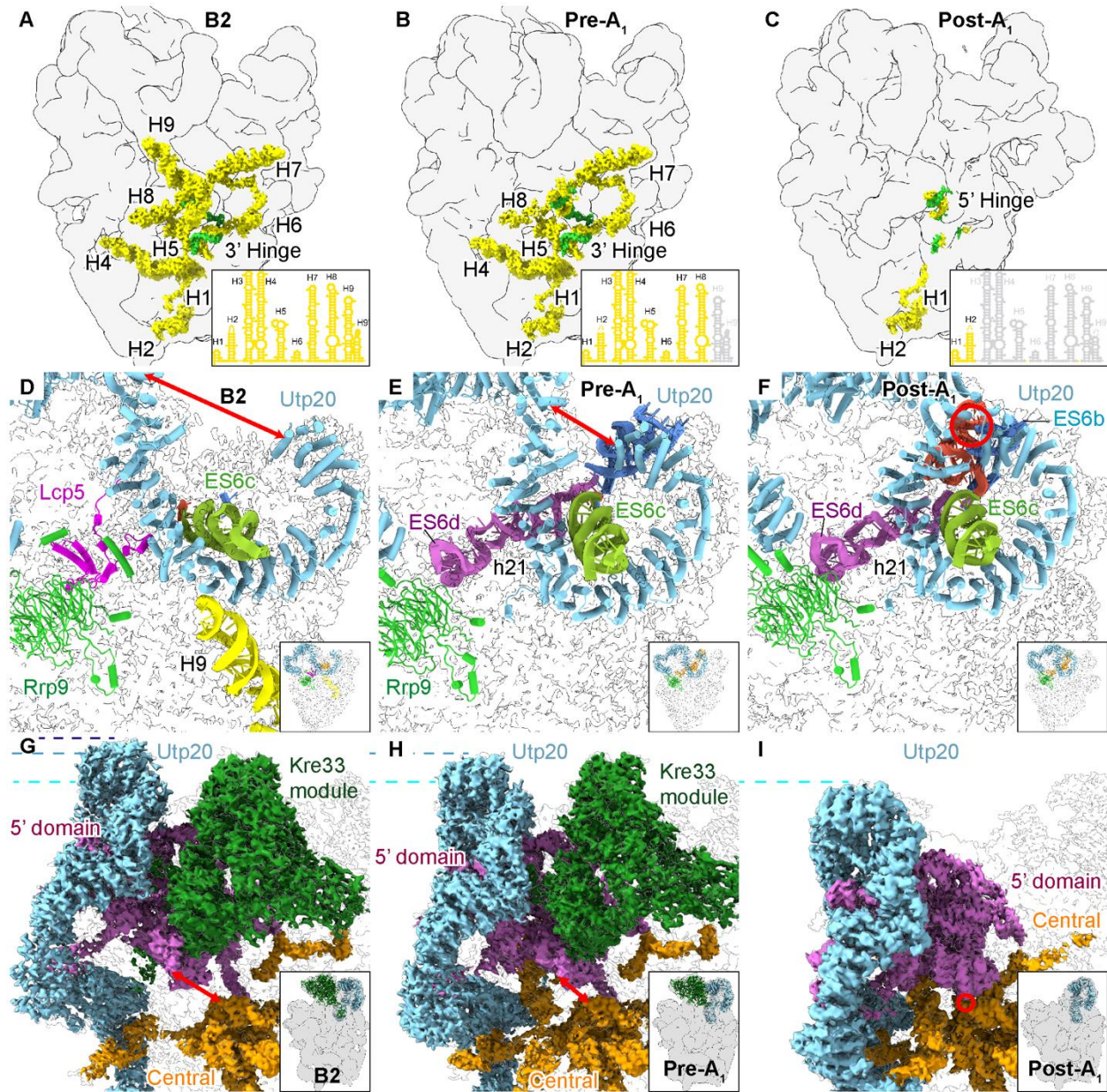
Fig. S7. Dismantling of the 5'-ETS RNA upon A$_1$ cleavage and compaction of the pre-18S rRNA.

(**A–C**) Cryo-EM maps of the 5'-ETS (yellow) in state B (**A**), Pre-A$_1$ (**B**) and Post-A$_1$ (**C**) with the particle shown in gray. Secondary structure diagrams of the 5'-ETS with remaining (yellow) and dismantled (gray) helices shown in boxes. Remnants of U3 snoRNA included in green. (**D–F**) Models of Utp20, Lcp5 and Rrp9 bound to rRNA helices in state B2 (**D**), Pre-A$_1$ (**E**) and Post-A$_1$

(**F**), showing changes of the Lcp5 and Utp20 C-terminal region and concomitant positioning of rRNA expansion segment 6 (ES6, including ES6a, ES6b, ES6c, and ES6d) and h21 of the 18S pre-rRNA, as well as disappearance of H9 of the 5'-ETS rRNA. (**G–I**) Upon A_1 cleavage, the Kre33 module (Kre33 dimer, Enp2, Bfr2, Lc5) is released, while compaction of the 5' and central domains of the 18S pre-RNA to a near-mature conformation takes place between state B2 (**D**), Pre-A_1 (**E**) and Post-A_1 (**F**). Dashed lines indicate the movement of Utp20. (**G–I**) Cryo-EM maps of the Kre33 module and Utp20 bound to the 5' and central domains in states B2 (**G**), Pre-A_1 (**H**) and Post-A_1 (**I**) highlighting the compaction of pre-18S rRNA segments (red arrow and circle).
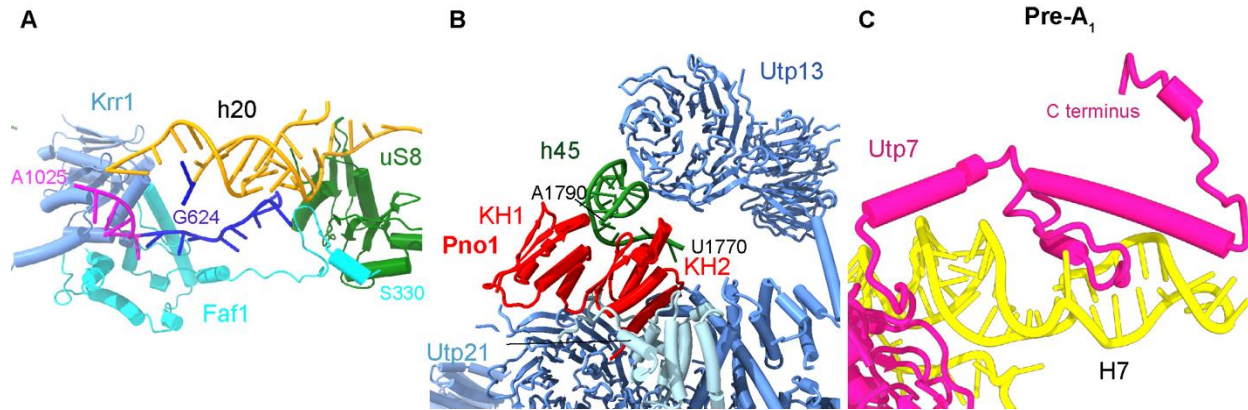
Fig. S8. Conformational, positional and compositional changes upon A_1 cleavage.

(**A**) Ribbon representation of the platform region in state Pre-A_1. Related to Fig, 3H. Nucleotides 1021–1025 of the 18S rRNA are labeled in magenta. (**B**) Model of the helical bundle of Utp21 bound to Pno1 and h45 in state Pre-A_1. Related to Fig. 3M. The UTP-B complex is shown in blue and Utp21 in light blue. (**C**) Ribbon representation of the very C terminus of Utp7 that adopts a helix and loop structure to interact with H7 of the 5'-ETS RNA.
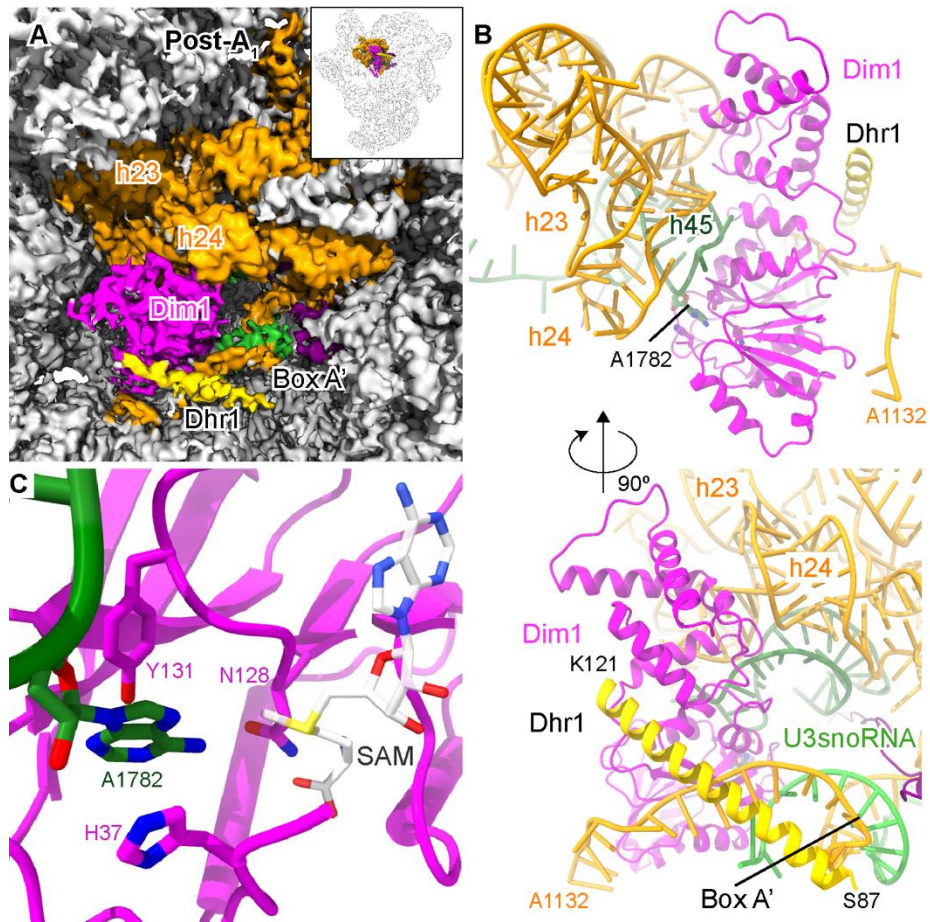
Fig. S9. Dim1 methylates A1782 of the 18S pre-rRNA on the 90S pre-ribosome.

(**A**) EM density in the platform region in state Post-A$_1$, showing Dim1 methyltransferase (magenta) is bound mainly to rRNA helix h24. (**B**) Two views of Dim1 interacting with the central domain region in state Post-A$_1$, including rRNA helices h24 and h45 as well as Box A' and an N-terminal helix of Dhr1 (yellow). (**C**) Molecular model of Dim1 bound to A1782 of the 18S rRNA in state Post-A$_1$. Although clear density for S-Adenosyl methionine (SAM) was missing, its position was superimposed from human DIMT1 (PDB ID: 1ZQ9) and all key residues are shown as sticks.
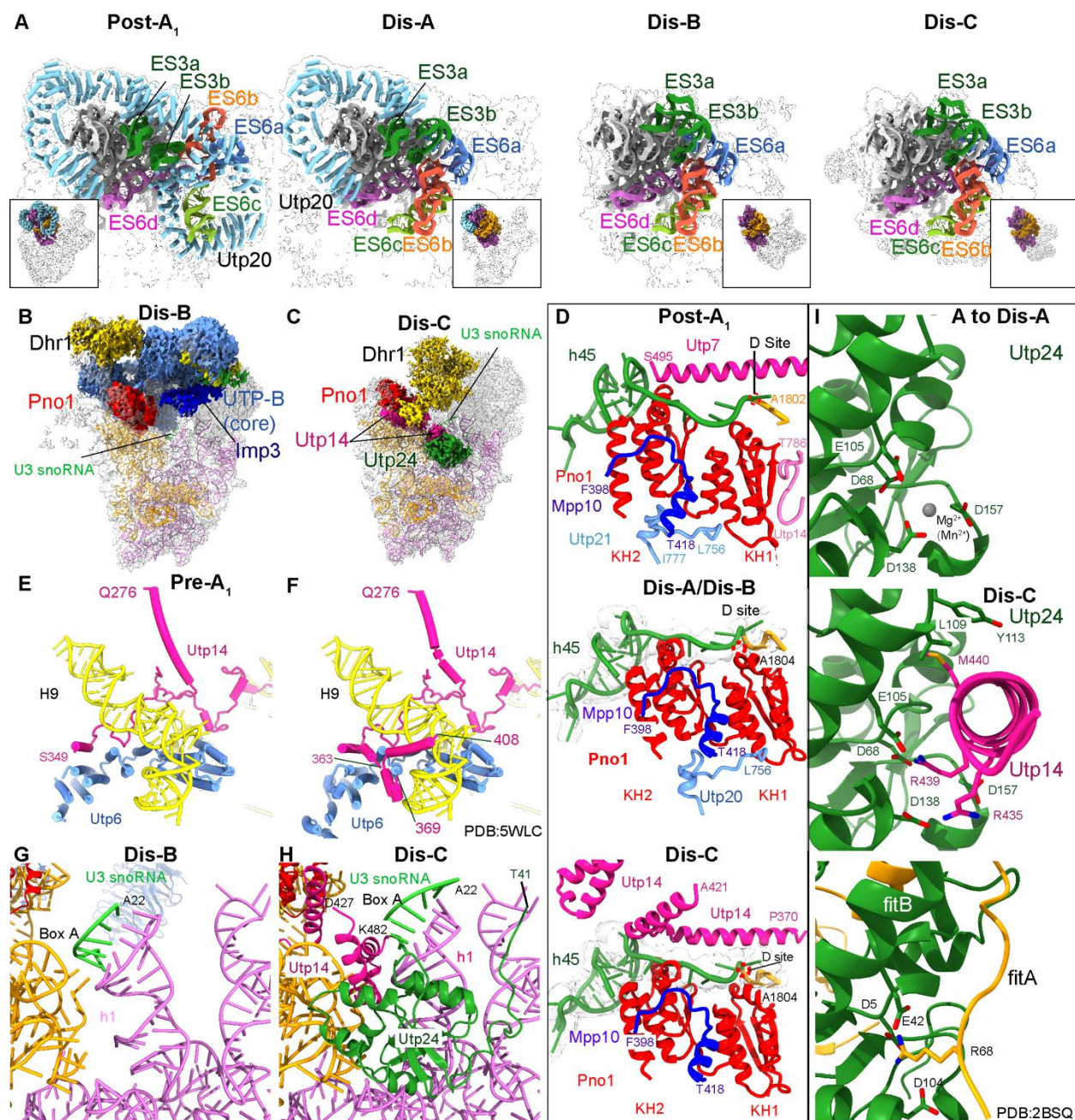
Fig. S10. The rearrangement from state Dis-B to Dis-C.

(**A**) Ribbon representation of the expansion segments 3 and 6 (ES3, ES6) during the transition from state Post-A₁ to Dis-C, highlighting the maturation process. ES3 is shown in green and the four ES6 helices are colored blue, orange, light green and magenta, respectively. Utp20 is shown in sky blue. (**B** and **C**) Assembly factors in state Dis-B (**B**), which are different from those in Dis-

C (**C**), are shown as a colored density map. (**D**) Different modes of interaction between Pno1 and the 3' end of the 18S pre-rRNA in state Post-A$_1$ (top), Dis-A/Dis-B (middle), and Dis-C (bottom). Notably, in contrast to Pno1, the C-terminal helix of Utp7 and one helix of Utp14 can interact with the 3' end of the 18S pre-rRNA in state Post-A$_1$ and Dis-C, respectively. (**E** and **F**) Ribbon representation of Utp14 associated to Utp6 and H9 of the 5'-ETS RNA. Models are from our state Pre-A$_1$ (**E**) or previously published (PDB ID: 5WLC) (**F**). (**G** and **H**) Model of the h1 region in in state Dis-B (**G**) and Dis-C (**H**), highlighting the reassociation of Utp24 together with Utp14. (**I**) Model of the catalytic center of Utp24 showing the coordination of a Mg^{2+} or Mn^{2+} ion in the states A to Dis-A (top). In state Dis-C (middle), the catalytic center is inhibited by R435 and R439 of Utp14, similar to the inhibitory arginine finger of the bacterial fitA protein which inserts into the catalytic center of the PIN nuclease fitB (bottom, PDB ID: 2BSQ).
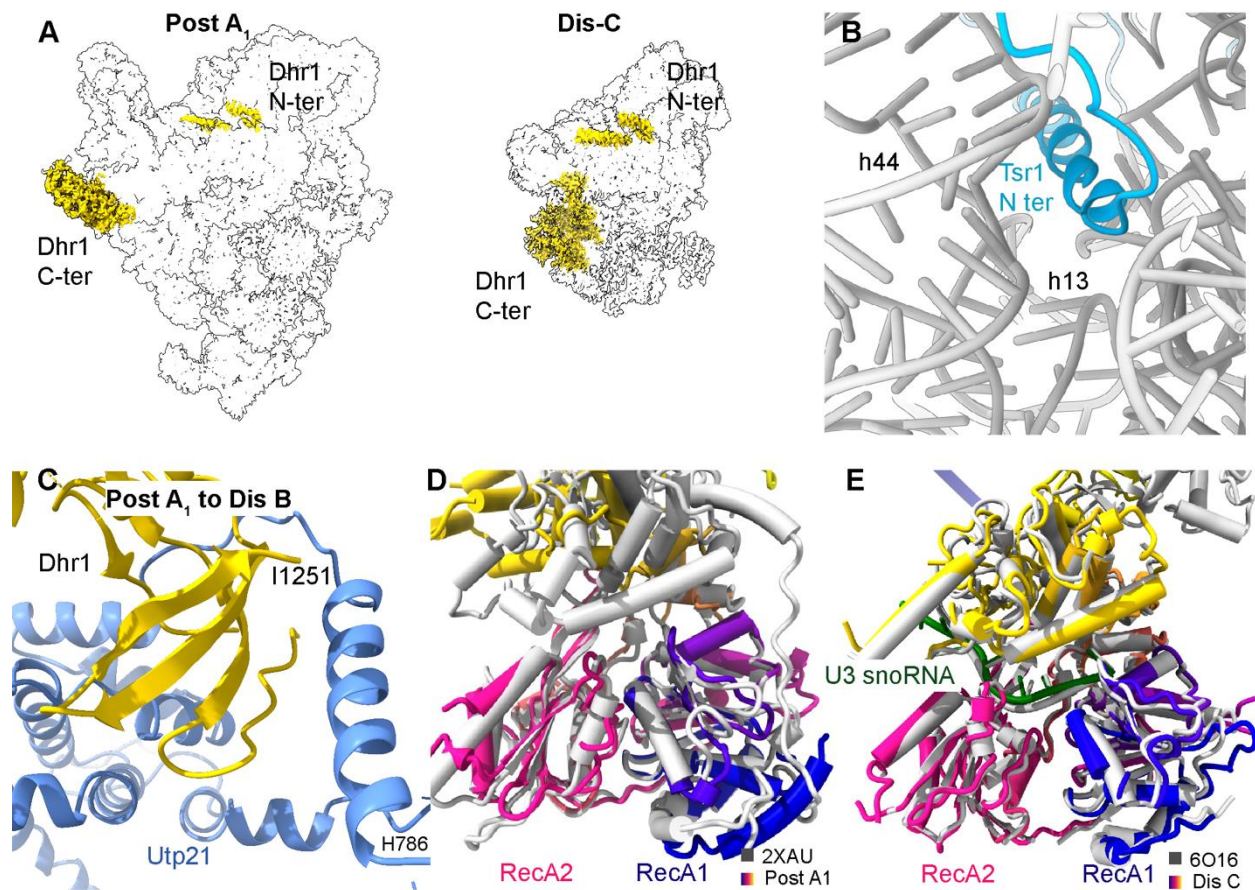
Fig. S11. Two different nucleotide binding conformations of Dhr1.

(**A**) The location of the N- and C-terminus (N-ter, C-ter) of Dhr1 in states Post-A$_1$ (left) and Dis-C (right). (**B**) Model of the N-terminus of Tsr1 interacting with the h13 region in the cytoplasmic pre-40S ribosome, which would clash with the N-terminus of Dhr1. Tsr1 from PDB ID 6EML is shown in blue and rRNA in gray. (**C**) Model of the C-terminal domain of Dhr1 interacting with the C-terminal helical bundle of Utp21 from state Post-A$_1$ to Dis-B (here state Post-A$_1$). (**D**) Superimposition of Dhr1 from state Post-A$_1$ with ADP-bound Prp43 (PDB ID: 2XAU, gray). (**E**) Superimposition of RNA-binding Dhr1 from state Dis-C with human nucleotide-free DHX37, the homolog of Dhr1 (PDB ID: 6O16, gray).

Table S1. Cryo-EM data collection, refinement and validation statistics.

	A	B2	Pre-A$_1$	Post-A$_1$	Dis-A	Dis-B	Dis-C
Data Collection and Processing							
Microscope	FEI Titan Krios (Falcon II)				FEI Titan Krios (K2)		
Magnification	129,151				47,214		
Voltage (kV)			300				
Electron exposure (e–/Å^2)		28				44	
Defocus range (µm)		-1 to -2.5				-0.8 to -2.5	
Pixel size (Å)		1.084				1.059	
Symmetry imposed			*C1*				
Initial particle images (no.)	1,242,390 (Noc4-Dhr1), 1,141,220 (Dhr1-Dim1)				753,312 (Noc4-Dhr1), 707,080 (Dhr1-Dim1)		
Final particle images (no.)	11,194	32,213	43,601	176,136	11,312	16,654	102,097
Map resolution (Å)	4.4	3.9	3.8	3.8	7.1	4.9	3.5
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Model refinement							
Initial model used (PDB code)			6RXU/6RXV/5WLC				
Model resolution (Å)		3.9	3.8	3.8			3.5
FSC threshold		0.5	0.5	0.5			0.5
Map sharpening *B* factor (Å^2)	-96	-88	-83	-154	-186	-116	-70
Model composition							
Non-hydrogen atoms		223,826	235,426	221,481			80,339
Protein residues		24,210	24,397	22,873			7,259
RNA		1,909	2,017	1,715			1,418
Ligands		4	5	4			43
B-factors							
B factors (Å^2)		47.70	60.28	48.82			60.32
Protein		41.79	53.87	46.08			49.09
RNA		73.31	87.63	60.81			78.36
Ligand		54.06	71.69	65.08			66.70
RMSDs							
Bond lengths (Å)		0.0218	0.0090	0.0073			0.0182
Bond angles (°)		1.67	0.97	0.92			1.40
Validation							
MolProbity score		2.10	1.92	1.92			1.92
Clashscore		11.37	8.57	9.28			7.62
Poor rotamers (%)		0.07	0.61	0.43			0.05
Ramachandran plot							
Favored (%)		90.70	92.83	93.55			91.47
Allowed (%)		9.22	7.13	6.40			8.42
Disallowed (%)		0.08	0.04	0.05			0.11

PDB codes		6ZQA	6ZQB	6ZQC	6ZQD	6ZQE	6ZQF	6ZQG
EMDB codes		11357	11358	11359	11360	11361	11362	11363

Movie S1.
The formation of h1 of the 18S pre-rRNA.

Data S1. (separate file)
Semiquantitative mass spectrometry results of the Noc4-Dhr1 sample.

References

1. F. Dragon, J. E. G. Gallagher, P. A. Compagnone-Post, B. M. Mitchell, K. A. Porwancher, K. A. Wehner, S. Wormsley, R. E. Settlage, J. Shabanowitz, Y. Osheim, A. L. Beyer, D. F. Hunt, S. J. Baserga, A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**, 967–970 (2002).

2. P. Grandi, V. Rybin, J. Bassler, E. Petfalski, D. Strauss, M. Marzioch, T. Schäfer, B. Kuster, H. Tschochner, D. Tollervey, A.-C. Gavin, E. Hurt, 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol. Cell* **10**, 105–115 (2002).

3. J. L. Woolford Jr., S. J. Baserga, Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics* **195**, 643–681 (2013).

4. J. Baßler, E. Hurt, Eukaryotic Ribosome Assembly. *Annu. Rev. Biochem.* **88**, 281–306 (2019).

5. S. Klinge, J. L. Woolford Jr., Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2019).

6. J. de la Cruz, D. Kressler, D. Tollervey, P. Linder, Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3′ end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J.* **17**, 1128–1140 (1998).

7. M. Thoms, E. Thomson, J. Baßler, M. Gnädig, S. Griesel, E. Hurt, The Exosome Is Recruited to RNA Substrates through Specific Adaptor Proteins. *Cell* **162**, 1029–1038 (2015).

8. M. Kornprobst, M. Turk, N. Kellner, J. Cheng, D. Flemming, I. Koš-Braun, M. Koš, M. Thoms, O. Berninghausen, R. Beckmann, E. Hurt, Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. *Cell* **166**, 380–393 (2016).

9. J. Barandun, M. Chaker-Margot, M. Hunziker, K. R. Molloy, B. T. Chait, S. Klinge, The complete structure of the small-subunit processome. *Nat. Struct. Mol. Biol.* **24**, 944–953 (2017).

10. M. Chaker-Margot, J. Barandun, M. Hunziker, S. Klinge, Architecture of the yeast small subunit processome. *Science* **355**, eaal1880 (2017).

11. J. Cheng, N. Kellner, O. Berninghausen, E. Hurt, R. Beckmann, 3.2-Å-resolution structure of the 90S preribosome before A1 pre-rRNA cleavage. *Nat. Struct. Mol. Biol.* **24**, 954–964 (2017).

12. Q. Sun, X. Zhu, J. Qi, W. An, P. Lan, D. Tan, R. Chen, B. Wang, S. Zheng, C. Zhang, X. Chen, W. Zhang, J. Chen, M.-Q. Dong, K. Ye, Molecular architecture of the 90S small subunit pre-ribosome. *eLife* **6**, e22086 (2017).

13. M. Chaker-Margot, M. Hunziker, J. Barandun, B. D. Dill, S. Klinge, Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis. *Nat. Struct. Mol. Biol.* **22**, 920–923 (2015).

14. L. Zhang, C. Wu, G. Cai, S. Chen, K. Ye, Stepwise and dynamic assembly of the earliest precursors of small ribosomal subunits in yeast. *Genes Dev.* **30**, 718–732 (2016).

15. J. Cheng, J. Baßler, P. Fischer, B. Lau, N. Kellner, R. Kunze, S. Griesel, M. Kallas, O. Berninghausen, D. Strauss, R. Beckmann, E. Hurt, Thermophile 90S Pre-ribosome Structures Reveal the Reverse Order of Co-transcriptional 18S rRNA Subdomain Integration. Mol. Cell 75, 1256–1269.e7 (2019).

16. T. Wegierski, E. Billy, F. Nasr, W. Filipowicz, Bms1p, a G-domain-containing protein, associates with Rcl1p and is required for 18S rRNA biogenesis in yeast. *RNA* **7**, 1254–1267 (2001).

17. S. Granneman, J. E. Gallagher, J. Vogelzangs, W. Horstman, W. J. van Venrooij, S. J. Baserga, G. J. Pruijn, The human Imp3 and Imp4 proteins form a ternary complex with hMpp10, which only interacts with the U3 snoRNA in 60-80S ribonucleoprotein complexes. *Nucleic Acids Res.* **31**, 1877–1887 (2003).

18. N. J. Krogan, W.-T. Peng, G. Cagney, M. D. Robinson, R. Haw, G. Zhong, X. Guo, X. Zhang, V. Canadien, D. P. Richards, B. K. Beattie, A. Lalev, W. Zhang, A. P. Davierwala, S. Mnaimneh, A. Starostine, A. P. Tikuisis, J. Grigull, N. Datta, J. E. Bray, T. R. Hughes, A. Emili, J. F. Greenblatt, High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell* **13**, 225–239 (2004).

19. J. Pérez-Fernández, A. Román, J. De Las Rivas, X. R. Bustelo, M. Dosil, The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell. Biol.* **27**, 5414–5429 (2007).

20. J. Pérez-Fernández, P. Martín-Marcos, M. Dosil, Elucidation of the assembly events required for the recruitment of Utp20, Imp4 and Bms1 onto nascent pre-ribosomes. *Nucleic Acids Res.* **39**, 8105–8121 (2011).

21. T. Schäfer, D. Strauss, E. Petfalski, D. Tollervey, E. Hurt, The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. *EMBO J.* **22**, 1370–1380 (2003).

22. T. Schäfer, B. Maco, E. Petfalski, D. Tollervey, B. Böttcher, U. Aebi, E. Hurt, Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature* **441**, 651–655 (2006).

23. B. S. Strunk, C. R. Loucks, M. Su, H. Vashisth, S. Cheng, J. Schilling, C. L. Brooks 3rd, K. Karbstein, G. Skiniotis, Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* **333**, 1449–1453 (2011).

24. E. Wyler, M. Zimmermann, B. Widmann, M. Gstaiger, J. Pfannstiel, U. Kutay, I. Zemp, Tandem affinity purification combined with inducible shRNA expression as a tool to study the maturation of macromolecular assemblies. *RNA* **17**, 189–200 (2011).

25. N. Larburu, C. Montellese, M.-F. O'Donohue, U. Kutay, P.-E. Gleizes, C. Plisson-Chastang, Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. *Nucleic Acids Res.* **44**, 8465–8478 (2016).

26. A. Heuer, E. Thomson, C. Schmidt, O. Berninghausen, T. Becker, E. Hurt, R. Beckmann, Cryo-EM structure of a late pre-40S ribosomal subunit from *Saccharomyces cerevisiae*. *eLife* **6**, e30189 (2017).

27. M. C. Johnson, H. Ghalei, K. A. Doxtader, K. Karbstein, M. E. Stroupe, Structural Heterogeneity in Pre-40S Ribosomes. *Structure* **25**, 329–340 (2017).

28. M. Ameismeier, J. Cheng, O. Berninghausen, R. Beckmann, Visualizing late states of human 40S ribosomal subunit maturation. *Nature* **558**, 249–253 (2018).

29. A. Scaiola, C. Peña, M. Weisser, D. Böhringer, M. Leibundgut, P. Klingauf-Nerurkar, S. Gerhardy, V. G. Panse, N. Ban, Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit. *EMBO J.* **37**, e98499 (2018).

30. V. Mitterer, R. Shayan, S. Ferreira-Cerca, G. Murat, T. Enne, D. Rinaldi, S. Weigl, H. Omanic, P.-E. Gleizes, D. Kressler, C. Plisson-Chastang, B. Pertschy, Conformational proofreading of distant 40S ribosomal subunit maturation events by a long-range communication mechanism. *Nat. Commun.* **10**, 2754 (2019).

31. D. Lafontaine, J. Vandenhaute, D. Tollervey, The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes Dev.* **9**, 2470–2481 (1995).

32. M. Sturm, J. Cheng, J. Baßler, R. Beckmann, E. Hurt, Interdependent action of KH domain proteins Krr1 and Dim2 drive the 40S platform assembly. *Nat. Commun.* **8**, 2213 (2017).

33. A. Colley, J. D. Beggs, D. Tollervey, D. L. Lafontaine, Dhr1p, a putative DEAH-box RNA helicase, is associated with the box C+D snoRNP U3. *Mol. Cell. Biol.* **20**, 7238–7246 (2000).

34. R. Sardana, J. Zhu, M. Gill, A. W. Johnson, Physical and functional interaction between the methyltransferase Bud23 and the essential DEAH-box RNA helicase Ecm16. *Mol. Cell. Biol.* **34**, 2208–2220 (2014).

35. R. Sardana, X. Liu, S. Granneman, J. Zhu, M. Gill, O. Papoulas, E. M. Marcotte, D. Tollervey, C. C. Correll, A. W. Johnson, The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLOS Biol.* **13**, e1002083 (2015).

36. J. Zhu, X. Liu, M. Anjos, C. C. Correll, A. W. Johnson, Utp14 Recruits and Activates the RNA Helicase Dhr1 To Undock U3 snoRNA from the Preribosome. *Mol. Cell. Biol.* **36**, 965–978 (2016).

37. M. Hunziker, J. Barandun, O. Buzovetsky, C. Steckler, H. Molina, S. Klinge, Conformational switches control early maturation of the eukaryotic small ribosomal subunit. *eLife* **8**, e45185 (2019).

38. F. Bleichert, S. Granneman, Y. N. Osheim, A. L. Beyer, S. J. Baserga, The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9464–9469 (2006).

39. R. Tomecki, A. Labno, K. Drazkowska, D. Cysewski, A. Dziembowski, hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. *RNA Biol.* **12**, 1010–1029 (2015).

40. G. R. Wells, F. Weichmann, D. Colvin, K. E. Sloan, G. Kudla, D. Tollervey, N. J. Watkins, C. Schneider, The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. *Nucleic Acids Res.* **44**, 5399–5409 (2016).

41. W. An, Y. Du, K. Ye, Structural and functional analysis of Utp24, an endonuclease for processing 18S ribosomal RNA. *PLOS ONE* **13**, e0195723 (2018).

42. K. Sharma, D. Tollervey, Base pairing between U3 small nucleolar RNA and the 5′ end of 18S rRNA is required for pre-rRNA processing. *Mol. Cell. Biol.* **19**, 6012–6019 (1999).

43. J. J. Black, Z. Wang, L. M. Goering, A. W. Johnson, Utp14 interaction with the small subunit processome. *RNA* **24**, 1214–1228 (2018).

44. F. M. Boneberg, T. Brandmann, L. Kobel, J. van den Heuvel, K. Bargsten, L. Bammert, U. Kutay, M. Jinek, Molecular mechanism of the RNA helicase DHX37 and its activation by UTP14A in ribosome biogenesis. *RNA* **25**, 685–701 (2019).

45. A. Roychowdhury, C. Joret, G. Bourgeois, V. Heurgué-Hamard, D. L. J. Lafontaine, M. Graille, The DEAH-box RNA helicase Dhr1 contains a remarkable carboxyl terminal domain essential for small ribosomal subunit biogenesis. *Nucleic Acids Res.* **47**, 7548–7563 (2019).

46. A. Narla, B. L. Ebert, Ribosomopathies: Human disorders of ribosome dysfunction. *Blood* **115**, 3196–3205 (2010).

47. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

48. M. Koš, D. Tollervey, The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in Saccharomyces cerevisiae. *Mol. Cell* **20**, 53–64 (2005).

49. S. V. Gnanasundram, I. C. Kos-Braun, M. Koš, At least two molecules of the RNA helicase Has1 are simultaneously present in pre-ribosomes during ribosome biogenesis. *Nucleic Acids Res.* **47**, 10852–10864 (2019).

50. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

51. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

52. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).

53. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

54. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

55. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

56. M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–W258 (2014).

57. I. Prokhorova, R. B. Altman, M. Djumagulov, J. P. Shrestha, A. Urzhumtsev, A. Ferguson, C. T. Chang, M. Yusupov, S. C. Blanchard, G. Yusupova, Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10899–E10908 (2017).

58. M. Popenda, M. Szachniuk, M. Antczak, K. J. Purzycka, P. Lukasiak, N. Bartol, J. Blazewicz, R. W. Adamiak, Automated 3D structure composition for large RNAs. *Nucleic Acids Res.* **40**, e112 (2012).

59. M. J. Tauchert, J. B. Fourmann, H. Christian, R. Lührmann, R. Ficner, Structural and functional analysis of the RNA helicase Prp43 from the thermophilic eukaryote Chaetomium thermophilum. *Acta Crystallogr. F* **72**, 112–120 (2016).

60. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart,

PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

61. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

62. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

63. D. Lafontaine, J. Delcour, A. L. Glasser, J. Desgrès, J. Vandenhaute, The DIM1 gene responsible for the conserved $m^6_2Am^6_2A$ dimethylation in the 3′-terminal loop of 18S rRNA is essential in yeast. *J. Mol. Biol.* **241**, 492–497 (1994).

64. C. Zorbas, E. Nicolas, L. Wacheul, E. Huvelle, V. Heurgué-Hamard, D. L. J. Lafontaine, The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol. Biol. Cell* **26**, 2080–2095 (2015).

65. E. Choque, C. Schneider, O. Gadal, C. Dez, Turnover of aberrant pre-40S pre-ribosomal particles is initiated by a novel endonucleolytic decay pathway. *Nucleic Acids Res.* **46**, 4699–4714 (2018).

66. M. Senissar, M. C. Manav, D. E. Brodersen, Structural conservation of the PIN domain active site across all domains of life. *Protein Sci.* **26**, 1474–1492 (2017).