TREX-2-like complexes couple UAP56 remodeling to alternative nuclear mRNA fates

Xing Gong[1,4], Xiaofei Ge[2,4], Mengqi Li[1,4], Ran Tao[1], Tianle Zhao[1], Fenghua Yang[1], and Xiaofeng Zhang[1,3,*]

[1]Center for Reproduction and Genetics, Department of Obstetrics and Gynecology, The First Affiliated Hospital of USTC; MOE Key Laboratory for Cellular Dynamics, Hefei National Research Center for Interdisciplinary Sciences at the Microscale; Center for Advanced Interdisciplinary Science and Biomedicine of IHM; Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

[2]Health and Wellness, City University of Macau, Macau 999078, China

[3]The USTC RNA Institute, University of Science and Technology of China, Hefei 230027, China

[4]These authors contributed equally to this work.

[*]Corresponding author. E-mail: xiaofengzhang@ustc.edu.cn

Abstract

Newly assembled messenger ribonucleoprotein particles (mRNPs) must be sorted between nuclear export and degradation. How a common mRNP intermediate is coupled to these opposing fates remains unclear. Here we identify two TREX-2-like complexes, LENG8-PCID2-DSS1 (LPD) and SAC3D1-PCID2-DSS1 (SPD). Like canonical TREX-2, both complexes bind UAP56 and its paralog URH49, stimulate ATPase activity, and promote RNA release. Cryo-electron microscopy structures of UAP56-SPD, URH49-SPD, and UAP56-LPD at approximately 3.1 Å resolution reveal a conserved Sac3-family architecture and a shared activating-loop mechanism for helicase remodeling. LENG8 is predominantly nuclear and associates with the poly(A)-tail exosome targeting complex (PAXT) and the nuclear RNA exosome. Its N-terminal region directly binds the ZFC3H1-MTR4 PAXT core, thereby coupling LPD-mediated RNA release from UAP56 to RNA capture by PAXT. Consistently, tethering LENG8 to a reporter RNA suppresses its nuclear export and expression. Together, our findings support a model in which TREX-2 and LPD act on a common UAP56-bound mRNP intermediate. TREX-2 directs the released RNA toward NXF1-NXT1-dependent export, whereas LPD channels it into PAXT-exosome-mediated decay. These findings provide a molecular framework for UAP56-dependent sorting of nuclear mRNPs.

Introduction

The nuclear fate of an mRNA is determined through a series of coordinated assembly and remodeling events. Nascent transcripts are capped, spliced, polyadenylated, and packaged with RNA-binding proteins to form messenger ribonucleoprotein particles (mRNPs). Export-competent mRNPs are recognized by the principal export receptor NXF1-NXT1 and transported through nuclear pore complexes (NPCs)[1-5]. By contrast, aberrant or incompletely processed transcripts are retained in the nucleus and degraded[6-12]. Nuclear export and surveillance are therefore not independent terminal events, but competing outcomes of an integrated RNA-sorting system. However, the molecular mechanisms that couple a common mRNP intermediate to these alternative fates remain poorly understood.

The DEAD-box ATPase UAP56, also known as DDX39B, occupies a central position in nuclear mRNP biogenesis[1,2]. Together with its close paralog URH49/DDX39A[13,14], UAP56 participates in pre-mRNA splicing, TREX assembly, and loading of export adaptors onto nascent transcripts[15-18]. UAP56 contains two RecA domains that close around ATP and RNA to form a stable RNA clamp[19,20]. Reopening of this clamp releases the RNA and permits helicase recycling, allowing the mRNP to engage downstream factors. TREX-2 promotes this transition during nuclear export. In metazoans, the UAP56-binding module of TREX-2 comprises the Sac3 domain of GANP (GANP^{Sac3}), PCID2 and DSS1[21-23]. A conserved loop from GANP^{Sac3} inserts between the two RecA domains and favors helicase opening[24-26]. The N-terminal region of GANP then facilitates capture of the released RNA by

NXF1-NXT1[27-29]. TREX-2 therefore couples UAP56 remodeling to NXF1-NXT1 loading and directs the mRNP toward nuclear export.

Beyond GANP, LENG8 and SAC3D1 contain related Sac3 domains[25,26]. LENG8 assembles with PCID2 and DSS1 and has recently been implicated in nuclear RNA retention or degradation[25,30]. Neither LENG8 nor SAC3D1 contains an N-terminal region analogous to the NXF1-NXT1-binding region of GANP (Fig. 1a). This architectural divergence raises the possibility that these Sac3-containing assemblies may use a conserved helicase-remodeling module while engaging distinct downstream factors. However, whether these assemblies remodel UAP56 or URH49 in the same manner as TREX-2, and which cellular pathways they serve, remain incompletely defined.

Here, we characterize two TREX-2-like complexes, LENG8-PCID2-DSS1 (LPD) and SAC3D1-PCID2-DSS1 (SPD). We show that both complexes bind UAP56 and URH49 and promote ATPase activation and RNA release. Cryo-EM structures define a conserved mechanism of helicase recognition and remodeling. We further demonstrate that the N-terminal region of LENG8 recruits ZFC3H1-MTR4 (hereafter referred to as PAXT^{core}) and couples LPD-mediated RNA release from UAP56 to RNA capture by PAXT. These findings reveal how a common Sac3-family remodeling module can be connected to distinct nuclear RNA fates.

Results

LPD and SPD form TREX-2-like complexes that remodel UAP56 and URH49

The UAP56-interacting module of canonical TREX-2 comprises GANP^{Sac3}, PCID2 and DSS1 and is referred to here as TREX-2^{M}. LENG8 and SAC3D1 contain related Sac3 regions but differ substantially in overall organization (Fig. 1a). LENG8 contains an extended N-terminal regions followed by a C-terminal Sac3 domain, whereas SAC3D1 consists largely of a Sac3-related core. We co-expressed LENG8 ^{Sac3} or full-length SAC3D1 with PCID2 and DSS1 and purified the resulting heterotrimeric complexes, designated LPD and SPD, respectively (Supplementary Fig. 1a-c).

We next tested whether LPD and SPD recognize UAP56 and URH49. In affinity pulldown assays, UAP56 co-eluted with both complexes at levels comparable to TREX-2^{M} (Fig. 1b; Supplementary Fig. 1d). URH49 was similarly retained by all three Sac3-family complexes (Fig. 1c; Supplementary Fig. 1e). Thus, recognition of UAP56 and URH49 is a shared property of the GANP-, LENG8-, and SAC3D1-containing assemblies. Microscale thermophoresis further confirmed direct binding to UAP56, with apparent K_d values of 0.72 ± 0.38 μM for TREX-2^{M}, 1.59 ± 0.82 μM for LPD, and 0.31 ± 0.13 μM for SPD (Supplementary Fig. 1f). These measurements indicate affinities spanning the submicromolar to low-micromolar range.

We then examined whether LPD and SPD also share the remodeling activity of TREX-2. TREX-2^{M} stimulated the RNA-dependent ATPase activity of UAP56 by approximately 39-fold. LPD and SPD produced stronger stimulation, increasing ATPase activity by approximately 114-fold and 103-fold, respectively (Fig. 1d). In native gel-shift assays, each complex caused a concentration-dependent decrease in

the preassembled UAP56-RNA complex and a corresponding increase in free RNA (Fig. 1e; Supplementary Fig. 1g). Thus, LPD and SPD are potent activators of UAP56 and promote release of bound RNA. These results establish a common biochemical activity among the three Sac3-family complexes.

A conserved mechanism for helicase remodeling

To define the molecular basis of this activity, we reconstituted UAP56-SPD, URH49-SPD, and UAP56-LPD complexes in the presence of RNA and determined their structures by cryo-EM analysis (Fig. 1f; Supplementary Figs. 2-5; Supplementary table 1). All three complexes share a conserved architecture: PCID2 and DSS1 support the Sac3-family subunit and together form a composite platform for helicase binding (Fig. 2a; Supplementary Fig. 6). UAP56 and URH49 adopt closed, RNA-clamped conformations in the SPD-bound structures (Fig. 2a; Supplementary Fig. 6b). By contrast, UAP56 adopts an open conformation in the LPD-bound structure, and RNA is absent. This conformation resembles that observed in the UAP56-TREX-2^{M} structure[24-26]. These structures therefore capture distinct conformational states of a conserved helicase-remodeling pathway.

The helicase-binding interface is featured by three structural elements: Loop-1, Loop-2, and the UAP56-activating loop (UAL). Loop-1 of UAP56 or URH49 engages a groove formed jointly by PCID2 and the Sac3-family subunit. The SPD- and LPD-bound complexes display similar networks of polar, electrostatic, and hydrophobic interactions at this interface (Fig. 2b-d), which are largely conserved in

UAP56-TREX-2^{M}. Intriguingly, Loop-2 adopts different conformations among UAP56-LPD, UAP56-SPD, URH49-SPD, and UAP56-TREX-2^{M} (Fig. 2e). This variation indicates a degree of complex-specific flexibility without altering the overall mode of helicase recognition.

The UAL constitutes the principal remodeling element. This loop projects from GANP, LENG8, or SAC3D1 toward the cleft between the two RecA domains and contains a conserved arginine residue (Fig. 2f,g). In the RNA-clamped state, the adenine moiety of the bound nucleotide is sandwiched between Phe65 and Phe381 in UAP56, or Phe64 and Phe380 in URH49. In the RNA-released state, the conserved UAL arginine approaches the nucleotide-binding pocket and occupies the position otherwise taken by Phe381 (Fig. 2h; Supplementary Fig. 5). This rearrangement alters nucleotide coordination and destabilizes the closed helicase conformation. Comparison of the SPD- and TREX-2^{M}-bound structures reveals an approximately 5.1 Å shift of the conserved arginine toward the nucleotide-binding pocket of UAP56 in the RNA-released state (Fig. 2i). Together, these structures reveal how Sac3-family complexes use conserved interaction interfaces and an activating-loop mechanism to engage and remodel UAP56 and URH49.

LENG8 associates with the nuclear PAXT-exosome machinery

We next examined the subcellular localization and interaction networks of LENG8 and SAC3D1. Immunofluorescence microscopy showed that LENG8 was predominantly nuclear, whereas SAC3D1 was enriched in the cytoplasm under the

same conditions (Fig. 3a). This distinct localization suggested that LPD and SPD function in different cellular pathways and prompted us to investigate whether LPD participates in nuclear RNA surveillance.

Affinity purification followed by mass spectrometry revealed that LENG8 associates with factors acting at multiple stages of nuclear mRNP metabolism (Fig. 3b). Among canonical TREX-2-associated proteins, only PCID2 was enriched; the GANP-interacting factors ENY2 and TPR[31] were not (Fig. 3c). This pattern is consistent with LENG8 forming an alternative Sac3-family complex rather than canonical TREX-2. In contrast, LENG8 strongly enriched the PAXT core components ZFC3H1 and MTR4, together with multiple subunits of the nuclear RNA exosome (Fig. 3c). PABPN1 was also detected, but showed weaker enrichment. Proteins involved in early spliceosome assembly, including components of the U1 and U2 snRNPs, were also prominently enriched. Early spliceosomal factors can promote the nuclear retention of incompletely processed transcripts, whereas PAXT targets unstable nuclear RNAs for retention and degradation[9-12,32]. This interaction profile supports a role of LENG8 in nuclear RNA quality control.

The association of LENG8 with the RNA surveillance machinery was validated by co-immunoprecipitation. Strep-LENG8 recovered endogenous ZFC3H1 and MTR4, as well as the exosome subunits EXOSC2 and EXOSC8. These interactions persisted after Benzonase treatment, indicating that they were not solely mediated by nucleic acids (Fig. 3e). By contrast, the SAC3D1 interactome was enriched for a distinct set of pathways, including protein processing in the

endoplasmic reticulum (Fig. 3d), consistent with its predominantly cytoplasmic localization. Together, these data link LPD to the nuclear PAXT-exosome pathway.

LENG8 couples UAP56 remodeling to PAXT-mediated RNA capture

LENG8 contains an extended N-terminal region (LENG8^{NTR}) and a C-terminal Sac3 domain that assembles with PCID2-DSS1 (Fig. 4a). To identify the region responsible for PAXT binding, we examined the interactions of full-length LENG8, LENG8^{NTR}, and LENG8^{Sac3} with a reconstituted PAXT core complex containing full-length ZFC3H1 and MTR4 (PAXT^{core}). Full-length LENG8 and LENG8^{NTR} efficiently recovered PAXT^{core} in MBP pulldown assays, whereas the isolated Sac3 domain did not (Fig. 4b; Supplementary Fig. 7a-d). An AlphaFold3 model places LENG8 residues 283-326 at an interface with ZFC3H1 and predicts four pairs of charge-assisted hydrogen bonds (Supplementary Fig. 7e). Thus, PAXT binding is mediated by the N-terminal region of LENG8 and is physically separable from the Sac3-PCID2-DSS1 module that remodels UAP56.

These findings define LENG8 as a bifunctional scaffold. Its C-terminal Sac3 domain forms the LPD complex and promotes RNA release from UAP56 (Fig. 1), whereas its N-terminal region recruits PAXT^{core} (Figs. 3 & 4). This organization parallels that of GANP, whose Sac3-PCID2-DSS1 module remodels UAP56 while its N-terminal region engages NXF1-NXT1[27,33]. The divergent N-terminal regions of GANP and LENG8 may therefore connect a conserved remodeling reaction to different downstream RNA fates.

We tested this model using a reconstituted RNA-release and RNA-capture assay (Fig. 4c). Addition of LPD to preassembled UAP56-RNA complexes caused a concentration-dependent loss of the UAP56-bound RNA species and an increase in free RNA. When the PAXT^{core} was included, a slower-migrating PAXT^{core}-RNA complex appeared concomitantly (Fig. 4d; Supplementary Fig. 7f). Thus, LPD-mediated remodeling of UAP56 can be directly coupled to capture of the released RNA by the PAXT^{core}.

Tethered LENG8 suppresses an mRNA export reporter

To test whether recruitment of LENG8 is sufficient to influence RNA fate in cells, we used a λN-BoxB RNA tethering assay[34]. A GFP reporter containing seven BoxB elements was co-expressed with the indicated λN-fusion proteins, each carrying a C-terminal mRuby3 transfection marker. Under control conditions, the reporter RNA is efficiently exported and translated. GFP fluorescence was therefore quantified within the mRuby3-positive population as a measure of reporter output (Fig. 4e, left).

Tethering LENG8 to the reporter RNA markedly reduced GFP fluorescence, whereas expression of LENG8 lacking the λN peptide had no detectable effect (Fig. 4e, right). Tethering ZFC3H1, a core component of PAXT, produced a similar reduction, whereas tethering the export receptor NXF1 increased reporter output. Thus, direct recruitment of LENG8 is sufficient to suppress the nuclear export of the tethered RNA.

Together, the biochemical, structural and cellular data support a model in which Sac3-family complexes couple UAP56 or URH49 remodeling to distinct nuclear RNA fates (Fig. 5). GANP-containing TREX-2 promotes RNA release from UAP56 and positions NXF1-NXT1 to receive the released RNA, thereby favoring nuclear export. LPD triggers a similar remodeling reaction, but the N-terminal region of LENG8 recruits the PAXT-exosome machinery and directs the RNA toward nuclear retention and decay. SPD also remodels UAP56 and URH49 in vitro, although its cellular function remains to be determined.

Discussion

We identify LPD and SPD as TREX-2-like complexes that remodel UAP56 and URH49, and define how LPD couples this reaction to nuclear RNA surveillance. Both complexes form Sac3-PCID2-DSS1 assemblies, bind UAP56 and URH49, stimulate UAP56 ATPase activity and promote RNA release (Fig. 1). The cryo-EM structures show that these activities are mediated by a conserved helicase-binding interface and an activating-loop that destabilizes the closed, RNA-clamped conformation (Fig. 2). TREX-2 is therefore one member of a broader family of Sac3-dependent helicase-remodeling complexes.

These structures reveal a common architectural principle. The conserved Sac3-PCID2-DSS1 module remodels UAP56, whereas the divergent regions of the Sac3-family proteins engage distinct downstream machineries. In TREX-2, the N-terminal region of GANP links UAP56 remodeling to NXF1-NXT1 and nuclear

export[27,33]. In LPD, the N-terminal region of LENG8 recruits PAXT^{core} and associates with the nuclear RNA exosome (Fig. 5). Thus, RNA fate is determined not solely by the remodeling reaction, which is shared by both pathways, but by the downstream acceptor positioned to receive the released RNA.

The association of LPD with PAXT suggests how nuclear RNA quality control may be coupled to mRNP biogenesis. UAP56 engages transcripts during splicing, TREX assembly and export-adaptor loading[1,2]. It is therefore positioned at a stage at which transcript maturation can be assessed. Defects in splicing, 3′-end formation, poly(A)-tail organization or export-adaptor assembly may favor engagement by LPD, whereas formation of an export-competent mRNP may favor TREX-2 and NXF1-NXT1. The molecular features that determine this choice remain unclear. Consistent with such a role, LENG8 co-purified with PAXT and exosome components, as well as proteins involved in splicing and cleavage and polyadenylation (Fig. 3b,c). This interaction profile places LPD at the interface between mRNP maturation and surveillance. Defining the endogenous RNA substrates of LENG8 and comparing them with GANP-, UAP56-, ZFC3H1- and exosome-associated transcriptomes should help identify the RNA features and processing states that specify the decay pathway.

Two recent studies independently established LENG8 as a nuclear RNA-surveillance factor. Tian et al. showed that LENG8 is recruited to pre-mRNAs by early splicing factors and forms a conserved complex with PCID2 and DSS1 (SEM1) that restricts export of misprocessed transcripts and promotes PAXT-dependent

degradation[30]. Bugai et al. identified the same TREX-2-like module as a component of PAXT and proposed that nucleoplasmic PAXT and nuclear pore-associated TREX-2 compete for UAP56-bound polyadenylated RNPs[35]. Our findings are consistent with this framework and provide complementary mechanistic information. We resolve a high-resolution, open UAP56-LPD state, define recognition of URH49 by LPD and SPD, and reconstitute capture of LPD-released RNA by the PAXT^{core}. Together, these studies support a shared model in which related Sac3-family complexes act on a common helicase-bound mRNP intermediate but couple remodeling to distinct RNA fates.

SPD broadens this family but raises a separate biological question. SAC3D1 uses the same activating-loop mechanism to remodel UAP56 and URH49 (Fig. 2), yet is predominantly cytoplasmic and has an interaction network distinct from LENG8 (Fig. 4). SPD may therefore act on a specialized pool of UAP56- or URH49-containing RNPs, but its substrates and downstream acceptor remain unknown.

Methods

Purification of TREX-2^{M}, LPD and SPD

The open reading frames encoding GANP^{Sac3} (residues 580-1000), LENG8^{Sac3} (540-800), SAC3D1, PCID2 and DSS1 were individually cloned into a modified pCAG mammalian expression vector[36]. To facilitate protein purification, PCID2 and DSS1 were each fused to an N-terminal Flag tag, whereas GANP^{Sac3}, LENG8^{Sac3} and SAC3D1 were each fused to an N-terminal Twin-Strep tag. HEK293F cells were maintained in SMM 293T-II medium (Sino Biological Inc.) at 37°C under 6% CO_2 with constant shaking. Plasmids encoding the three subunits of TREX-2^{M}, LPD or SPD were co-transfected into cells using 40-kD linear polyethylenimines (Yeasen). At 72 h after transfection, the cells were harvested by centrifugation at 4,000×g for 10 minutes. The cell pellets were resuspended in lysis buffer containing 50 mM HEPES-KOH, pH 7.9, 150 mM NaCl, 3 mM $MgCl_2$, 0.05% NP40, and 1 mM DTT, and lysed by sonication. Clarified cell lysates were applied by gravity flow to columns containing anti-Flag M2 affinity gel (Sigma). The resin was extensively washed with lysis buffer, and the bound proteins were eluted with lysis buffer supplemented with 0.4 mg/mL Flag peptide. The eluent was subsequently loaded to a column containing Strep-Tactin beads (IBA). After five washes with lysis buffer, the target complexes were eluted with 50 mM biotin and further purified by size exclusion chromatography (SEC) using a Superdex-200 Increase column (Cytiva) equilibrated with GF75 buffer (25 mM HEPES-KOH, pH 7.9, 75 mM NaCl, 1.5 mM $MgCl_2$). Peak fractions were analyzed on SDS-PAGE gels followed by Coomassie blue staining.

Purification of UAP56 and URH49

UAP56 was fused to an N-terminal 8×His tag and cloned into pET-21b expression vector. The protein was expressed in *E. coli* Rosetta (DE3) cells cultured in LB medium. When the cultures reached an OD_{600} of approximately 0.8, protein expression was induced with 0.2 mM IPTG, followed by incubation at 16 ℃ for 16 h. Cells were collected in the H300 buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1.5 mM $MgCl_2$, 25 mM imidazole, 0.5 mM PMSF) and lysed by sonication. After

clarification by centrifugation, the supernatant was applied to columns containing Ni-NTA agarose resin (Smart-Lifesciences). The resin was washed six times with H300 buffer, and the bound protein was eluted with GF75 buffer supplemented with 300 mM imidazole. The eluate was further purified by ion-exchange chromatography using a 5-mL HiTrap Heparin column (Cytiva). Fractions containing the target protein were pooled, concentrated using a 30-kDa Amicon centrifugal filter unit (Millipore), and subjected to SEC on a Superdex 200 Increase column (Cytiva) equilibrated with GF75 buffer. Peak fractions were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80 °C. URH49 was expressed and purified using the same procedure.

Purification of MBP-LENG8

Full-length LENG8 was fused to an N-terminal MBP tag and cloned into pET-15b expression vector. MBP-LENG8 was expressed in *E. coli* Rosetta (DE3) cells cultured in LB medium. When the cultures reached an OD_{600} of approximately 0.8, protein expression was induced with 0.2 mM IPTG, followed by incubation at 16 ℃ for 16 h. Cells were harvested by centrifugation, resuspended in lysis buffer, and disrupted by sonication. After clarification by centrifugation, the supernatant was applied to a column containing amylose resin (Smart-Lifesciences) with gravity flow. The resin was washed six times with lysis buffer, and the bound protein was eluted with GF75 buffer supplemented with 10 mM maltose. The eluate was further purified by SEC on a Superdex 200 Increase column (Cytiva). Peak fractions were pooled and concentrated, and the purified protein was flash-frozen in liquid nitrogen and stored at -80 °C. The MBP-tagged N-terminal region of LENG8 (NTR, residues 1-539) and its Sac3 domain (residues 540-800) were expressed and purified following the same procedure.

Purification of PAXT^{core} and PABPN1

The cDNAs encoding full-length ZFC3H1, MTR4 and PABPN1 were cloned separately into a modified pCAG mammalian expression vector. ZFC3H1 and MTR4 were each fused to an N-terminal Twin-Strep tag, whereas PABPN1 was fused to an

N-terminal Flag tag. To purify PAXT^{core} complex, plasmids encoding MTR4 and ZFC3H1were co-transfected into HEK293F cells at a 1:1 ratio. After 72 h, cells were harvested and lysed by sonication in lysis buffer. The clarified lysate was incubated with 2 mL of Strep-Tactin resin (IBA), and the resin was washed six times with lysis buffer before elution with lysis buffer containing 50 mM biotin. The PAXT^{core} complex was further purified by SEC on a Superose 6 Increase column (Cytiva) equilibrated with GF75 buffer. Fractions containing the complex were identified by SDS-PAGE and Coomassie blue staining. Flag-PABPN1 was expressed in HEK293F cells under the same conditions and purified by anti-Flag affinity chromatography.

Pulldown Assay

To assess the interactions of TREX-2^{M}, LPD, and SPD with UAP56 or URH49, each purified complex was mixed at a final concentration of 1 μM with 2 μM UAP56 or URH49 in GF75 buffer. The mixtures were incubated at 4 °C for 2 h to allow complex formation. The resulting complexes were captured on Strep-Tactin resin via the N-terminal Twin-Strep tags fused to GANP^{Sac3}, LENG8^{Sac3}, or SAC3D1. After extensive washing with GF75 buffer, bound proteins were eluted with GF75 buffer supplemented with 50 mM biotin and analyzed by SDS-PAGE.

Binding-affinity measurements using microscale thermophoresis

The binding affinities of TREX-2^{M}, LPD, and SPD for UAP56 were determined by microscale thermophoresis (MST) using a Monolith NT.115 instrument (NanoTemper Technologies GmbH, Germany). GFP-tagged UAP56 was expressed in *E. coli* Rosetta (DE3) cells, purified by affinity chromatography, and diluted to 10 nM in MST buffer (25 mM HEPES-KOH, pH 7.9, 150 mM NaCl, 0.05% tween-20, 2 mM $MgCl_2$). For each binding assay, a 16-point, twofold serial dilution of the indicated ligand was prepared and mixed with an equal volume of purified GFP-UAP56. After incubation at room temperature for 30 min, the samples were loaded into premium-coated capillaries (NanoTemper Technologies). MST measurements were performed at 25 °C using 60% excitation power and medium MST power. Data were analyzed using

MO.Affinity Analysis software (NanoTemper Technologies). Normalized fluorescence (F_norm) was plotted against ligand concentration and fitted using the equilibrium dissociation constant (K_d) model implemented in the software. All measurements were performed in three independent experiments.

ATPase activity assay

The RNA-stimulated ATPase activity of UAP56 was measured using an ATPase Colorimetric Assay Kit (Innova Biosciences). Purified UAP56 at a final concentration of 0.5 μM was added to reaction buffer containing 25 mM HEPES-KOH, pH 7.9, 150 mM NaCl, 5 mM $MgCl_2$, and 40 μM 15-nt poly(U) RNA. The reaction mixtures were pre-incubated on ice for 20 min. ATP hydrolysis was initiated by the addition of ATP to a final concentration of 1 mM, followed by incubation at 37 °C for 10 min. The amount of released inorganic phosphate was quantified using the ATPase Colorimetric Assay Kit by measuring absorbance at 650 nm with a BioTek Synergy H1 multimode microplate reader. To assess the effects of TREX-2^M, LPD, and SPD on UAP56 ATPase activity, the assays were performed in parallel in the presence of 1 μM TREX-2^M, LPD, or SPD. Each condition was tested in three independent experiments. Statistical analyses were performed using GraphPad Prism 8.

Electrophoretic mobility shift assay

40 nM 5'-FAM-labeled poly(U) RNA was incubated with 4 μM UAP56 in binding buffer containing 25 mM HEPES-KOH, pH 7.9, 75 mM NaCl, 1.5 mM $MgCl_2$, 0.5 U/μL RNase inhibitor (Yeasen), and 2 mM ATPγS (Sigma). The mixtures were incubated at room temperature for 30 min to allow formation of the UAP56-RNA complex. Increasing concentrations of TREX-2^M, LPD, SPD, or LPD-$PAXT^{core}$ were then added to the preassembled UAP56-RNA complexes, followed by incubation at room temperature for an additional 30 min. Samples were resolved by electrophoresis on 6% native polyacrylamide gels in 0.5× Tris-borate-EDTA (TBE) running buffer at room temperature. Fluorescent RNA signals were visualized using an Amersham Imager 680 system (GE Healthcare).

Immunofluorescence

HeLa cells were cultured in 12-well glass-bottom plates in DMEM supplemented with 10% fetal bovine serum at 37 °C. For immunofluorescence staining, the culture medium was removed, and the cells were washed with PBS before fixation with 4% paraformaldehyde in PBS for 20 min. After a brief PBS wash, the cells were permeabilized with 0.3% Triton X-100 in PBS for 10 min and then blocked with 5% BSA and 0.1% Triton X-100 in PBS for 1 h. The cells were then incubated with primary antibodies for 1 h at 37 °C. The primary antibodies used were rabbit anti-LENG8 (HPA061571, Sigma-Aldrich; 1:50 dilution) and rabbit anti-SAC3D1 (25857-1-AP, Proteintech; 1:250 dilution). After two 5-min washes with PBS containing 0.1% Triton X-100, the cells were incubated with Alexa Fluor 594-conjugated goat anti-rabbit IgG (ab150080, Abcam; 1:200 dilution) for 1 h at 37 °C. After two washes with PBS, nuclei were stained with DAPI for 10 min. The cells were then washed twice with PBS, stained with DAPI for 10 min, and washed again with PBS. Fluorescence images were acquired using a THUNDER Imager Cell Culture microscope (Leica Microsystems). Computational clearing and image processing were performed using LAS X software (Leica Microsystems).

Cryo-EM sample preparation and data collection

The UAP56-LPD, UAP56-SPD, and URH49-SPD complexes were reconstituted in G75 buffer supplemented with 2 mM AMPPNP, 10 μM of 20-nt poly(U) RNA. The samples were concentrated to approximately 1.2 mg/mL and used for cryo-EM grid preparation. Aliquots of each sample were applied to glow-discharged amorphous NiTi foil grids (Au300-1.2/1.3, Nanodim Tech.), blotted for 3 seconds and rapidly plunged into liquid ethane using Vitrobot Mark IV operating at 8 °C and 100% humidity. Cryo-EM data was collected with EPU on a 300-kV Titan Krios electron microscope (Thermo Fisher) at a normal magnification of 105,000×. Movies were recorded using a Falcon 4 direct electron detector equipped with a GIF Quantum energy filter (slit width 20 eV). The calibrated pixel size was 0.82 Å, and the nominal

defocus range was -1.5 to -2.3 μm.

Cryo-EM data processing

All cryo-EM datasets were processed using CryoSPARC v4.2[37]. For the UAP56-SPD complex, 1,657 micrographs were collected, from which 1,470,555 particles were identified by template-based picking and extracted. After two-dimensional (2D) classification, 785,269 particles were retained. Three initial models were generated by Ab initio reconstruction, and particles assigned to class 3 were selected for further processing. Subsequent rounds of heterogeneous refinement, non-uniform refinement, and three-dimensional classification yielded a final set of 272,959 particles, producing a reconstruction at an overall resolution of 3.08 Å.

For the URH49-SPD complex, 9,896 micrographs were collected, and 9,360,453 particles were identified by template-based picking. Following particle extraction and two rounds of two-dimensional classification, 2,298,323 particles were retained. Four initial models were generated by ab initio reconstruction, and particles assigned to classes 2 and 3 were combined for subsequent processing. After several rounds of heterogeneous refinement, non-uniform refinement, and three-dimensional classification, 249,215 particles contributed to the final reconstruction, which reached an overall resolution of 3.07 Å.

For the UAP56-LPD complex, 6,303 micrographs were imported for processing. Template-based particle picking yielded 3,810,247 particles, which were extracted and divided into six subsets. Each subset was subjected to two rounds of two-dimensional (2D) classification, resulting in a combined set of 1,018,869 high-quality particles. These particles were used for Ab initio reconstruction. Following multiple rounds of heterogeneous refinement, non-uniform refinement, and three-dimensional classification, 268,320 particles were selected for the final reconstruction, which reached an overall resolution of 3.13 Å. Overall resolutions were estimated using the Fourier shell correlation criterion of 0.143. Local resolution variations were estimated in cryoSPARC.

Model building

AlphaFold3-predicted models of the UAP56-LPD, UAP56-SPD, and URH49-SPD complexes were docked into the corresponding cryo-EM density maps using ChimeraX[38] to generate initial atomic models. The models were manually inspected and adjusted in Coot[39] based on the local density features. The resulting models were subjected to real-space refinement against the cryo-EM maps using phenix.real_space_refine[40], with secondary-structure restraints applied throughout refinement. Model quality was assessed using MolProbity[41] statistics and Ramachandran plot analysis.

Immunoprecipitation and mass spectrometry analysis

Immunoprecipitations of Strep-LENG8, Strep-SAC3D1 and Strep controls were performed in the presence of 2U/mL Benzonase (Vazyme) and in triplicates. Proteins recovered from each immunoprecipitation were precipitated with 15% trichloroacetic acid and dried under vacuum. The resulting pellets were solubilized in 8 M urea and 100 mM Tris-HCl, pH 8.5, followed by reduction with TCEP, alkylation with iodoacetamide, and digestion with trypsin. Tryptic peptides were dissolved in 0.1% formic acid, desalted using Pierce C18 spin columns, and separated on an EASY-nLC 1200 liquid chromatography system using a 120-min gradient of 0-80% acetonitrile in 0.1% formic acid. The liquid chromatography system was coupled directly to an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher), and spectra were acquired in data-dependent acquisition mode. Raw mass spectrometry data were processed with MaxQuant and searched against the UniProt protein sequence database. The "match between runs" and label-free quantification options were enabled. Protein-group intensities generated by MaxQuant were subsequently analyzed using the DEP package, including data normalization, missing-value imputation, and differential-enrichment analysis. For visualization, the $\log_2$ fold change in LFQ intensity for each protein in the LENG8 or SAC3D1 immunoprecipitate relative to the control was plotted against the corresponding -$\log_{10}$-transformed Student's *t*-test *P* value calculated from the three independent experiments.

Immunoprecipitation followed by western blotting

Strep-LENG8 immunoprecipitation was performed as described above. The recovered proteins were resolved by SDS-PAGE, transferred to membranes, and analyzed by standard western blotting. Membranes were probed with antibodies against Strep (HA722229, HUABIO; 1:10,000), ZFC3H1 (86822-1-RR, Proteintech; 1:500), MTR4 (86536-1-RR, Proteintech; 1:2,000), EXOSC2 (14805-1-AP, Proteintech; 1:1,000), and EXOSC8 (11979-1-AP, Proteintech; 1:1,000). HRP-conjugated goat anti-rabbit IgG (RGAR001, Proteintech; 1:10,000) was used as the secondary antibody.

Pulldown assay between LENG8 and the PAXT^{core} complex

To examine the interaction between LENG8 and the PAXT^{core} complex, purified full-length LENG8, LENG8^{NTD} or LENG8^{Sac3} was mixed at a final concentration of 0.5 μM with 1.5 μM of PAXT^{core} complex in GF75 buffer (25 mM HEPES-KOH, pH 7.9, 75 mM NaCl, and 1.5 mM $MgCl_2$). The mixtures were incubated at 4 °C for 2 h to allow complex formation. LENG8 proteins and their associated binding partners were captured on amylose resin via the N-terminal MBP tag. After four washes with GF75 buffer, bound proteins were eluted with GF75 buffer supplemented with 20 mM maltose and analyzed by SDS-PAGE. Each pulldown experiment was independently repeated three times.

RNA tethering assay

HeLa cells were seeded in six-well plates and cultured overnight before transfection. Cells were co-transfected with a GFP reporter plasmid containing seven BoxB elements[34] and plasmids encoding the indicated λN-fusion proteins using Lipofectamine 3000 (Thermo Fisher). The λN-fusion constructs also expressed mRuby3 as a marker of successful transfection. At 24 h after transfection, cells were harvested and subjected to flow cytometry (Aurora, Cytek Biosciences). Transfected cells were first identified by gating on the mRuby3-positive population, after which GFP fluorescence was quantified within this population. The effects of tethering the

indicated proteins to the reporter RNA were determined by comparing GFP fluorescence intensities between the experimental and control groups.

Statistical analysis

Statistical analyses were performed using GraphPad Prism 8. Statistical significance was determined using one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test. A P value of < 0.05 was considered statistically significant. Sample sizes and numbers of independent experiments are indicated in the corresponding figure legends.

Data availability: The atomic coordinates and corresponding cryo-EM maps have been deposited in the Protein Data Bank and Electron Microscopy Data Bank, respectively. The accession codes are 28FJ and EMD-81704 for the human UAP56-LPD complex, 28FK and EMD-81705 for the UAP56-SPD complex, and 28ER and EMD-81691 for the URH49-SPD complex. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the iProX partner repository[42] with the dataset identifier PXD082111. Source Data are provided with this paper.

References

1. Faraway, R., Zenklusen, D. & Plaschka, C. Mechanisms of Messenger RNA Packaging and Export. *Annual Review of Cell and Developmental Biology* **41**, 479-504 (2025).
2. Stewart, M. From transcription to export: mRNA's winding path to the cytoplasm. *Trends in Biochemical Sciences* **50**, 748-765 (2025).
3. Chen, S., Jiang, Q., Fan, J. & Cheng, H. Nuclear mRNA export. *Acta Biochim Biophys Sin (Shanghai)* (2024).
4. Wickramasinghe, V.O. & Laskey, R.A. Control of mammalian gene expression by selective mRNA export. *Nat Rev Mol Cell Biol* **16**, 431-42 (2015).
5. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology* **8**, 761-773 (2007).
6. Fan, J. et al. mRNAs are sorted for export or degradation before passing through nuclear speckles. *Nucleic Acids Res* **46**, 8404-8416 (2018).
7. Wegener, M. & Muller-McNicoll, M. Nuclear retention of mRNAs - quality control, gene regulation and human disease. *Semin Cell Dev Biol* **79**, 131-142 (2018).
8. Soles, L.V. et al. A nuclear RNA degradation code is recognized by PAXT for eukaryotic transcriptome surveillance. *Molecular Cell* **85**, 1575-1588 (2025).
9. Ogami, K. et al. An Mtr4/ZFC3H1 complex facilitates turnover of unstable nuclear RNAs to prevent their cytoplasmic transport and global translational repression. *Genes Dev* **31**, 1257-1271 (2017).
10. Fan, J. et al. Dual modes of ZFC3H1 confer selectivity in nuclear RNA sorting. *Mol Cell* **84**, 4297-4313 e7 (2024).
11. Lee, E.S. et al. ZFC3H1 and U1-70K promote the nuclear retention of mRNAs with 5' splice site motifs within nuclear speckles. *RNA* **28**, 878-894 (2022).
12. Wang, Y. et al. ZFC3H1 prevents RNA trafficking into nuclear speckles through condensation. *Nucleic Acids Res* **49**, 10630-10643 (2021).
13. Yamazaki, T. et al. The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol Biol Cell* **21**, 2953-65 (2010).
14. Fujita, K.I. et al. Structural differences between the closely related RNA helicases, UAP56 and URH49, fashion distinct functional apo-complexes. *Nat Commun* **15**, 455 (2024).
15. Luo, M.L. et al. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413**, 644-7 (2001).
16. Dufu, K. et al. ATP is required for interactions between UAP56 and two conserved mRNA export proteins, Aly and CIP29, to assemble the TREX complex. *Genes Dev* **24**, 2043-53 (2010).
17. Cheng, H. et al. Human mRNA export machinery recruited to the 5' end of mRNA. *Cell* **127**, 1389-400 (2006).
18. Reed, R. & Hurt, E. A conserved mRNA export machinery coupled to pre-mRNA splicing. *Cell* **108**, 523-31 (2002).
19. Zhao, R., Shen, J., Green, M.R., MacMorris, M. & Blumenthal, T. Crystal structure of UAP56, a DExD/H-box protein involved in pre-mRNA splicing and mRNA export. *Structure* **12**, 1373-81 (2004).
20. Xie, Y. et al. Structural basis for high-order complex of SARNP and DDX39B to facilitate

mRNP assembly. *Cell Rep* **42**, 112988 (2023).

21. Ellisdon, A.M., Dimitrova, L., Hurt, E. & Stewart, M. Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. *Nat Struct Mol Biol* **19**, 328-36 (2012).
22. Valkov, E., Dean, J.C., Jani, D., Kuhlmann, S.I. & Stewart, M. Structural basis for the assembly and disassembly of mRNA nuclear export complexes. *Biochim Biophys Acta* **1819**, 578-92 (2012).
23. Umlauf, D. et al. The human TREX-2 complex is stably associated with the nuclear pore basket. *J Cell Sci* **126**, 2656-67 (2013).
24. Hohmann, U. et al. An ATP-gated molecular switch orchestrates human mRNA export. *Nature* (2025).
25. Clarke, B.P. et al. Structural mechanism of DDX39B regulation by human TREX-2 and a related complex in mRNP remodeling. *Nature Communications* **16**(2025).
26. Gong, X. et al. Molecular insights into mRNA export regulation by the human TREX-2 complex. *Nat Commun* **17**(2026).
27. Wickramasinghe, V.O. et al. mRNA export from mammalian cell nuclei is dependent on GANP. *Curr Biol* **20**, 25-31 (2010).
28. Dimitrova, L. et al. Structural Characterization of the Chaetomium thermophilum TREX-2 Complex and its Interaction with the mRNA Nuclear Export Factor Mex67:Mtr2. *Structure* **23**, 1246-57 (2015).
29. Fischer, T. et al. The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J* **21**, 5843-52 (2002).
30. Tian, L. et al. LENG8 mediates RNA nuclear retention and degradation in eukaryotes. *Mol Cell* **86**, 1478-1494 e8 (2026).
31. Aksenova, V. et al. Nucleoporin TPR is an integral component of the TREX-2 mRNA export pathway. *Nat Commun* **11**, 4577 (2020).
32. Silla, T., Karadoulama, E., Makosa, D., Lubas, M. & Jensen, T.H. The RNA Exosome Adaptor ZFC3H1 Functionally Competes with Nuclear Export Activity to Retain Target Transcripts. *Cell Rep* **23**, 2199-2210 (2018).
33. Wickramasinghe, V.O., Stewart, M. & Laskey, R.A. GANP enhances the efficiency of mRNA nuclear export in mammalian cells. *Nucleus* **1**, 393-6 (2010).
34. Coller, J. & Wickens, M. Tethered function assays: an adaptable approach to study RNA regulatory proteins. *Methods Enzymol* **429**, 299-321 (2007).
35. Andrii Bugai, U.H., Ana Lorenzo, Max Graf, Laura Fin, Jérôme O. Rouvière, Laszlo Tirian, Yuhui Dou, Patrik Polák, Dennis Johnsen, Lis Jakobsen, Jens Skorstengaard Andersen, Julius Brennecke, Clemens Plaschka, Torben Heick Jensen. Molecular basis of polyadenylated RNA fate determination in the nucleus. *bioRxiv* (2025).
36. Zhang, X. et al. Structural basis of pre-tRNA intron removal by human tRNA splicing endonuclease. *Mol Cell* **83**, 1328-1339 e4 (2023).
37. DiIorio, M.C. & Kulczyk, A.W. A Robust Single-Particle Cryo-Electron Microscopy (cryo-EM) Processing Workflow with cryoSPARC, RELION, and Scipion. *J Vis Exp* (2022).
38. Meng, E.C. et al. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci* **32**, e4792 (2023).
39. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta*

Crystallographica Section D-Structural Biology **60**, 2126-2132 (2004).

40. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D-Structural Biology* **66**, 213-221 (2010).
41. Davis, I.W. et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* **35**, W375-W383 (2007).
42. Chen, T. et al. iProX in 2021: connecting proteomics data sharing with big data. *Nucleic Acids Research* **50**, D1522-D1527 (2022).

Acknowledgements

We thank Prof. Hong Cheng for valuable suggestions. We thank the Cryo-EM Center at the University of Science and Technology of China for access to cryo-electron microscopy facilities and technical support, and the Core Facility Center for Life Sciences for access to shared instrumentation and technical services. We also thank Mingzhu Fan from the Mass Spectrometry and Metabolomics Core Facility at Westlake University for assistance with LC-MS/MS analysis. **Funding**: This work was supported by funds from the National Natural Science Foundation of China (32471302 and 32671709 to X.Z.; 32601783 to F.Y.), the Natural Science Foundation of Anhui Province (2408085MC068 to X.Z.), the Research Funds of Center for Advanced Interdisciplinary Science and Biomedicine of IHM (QYPY20230023 to X.Z.), and the Major Program of Shenzhen Bay Laboratory (C1012523007).

Author contributions

Xing Gong, M.L., R.T., and T.Z. designed and performed the experiments. Xin Gong and F.Y. prepared cryo-EM samples and collected the cryo-EM data. Xiaofei Ge and X.Z. processed the cryo-EM data and built the atomic model. All authors contributed to data analysis and interpretation. X.Z. supervised the project and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Figure legends

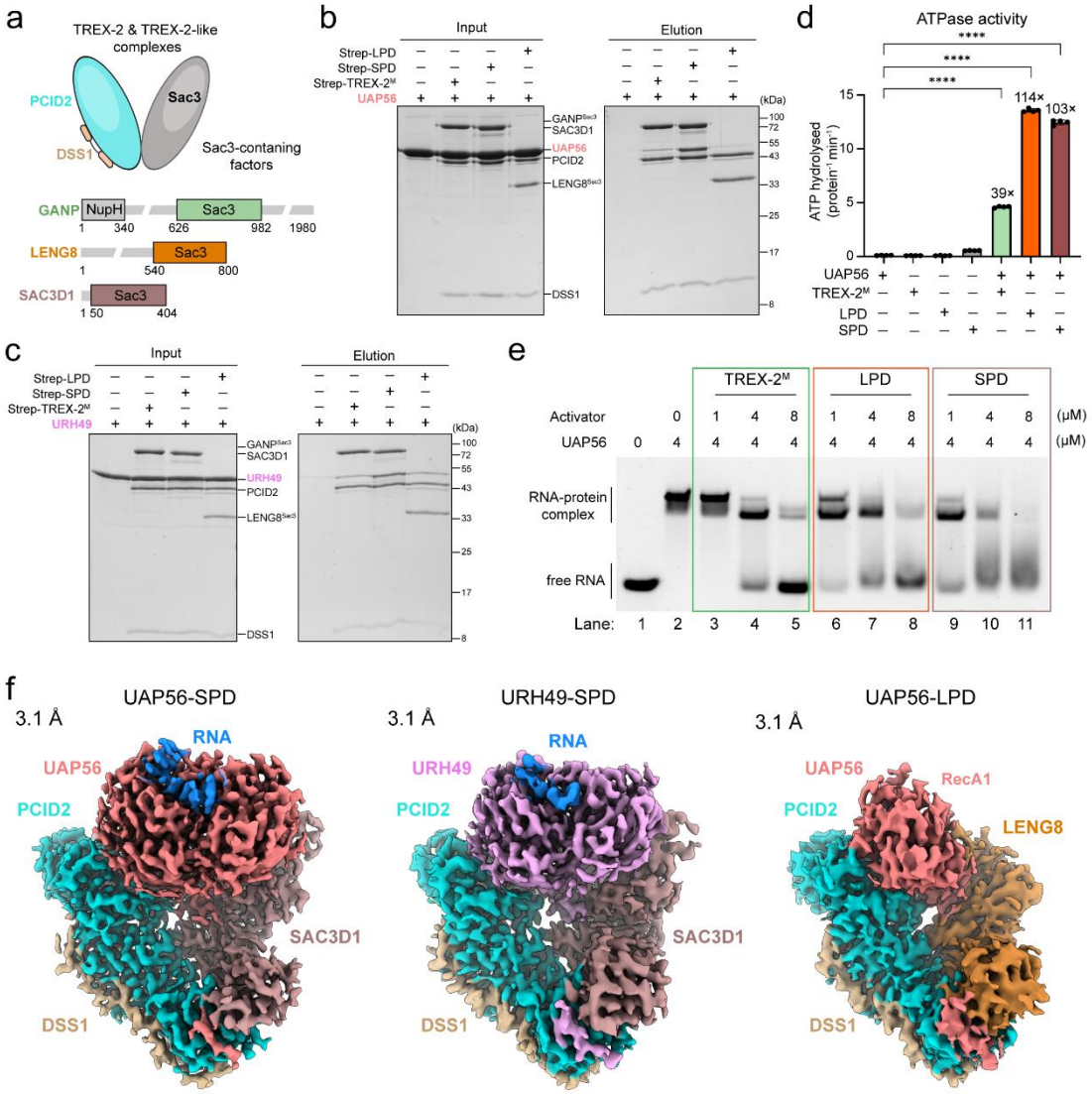

Figure 1 | Biochemical and structural characterization of TREX-2-like complexes

a, Schematic representation of the canonical TREX-2 module and the TREX-2-like LPD and SPD complexes. The domain organizations of GANP, LENG8, and SAC3D1 are shown below; the Sac3 domains associate with PCID2 and DSS1. GANP contains an N-terminal nucleoporin-homology domain (NupH). **b**, Strep-affinity pulldown assays examining the interactions of UAP56 with TREX-2^{M}, LPD, and SPD. Input and elution fractions were analyzed by SDS-PAGE. **c**, Strep-affinity pulldown assays examining the interactions of URH49 with the indicated Sac3-family complexes.

Experiments were independently repeated three times with similar results. **d**, RNA-dependent ATPase activity of UAP56 alone or in the presence of TREX-2^{M}, LPD, or SPD. Values above the bars indicate fold stimulation relative to UAP56 alone. Data are mean ± s.d. from four independent experiments. Statistical significance was determined using one-way ANOVA followed by Dunnett's multiple comparisons test, with UAP56 alone serving as the control; **** $P < 0.0001$. **e**, Native electrophoretic mobility shift assays showing concentration-dependent release of RNA from preassembled UAP56-RNA complex upon addition of TREX-2^{M}, LPD, or SPD. **f**, Cryo-EM density maps of UAP56-SPD, URH49-SPD, and UAP56-LPD at the indicated overall resolutions. Individual components are colored as indicated.

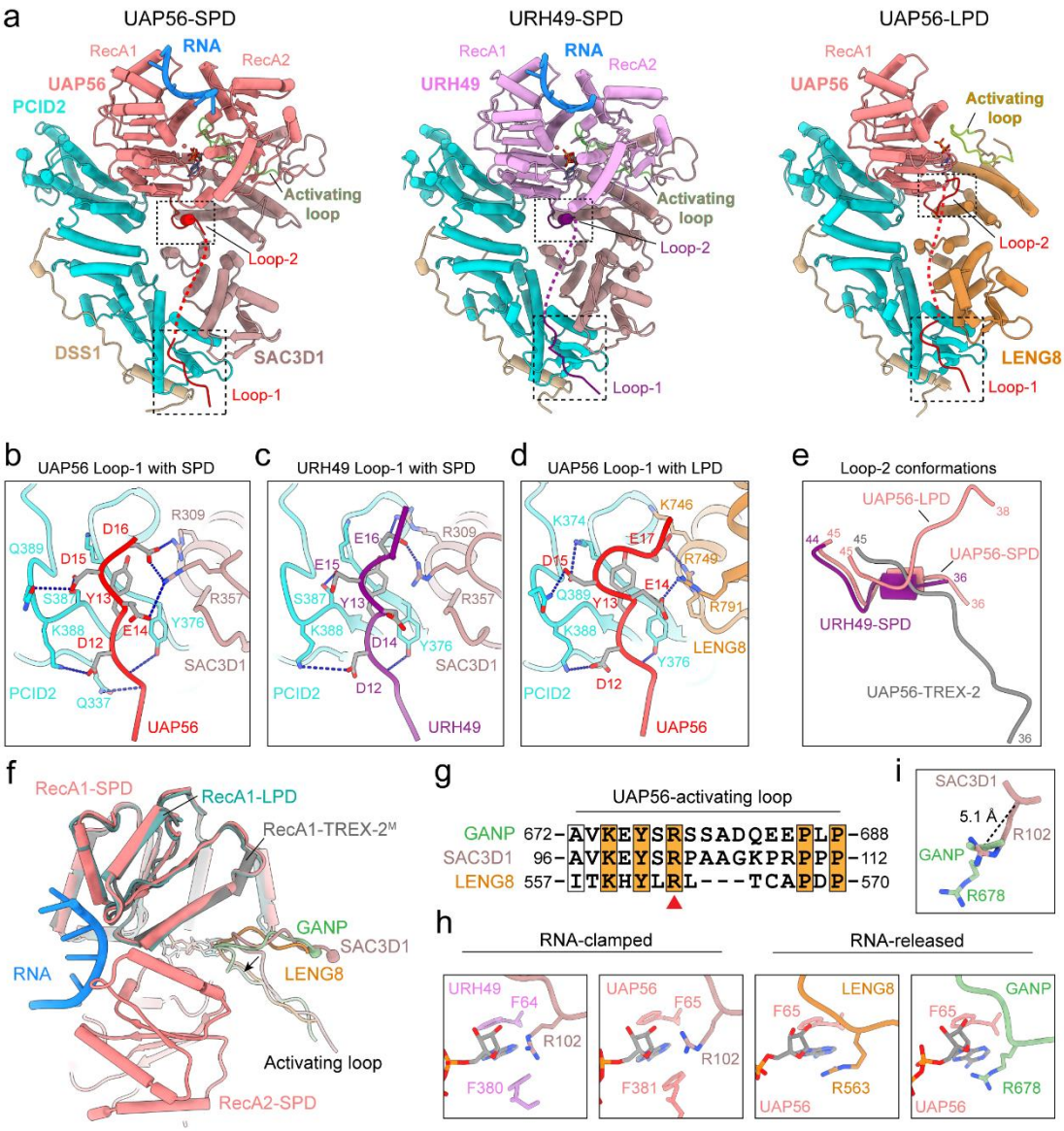

Figure 2 | A conserved activating-loop mechanism remodels UAP56 and URH49

a, Overall structures of UAP56-SPD, URH49-SPD, and UAP56-LPD complexes. Boxed regions highlight three conserved structural elements: Loop-1, Loop-2, and the UAP56-activating loop (UAL). The ATPase/helicase UAP56 and URH49 adopt closed conformations and clamp the bound RNA in the UAP56-SPD and URH49-SPD structures. By contrast, UAP56 adopts an open conformation in UAP56-LPD complex, accompanied by RNA release. **b-d**, Close-up views of the Loop-1 interfaces in UAP56-SPD, URH49-SPD, and UAP56-LPD, respectively. Selected interacting

residues are shown as sticks, and hydrogen bonds or electrostatic contacts are indicated by dashed lines. **e**, Comparison of the Loop-2 conformations in the UAP56-LPD, UAP56-SPD, URH49-SPD, and the UAP56-TREX-2^{M} (PDB: 9UPB[26]). **f**, Superposition of the RecA1 domains and UAP56-activating-loops in SPD-, LPD-, and TREX-2^{M}-bound complexes. RNA is clamped by the RecA1 and RecA2 domains in the UAP56-SPD structure, but is absent from the UAP56-LPD and UAP56-TREX-2^{M} structures. **g**, Sequence alignment of the UALs from GANP, SAC3D1, and LENG8. The conserved arginine residue is indicated by a red triangle. **h**, Comparison of nucleotide coordination in the RNA-clamped and RNA-released states. In the clamped state, the adenine moiety is sandwiched between Phe65 and Phe381 in UAP56, or between Phe64 and Phe380 in URH49. In the released state, the conserved UAL arginine occupies the position otherwise taken by Phe381. **i**, Comparison of the conserved UAL arginine residues in SAC3D1 and GANP. The approximately 5.1 Å displacement is indicated.

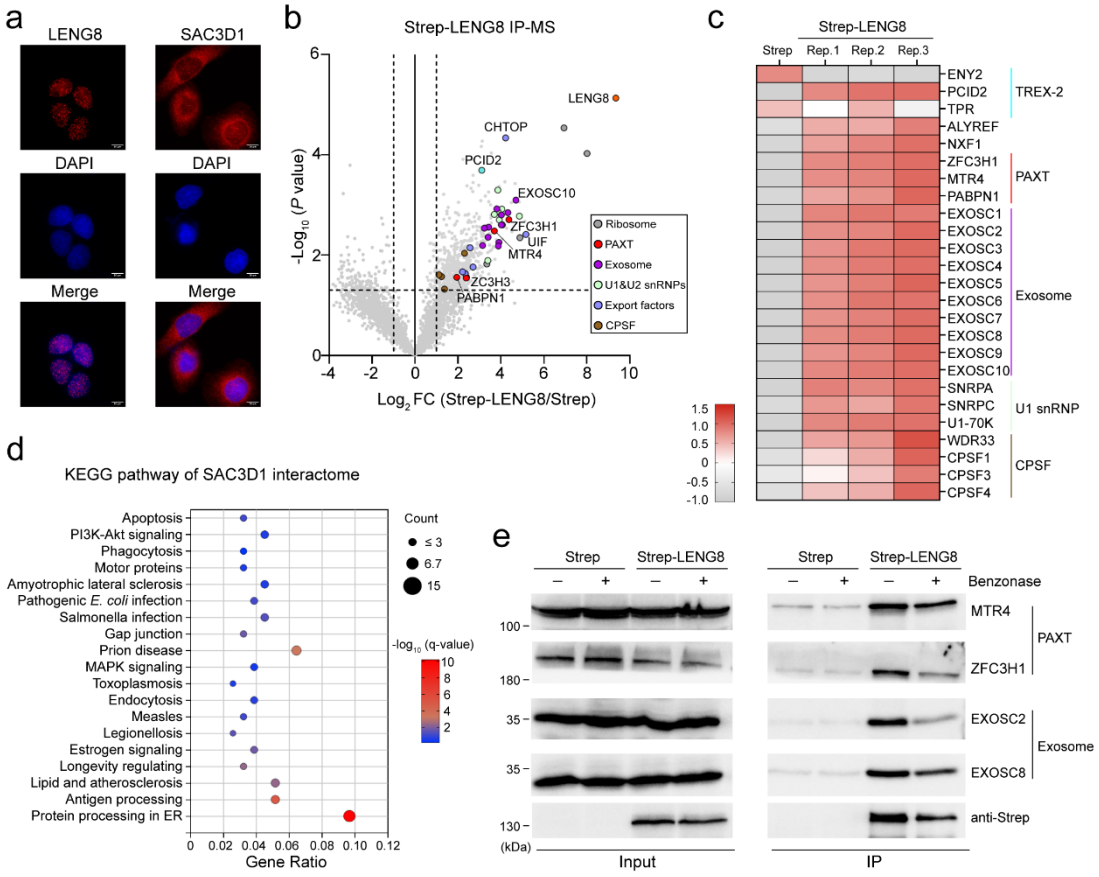

Figure 3 | LENG8 associates with the nuclear PAXT-exosome machinery

a, Representative fluorescence microscopy images showing the subcellular localization of LENG8 and SAC3D1. DAPI marks nuclei. Scale bars, 10 μm. **b**, Volcano plot showing proteins enriched by Strep-LENG8 affinity purification coupled to mass spectrometry relative to the Strep control. Selected factors associated with mRNA export, PAXT, the nuclear RNA exosome, the spliceosomal complexes, and cleavage and polyadenylation (CPSF) are labeled. **c**, Heat map showing enrichment of representative LENG8-interacting proteins across biological replicates. Proteins are grouped by functional complex. **d**, KEGG pathway enrichment analysis of the SAC3D1 interactome. Dot size represents protein count, the x axis represents gene ratio, and color denotes enrichment significance. **e**, Co-immunoprecipitation of

endogenous MTR4, ZFC3H1, EXOSC2, and EXOSC8 with Strep-LENG8 in the absence or presence of Benzonase. Input and immunoprecipitated fractions were analyzed by immunoblotting.

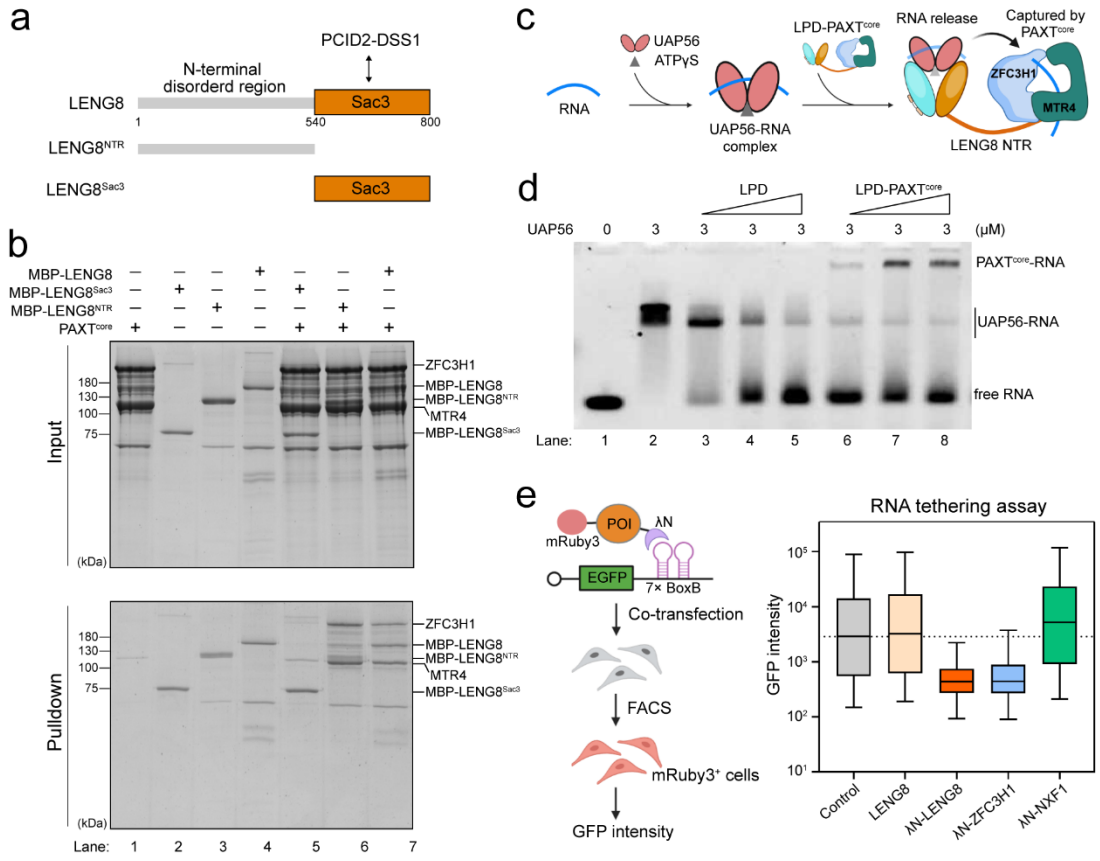

Figure 4 | LENG8 couples UAP56 remodeling to PAXT-mediated RNA capture

a, Domain organization of LENG8 and schematic representation of the truncation constructs used in this study. The N-terminal region (LENG8^{NTR}) and the C-terminal Sac3 domain (LENG8^{Sac3}) are indicated. The Sac3 domain associates with PCID2-DSS1. **b**, MBP pulldown assays examining the interactions of full-length LENG8, LENG8^{NTR}, or the LENG8^{Sac3} with the PAXT^{core} complex comprising full-length ZFC3H1 and MTR4. Input and pulldown fractions were analyzed by SDS-PAGE. Experiments were independently repeated three times with similar results. **c**, Schematic representation of the reconstituted RNA-release and RNA-capture assay. The Sac3 domain of LENG8 assembles with PCID2-DSS1 to form the TREX-2-like LPD complex, which promotes RNA release from UAP56, whereas the N-terminal region (NTR) of LENG8 recruits the PAXT^{core} complex to capture the released RNA.

d, Native gel analysis of the reconstituted RNA-release and RNA-capture reaction. Bands corresponding to free RNA, the UAP56-RNA complex, and the PAXT^{core}-RNA complex are indicated. Addition of LPD into preassembled UAP56-RNA complex promotes RNA release from UAP56, and the liberated RNA was captured by PAXT^{core}. Experiments were independently repeated three times with similar results. **e**, Schematic representation of the λN-BoxB RNA tethering assay. Seven BoxB RNA elements were inserted downstream of the EGFP coding sequence. The indicated λN-fusion proteins, carrying a C-terminal mRuby3 transfection marker, were tethered to the reporter RNA through λN-BoxB interactions. POI: protein of interest. GFP fluorescence was quantified by flow cytometry within the mRuby3-positive cell population. Efficient nuclear export and expression of the reporter RNA result in GFP fluorescence, while nuclear retention and/or degradation of the reporter RNA are expected to reduce the GFP fluorescence. Tethering of LENG8 to the reporter significantly decreased GFP fluorescence, whereas expression of LENG8 lacking the λN tether had no detectable effect. Tethering of ZFC3H1 similarly reduced GFP fluorescence. By contrast, tethering of the mRNA export receptor NXF1 increased GFP fluorescence. Box-and-whisker plots show the median (center line), interquartile range (box), and 5th-95th percentiles (whiskers).

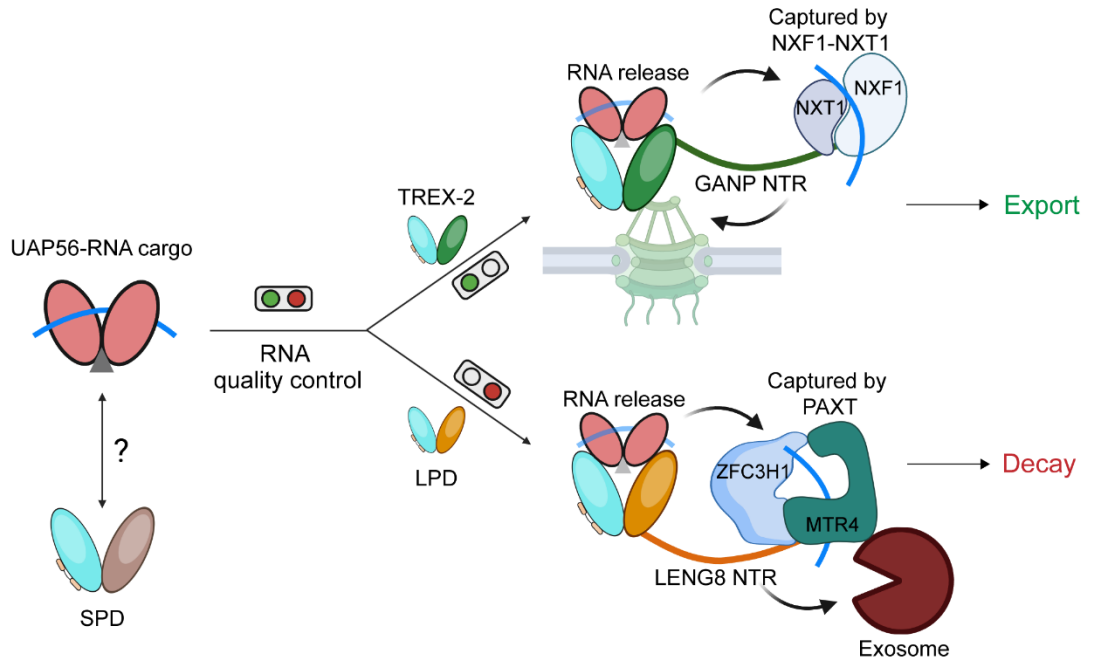

Figure 5 | Model for UAP56-dependent sorting of nuclear mRNP between export and decay

During nuclear mRNP assembly, UAP56 binds the RNA cargo. Engagement of UAP56 by the GANP-containing TREX-2 complex stimulates its ATPase activity and promotes RNA release. The released RNA is subsequently captured by NXF1-NXT1 through the bridge of N-terminal region of GANP, thereby directing the transcript toward nuclear export. By contrast, engagement by LPD triggers a similar UAP56-remodeling and RNA-release reaction, but the N-terminal region of LENG8 recruits PAXT and the nuclear RNA exosome, directing the transcript toward nuclear degradation. SPD also remodels UAP56 and URH49 *in vitro*, but its downstream pathway and biological significance remain to be defined.