REPORT

STRUCTURAL BIOLOGY

Cryo-EM structure of 90*S* small ribosomal subunit precursors in transition states

Yifei Du[1]*, Weidong An[1]*, Xing Zhu[1], Qi Sun[1,2], Jia Qi[1,2,3], Keqiong Ye[1,4]†

The 90*S* preribosome is a large, early assembly intermediate of small ribosomal subunits that undergoes structural changes to give a pre-40*S* ribosome. Here, we gained insight into this transition by determining cryo–electron microscopy structures of *Saccharomyces cerevisiae* intermediates in the path from the 90*S* to the pre-40*S*. The full transition is blocked by deletion of RNA helicase Dhr1. A series of structural snapshots revealed that the excised 5′ external transcribed spacer (5′ ETS) is degraded within 90*S*, driving stepwise disassembly of assembly factors and ribosome maturation. The nuclear exosome, an RNA degradation machine, docks on the 90*S* through helicase Mtr4 and is primed to digest the 3′ end of the 5′ ETS. The structures resolved between 3.2- and 8.6-angstrom resolution reveal key intermediates and the critical role of 5′ ETS degradation in 90*S* progression.

Ribosomes, large RNA-protein complexes composed of a 40*S* small subunit (SSU) and a 60*S* large subunit, are responsible for protein synthesis. Assembly of eukaryotic ribosomes is a complicated and dynamic process that involves many preribosomal assembly intermediates (*1*–*3*). More than 200 trans-acting assembly factors (AFs) and many small nucleolar RNAs (snoRNAs) function in modification and processing of ribosomal RNA (rRNA) and assembly of ribosomal proteins (RPs).

The SSU is composed of an 18*S* rRNA and 33 RPs in the yeast *Saccharomyces cerevisiae*. The 90*S* preribosome (the SSU processome) is the earliest assembly intermediate of the SSU (*4*, *5*) and forms cotranscriptionally on the 5′ region of a nascent pre-rRNA transcript containing the 5′ external transcribed spacer (5′ ETS), the 18*S* rRNA sequence, and the internal transcribed spacer 1 (ITS1) (*6*–*8*). The 5′ ETS initiates the assembly of 90*S* by recruiting a large number of AFs, including the large UTPA, UTPB, and U3 snoRNP complexes. After the assembly of 18*S* rRNA and release of labile factors, the 90*S* is compacted into a large structure comprising pre-rRNA, U3 snoRNA, ~50 AFs, and ~20 RPs. After the A0 and A1 sites of the 5′ ETS and the A2 site of ITS1 are sequentially cleaved, the 90*S* is converted into a pre-40*S* particle that is rapidly exported to the cytoplasm for final maturation (Fig. 1A).

[1]Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. [2]National Institute of Biological Sciences, Beijing 102206, China. [3]Department of Biochemistry and Molecular Biology, College of Life Sciences, Beijing Normal University, Beijing 100875, China. [4]University of Chinese Academy of Sciences, Beijing 100049, China.
*These authors contributed equally to this work.
†Corresponding author. Email: yekeqiong@ibp.ac.cn

Several cryo–electron microscopy (cryo-EM) structures have been determined for the 90*S* of *S. cerevisiae* and *Chaetomium thermophilum* before A1 site cleavage (*9*–*15*). In the 90*S* structure, the 5′, central, 3′ major, and 3′ minor domains of 18*S* rRNA are partially assembled and not yet packed. By contrast, the pre-40*S* that contains a 20*S* pre-rRNA without the 5′ ETS, 31 RPs, and a few AFs has established the global architecture of mature SSUs (*16*–*18*). No intermediate states between 90*S* and pre-40*S* have been structurally analyzed, so how the transition occurs is not known. The cleaved 5′ ETS fragment was proposed to be released together with associated factors during the transition (*9*). By cryo-EM analysis of 90*S* transition states, we show that the 5′ ETS is degraded within the 90*S*, which further triggers AF release and maturation of the ribosome.

Several RNA helicases are required for ribosome assembly, potentially driving key structural reorganization of preribosomes (*19*). To stall the assembly of the SSU, we depleted the essential RNA helicase Dhr1, which acts after A0 cleavage (*20*, *21*), and purified the SSU preribosomes from *S. cerevisiae* by Enp1, which is present in both 90*S* and pre-40*S* ribosomes (*22*) and should capture all transition states. Three states were previously determined at 8- to 9-Å resolution by cryo-EM. State 1, representing a fully assembled 90*S*, was described previously (*13*). To improve the structures, 18,028 images were collected in a Titan Krios 300-kV electron microscope with a K2 Summit camera (fig. S1, A and B, and table S1). Three-dimensional (3D) classification of 382,298 particles yielded seven distinct states, designated as A to E, A1, and C1 (Fig. 1B and fig. S1C). Local density was further improved with focused classification and refinement (fig. S1D). States A and D were reconstructed at 3.2- and 3.8-Å resolution, respectively, and the other states were determined to 4- to 8.6-Å resolution (figs. S2 to S4 and data S1).

In all of these structures, the A1 site of pre-rRNA is intact and the A0 site is not visible. The A0 site should have been mostly processed in Dhr1-depleted 90*S* (*13*). States A to E can be ordered chronologically because structural changes occurring in one state are preserved in its following states, and the 5′ ETS density is increasingly reduced in these states. State A1 is similar to state A but adopts an open conformation at the central domain, as shown in the 90*S* from starved yeast cells and state a/b of *C. thermophilum* 90*S* (*10*, *14*). State C1 completely lacks the central domain and its surrounding structures and is likely a degradation product of state C.

State A resembles the previously described state 1 and represents a fully assembled 90*S* (*13*). The 5′ ETS and its associated factors constitute the bulk of the structural core. The 5′ and central domains of 18*S* rRNA with bound RPs and AFs form two bulky projections on the top, and the 3′ major domain is bound at the center of the structural core (Fig. 1B and fig. S4).

The 5′ ETS RNA folds into 10 helices (H1 to H10) and forms two intermolecular helices (Ha and Hb) with U3 snoRNA. The density of the 5′ ETS is gradually and unevenly reduced in states B to E (Figs. 2, A to F, and 3A). The 5′ ETS should be degraded, rather than adopting a relaxed conformation, because its level was greatly reduced in Dhr1-depleted 90*S* compared with wild-type 90*S* (*13*). The H2 to H4 and H6 and H7 regions show almost no density in state D, whereas the H1, Ha to H5, and Hb to H9 regions still retain 60 to 90% of their original densities and appear to be more resistant to degradation.

RNA helices H2, H3, H4, and H7 are located at an opening near the base of the 90*S* structure (Fig. 2C). Accompanying their degradation, the nearby regions also become disordered. Utp30, Rrt14, Rps18, an α-helix of Sas10 that binds Utp30, and an α-helix of Utp11 that binds H7 disappear altogether in state B (Fig. 2D). In state E, the base of 90*S* shows greatly reduced densities, suggesting that the UTPA complex and other nearby AFs begin to dissociate (Fig. 2F). Therefore, degradation of the 5′ ETS destabilizes and promotes disassembly of the base of 90*S*.

The mature SSU structure can be divided into hallmark features of body, platform, and head, which are made up mainly by the 5′, central, and 3′ major domains of 18*S* rRNA, respectively (*23*, *24*). The global architecture of SSU is further determined by quaternary interactions between individual domains (Fig. 1A and fig. S5A). For example, the 3′ minor domain comprises helices h44 and h45 that interact with the platform and body, respectively. Extension segment 6 (ES6) in the central domain mediates two long-range interactions

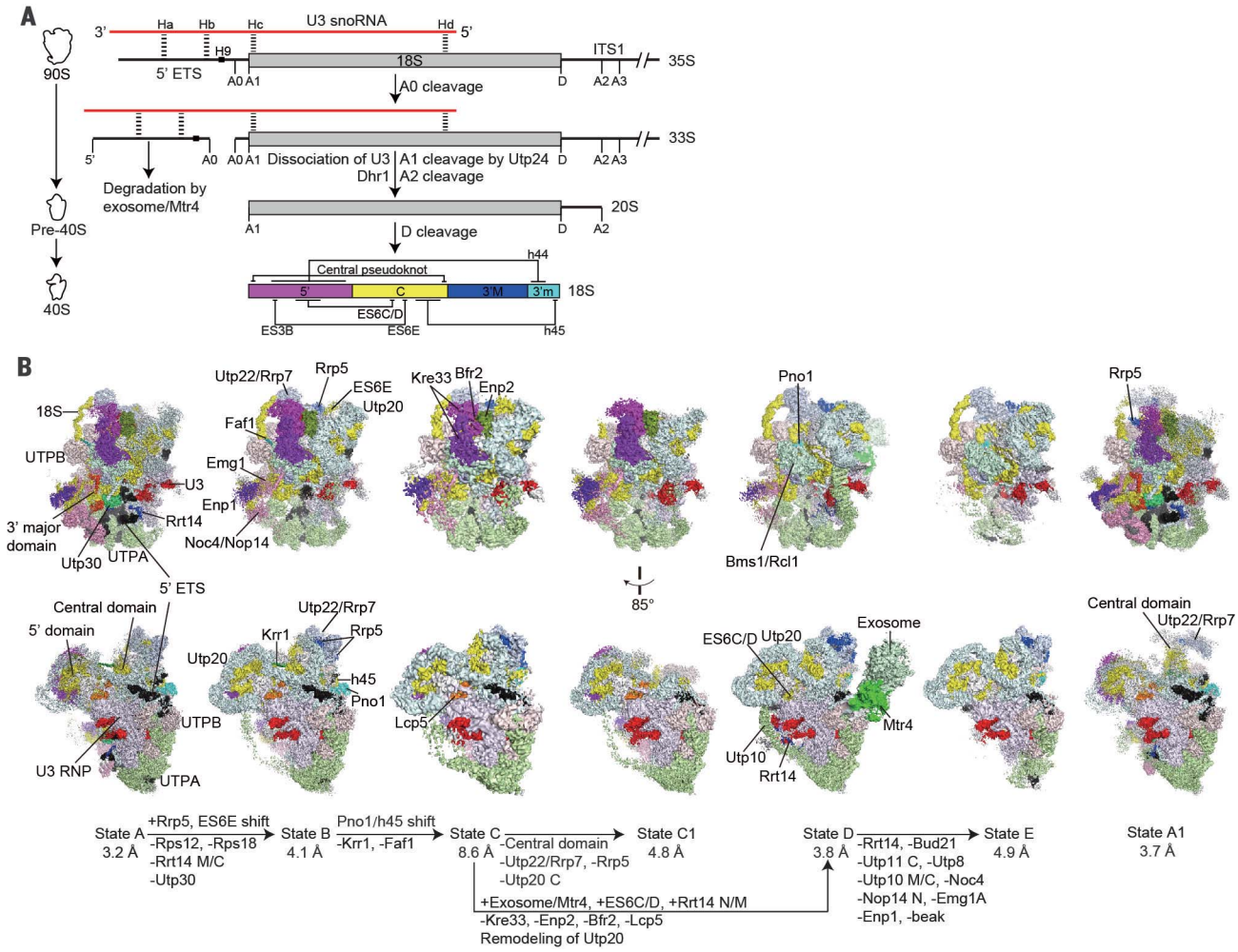

Fig. 1. Cryo-EM structures of Dhr1-depleted 90*S* preribosomes. (**A**) Schematic of pre-rRNA processing. Processing sites and interactions between the U3 snoRNA and the 5′ ETS are marked. Long-range interactions among 5′, central (C), 3′ major (3′M), and 3′ minor (3′m) domains of mature 18*S* rRNA are indicated. (**B**) Density maps of seven states are shown in two views. 18*S*, 5′ ETS, and U3 RNAs are shown in yellow, black, and red, respectively. The large structural modules UTPA, UTPB, U3 snoRNP, Noc4/Nop14/Emg1, Mpp10/Imp3/Imp4, Bms1/Rcl1, and Utp22/Rrp7 are shown in light colors. Proteins with major structural changes are shown in bright colors. Other AFs and RPs are shown in light cyan and white. Unmodeled densities are shown in white, and modeled but unassigned densities are shown in gray. The color theme is generally applied in the other figures. State D is a composite map with the exosome reconstructed from exosome-enriched particles. Compositional and conformational changes are indicated for the connected states.

with the 5′ domain. These interdomain interactions are absent in state A but start to emerge in the subsequent states.

In states A and B, the ribosome assembly factor Krr1 binds to the central domain at the same position as h45 does in the mature SSU, preventing its assembly. Concomitant with removal of helix H6 of the 5′ ETS, Krr1 and its binding protein Faf1, which bind to H6, are released in state C, allowing for integration of h45 (Fig. 3, A to C, and fig. S7A). In addition, Utp7, Mpp10, Utp12, Utp21, and Imp4, which are located near H6, become partially unstructured in state C, and Imp3, in complex with an α-helix of Mpp10, rotates by ~15° (Fig. 3B). Moreover, h45 and its binding protein Pno1, anchored on the UTPB complex, rotate as a whole unit by ~60°, which brings h45 closer to its binding site in the central domain. The movement of h45 also depends on prior removal of H6 because the new position of h45 was previously occupied by H6 and the H5 to H6 linker (Fig. 3A). Removal of H6 thus appears to drive maturation of the central domain.

Helix E of ES6 (ES6E) in the central domain of 18*S* rRNA makes a long-range interaction with helix B of ES3 (ES3B) in the 5′ domain in mature SSU (Fig. 1A and fig. S5A). ES6E initially binds Utp22 in state A and is resituated to bind Utp20 in state B (Fig. 3C and fig. S6, A to C). The structural remodeling brings ES6E closer to, but not yet in contact with, the 5′ domain (Fig. 3E). The original position of ES6E on Utp22 is occupied by the TPR domain of Rrp5, suggesting that Rrp5 drives the remodeling of ES6E. Rrp5 additionally contains 12 S1 RNA-binding domains, some of which can be visualized (fig. S6, B to D). The TPR domain of Rrp5 was mistakenly modeled in previous state 1 because of insufficient classification.

In state D, Kre33, Enp2, Lcp5, and a short segment of Brf2, which are bound to the apical part of the 5′ domain in earlier states, are missing (Figs. 2, D and E, and 3, C and D). Release of these AFs is probably facilitated by prior release of Krr1 in state C because the extended C-terminal tail of Krr1 binds Enp2 and Bfr2 (Fig. 3C). Utp20 is an extended superhelical structure that wraps the 5′ and central domains of 18*S* rRNA with its N- and C-terminal half, respectively (Fig. 3D and fig. S7B). From state C to D, the N-half moves toward the central domain as a rigid body, whereas the C-half is greatly compressed (Fig. 3D and fig. S7, C to E). As a result of release of

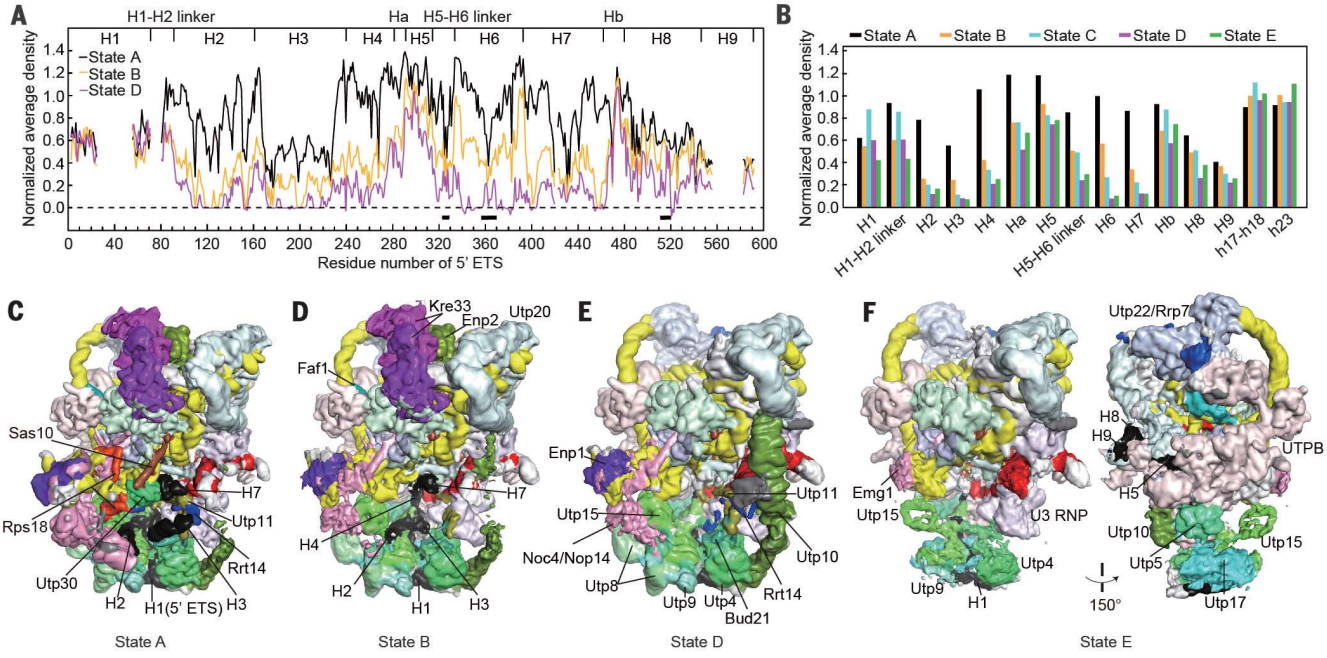

Fig. 2. Degradation of 5′ ETS in 90S. (**A**) Normalized average density for the nucleotide of the 5′ ETS. Density was normalized against a segment of 18S rRNA that is well ordered in all states. The secondary structures of the 5′ ETS are shown on the top. Three short regions marked by lines are occupied by h45 and Utp20 in state D and show an artificial increase of density. (**B**) Normalized average density for RNA secondary structures. Helices 17 to 18 and 23 are from the 18S rRNA. (**C** to **F**) Density maps of states A, B, D, and E illustrating structural changes at the base of 90S low-pass filtered to 10 Å. Two views are shown for state E.

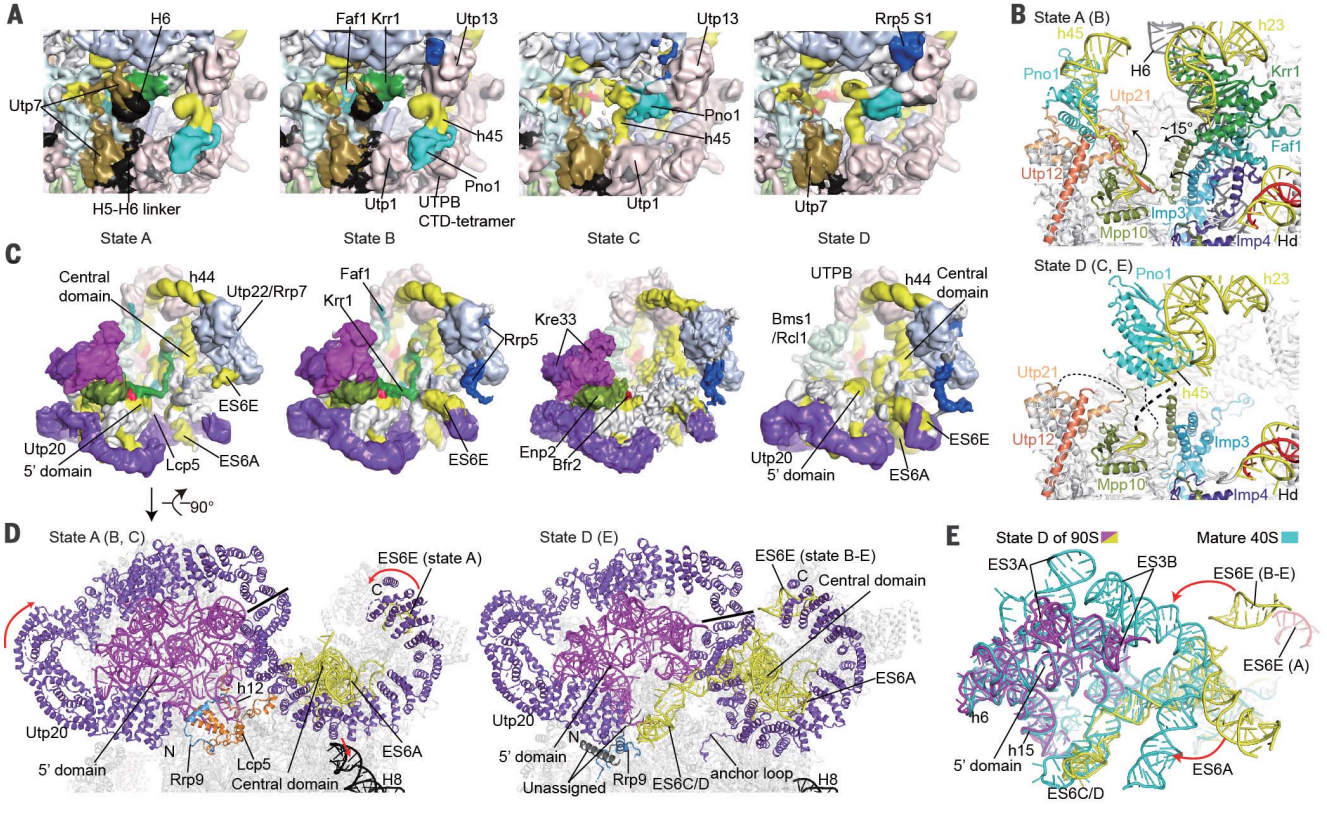

Fig. 3. Maturation of the 5′ and central domains. (**A**) Density maps of states A to D illustrating structural changes around the central domain. (**B**) Models of state A and D around the central domain. Some obstructing structures are omitted for clarity. Other states with similar features are noted in parentheses. (**C**) Density maps of states A to D illustrating structural changes around the 5′ and central domains. The maps shown in (A) and (C) were low-pass filtered to 10 Å. (**D**) Models of state A and D at the 5′ and central domains. This view is related to (C) by a 90° rotation. The N- and C-halves of Utp20 are divided by a line. (**E**) Structures of 18S rRNA in state D of 90S (5′ domain, magenta; central domain, yellow) and a mature 40S ribosome (cyan) are aligned by the apical part of the 5′ domain. ES6E of state A (red) is also displayed.

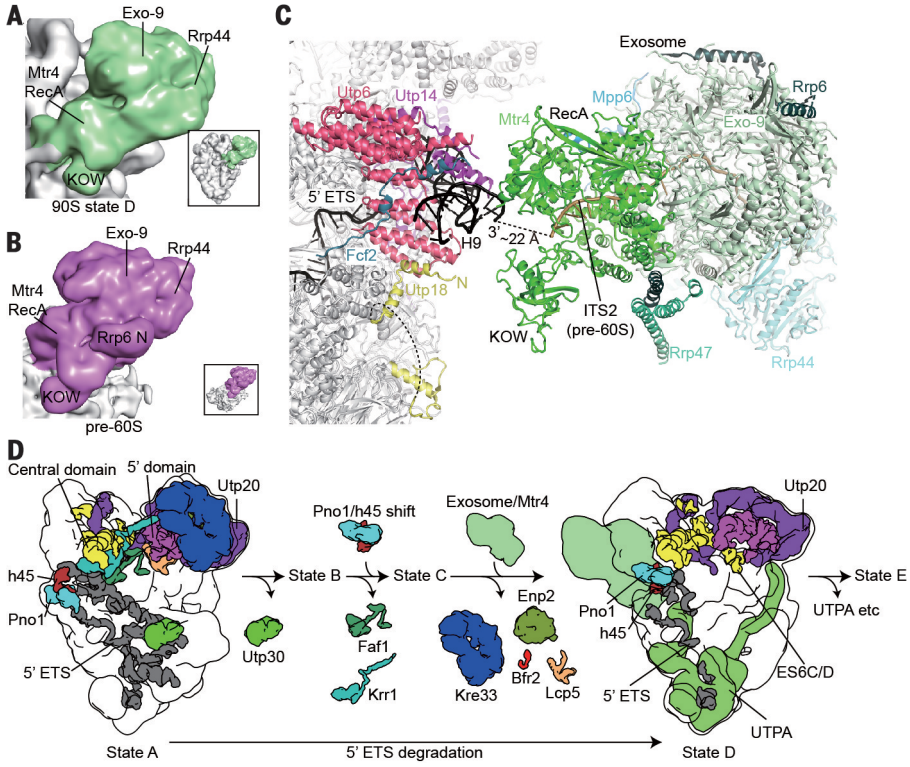

Fig. 4. Binding of the exosome to 90S. (**A** and **B**) Density maps of exosome-bound 90*S* (A) and pre-60*S* (B) ribosomes low-pass filtered to 20 Å. Exo-9 refers to the nine-subunit core of the exosome. Overall maps are shown in inserts. (**C**) Interaction between the exosome-Mtr4 and 90*S*. (**D**) Diagram of 90*S* progression showing major remodeling events.

the AFs bound to the 5′ domain and remodeling of the Utp20 structure, the apical part of the 5′ domain is pulled closer to the central domain by ~10 Å. By contrast, the basal part of the 5′ domain, which is buried in the structural core of the 90*S*, and the central domain are static (fig. S7F). The apical part of the 5′ domain maintains its core structure through the transition, but a few peripheral RNA elements (ES3B and h6) are shifted and helices 12 to 14 become disordered (fig. S7G). Helix C/D of ES6 (ES6C/D), which is invisible in states A to C, packs into the 5′ domain in state D (Fig. 3D). ES6C/D mainly adopts a mature-like conformation (Fig. 3E) with only its tip, bound by an unassigned peptide, deviating (Fig. 3D). The binding site of ES6C/D on the 5′ domain was originally occupied by Lcp5, Rrp9, and h12. Assembly of ES6C/D is only feasible after Lcp5 has been removed and the N-terminal residues of Rrp9 are remodeled to bind Utp20 (Fig. 3D).

The remodeling of Utp20 in state D also affects the core structure of 90*S*. The C-half of Utp20 moves closer to the structural core and projects an anchor loop to bind Rps22, Utp24, and Rps9 (Fig. 3D and fig. S7H). The N-half of Utp20 contacts Utp10 more extensively and straightens the C domain of this long superhelical protein (fig. S8C). This stabilizes several structures around the C domain of Utp10 so that they are visualized for the first time, including a few N-terminal helices of Rrt14 that bind Utp10 and the P5 helix of U3 snoRNA, the P3 helix of U3 sandwiched between Utp10 and Rrp9, and an unassigned protein domain bound to Utp10 (Fig. 2E and fig. S8, A to G).

The exosome and RNA helicase Mtr4 are involved in degradation of the excised 5′ ETS fragment and the 3′ processing of the 5.8*S* rRNA (*25*–*27*). The nuclear exosome contains a nine-subunit ring structure (Exo-9), Rrp44 with 3′ to 5′ exoribonuclease and endoribonuclease domains, and cofactors Mpp6, Rrp6, and Rrp47. As shown by semiquantitative mass spectrometric analysis, Mtr4 and the exosome proteins are threefold more abundant in Dhr1-depleted 90*S* compared with wild-type 90*S* (fig. S9 and data S2), suggesting that the exosome-bound 90*S* is enriched in our sample. A block of very weak density was observed to project out from the H8 to H9 regions in states D and E. The density was improved when the particles of state D enriched with the extra density were selected for reconstruction (figs. S1D and S2B). The resulting density accommodates a structure of exosome-Mtr4 complex (Fig. 4A), which was determined in the context of a pre-60*S* ribosome (Fig. 4B) (*28*). The model shows that Mtr4 mediates the binding of the exosome to 90*S* (Fig. 4C). Confirming the structural arrangement, the binding of the exosome was blocked upon depletion of Mtr4 (fig. S9). Two RecA domains of Mtr4 dock on a composite surface of Utp6, Utp14, Fcf2, and the base of H9. Utp14 at the interface is important for exosome binding (*29*). The N-terminal region of Utp18 that binds the KOW domain of Mtr4 is also close to the interface (*26*). The 3′ end of H9 is ~22 Å away from the entry of the RNA channel in the exosome. This arrangement would allow the exosome to degrade the exposed 3′ end of the 5′ ETS after site A0 is cleaved. The exosome is visualized in states D and E, yet the 5′ ETS is already degraded in state B. The exosome may be bound in earlier states but is too flexible to be visualized. Its conformation is somehow restricted after structural remodeling in state D.

On the basis of our structures, the excised 5′ ETS is degraded in situ within the 90*S* structure, rather than being released together with associated 5′ ETS AFs (*9*). We have captured the structure of 90*S* associated with the exosome that is primed to digest the 3′ end of the excised 5′ ETS fragment using the 3′ to 5′ exonuclease activity of Rrp44. However, the possibility that the exosome is recruited to degrade stalled precursors cannot be excluded at present. Three discontinuous regions (H1, Ha to H5, and Hb to H9) of the 5′ ETS are persistently bound to 90*S*, suggesting that the digestion also occurs internally, likely involving the endonuclease activity of Rrp44 (*30*, *31*).

States A to E show consecutive structural changes and features of maturing preribosomes, suggesting that they are evolvable physiological intermediates rather than dead-end products. Moreover, the RNA degradation is specific because only the 5′ ETS, not the 18*S* region, is affected. States A1 and C1, which have an alternative or disrupted central domain structure and no detectable successive states, may have resulted from a quality control process.

These structural snapshots illustrate the order of molecular events during the early transition of the 90*S* to the pre-40*S* ribosome (Fig. 4D). With the release of Krr1, the central domain is ready to accommodate h45. Additional release of AFs bound to the 5′ domain and remodeling of Utp20 bring the 5′ and central domains closer, allowing helix ES6C/D to make the first interdomain interaction. The extended and flexible structure of Utp20 is well suited to scaffold the two domains during the dynamic packing process. At state E, the UTPA complex begins to dissociate. Further release of the U3 snoRNP probably requires the helicase activity of Dhr1 (*20*) and cannot proceed in the Dhr1-depleted 90*S*. From a structural point of view, digestion of 5′ ETS is critical for starting the chain of remodeling events by destabilizing local structures of 90*S*

and making space for repositioned components. Supporting the driving role of 5′ ETS digestion for 90*S* progression, the exosome components and Mtr4 are required for 18*S* rRNA processing (*32*). Our study reveals an unexpected strategy in remodeling of a large RNA-protein complex by digestion of its RNA and sets the stage for further investigation of the transition of the 90*S* to the pre-40*S* ribosome.

REFERENCES AND NOTES

1. J. L. Woolford Jr., S. J. Baserga, *Genetics* **195**, 643–681 (2013).
2. J. Baßler, E. Hurt, *Annu. Rev. Biochem.* **88**, 281–306 (2019).
3. S. Klinge, J. L. Woolford Jr., *Nat. Rev. Mol. Cell Biol.* **20**, 116–131 (2019).
4. F. Dragon *et al*., *Nature* **417**, 967–970 (2002).
5. P. Grandi *et al*., *Mol. Cell* **10**, 105–115 (2002).
6. Y. N. Osheim *et al*., *Mol. Cell* **16**, 943–954 (2004).
7. M. Chaker-Margot, M. Hunziker, J. Barandun, B. D. Dill, S. Klinge, *Nat. Struct. Mol. Biol.* **22**, 920–923 (2015).
8. L. Zhang, C. Wu, G. Cai, S. Chen, K. Ye, *Genes Dev.* **30**, 718–732 (2016).
9. M. Kornprobst *et al*., *Cell* **166**, 380–393 (2016).
10. J. Barandun *et al*., *Nat. Struct. Mol. Biol.* **24**, 944–953 (2017).
11. M. Chaker-Margot, J. Barandun, M. Hunziker, S. Klinge, *Science* **355**, eaal1880 (2017).
12. J. Cheng, N. Kellner, O. Berninghausen, E. Hurt, R. Beckmann, *Nat. Struct. Mol. Biol.* **24**, 954–964 (2017).
13. Q. Sun *et al*., *eLife* **6**, e22086 (2017).
14. J. Cheng *et al*., *Mol. Cell* **75**, 1256–1269.e7 (2019).
15. M. Hunziker *et al*., *eLife* **8**, e45185 (2019).
16. A. Heuer *et al*., *eLife* **6**, e30189 (2017).
17. M. Ameismeier, J. Cheng, O. Berninghausen, R. Beckmann, *Nature* **558**, 249–253 (2018).
18. A. Scaiola *et al*., *EMBO J.* **37**, e98499 (2018).
19. O. Rodríguez-Galán, J. J. García-Gómez, J. de la Cruz, *Biochim. Biophys. Acta* **1829**, 775–790 (2013).
20. R. Sardana *et al*., *PLOS Biol.* **13**, e1002083 (2015).
21. A. Colley, J. D. Beggs, D. Tollervey, D. L. Lafontaine, *Mol. Cell. Biol.* **20**, 7238–7246 (2000).
22. T. Schäfer, D. Strauss, E. Petfalski, D. Tollervey, E. Hurt, *EMBO J.* **22**, 1370–1380 (2003).
23. A. Ben-Shem *et al*., *Science* **334**, 1524–1529 (2011).
24. J. Rabl, M. Leibundgut, S. F. Ataide, A. Haag, N. Ban, *Science* **331**, 730–736 (2011).
25. J. de la Cruz, D. Kressler, D. Tollervey, P. Linder, *EMBO J.* **17**, 1128–1140 (1998).
26. M. Thoms *et al*., *Cell* **162**, 1029–1038 (2015).
27. C. Allmang *et al*., *EMBO J.* **18**, 5399–5410 (1999).
28. J. M. Schuller, S. Falk, L. Fromm, E. Hurt, E. Conti, *Science* **360**, 219–222 (2018).
29. J. J. Black, Z. Wang, L. M. Goering, A. W. Johnson, *RNA* **24**, 1214–1228 (2018).
30. D. Schaeffer *et al*., *Nat. Struct. Mol. Biol.* **16**, 56–62 (2009).
31. A. Lebreton, R. Tomecki, A. Dziembowski, B. Séraphin, *Nature* **456**, 993–996 (2008).
32. C. Allmang, P. Mitchell, E. Petfalski, D. Tollervey, *Nucleic Acids Res.* **28**, 1684–1691 (2000).

ACKNOWLEDGMENTS

We thank the Center for Biological Imaging (CBI), Institute of Biophysics, Chinese Academy of Science (CAS), for cryo-EM study; the HPC Service Station in the CBI for image processing; and F. Sun, X. Huang, G. Ji, B. Zhu, and D. Fan at the CBI for help in EM sample preparation and data collection. **Funding:** This work was supported by the National Key R&D Program of China (grant no. 2017YFA0504600), the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDB37010201), and the National Natural Science Foundation of China (grant nos. 91940302, 91540201, and 31430024). **Author contributions:** K.Y. initiated the project; Y.D. collected cryo-EM data and determined the structures; W.A. prepared samples; X.Z, Q.S., and J.Q. carried out the early-stage study; and Y.D. and K.Y. analyzed the structure and wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. The cryo-EM density maps and models have been deposited in EMDB and PDB with accession codes: EMD-0949, EMD-0950, EMD-0951, EMD-0952, EMD-0953, EMD-0954, EMD-0955, 6LQP, 6LQQ, 6LQR, 6LQS, 6LQT, 6LQU and 6LQV. Materials are available from K.Y. on request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/369/6510/1477/suppl/DC1
Materials and Methods
Figs. S1 to S9
Table S1
References (*33*–*49*)
Data S1 and S2
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

10 February 2020; accepted 13 July 2020
10.1126/science.aba9690