Title

Visualizing the nucleoplasmic maturation of human pre-60S ribosomal particles.

Journal

Cell Research, 33(11)

Authors

Zhang, Yunyang
Liang, Xiaomeng
Luo, Sha
et al.

Publication Date

2023-11-01

DOI

10.1038/s41422-023-00853-9

Peer reviewed

ARTICLE

Visualizing the nucleoplasmic maturation of human pre-60S ribosomal particles

Yunyang Zhang[1], Xiaomeng Liang[1], Sha Luo[1], Yan Chen[1], Yu Li[1], Chengying Ma[1,2], Ningning Li[1,2] and Ning Gao[1,2,3]✉

Eukaryotic ribosome assembly is a highly orchestrated process that involves over two hundred protein factors. After early assembly events on nascent rRNA in the nucleolus, pre-60S particles undergo continuous maturation steps in the nucleoplasm, and prepare for nuclear export. Here, we report eleven cryo-EM structures of the nuclear pre-60S particles isolated from human cells through epitope-tagged GNL2, at resolutions of 2.8–4.3 Å. These high-resolution snapshots provide fine details for several major structural remodeling events at a virtual temporal resolution. Two new human nuclear factors, L10K and C11orf98, were also identified. Comparative structural analyses reveal that many assembly factors act as successive place holders to control the timing of factor association/dissociation events. They display multi-phasic binding properties for different domains and generate complex binding inter-dependencies as a means to guide the rRNA maturation process towards its mature conformation. Overall, our data reveal that nuclear assembly of human pre-60S particles is generally hierarchical with short branch pathways, and a few factors display specific roles as rRNA chaperones by confining rRNA helices locally to facilitate their folding, such as the C-terminal domain of SDAD1.

Cell Research (2023) 33:867–878; https://doi.org/10.1038/s41422-023-00853-9

INTRODUCTION

Eukaryotic ribosome biogenesis is the most energetically expensive process in the cell, involving the transcription, processing, modification and folding of rRNAs, and the incorporation of ~80 ribosomal proteins. Over two hundred protein factors, organized in temporal/spatial groups, participate in the making of ribosomal subunits.[1–4] The assembly of ribosomal subunits starts co-transcriptionally in the nucleolus on the nascent elongating pre-rRNA transcripts. The primary pre-rRNA contains rRNA sequences for both the large and small subunits as well as the external and internal transcribed spacers.[5,6] After the cleavages in the internal transcribed spacer 1 (ITS1), the pre-40S and pre-60S particles undergo separate maturation pathways in the nucleolus and nucleoplasm, and are subsequently exported to the cytoplasm for the final maturation.[7] In the past ten years, our knowledge on eukaryotic ribosome assembly has been greatly expanded by structural studies of the yeast endogenous pre-ribosomal intermediates in various assembly stages (reviewed in refs. [3,4]), and major assembly events, general assembly pathways, and quality control checkpoints have been identified. The rich structural information from these high-resolution snapshots in nearly all major assembly stages[8–21] has enabled the formulation and experimental validation of various functional models for individual assembly factors (e.g., see refs. [22–29]).

Despite these progresses on yeast ribosome biogenesis, the assembly of human ribosomal subunits is less characterized. The length of human pre-rRNA is twice the size of its yeast counterpart, and human ITS2 alone is ~5 times the size of yeast ITS2.[5] In addition, previous studies have suggested a largely increased complexity for human ribosome assembly, and potentially much more factors are required in the pre-rRNA processing and assembly processes.[1,30–35] Furthermore, an increasing body of evidence shows that human ribosome biogenesis is tightly coupled with other cellular regulatory processes, such as c-Myc, mTOR and p53 pathways.[36–40] On one side, defective human ribosome biogenesis caused by mutations on assembly factors and ribosomal proteins, results in a reduction of functional ribosomes and/or an accumulation of functionally impaired ribosomes, which underlies several genetic diseases collectively known as ribosomopathies that are characterized by impaired hematopoiesis and increased cancer susceptibility.[41] On the other side, hyperactivated ribosome biogenesis is also a common hallmark of cancer cells,[37] and abnormally upregulated expression levels of various assembly factors and/or certain ribosomal proteins have been reported in many types of clinical tumor samples (e.g., see refs. [42,43]).

In the past five years, structures of the human pre-40S and 90S (small subunit processome) have been reported,[19,44,45] but our understanding towards human large subunit assembly is limited to stages before and after nuclear export.[20] Here we report the structural characterization of human pre-60S particles purified through epitope-tagged nucleolar guanine nucleotide binding protein-like 2 (GNL2). This collection of high-resolution structures nearly spans the complete nucleoplasmic assembly stages and reveals the details of several major structural remodeling events. Moreover, structural analyses demonstrate that local folding events of rRNA helices facilitated by assembly factors are key steps of the conformational maturation of pre-60S particles and

[1]State Key Laboratory of Membrane Biology, Peking-Tsinghua Joint Center for Life Sciences, School of Life Sciences, Peking University, Beijing, China. [2]Changping Laboratory, Beijing, China. [3]National Biomedical Imaging Center, Peking University, Beijing, China. ✉email: gaon@pku.edu.cn

Received: 10 May 2023 Accepted: 7 July 2023
Published online: 25 July 2023

that assembly factors act as successive place holders to direct the assembly process.

RESULTS

Structures of human pre-60S particles obtained through epitope-tagged GNL2

Native nuclear pre-60S particles were purified from an engineered HEK293 cell line, in which affinity tags were inserted to the C-terminus of GNL2 (Nog2 in yeast) using CRISPR/Cas9 (Supplementary information, Fig. S1a). Affinity-purified samples were first subjected to SDS-PAGE and silver staining, and multiple non-ribosomal proteins could be clearly identified (Supplementary information, Fig. S1b). Mass spectrometry analysis indicated that the sample contained many human proteins with homology to known yeast ribosomal assembly factors (Supplementary information, Table S1). Cryo-EM was next employed to characterize these pre-60S particles (Supplementary information, Figs. S1c, d and S2). Through hierarchical 3D classifications, we obtained eleven density maps for the nuclear pre-60S particles at resolution range between 2.8 Å and 4.3 Å, representing a collection of distinct assembly intermediates at nucleolar and nucleoplasmic assembly stages (Fig. 1; Supplementary information, Fig. S3). From yeast pre-60S structures at similar stages[8,12,21] and AlphaFold-predicted models of human factors,[46] we built atomic models for these density maps (Supplementary information, Fig. S4).

The first observation from these structures is that the ITS2 removal shows no coordination with the maturation events in the central protuberance (CP) and peptidyl transferase center (PTC) regions, which was also observed in yeast pre-60S structures.[8,12] For clarity and consistency, the states with and without the ITS2 (foot structure) are named state N′ and state N, respectively. This structural observation indicates that the ITS2 processing proceeds relatively independently of the GNL2-particles and spans the entire lifetime of GNL2 (Fig. 1). If we ignore the differences in the ITS2 region, these structures could be assigned into a generally continuous and hierarchical maturation pathway, based on conformational difference of the CP and known temporal relationships of corresponding yeast factors,[3,4] and they have covered a few major structural remodeling events on the GNL2-particles.

The first is the rotation of the 5S RNP (consisting of the 5S rRNA, ribosomal protein L5/uL18 and L11/uL5),[12,13,47] which separates these eleven GNL2-structures into two groups (pre-rotation vs post-rotation). Two states (A and B′) are in the pre-rotation group (Fig. 1a), and similar to the yeast pre-60S particles purified through Nog2,[12] they share a common set of human counterparts of yeast assembly factors, such as GNL2, RPF2, RRS1, NLE1 (Rsa4 in yeast), NSA2, GNL3 (Nug1 in yeast), CCDC86 (Cgr1 in yeast), MRTO4, GTPBP4 (Nog1 in yeast), LLPH (YBL028C in yeast), eIF6 and RLP24.

After the 5S RNP rotation, the next major remodeling event is the removal of Rsa4 from the pre-60S particles, which is catalyzed by an AAA+ assembly factor Rea1.[48,49] Yeast pre-60S structures[8,12,13] have demonstrated that the docking of Rea1 requires a prior binding of Sda1 and the Rix1 complex (Ipi1, Ipi3 and Rix1). Although the removal of NLE1 during human pre-60S assembly has not been fully characterized, it should involve a similar set of factors, including MDN1 (Rea1 in yeast),[50] SDAD1 (Sda1 in yeast), and the equivalent PELP1–TEX10–WDR18 complex.[51] Based on the presence or absence of NLE1, the rest nine structures could be further divided into two groups (NLE1+ vs NLE1–) (Fig. 1b, c). Although MDN1, SDAD1 and the PELP1–TEX10–WDR18 complex are all present in our samples (Supplementary information, Table S1), 3D classification failed to enrich an MDN1-bound or PELP1–TEX10–WDR18-bound population, indicating that the MDN1-catalyzed process could be very transient. However, stable binding of SDAD1 was found in four structures, including both the NLE1+ states (states D/D′) and NLE1– states (states F/F′).

In the NLE1– group of the pre-60S structures, concomitant with the NLE1 removal, CCDC86 and the remaining N-terminal segment of GNL3 have dissociated, and NSA2 begins to depart. Importantly, a new finding is the discovery of TMA16 and a previously uncharacterized factor L10K (Leydig cell tumor 10 kDa protein homolog, C19orf53) in the latest states of the NLE1– group. While TMA16 and its yeast homolog were recently identified to be a component of late nuclear pre-60S particles,[20,26,52] the involvement of L10K in ribosome assembly was previously unknown.

Overall, with deep 3D classification, we were able to generate multiple high-resolution structures for the continuous nucleoplasmic assembly steps on the GNL2-particles that involve multiple association/dissociation events of assembly factors (Fig. 1d). These structures provide successive snapshots for visualizing the CP maturation at a virtual “temporal” resolution.

Replacement of SPB1 by ZNF593 marks the completion of nucleolar assembly steps

In states A and B′, the construction of the CP has completed and the 5S RNP in pre-rotational state stably associates with the CP. A major difference between the two states is the presence of either SPB1 or ZNF593 (Fig. 1a, CP pre-rotation). These two factors are mutually exclusive on the pre-60S particles, as they share overlapping binding sites, similar to the yeast counterparts Spb1 and Bud20.[8,11,12] Yeast Spb1 is a nucleolar factor that binds during early stages of peptide exit tunnel construction, with an enzymatic activity as an RNA 2′-*O*-methyltransferase for G2922 in the ribosomal A-loop (Helix 92).[8,11,53,54] Per analogy with yeast Spb1, human SPB1 is expected to modify G4499 of the A-loop.[55,56] Both Spb1 and SPB1 contain an N-terminal methyltransferase domain, a middle DUF3381 and a long flexible C-terminal tail. In state A, only the middle domain (residues 249–319) of SPB1 is found, and the catalytic domain and the C-terminal tail are completely invisible. This is consistent with the yeast pre-60S structures in which the C-terminal domain (CTD) of Spb1 interferes with the assembly of domain V of the 25S rRNA, including the CP and L1 stalk,[8] and the catalytic domain of Spb1 overlaps with Nog2 on the pre-60S particle.[4,11,12] However, unlike previous yeast pre-60S structures, which have suggested that the complete release of Spb1 occurs before Nog2 enters the nucleolar pre-60S particle,[4,8] our structures show that DUF3381 of SPB1 could stay throughout the whole construction process of domain V and that the replacement of DUF3381 by ZNF593 occurs on the GNL2-particles.

Bud20, the yeast homolog of ZNF593, has been demonstrated to be an export factor required for pre-60S export through the nuclear pore complex.[57] Both ZNF593 and Bud20 possess a long N-terminal loop, which inserts into the junctional region of domains 0, II, IV and V of the 28S or the 25S rRNA.[12,20] Given their strategic position in a multi-domain interface, it is possible that the binding of ZNF593/Bud20 is a checkpoint event, which integrates subtle conformational signals from major domains of the 28S/25S rRNA. Consistent with this notion, yeast genetic and structural data showed that the disruption of the early nucleolar stages of peptide exit tunnel construction by *rpl4*Δ63–87 or by *nog1*Δ595–647 resulted in misassembled L1 stalk (H74–H76 of domain V) and central helices H68–H69, and also the absence of Bud20 in Nog2-particles.[23] Our structures indicate that ZNF593 is the first export factor recruited to the GNL2-particles. Thus, the replacement of SPB1 by ZNF593 signifies a completion of nucleolar maturation events and a start to prepare for nuclear export.

Coordinated rotation of the 5S RNP and CP upon RPF2/RRS1 release

The remodeling of the 5S RNP, characterized by a nearly 180° rotation, is a major nucleoplasmic maturation event.[12,13,47] After rotation, the 5S RNP assumes an orientation similar to the one seen in the mature 60S subunit. Previously, it was noted in the pre-rotational structure of the yeast Nog2-particles that the CP helices (H80–H88, Fig. 2a) are in completely different arrangements

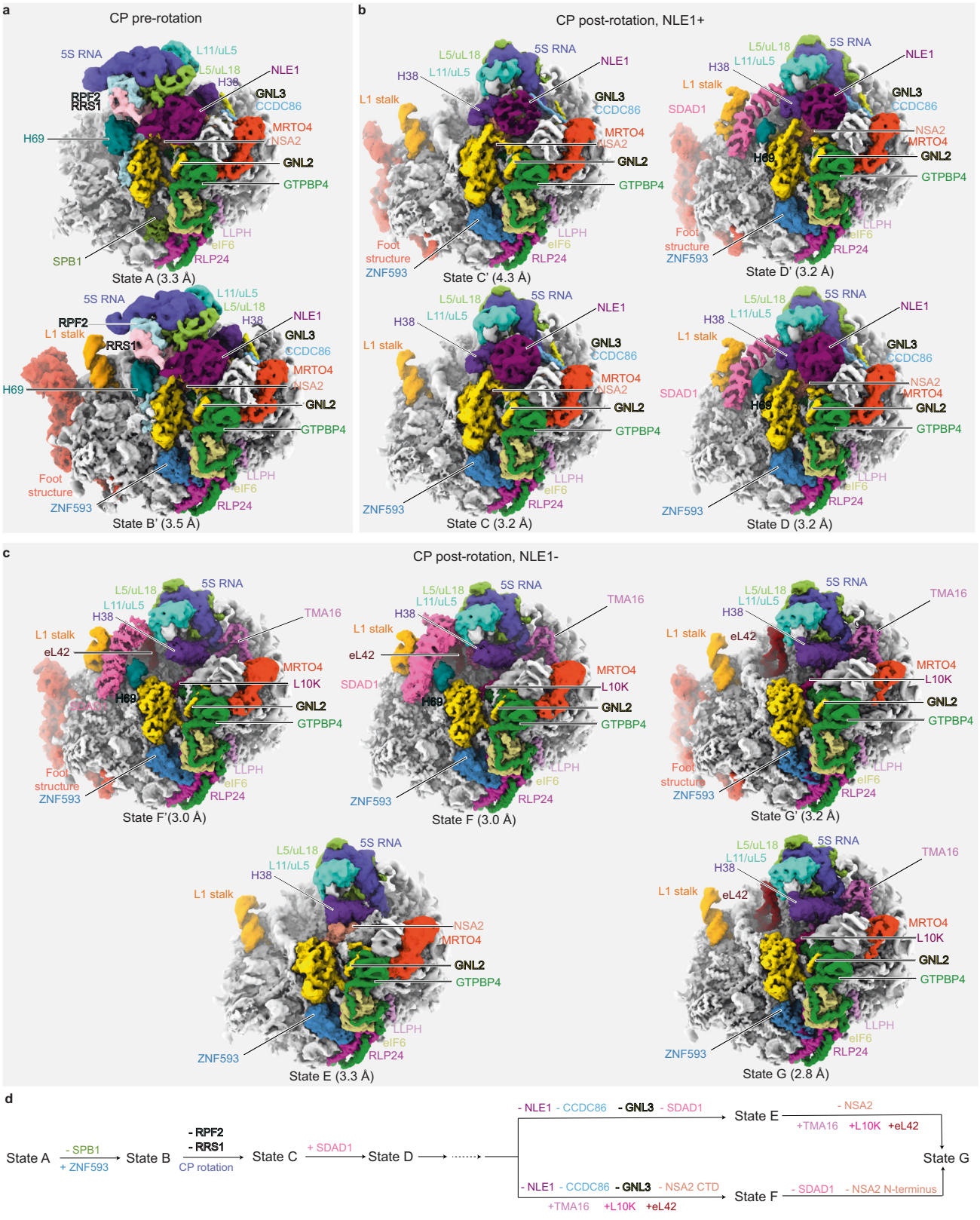

Fig. 1 Cryo-EM structures of human pre-60S particles obtained through epitope-tagged GNL2. a–c Based on the conformational difference of the CP and the presence or absence of NLE1, cryo-EM density maps of the eleven structures of GNL2-particles were divided into 3 groups: CP pre-rotation (**a**), CP post-rotation, NLE1+ (**b**) and CP post-rotation, NLE1– (**c**). The maps were viewed from the subunit interface, and Gaussian filtered with sDev of 1 Å in ChimeraX. Individual assembly factors, RNA/RNPs (including 5S RNP, L1 stalk, H38, H69, ITS2 and associated factors) and ribosomal protein eL42 are color-coded. Below each map, the overall resolution is shown. **d** A schematic diagram for the assembly pathways highlighting the temporal relationships of these states and changes in composition and conformation. For simplicity, the foot structure is ignored in the schematic diagram.

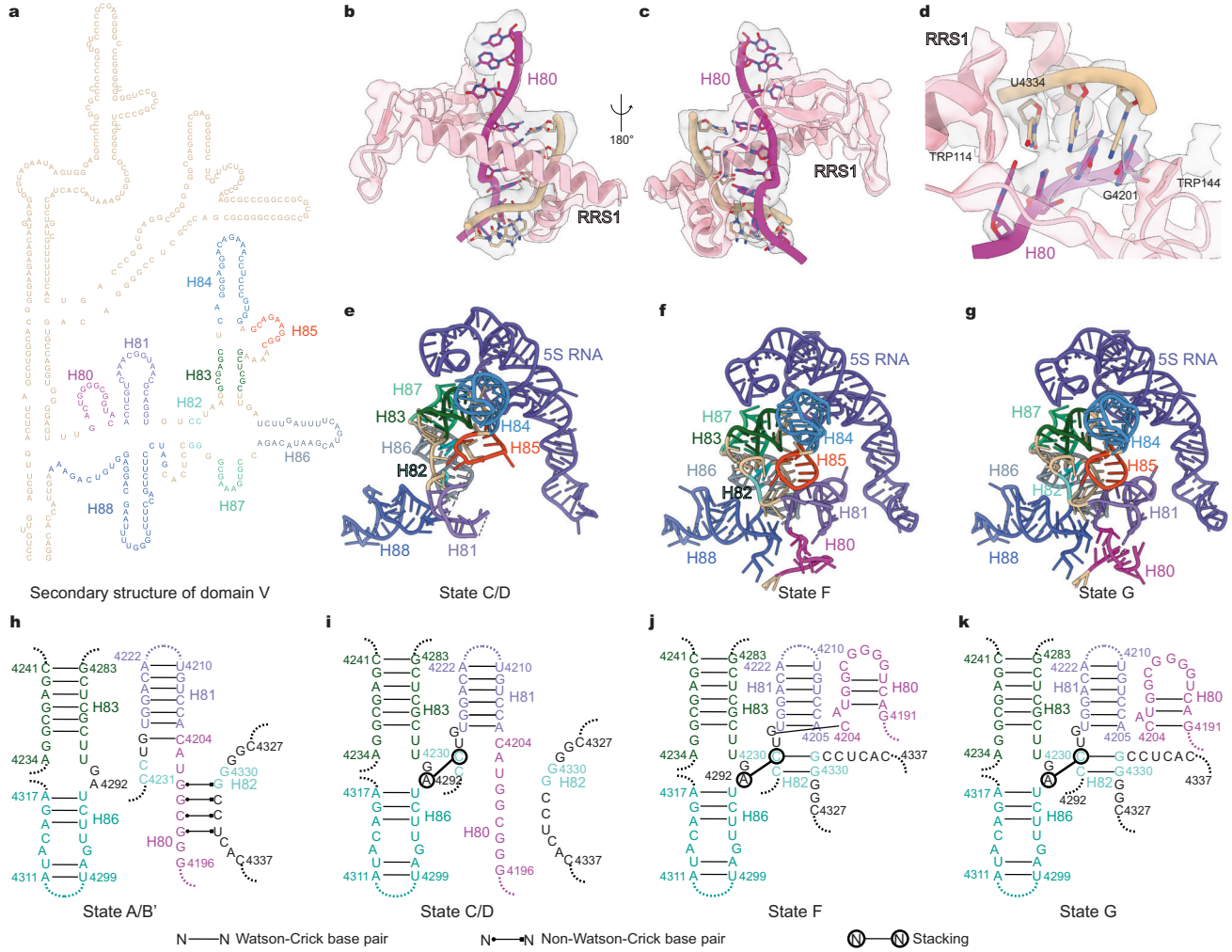

Fig. 2 Continuous conformational maturation of the CP in GNL2-particles. a Secondary structure of domain V (H74–H88) from the mature 28S rRNA, with helices H80–H88 color-coded. **b**, **c** Two 180°-rotated views show that the C-terminus of RRS1 (from state B) encircles the non-canonical helix formed by H80 and the H87–H88 linker. **d** Two tryptophan residues, W114 and W144, of RRS1 (from state B) stack on the two terminal bases (U4334 and G4201) of the non-canonical helix. **e**–**g** Zoom-in view of the rRNA helices in the CP from states C/D, F and G, illustrating the continuous, subtle conformational maturation of the CP (from states C/D to F). See also Supplementary information, Fig. S5. **h**–**k** Diagrams of secondary structures of selected CP rRNA helices (including H80–H83 and H86). Different types of base interactions are indicated.

from their mature conformation.[12] Similar to the yeast structures, the CP helices in states A/B' (pre-rotational conformation for the 5S RNP) are orientated in dramatically different positions, compared to those in states C/D (post-rotational state for the 5S RNP) (Supplementary information, Figs. S5a, b and S6a–d). Consistent with the observation from yeast structures,[12,13,47] all the CP helices, except H88, are seen to undergo large conformational changes from pre-rotational to post-rotational conformations. Previously, the structural details regarding 5S RNP rotation have not been clearly illustrated on yeast pre-60S particles due to the lack of sequential high-resolution structures. In the current study, we found that when the CP helices in the two conformations were structurally aligned using the 5S RNA as reference, the relative anchor points of the 5S RNA to the CP remain largely unaltered (Supplementary information, Fig. S6c, e and Video S1). This was also noted in the yeast structural data when a mature conformation of the CP was used as the post-rotational state.[13] Thus, our findings clarify that there is no separate 5S RNP rotation, which in fact arises from the CP remodeling upon RPF2/RRS1 release.

Structural comparisons indicate that such a drastic rearrangement of the CP-5S RNP could be attributed to the conformational changes of two linker regions of domain V. One is the linker between H88 and H87 (including H82), and the other is H80. In the pre-rotational state, H80 is stretched as a single strand (Supplementary information, Fig. S6c), with bases of 4198-GCGG-4201 forming a helix with the H87–H88 linker (4331-GCCU-4334) (Fig. 2h). This non-canonical helix between two remote rRNA fragments is stabilized by the C-terminal region of RRS1, which tightly encircles the two strands (Fig. 2b, c). Importantly, two tryptophan residues of RRS1, W114 and W144, perfectly stack on the two terminal bases of this helix, U4334 and G4201, respectively (Fig. 2d). Apparently, the conserved C-terminal region of RRS1 is crucial to maintain the pre-rotational conformation of the CP (Supplementary information, Fig. S6f).

It is unclear what triggers the release of RRS1 and RPF2. Based on the yeast structural data, it was proposed that the binding of Sda1 could be responsible for the release of Rpf2 and Rrs1 due to the overlapping binding sites between Sda1 and Rpf2.[4] However, in state C, RPF2 and RRS1 have been released, but SDAD1 is yet to be recruited (Fig. 1b), indicating that the SDAD1 binding is not a determining factor for RPF2/RRS1 release. Nevertheless, upon their departure in states C/D, this non-canonical helix is freed from such

a restricted conformation, and the two strands separate and become disordered in the density maps of states C/D (Fig. 2e, i). Upon the rotation of the CP, NLE1 displays a large rotation towards CCDC86, coupled with the reposition of H38 and a conformational change of CCDC86 (Supplementary information, Fig. S7c, d). Due to this rotation, it could be observed from the structures that the interactions of NLE1 with the CP and other factors in the pre-rotational state are mostly gone, resulting in a significantly weakened ribosomal association (Supplementary information, Fig. S7a, b, e–h).

These analyses indicate that the first major rRNA remodeling event on the GNL2-particles is the rotation of the CP-5S rRNA as a whole (Supplementary information, Video S1). This event is triggered by conformational rearrangements of long-range tertiary interactions in the H80 region after RRS1 release. Together with the yeast pre-60S structures,[12,13,47] our data suggest that this CP remodeling process is highly conserved between yeast and human cells.

Assembly factors act as successive place holders to facilitate conformational maturation of the CP helices

After the CP rotation, the CP helices undergo further subtle maturation steps, particularly in the linker sequences connecting the rRNA helices. The general conformations of the CP in these intermediate states are characterized by a deflection relative to the main body of the pre-60S particle (Supplementary information, Fig. S5b–f). A continuous rotational conformational change of the CP is seen from state D to state E, and from state E to state G. The CP eventually assumes a mature-like conformation in state G after the incorporation of the CP-binding protein eL42. On the level of individual rRNA helices, the most outstanding conformational changes take place in the regions of H80 and H82, which become ordered in states E/F and their native helical conformations start to form (Fig. 2f, j). During these maturation steps, certain nucleotide bases switch between alternative base pair partners. For example, C4204 of H80 forms a canonical base pair with G4228 of the H81–H82 linker in states F/F′. However, this base pair is gone in states G/G′. Instead, C4204 forms a new G:C pair with G4191 within H80 (Fig. 2f, g, j, k).

These rRNA maturation steps are coupled with the dissociation and association of different assembly factors, suggesting that some factors have roles as place holders to control the timing of next assembly events. An outstanding example is the coordinated release of NSA2, which occurs in several steps. NSA2 contains three helices in the N-terminal region (H1, H2 and H3), a middle linker region (residues 73–114), and a β-barrel CTD (Supplementary information, Fig. S8e, h). From states B′ to C/D, the middle region of NSA2 loses its interaction with NLE1 and GNL2, and becomes disordered (Supplementary information, Figs. S7a, b, g and S8a, b, h), due to the repositioning of NLE1 (Supplementary information, Fig. S7c, d). From states D to F, the release of NLE1 is coupled with the dissociation of the H3 and CTD of NSA2, and only the N-terminal H1 and H2 remain bound to the pre-60S particle (Supplementary information, Fig. S8b, c, e). It is worth mentioning that the N-terminus of yeast Nsa2 was demonstrated to be essential for its association with the pre-60S particles.[58] The dissociation of the H3 and CTD of NSA2 is necessary for the binding of L10K, as they share overlapping binding sites with L10K (Supplementary information, Fig. S8e–h). After L10K associates in states F/F′, L10K takes turn to act as a place holder in the same region on the pre-60S particle. As suggested in the structures of human NMD3-particles, L10K presents a steric hindrance for the accommodation of the N-terminal domain of NMD3.[20]

Furthermore, some factors also display a chaperone function to guide the rRNA conformational maturation. Although we have not obtained an MDN1-containing structure, two sets of pre-60S structures contain SDAD1, representing the snapshots after the initial binding of SDAD1 (states D/D′) and the moment immediately before its departure (states F/F′). An inspection of the SDAD1-containing states F/F′ reveals a highly specific role of SDAD1 in chaperoning the folding of H80. The C-terminal region of SDAD1, which is disordered in states D/D′, is stabilized in states F/F′ and displays extensive interactions with the CP helices. Interestingly, the C-terminal region of SDAD1 closely circles around H80 (Fig. 3a–d), and in this conformation three base pairs in the stem region of H80 have formed (G4200:U4194, G4201:C4193 and U4202:A4192) (Fig. 2j). Probably to avoid undesired base pairs in its G-rich stem-loop region, an evolutionarily invariant phenylalanine residue of SDAD1, F651 (Fig. 3f), is inserted into the terminal loop of H80 (Fig. 3e). After SDAD1 dissociation, one non-canonical G4200:U4194 base pair is gone, replaced by a native base pair of G4191:C4204 on the other end of the stem (Fig. 2j, k). The terminal loop of H80 is functionally known as the P-loop, which interacts with the CCA end of the P-site tRNA during translation.[59,60] Of note, the C-terminal region of SDAD1 is highly conserved from yeast to mammalian species (Fig. 3f). Although not functionally tested, it is highly likely that a specific role of SDAD1 is to facilitate the conformational maturation of the P-loop as an rRNA chaperone.

Altogether, our structural data indicate that these CP-binding factors act as successive place holders to control the association timing of the next factor(s), and in this way to facilitate the maturation of the CP helices towards the native conformation.

Novel assembly factors identified in the human pre-60S structures

The structure of state G was solved at 2.8-Å resolution, enabling the identification and modeling of L10K, which is the previously unidentified protein X found in the early states of NMD3-particles.[20] L10K was first reported as a tumor-associated protein in rat three decades ago,[61] and was later shown to be a nuclear factor.[62] However, its exact function has remained elusive. L10K is recruited to the space between the CP and the P0 stalk together with TMA16 in state F (Fig. 4a–c). L10K consists of a central helix and two terminal loops (Fig. 4d, e). The N-terminal loop is located between H42 and H89 of the 28S rRNA, and interacts with both rRNA helices through a few arginine or lysine residues (Fig. 4f, g). The N-terminal portion of the middle helix is roughly parallel with the N-terminal helix of TMA16, and establishes hydrophobic interactions with both TMA16 and GTPBP4 (Fig. 4h; Supplementary information, Fig. S8c, d). The tight hydrophobic packing between the two helices of TMA16 and L10K suggests that these two factors may join together and be co-recruited to pre-60S particle at the same time. The C-terminal loop of L10K is located between GNL2 and GTPBP4, with the 20 terminal residues floating in the PTC region (Fig. 4d; Supplementary information, Fig. S8d). An important finding is that the N-terminal helices of TMA16, L10K and GTPBP4 together encircle H89 (Supplementary information, Fig. S8c, d), in a similar manner to the CTD of SDAD1 does to H80. In fact, this helical segment of H89 is isolated from the rest of the 28S rRNA and its terminal region is still flexible. Thus, this spatial segregation is probably also a means to avoid non-native folding. Homology search of L10K identified its homologs in other species, including *R. norvegicus*, *D. rerio*, *S. pombe*, *S. cerevisiae* and *C. elegans*, and many of its pre-60S-interacting residues are highly conserved (Fig. 4d). Although the yeast equivalent of L10K, YLR363W-A has not been found in yeast pre-60S structures of similar assembly stages, it should presumably work in a similar way in the assembly of yeast pre-60S particles.

In the solvent face of the pre-60S particle, another new factor was found in the map of state F′ (Supplementary information, Fig. S9a). Tentative modeling of potential candidates from our mass spectrometry data identified a previously uncharacterized protein, C11orf98 (123 residues in full length). This factor binds at a generally equivalent position of the yeast factor Alb1 (Supplementary information, Fig. S9b), but the sequence homology

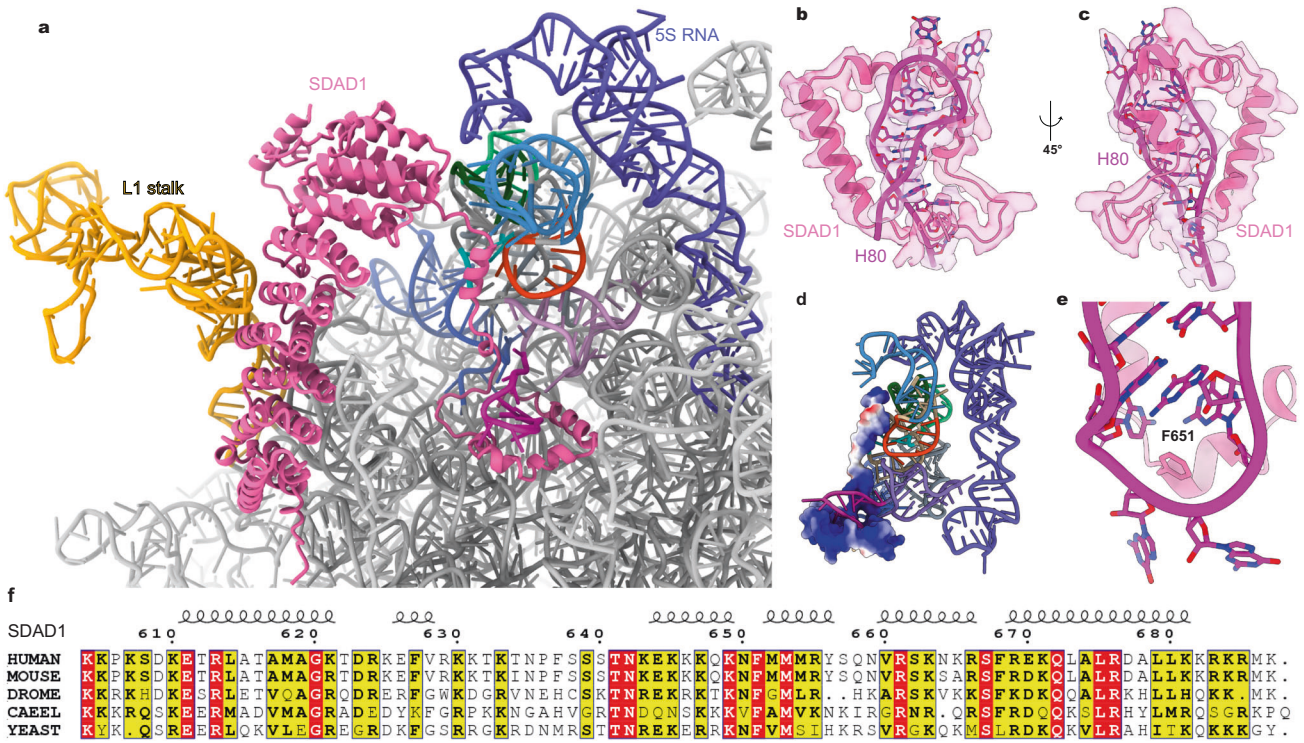

Fig. 3 The C-terminal region of SDAD1 functions in chaperoning the folding of H80. a Overview of the SDAD1-interacting rRNA components (including L1 stalk, 5S RNA and H80–H88) from state F. The rRNA helices are colored in the same scheme as in Fig. 2. **b**, **c** Two views show that SDAD1 intertwines with H80. **d** The C-terminal region of SDAD1 interacts with the CP rRNA through electrostatic interactions. The C-terminal region of SDAD1 is shown in electrostatic surface potential. H80 is tightly encircled by positively charged residues of SDAD1. **e** Zoom-in view shows that the residue F651 of SDAD1 inserts into the terminal loop of H80. **f** Sequence alignment of the C-termini of SDAD1/SDA1 from different species.

between them is very limited (only 28.8% sequence identity, Supplementary information, Fig. S9g). Structurally, a middle helix of C11orf98 (residues 67–83) aligns well with that of Alb1 (residues 90–113), and both of them interact with H19 and H24 in a similar way (Supplementary information, Fig. S9c, d). Other than this, although the N-termini of the two factors occupy equivalent positions underneath H94, H98 and ES39b, they adopt different secondary structures and show distinct atomic contacts with the pre-60S particles. Sequence alignment shows that C11orf98 lacks the Arx1-interacting motif of Alb1, and the C-terminal sequence beyond residue K83 is completely flexible (Supplementary information, Fig. S9c, g). This observation might provide an explanation for the fact that no PA2G4 (Arx1 in yeast) is observed in any of our GNL2-containing structures.

We also identified a species-specific feature for the assembly factor LLPH, which interacts with the N-terminal region of C11orf98 (Supplementary information, Fig. S9a). LLPH is highly conserved in its very N-terminus with its yeast counterpart YBL028C (Supplementary information, Fig. S9h). However, their C-terminal sequences fold differently, but are still embedded in the same rRNA pocket surrounded by H94, H98 and ES39b (Supplementary information, Fig. S9a, b, e, f).

These findings indicate that human ribosome assembly factors do possess species-specific features, different from the yeast model system, and some of these features may reflect a subtle difference in the kinetics of certain steps of the 28S/25S assembly.

ITS2 and associated factors in the human pre-60S structures

The human ITS2 is 1156 nt, significantly longer than that of the budding yeast (232 nt) (Supplementary information, Fig. S10c).[5] Despite this difference, in the ITS2-containing structures of the GNL2-particles, the resolved ITS2 region displays a similar size to the yeast one (Supplementary information, Fig. S10d), with the 3′-end of the 5.8S containing roughly 63 nucleotides (helix 1 and helix 2, Supplementary information, Fig. S10a, b). Therefore, most of the ITS2 sequences in these states might have been cleaved, suggesting that the precursor form is in the processing stages of the 8S species.[63] In our maps of the ITS2 region, except the connecting region between ITS2 and the 5.8S rRNA, ITS2 nucleotides were not resolved at atomic resolution (Supplementary information, Fig. S10d). However, two helices could be easily identified (Supplementary information, Fig. S10d). This observation is consistent with the secondary structural model of the human ITS2 sequence. Following the 3′-end of the 5.8S rRNA, there are two predicted helices, with an overall similarity to the yeast one (Supplementary information, Fig. S10a, b).

Five ITS2-associated factors were identified, including PES1 (Nop7 in yeast), NIFK (Nop15 in yeast), NOP53, RPL7L1 (Rlp7 in yeast), and RSL1D1 (Cic1 in yeast) (Fig. 5a; Supplementary information, Figs. S11, S12). Some of these factors possess species-specific sequence or structural features (Supplementary information, Figs. S11, S12). In yeast, Nop7, Nop15, Rlp7 and Cic1 remain associated with ITS2 from nucleolus to nucleus, and the replacement of Erb1 by Nop53 determines the timing of ITS2 processing,[8,11,12] because Nop53 is responsible for the recruitment of the exosome. With focused classification, we were able to trace the nearly full-length sequence of NOP53 on the pre-60S particle (Fig. 5e, f). The N-terminal region of NOP53 is inserted in the major groove of the L1 stalk (H76), and its C-terminus ends in the region next to the stem base of ES31a. Both of these two rRNA helices are part of domain V, explaining the previous observation that yeast Nop53 starts to associate with the pre-60S particles when the construction of domain V is in progress.[8] Starting from the N-terminus, NOP53 makes contacts with multiple rRNA helices and proteins, including eL36, RSL1D1, eL8, PES1, uL23, uL29, RPL7L1, eL27, eL30, eL19, eL38, rRNA helices H15, ES20a, H58, ES31a, ES31b and the 28S–5.8S duplex (Fig. 5e, f; Supplementary

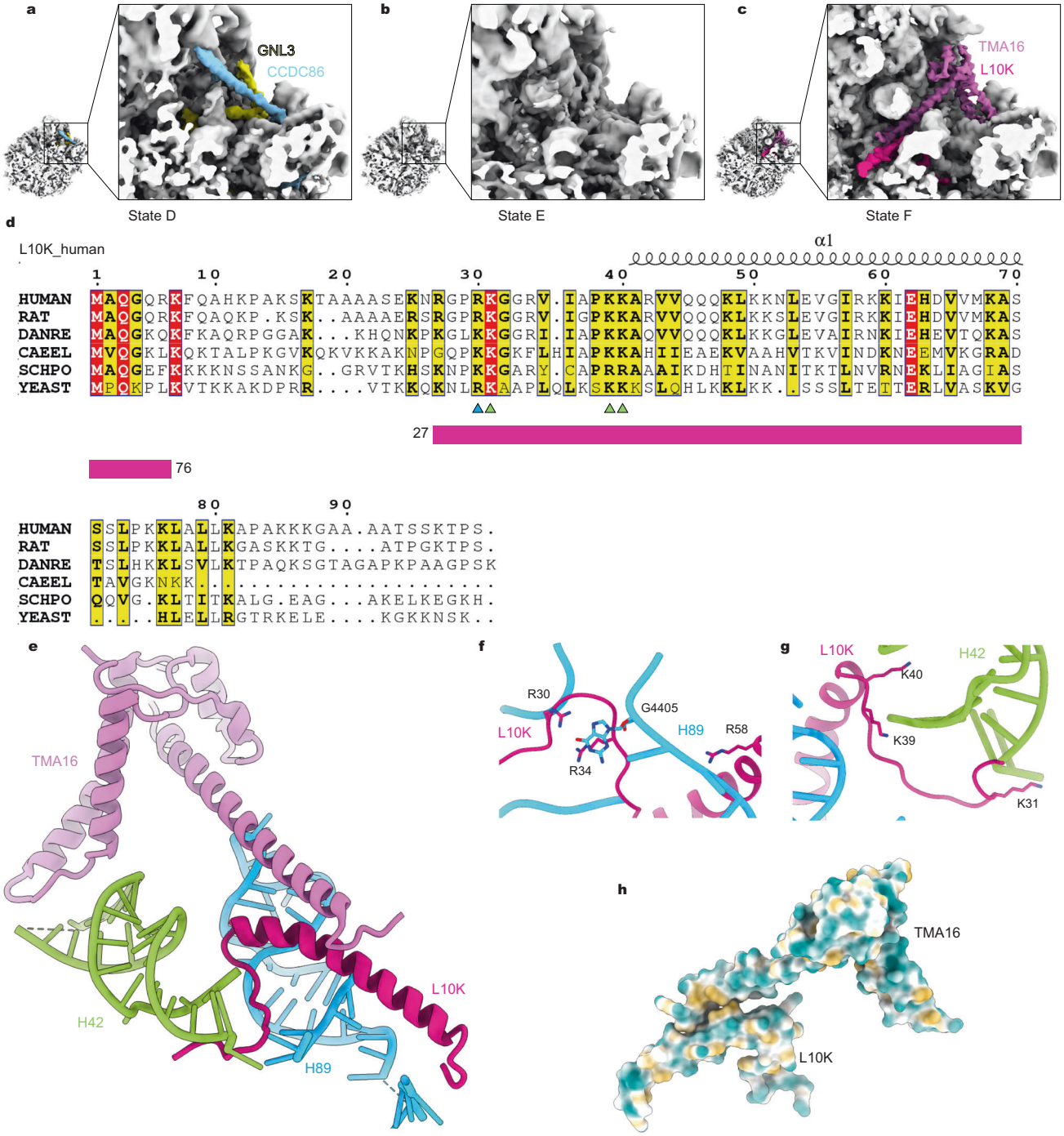

Fig. 4 Structure and interaction of human L10K in state F. a–c Cryo-EM density maps (Gaussian filtered with sDev of 1 Å) of states D, E and F, with CCDC86, GNL3, TMA16 and L10K color-coded. **d** Sequence alignment of human L10K. The resolved part (residues 27–76) is shown, and the residues that interact with rRNA are indicated in triangles (blue for H89-interacting residues and green for H42-interacting residues) under the alignment. **e** Zoom-in view shows the interacting components of L10K. TMA16 and surrounding rRNAs are shown. **f** Interactions between L10K and H89. **g** Interactions between L10K and H42. **h** The surface colored by hydrophobicity shows the hydrophobic interactions between TMA16 and L10K. Yellow means hydrophobic, and blue means hydrophilic.

information, Fig. S11d, e). It is distributed on a large surface area of the pre-60S particle, involving Domains I, II, III and V of the 28S rRNA, the 5.8S rRNA and ITS2. Of note, a helix of NOP53 is seen to be parallel with ES27 and to stabilize its conformation (Fig. 5e, f). This observation was also recently reported for yeast Nop53, which stabilizes ES27 after the release of Spb4 during late nucleolar stages of yeast pre-60S assembly.[25] These observations suggest that NOP53 could integrate conformational signals from different regions of the pre-60S particle to control the timing of the ITS2 processing.

One important observation from the human GNL2-containing pre-60S structures is that the ITS2 processing spans the lifetime of GNL2 until very late stages such as state G′. This is in contrast to the yeast structures, in which the ITS2 removal completes at earlier stages. Given the length and the complexity of human ITS2 (Supplementary information, Fig. S10a), the removal of human

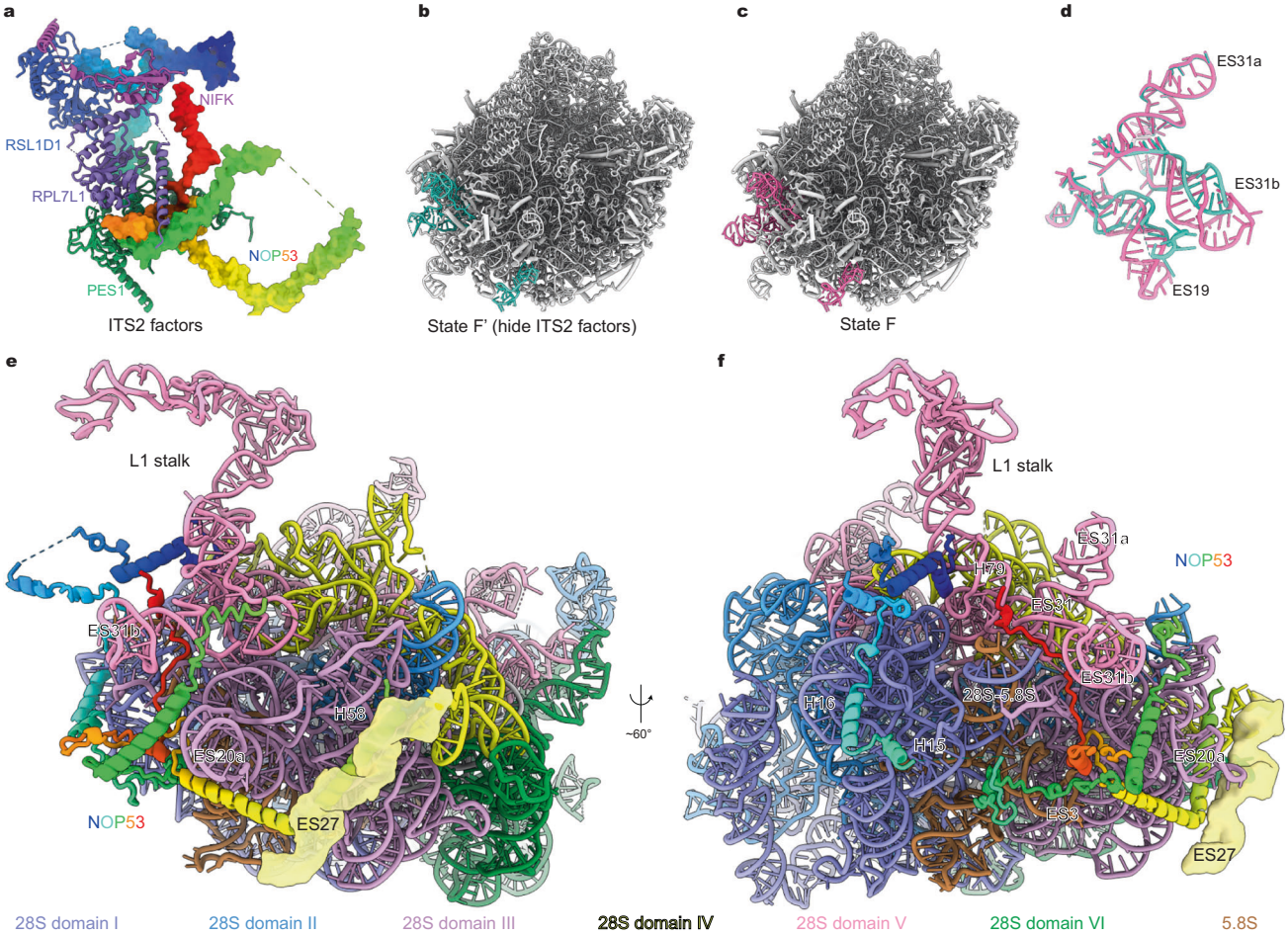

Fig. 5 Structure of the human ITS2-associated factors. a The atomic models of the ITS2-associated factors. RSL1D1, NIFK, RPL7L1 and PES1 are shown in ribbon representation, and NOP53 is shown in surface representation with rainbow color from N-terminus to C-terminus (from blue to red). **b**, **c** Atomic models of states F′ and F. The models are shown in gray, with rRNA helices next to the ITS2 foot structure colored in cyan (state F′) and hot pink (state F). For clarity, the foot structure of state F′ is omitted for display. **d** Superimposition of rRNA helices (including ES31a, ES31b and ES19) from states F and F′, showing the conformational changes of ES31b and ES19 after removal of the foot structure. **e**, **f** NOP53 interacts with multiple distant rRNA helices of the 28S rRNA in state F′. The different domains of rRNAs are colored. The unmodeled ES27 is exhibited using split density map. The two views show that NOP53 (with rainbow color and ribbon representation) is distributed on a large area of the pre-60S particle and contacts with multiple RNA helices, including H79, H16, H15, ES3, ES27, ES20a, H58 and ES31.

ITS2 could be more time-consuming. After the full processing of ITS2, significant conformational changes occur on a few rRNA helices that are next to the ITS2. A comparison of the 28S rRNA in states F and F′ (Fig. 5b, c) shows that ES19 and ES31 are the two most affected regions (Fig. 5d).

The PTC maturation is initiated in H92 and coordinated by the GTP hydrolysis on GNL2

GNL2 is a highly conserved nuclear GTPase among species (Supplementary information, Fig. S13b).[64] It is stably bound in all the states we obtained. The major difference among these states is that a gradual loss of interactions with the ribosomal components and other factors was seen at the N-terminal extension (NTE) of GNL2 (Fig. 6a–d). In states A and B′, the NTE of GNL2 interacts with H87–H88 linker, RPF2, NSA2, NLE1, and GTPBP4 (Fig. 6a; Supplementary information, Fig. S13a). After the CP rotation, GNL2 no longer interacts with H87–H88 linker (due to the CP rotation), RPF2 (due to RPF2 release) and NLE1 (due to the repositioning of NLE1), and the corresponding parts of the NTE of GNL2 become flexible in states C/D (Fig. 6b). After the subsequent NSA2 release in state F, the N-terminus of GNL2 (before residue 133) becomes completely invisible, and only the GTPBP4 interacting motifs are stabilized (Fig. 6c, d; Supplementary information, Fig. S13a).

In all the states, the G-domain and CTD of GNL2 remain largely as a rigid body, and no significant inter-domain conformational change was observed. A close inspection of the GTPase center of GNL2 reveals a major change in the ribosomal contact for the G-domain. In states A–D, G4499 of H92 is in a flipped position and the base interacts with a few residues close to the active center of GNL2, including R205, V381 and Q383. In contrast, in states F and G, H92 adopts a different conformation, with G4499 retracted from GNL2 (Fig. 6h). These structural observations agree well with recent structural studies on yeast Nog2,[28,29] which reported two similar conformations of H92, and discovered that one function of Nog2 is directly linked to Spb1, the enzyme responsible for the 2′-*O*-methylation of G2922 (the yeast counterpart of G4499 of the human 28S rRNA). The flipping of G4499 is coupled to the release of NSA2 in states F/G, and as a result, there are also drastic conformational changes of H92 (Supplementary information, Fig. S14). After this transition, H92 interacts extensively with the linker between H89 and H90, but part of the single stranded H89 (4438–4447) now becomes relatively disordered (Supplementary information, Fig. S14b, d). Based on these data, it could be

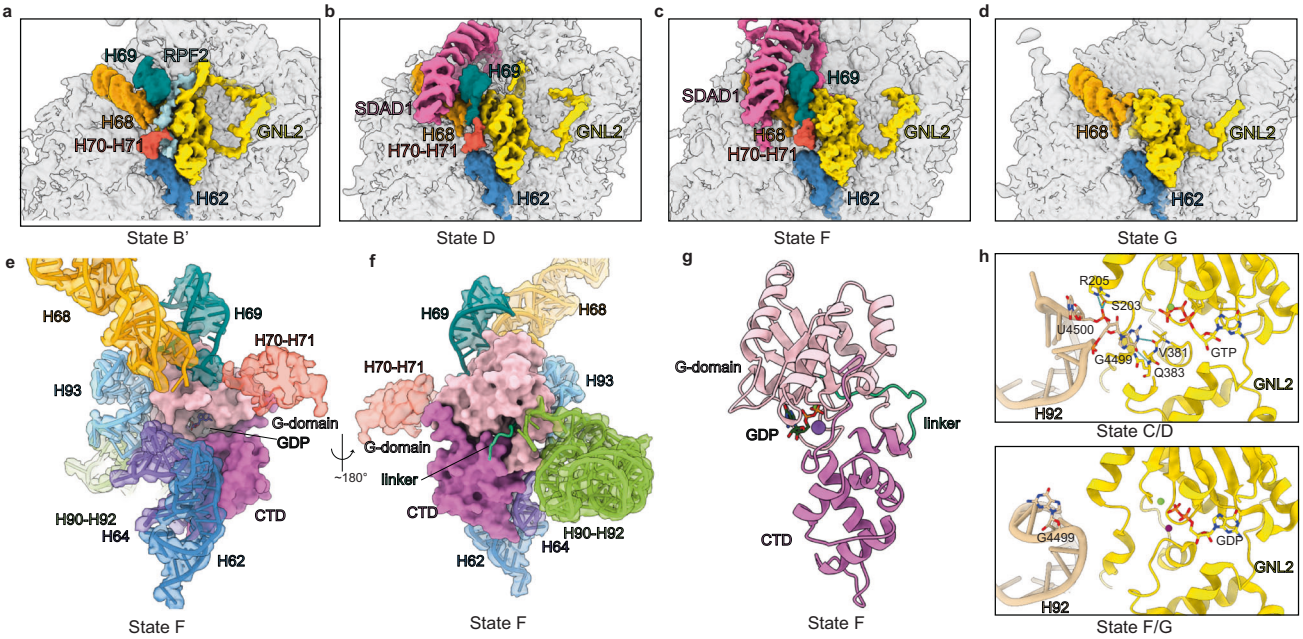

Fig. 6 Structural dynamics and interactional changes of GNL2 in GNL2-particles. a–d Cryo-EM density maps (Gaussian filtered with sDev of 1 Å in ChimeraX) show the changes in the interaction partners of GNL2 in states B′, D, F and G. The rRNA components (H68–H71 and H62) and assembly factors (GNL2, RPF2 and SDAD1) are color-coded. **e**, **f** Two 180°-rotated views show the interactions between rRNA helices and GNL2 in state F, including H62, H68, H69, H70–H71, H90–H92 and H93. The GTPase domain, CTD and the linker between them are colored pink, magenta and green, respectively. The GDP is also shown in **e**. **g** The atomic model of the G-domain and CTD of GNL2 in state F, with the GDP and the linker highlighted. The purple ball represents K^+, and the green ball represents Mg^{2+}. **h** The repositioning of H92 from states C/D (upper panel) to states F/G (lower panel). Residues of GNL2 involved in interaction with the flipped G4499 are highlighted in stick models (upper panel).

concluded that this structural remodeling on H92 is the first major step of the PTC maturation.

An important finding is that in the GTPase centers of GNL2 in states F and G, a GDP molecule could be identified (Fig. 6e–h; Supplementary information, Fig. S13e, f), and in the best-resolved state G, one K^+ and one Mg^{2+} could be clearly seen (Supplementary information, Fig. S13f). In contrast, the examination of the GTPase centers in earlier states, including A–D, suggests the presence of a γ-phosphate moiety, indicating that they are in pre-catalytic state. Although the active centers in these states are not fully resolved in atomic details as in state G, it is clear that there is no density for K^+ (Supplementary information, Fig. S13c, d). Given the recent data on the functional coupling between yeast Nog2 and Spb1,[28,29] it could be possible that Spb1/SPB1 or methylated G2922/G4499 might have a role in regulating the catalytic activity of Nog2/GNL2. Regarding the maturation of PTC, our structural data show that the release of NSA2 CTD and the disengagement of G4499 from the active center of GNL2 (Supplementary information, Fig. S14) are correlated with the post-catalytic GDP-bound state of GNL2 (Fig. 6h). Therefore, we speculate that during the PTC maturation, the release of NSA2 CTD causes the disengagement of H92 from the GNL2 active center, which subsequently activates the GTPase of GNL2. Ours and recent yeast data show that GNL2/Nog2 remains bound with the pre-60S particles after GTP hydrolysis, indicating that unlike translational GTPases, the GTP hydrolysis of GNL2/Nog2 is not employed to regulate its own affinity with the pre-60S particle. Instead, Nog2/GNL2 in the GDP-bound state continues to act as a place holder for nuclear export adaptor Nmd3/NMD3.[10,12,20,65] Given that GNL2 on the pre-60S particle is surrounded by the PTC helices (H89–H93) and central helices (H68–H71), it is reasonable to speculate that the conformational maturation of these helices could have a determining role in the release of GNL2. In support of this hypothesis, in the state pre-A structure of subsequent NMD3-particles,[20] both the G-domain and CTD of GNL2 have dissociated (only the GTPBP4-interacting helix of GNL2 remained stably bound), and H69–H71 already acquires a native-like conformation, indicating that the release of the core domains of GNL2 might be triggered by the binding of NMD3. In further support, a very recent study on yeast NPC-trapped pre-60S particles reported an assembly intermediate with simultaneous stable binding of Nog2-CTD and Nmd3-CTD.[66] In this structure, the G-domain of Nog2 has been dislodged and become flexible due to the insertion of the Nmd3-CTD. Therefore, these results together emphasize a possible checkpoint on the maturation of the central helices and a crucial role of GNL2 in monitoring the status of the central helices.

DISCUSSION

In the present work, we established a human cell line with epitope-tagged assembly factor GNL2 for the purification of endogenous nuclear pre-60S particles. To minimize any potential interference with the assembly process, we have inserted a long linker, 3*(GGGGS), between the C-terminus of GNL2 and the FLAG tag (Supplementary information, Fig. S1a). Subsequently, we obtained eleven cryo-EM structures for the pre-60S particles isolated from this cell line. These structures span the whole lifetime of GNL2 on the pre-60S particles and covers the majority of assembly events during the late nucleolar and early to middle nucleoplasmic stages. Compared to our previous study on yeast pre-60S particles isolated through epitope-tagged Nog2,[12] which provides a high-resolution pre-rotation state (Nog2-particle state 1, PDB: 3JCT), the current study provides not only the pre-rotational state B′ (an equivalent state to Nog2-particle state 1), but also fine structural details for more subsequent assembly states for human pre-60S particles. For some of these states, similar or equivalent yeast pre-60S states have been previously decribed, e.g., states D/D′ (similar to Rix1–Rea1 particle regarding the pre-60S conformation, PDB: 6YLH[8]) and state E (equivalent to the yeast late nuclear state, PDB: 6N8J[21]). Importantly, we have

also provided details on assembly states that are not previously characterized in yeast, such as states A, C/C′, F/F′ and G/G′. The rich structural information in our structures virtually provides a temporal resolution for understanding the continuous pre-60S maturation as they gradually acquire nuclear export competency. Together with our previous work on human NMD3-containing pre-60S particles,[20] our data show that the general assembly pathway in eukaryotes is highly conserved.

An important conclusion from these structural data is the presence of parallel assembly pathways (Supplementary information, Fig. S15). The eleven structures could not be assigned in a single linear assembly pathway, although they display a general hierarchy. The first is that the ITS2 processing is relatively independent of the maturation of the CP. For each major state, both the ITS2-containing and -free maps were found, even for the very late states G and G′. The second is that the release of certain assembly factors does not follow a linear order. State F still contains SDAD1 and the N-terminus of NSA2, indicating that it is between states D and G. However, state E, which lacks SDAD1, does not contain TMA16, L10K and eL42 as seen in state G (Fig. 1d), also suggesting an intermediate state earlier than state G. In particular, SDAD1 and NSA2 display two different temporal relationships. The transition from state D to state E represents a maturation pathway where SDAD1 is released in the first place before the departure of the H3 and CTD of NSA2. In contrast, the transition from states D to F defines a parallel pathway where the dissociation of the H3 and CTD of NSA2 occurs first. These two branches eventually converge at state G, which represents the latest state of GNL2-particles ready for the binding of NMD3. Therefore, these data indicate that the nucleoplamic assembly of the pre-60S particles is generally sequential but also with limited short branches.

The rich structural information on various assembly factors greatly facilitates the understanding of their roles in ribosome assembly. Consistent with the previous yeast models based on genetic, biochemical and structural data,[2–4] assembly factors are organized in temporal groups. Structural analyses indicate that the key to keep this temporal setup is their molecular roles as place holders. For instance, SPB1 contains two structured domains; the N-terminal catalytic domain is mutually exclusive with GNL2 and the middle DUF3381 is in conflict with ZNF593. For another example, NSA2, TMA16/L10K, and NMD3 are successive factors which share partially overlapped binding site in the central region. In addition, the binding of a late factor may require the release of more than one set of factors in different stages. The binding of TMA16/L10K requires the release of NLE1 and CCDC86 in the CP and the dissociation of H3 and CTD of NSA2 in the PTC region. Therefore, the complex binding interdependencies among these factors determine the directionality of the pre-60S assembly.

In terms of the rRNA conformational maturation, some assembly factors display apparent roles as rRNA chaperones, such as RRS1 and SDAD1 in maintaining the conformations of H80 in pre-mature (Fig. 2b–d) and mature-like (Fig. 3) states, respectively. These RNA chaperone functions are largely executed through physical isolation of certain rRNA fragments from their surroundings, as seen for H80 (Figs. 2b–d and 3) and H89 (Supplementary information, Fig. S8a–d). In fact, H89 is always in a relatively isolated state throughout although different factors around it have undergone association/dissociation events (Supplementary information, Fig. S8a–d). These observations indicate that certain assembly factors function to prevent potential non-native rRNA folding traps.

MATERIALS AND METHODS

Generation of a C-terminally tagged GNL2 cell line

HEK293 cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS) (BBI), at 37 °C in a 5.5% CO_2 humidified atmosphere. Genomic insertion of triple Flag-twin Strep-p2A-acGFP1 ORF at the C-terminus of GNL2 was performed by standard CRISPR/Cas9.[67] The design of the guide RNAs (gRNAs) was performed by a web-based tool (http://crispor.tefor.net/) and the selected gRNA (5′-GCAAAAGCAGTAATGTTTAAA-3′) was cloned into pX330 vector (a gift from Dr. Jiazhi Hu). A linear donor fragment carrying the knock-in sequence surrounded by ~800-bp homology arms was amplified by PCR. HEK293 cells, in 6-well plates at 80% confluency, were transfected with the gRNA vector and the donor fragment at mass ratio of 1:1 using Polyethylenimine (PEI) (Polysciences Inc., 23966–1). After two weeks, GFP-positive single cells were sorted and seeded on 96-well plates by flow cytometry (MoFlo). After two weeks, clones were screened under fluorescence microscope to remove multiclones and GFP-false-positive clones. The GFP-positive monoclones were selected for expansion. These positive monoclones were validated by PCR on extracted genomic DNA. Subsequently, the homozygous clones were validated by DNA sequencing and western blotting analysis. The validated clones were expanded to four 10-cm dishes, and adapted to suspension growth in SMM 293-TII (Sino Biological Inc., M293TII).

Purification of pre-60S particles

The cells with homozygous 3×FLAG-tagged GNL2 were collected by centrifugation and washed with PBS twice. Cell pellets were resuspended in lysis buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $Mg(OAc)_2$, supplemented with 1× cocktail protease inhibitor, 0.1% Triton X-100, 1 mM NaF and 1 mM DTT), and lysed using high pressure cell disruptor (JNBIO). The lysate was centrifugated for 60 min at 8000× *g*. Subsequently, the supernatants were collected and incubated with equilibrated ANTI-FLAG Affinity Agarose Gel (Sigma) for 2 h, followed by washing with 100 c.v. washing buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $Mg(OAc)_2$, 0.05% NP-40, 1 mM NaF and 1 mM DTT) and eluting with 1 mg/mL 3× FLAG peptide in 5 c.v. washing buffer (50 mM HEPES, pH 7.5, 150 mM KCl, 5 mM $Mg(OAc)_2$, 1 mM NaF and 1 mM DTT). The eluate was concentrated on 100-kDa molecular mass cut-off filters (Merck Millipore) for EM sample preparation. The samples were also subjected to mass spectrometry and SDS-PAGE analyses.

Mass spectrometry analysis

For protein identification, 10 µL of GNL2 samples (12 of A260 units/mL) were subjected to SDS-PAGE. Subsequently, the Coomassie-stained total aggregated proteins were cut out of the gel and destained. After dithiothreitol reduction and iodoacetamide alkylation, the proteins were digested with porcine trypsin. The resulting tryptic peptides were extracted, dried and resuspended in 10 µL of 0.1% formic acid (FA)/H_2O.

Using an Easy-nLC 1200 system, 5 µL of samples were loaded at a speed of 280 nL/min in 0.1% FA onto a trap column and eluted across a fritless analytical resolving column with a 75-min gradient. Buffer A consisted of 0.1% (v/v) FA in H_2O and Buffer B consisted of 0.1% (v/v) FA in 80% acetonitrile. The gradient was set as follows: 4%–8% B in 4 min; 8%–20% B in 46 min; 25%–35% B in 10 min; 35%–90% B in 12 min; 90% B in 3 min.

Data-dependent tandem mass spectrometry (MS/MS) analysis was performed with a Thermo Orbitrap Fusion Lumos (Thermo Fisher Scientific). Full MS and tandem mass spectra were extracted from raw files, and the tandem mass spectra were searched against a *Homo sapiens* protein database by Proteome Discoverer 2.2 software. The false discovery rate applied at the peptide and protein levels was 1%.

Cryo-EM sample preparation

The GNL2 sample was diluted to a concentration of 12 of A260 units/mL. Prior to sample freezing, Quantifoil R1.2/1.3 grids were coated with a thin layer of carbon and glow-discharged for 30 s in middle level with a plasma cleaner (PDC-32G-2, Harrick Plasma). 3.5 µL of GNL2 samples were loaded on grids and blotted for 1 s with −1 blot force using an FEI Vitrobot at 4 °C and 100% humidity.

Data collection and image processing

Cryo-samples were screened in an FEI Talo Arctica at 200 kV and transferred to a FEI Titan Krios (with Gatan K2 summit) at 300 kV for data collection. Movies were acquired using SerialEM[68] at a magnification of 92,000× (pixel size of 1.37 Å) with the defocus range varying from −1.2 µm to −1.6 µm. Each movie contained 32 frames with a dose rate of ~10 electrons/$Å^2$/s for a total exposure time of 6.4 s. Four batches of movies were collected and processed using the same procedures in RELION 4.0[69] as follows.

Movie stacks were motion-corrected and dose-weighted via MotionCorr2.[70] The CTF parameters were estimated using CTFind4 program.[71] After particle picking and extraction, 2D classification was performed to remove non-

ribosomal and noisy classes. After further 3D classification, all 60S or pre-60S classes were selected and four batches of selected particles were combined for further processing (Supplementary information, Fig. S2).

The combined ~600,000 particles were split into two halves to lower the computation burden. For each half, one round of global 3D classification was performed with the state-1 map of the yeast Nog2 particle[12] (low-pass filtered to 40 Å) as reference. Based on the structural features of the resulting maps, similar classes were pooled, including 1) 5S RNP-unrotated state (class 1, 71,235 particles), 2) 5S RNA-rotated state with mixed conformations of the L1 stalk and with weak densities of NLE1 and ITS2-containing foot structure (class 2, 335,375 particles), and 3) 5S RNA-rotated state with open L1 stalk and without NLE1 and ITS2-containing foot structure (class 3, 166,757 particles). Subsequently, these three combined classes were subjected to several rounds of mask-based 3D classification.

Regarding the 5S RNP-unrotated state, the densities of ITS2-containing foot structure were not solid, indicating the presence of structural heterogeneity. A mask-based 3D classification on this region generated two major classes, with or without ITS2-containing foot structure (hereafter abbreviated as "foot + " and "foot–"). Subsequently, both "foot + " and "foot–" classes were further subjected to another round of mask-based 3D classification on NLE1 region (including the 5S RNP) due to structural heterogeneity in this region. After that, low-resolution and post-rotation classes were discarded, resulting in two subsets of particles for high-resolution refinement. After CTF refinement in RELION 4.0, the two maps were further improved, state A (3.3 Å) and state B′ (3.5 Å).

Regarding the class 2, three rounds of mask-based 3D classifications on SDAD1 region (including the L1 stalk and SDAD1), NLE1 region (including NLE1 and the surrounding 5S RNP) and ITS2 region were performed separately. With the resolution improvement, more structural heterogeneities were observed and several more rounds of mask-based 3D classifications on these three regions were performed. As a result, eight distinct subclasses were obtained. Among these subclasses, six maps were further improved by CTF refinement, including state C′ (4.3 Å), state D′ and state D (3.2 Å), state F′ and state F (3.0 Å) and state G′ (3.2 Å).

The remaining two subclasses were similar to the class 3, and they were combined with the class 3 for further analysis. As a result, three states with distinct features were obtained, state C (3.2 Å), state E (3.3 Å) and state G (2.8 Å).

For map analysis, all the maps were sharpened using DeepEMhancer[72] and post-processing options in RELION 4.0.

Atomic model building and refinement

For modeling of the assembly factors, initial templates of TMA16, GTPBP4, eIF6, RLP24, ZNF593 and LLPH were from the model of state pre-A pre-60S NMD3-particles (PDB: 6LSS).[20] For each of these factors, the template was fitted into the maps by rigid-body fitting in UCSF Chimera,[73] followed by manual rebuilding in COOT.[74] The starting model of GNL2, GNL3, NLE1, CCDC86, NSA2, MRTO4, RPF2, RRS1, RSL1D1, RPL7L1, PES1 and NOP53 were derived from AlphaFold protein structure database (https://alphafold.ebi.ac.uk).[46] The predicted 3D models were docked into the density maps by rigid-body fitting, followed by extensive manual rebuilding in COOT.

Most of the assembly factors were well resolved, which enabled atomic modeling for majority of their sequences. For many factors, AlphaFold-predicted models were highly similar to the final refined models in the visible portions of the assembly factors. For modeling of L10K and C11orf98, tentative modeling of all possible candidates included in mass spectrometry data was attempted to be built into the density map until the correct protein was identified. Specifically, a poly-alanine model was first built manually, the AlphaFold-predicated models were first screened manually by secondary structural features and then by density fitting and side-chain density matching.

For modeling of the ribosomal components, starting models of ribosomal proteins and rRNAs were from the published model of the human state pre-A NMD3-particles (PDB: 6LSS[20]) and the yeast pre-60S structures (PDB: 3JCT[12] and 6YLH[8]). The model of the 5S rRNA was fitted in the maps of states A/B′ in Chimera and manually adjusted in COOT. Domain V of the 28S rRNA displays large conformational differences among these classes, and certain regions are highly flexible in different maps. These flexible regions were removed from the initial model and not modeled.

For each assembly state, multiple rounds of model refinement using real-space refinement in PHENIX[75] (with secondary structure, geometry and base pair constraints applied) and manual adjustment in COOT were applied. Models were evaluated (Supplementary information, Table S2) using Molprobity.[76] The maps and models were visualized in UCSF Chimera and ChimeraX[77] for analysis. The figures were generated using ChimeraX.

DATA AVAILABILITY

The cryo-EM maps and atomic coordinates of the states A, B′, C′, C, D′, D, E, F′, F, G′ and G have been deposited in the EMDB and PDB databases with accession codes EMD-35672, EMD-35673, EMD-35649, EMD-35639, EMD-35651, EMD-35599, EMD-35375, EMD-35597, EMD-35371, EMD-35596, EMD-35370 and PDB 8IR1, 8IR3, 8IPX, 8IPD, 8IPY, 8INK, 8IE3, 8INF, 8IDY, 8INE, 8IDT, respectively.

REFERENCES

1. Bohnsack, K. E. & Bohnsack, M. T. Uncovering the assembly pathway of human ribosomes and its emerging links to disease. *EMBO J.* **38**, e100278 (2019).
2. Woolford, J. L. Jr & Baserga, S. J. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* **195**, 643–681 (2013).
3. Bassler, J. & Hurt, E. Eukaryotic ribosome assembly. *Annu. Rev. Biochem.* **88**, 281–306 (2019).
4. Klinge, S. & Woolford, J. L. Jr Ribosome assembly coming into focus. *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2019).
5. Mullineux, S. T. & Lafontaine, D. L. Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie* **94**, 1521–1532 (2012).
6. 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).
7. Pena, C., Hurt, E. & Panse, V. G. Eukaryotic ribosome assembly, transport and quality control. *Nat. Struct. Mol. Biol.* **24**, 689–699 (2017).
8. Kater, L. et al. Construction of the central protuberance and L1 stalk during 60S subunit biogenesis. *Mol. Cell* **79**, 615–628.e5 (2020).
9. Sanghai, Z. A. et al. Modular assembly of the nucleolar pre-60S ribosomal subunit. *Nature* **556**, 126–129 (2018).
10. Ma, C. et al. Structural snapshot of cytoplasmic pre-60S ribosomal particles bound by Nmd3, Lsg1, Tif6 and Reh1. *Nat. Struct. Mol. Biol.* **24**, 214–220 (2017).
11. Kater, L. et al. Visualizing the assembly pathway of nucleolar pre-60S ribosomes. *Cell* **171**, 1599–1610.e14 (2017).
12. Wu, S. et al. Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature* **534**, 133–137 (2016).
13. Barrio-Garcia, C. et al. Architecture of the Rix1-Rea1 checkpoint machinery during pre-60S-ribosome remodeling. *Nat. Struct. Mol. Biol.* **23**, 37–44 (2016).
14. Cheng, J. D. et al. 90S pre-ribosome transformation into the primordial 40S subunit. *Science* **369**, 1470–1476 (2020).
15. Kornprobst, M. et al. Architecture of the 90S pre-ribosome: a structural view on the birth of the eukaryotic ribosome. *Cell* **166**, 380–393 (2016).
16. Barandun, J. et al. The complete structure of the small-subunit processome. *Nat. Struct. Mol. Biol.* **24**, 944–953 (2017).
17. Sun, Q. et al. Molecular architecture of the 90S small subunit pre-ribosome. *Elife* **6**, e22086 (2017).
18. Du, Y. et al. Cryo-EM structure of 90S small ribosomal subunit precursors in transition states. *Science* **369**, 1477–1481 (2020).
19. Ameismeier, M., Cheng, J., Berninghausen, O. & Beckmann, R. Visualizing late states of human 40S ribosomal subunit maturation. *Nature* **558**, 249–253 (2018).
20. Liang, X. et al. Structural snapshots of human pre-60S ribosomal particles before and after nuclear export. *Nat. Commun.* **11**, 3542 (2020).
21. Zhou, Y., Musalgaonkar, S., Johnson, A. W. & Taylor, D. W. Tightly-orchestrated rearrangements govern catalytic center assembly of the ribosome. *Nat. Commun.* **10**, 958 (2019).
22. Thoms, M. et al. Suppressor mutations in Rpf2-Rrs1 or Rpl5 bypass the Cgr1 function for pre-ribosomal 5S RNP-rotation. *Nat. Commun.* **9**, 4094 (2018).
23. Wilson, D. M. et al. Structural insights into assembly of the ribosomal nascent polypeptide exit tunnel. *Nat. Commun.* **11**, 5111 (2020).
24. Micic, J. et al. Coupling of 5S RNP rotation with maturation of functional centers during large ribosomal subunit assembly. *Nat. Commun.* **11**, 3751 (2020).
25. Cruz, V. E. et al. Sequence-specific remodeling of a topologically complex RNP substrate by Spb4. *Nat. Struct. Mol. Biol.* **29**, 1228–1238 (2022).
26. Prattes, M. et al. Visualizing maturation factor extraction from the nascent ribosome by the AAA-ATPase Drg1. *Nat. Struct. Mol. Biol.* **29**, 942–953 (2022).
27. Ugolini, I. et al. Chromatin localization of nucleophosmin organizes ribosome biogenesis. *Mol. Cell* **82**, 4443–4457.e9 (2022).
28. Yelland, J. N., Bravo, J. P. K., Black, J. J., Taylor, D. W. & Johnson, A. W. A single 2'-O-methylation of ribosomal RNA gates assembly of a functional ribosome. *Nat. Struct. Mol. Biol.* **30**, 91–98 (2023).
29. Sekulski, K., Cruz, V. E., Weirich, C. S. & Erzberger, J. P. rRNA methylation by Spb1 regulates the GTPase activity of Nog2 during 60S ribosomal subunit assembly. *Nat. Commun.* **14**, 1207 (2023).
30. 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).

31. Dorner, K. et al. Genome-wide RNAi screen identifies novel players in human 60S subunit biogenesis including key enzymes of polyamine metabolism. *Nucleic Acids Res.* **50**, 2872–2888 (2022).
32. 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).
33. Farley-Barnes, K. I. et al. Diverse regulators of human ribosome biogenesis discovered by changes in nucleolar number. *Cell Rep.* **22**, 1923–1934 (2018).
34. Wild, T. et al. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol.* **8**, e1000522 (2010).
35. Ogawa, L. M. et al. Increased numbers of nucleoli in a genome-wide RNAi screen reveal proteins that link the cell cycle to RNA polymerase I transcription. *Mol. Biol. Cell* **32**, 956–973 (2021).
36. Ni, C. & Buszczak, M. The homeostatic regulation of ribosome biogenesis. *Semin. Cell Dev. Biol.* **136**, 13–26 (2023).
37. Pelletier, J., Thomas, G. & Volarević, S. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat. Rev. Cancer* **18**, 51–63 (2018).
38. Iadevaia, V., Liu, R. & Proud, C. G. mTORC1 signaling controls multiple steps in ribosome biogenesis. *Semin. Cell Dev. Biol.* **36**, 113–120 (2014).
39. Liu, Y., Deisenroth, C. & Zhang, Y. RP-MDM2-p53 pathway: linking ribosomal biogenesis and tumor surveillance. *Trends Cancer* **2**, 191–204 (2016).
40. van Riggelen, J., Yetil, A. & Felsher, D. W. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev. Cancer* **10**, 301–309 (2010).
41. Aspesi, A. & Ellis, S. R. Rare ribosomopathies: insights into mechanisms of cancer. *Nat. Rev. Cancer* **19**, 228–238 (2019).
42. Cao, P. et al. Genomic gain of RRS1 promotes hepatocellular carcinoma through reducing the RPL11-MDM2-p53 signaling. *Sci. Adv.* **7**, eabf4304 (2021).
43. Ebright, R. Y. et al. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. *Science* **367**, 1468–1473 (2020).
44. Singh, S., Vanden Broeck, A., Miller, L., Chaker-Margot, M. & Klinge, S. Nucleolar maturation of the human small subunit processome. *Science* **373**, eabj5338 (2021).
45. Ameismeier, M. et al. Structural basis for the final steps of human 40S ribosome maturation. *Nature* **587**, 683–687 (2020).
46. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
47. Leidig, C. et al. 60S ribosome biogenesis requires rotation of the 5S ribonucleoprotein particle. *Nat. Commun.* **5**, 3491 (2014).
48. Bassler, J. et al. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Mol. Cell* **38**, 712–721 (2010).
49. Ulbrich, C. et al. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell* **138**, 911–922 (2009).
50. Raman, N., Weir, E. & Muller, S. The AAA ATPase MDN1 acts as a SUMO-targeted regulator in mammalian pre-ribosome remodeling. *Mol. Cell* **64**, 607–615 (2016).
51. Finkbeiner, E., Haindl, M. & Muller, S. The SUMO system controls nucleolar partitioning of a novel mammalian ribosome biogenesis complex. *EMBO J.* **30**, 1067–1078 (2011).
52. Klingauf-Nerurkar, P. et al. The GTPase Nog1 co-ordinates the assembly, maturation and quality control of distant ribosomal functional centers. *Elife* **9**, e52474 (2020).
53. Kressler, D., Rojo, M., Linder, P. & Cruz, J. Spb1p is a putative methyltransferase required for 60S ribosomal subunit biogenesis in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **27**, 4598–4608 (1999).
54. Lapeyre, B. & Purushothaman, S. K. Spb1p-directed formation of Gm(2922) in the ribosome catalytic center occurs at a late processing stage. *Mol. Cell* **16**, 663–669 (2004).
55. Krogh, N. et al. Profiling of 2'-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res.* **44**, 7884–7895 (2016).
56. Simabuco, F. M., Morello, L. G., Aragao, A. Z., Paes Leme, A. F. & Zanchin, N. I. Proteomic characterization of the human FTSJ3 preribosomal complexes. *J. Proteom. Res.* **11**, 3112–3126 (2012).
57. Bassler, J. et al. The conserved Bud20 zinc finger protein is a new component of the ribosomal 60S subunit export machinery. *Mol. Cell. Biol.* **32**, 4898–4912 (2012).
58. Paternoga, H. et al. Mutational analysis of the Nsa2 N-terminus reveals its essential role in ribosomal 60S subunit assembly. *Int. J. Mol. Sci.* **21**, 9108 (2020).
59. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
60. Moazed, D. & Noller, H. F. Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E sites. *Cell* **57**, 585–597 (1989).
61. Rabbani, S. A., Yasuda, T., Bennett, H. P., Hendy, G. N. & Banville, D. Nucleotide and (derived) amino acid sequence of a novel peptide from a rat (hypercalcemic) Leydig cell tumor. *Biochim. Biophys. Acta* **1171**, 229–230 (1992).
62. 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–e00317 (2018).
63. Pirouz, M., Munafo, M., Ebrahimi, A. G., Choe, J. & Gregory, R. I. Exonuclease requirements for mammalian ribosomal RNA biogenesis and surveillance. *Nat. Struct. Mol. Biol.* **26**, 490–500 (2019).
64. Saveanu, C. et al. Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps. *EMBO J.* **20**, 6475–6484 (2001).
65. Matsuo, Y. et al. Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. *Nature* **505**, 112–116 (2014).
66. Li, Z. et al. Nuclear export of pre-60S particles through the nuclear pore complex. *Nature* **618**, 411–418 (2023).
67. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
68. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
69. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).
70. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
71. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
72. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
73. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
74. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
75. Adams, P. D. et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
76. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
77. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

ACKNOWLEDGEMENTS

We thank Drs. W. Wei, X. Ji and J. Hu at Peking University for their help with CRISPR Knock-In experiments. We also thank the Core Facilities at the School of Life Sciences, Peking University for help with negative staining EM; the Cryo-EM Platform of Peking University for help with data collection; the High-performance Computing Platform of Peking University for help with computation; the National Center for Protein Sciences at Peking University for assistance with flow cytometry and mass spectrometry. The work was supported by the National Natural Science Foundation of China (32230051 to N.G., 31922036 to N.L.), the National Key R&D Program of China (2019YFA0508904 to N.G.), and the Qidong-SLS Innovation Fund to N.G.

AUTHOR CONTRIBUTIONS

Y.Z. and S.L. established the CRISPR-KI cell line. Y.Z. prepared the pre-ribosomal samples, and collected the cryo-EM data. Y.Z. and X.L. performed EM analysis (with the help of Y.L., Y.C., N.L. and C.M.). N.G., Y.Z. and X.L. performed cryo-EM model building and wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41422-023-00853-9.

Correspondence and requests for materials should be addressed to Ning Gao.

Reprints and permission information is available at http://www.nature.com/reprints