Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2023 June 01.

Published in final edited form as:

Nat Chem Biol. 2022 June ; 18(6): 643–651. doi:10.1038/s41589-022-01003-9.

Nascent alt-protein chemoproteomics reveals a pre-60S assembly checkpoint inhibitor

Xiongwen Cao[1,2], **Alexandra Khitun**[1,2], **Cecelia M. Harold**[3], **Carson J. Bryant**[4], **Shu-Jian Zheng**[1,2], **Susan J. Baserga**[3,4,5], **Sarah A. Slavoff**[1,2,4]

[1]Department of Chemistry, Yale University, New Haven, Connecticut 06520, United States

[2]Institute of Biomolecular Design and Discovery, Yale University, West Haven, Connecticut 06516, United States

[3]Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06520, United States

[4]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, United States

[5]Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, Connecticut 06520, United States

Abstract

Many unannotated microproteins and alternative proteins (alt-proteins) are co-encoded with canonical proteins, but few of their functions are known. Motivated by the hypothesis that alt-proteins undergoing regulated synthesis could play important cellular roles, we developed a chemoproteomic pipeline to identify nascent alt-proteins in human cells. We identified 22 actively translated alt-proteins or N-terminal extensions, one of which is post-transcriptionally upregulated by DNA damage stress. We further defined cell cycle-regulated MINAS-60 (*MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60*S ribosomal subunit), a nucleolar alt-protein co-encoded with human RBM10. Depletion of MINAS-60 increases the amount of cytoplasmic 60S ribosomal subunit, upregulating global protein synthesis and cell proliferation. Mechanistically, MINAS-60 represses the rate of late-stage pre-60S assembly and export to the cytoplasm. Together, these results implicate MINAS-60 as a potential checkpoint inhibitor of pre-60S assembly, and demonstrate that chemoproteomics enables hypothesis generation for uncharacterized alt-proteins.

Graphical Abstract

*Correspondence: sarah.slavoff@yale.edu.

Author contributions

X.C. and A.K. designed and performed the chemoproteomic profiling, X.C. performed functional assays, C.M.H. and C.J.B. performed northern blotting, X.C. and S.Z. performed LC-MS/MS experiments. S. A. S. and S. J. B. designed experiments and analyzed data. X. C. and S. A. S. wrote the manuscript, and all authors edited and approved the final version of the manuscript.

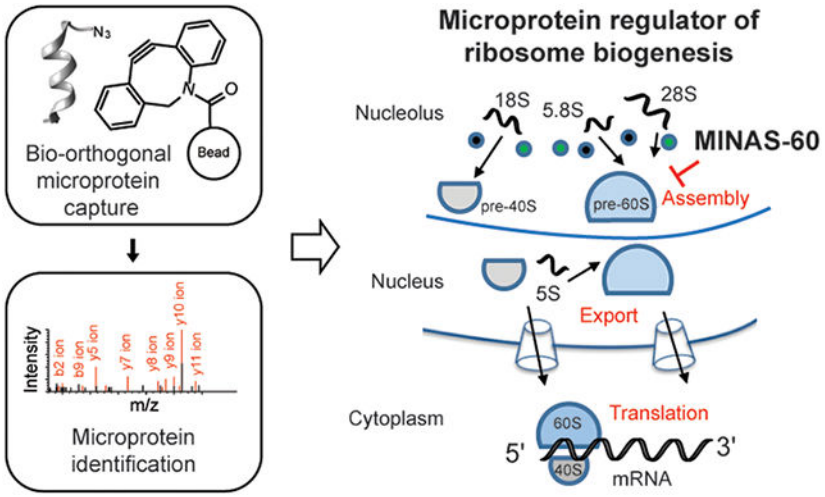
Competing interests

The authors declare no competing interests.

Code availability

All custom code for alt- and microprotein identification and database construction are available at Zenodo (DOI: 10.5281/zenodo.5921116), at https://slavofflab.yale.edu/code or via request.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Introduction

Expression of thousands of unannotated small open reading frames (smORFs, fewer than 150 codons) has recently been revealed in mammalian cells[1]. These smORFs are found in long non-coding RNAs, 5' and 3' untranslated regions (UTRs) of mRNAs, and frame-shifted ORFs overlapping protein coding sequences (CDS), the latter of which are termed alternative ORFs (alt-ORFs)[2]. Some smORF-encoded proteins (SEPs, micropeptides or microproteins) and alt-ORF-encoded proteins (alt-proteins) play important roles in vertebrate biology[3,4]. Fewer alt-proteins have been well-defined, including alt-FUS, which cooperates with FUS in formation of cytotoxic aggregates[5], and alt-RPL36, which regulates PI3K-AKT-mTOR signaling[6]. Defining the functions of SEPs and alt-proteins therefore represents a major opportunity to gain insights into biology.

Currently, the overwhelming majority of smORFs and alt-ORFs remain undefined because their short lengths and intermediate conservation render homology-based annotation challenging[7]. Furthermore, it is challenging to separate the functions of alt-ORFs from the protein CDS in which they are nested. It therefore remains unclear whether alt-proteins are broadly functional. We hypothesize that bio-orthogonal protein labeling strategies combined with proteomic analysis – collectively termed “chemoproteomics” – can be leveraged to identify alt-proteins with properties akin to canonical proteins, which may have important cellular roles. Providing precedent, chemoproteomic profiling of cysteine-containing microproteins, a feature of some enzyme active sites, revealed 16 novel human microproteins[8]. In this work, we test the hypothesis that alt-proteins undergoing high or regulated synthesis are likely functional, and develop a chemoproteomic approach to identify them.

Using this method, we identify an alt-ORF overlapping human RBM10 that we propose encodes a checkpoint inhibitor of ribosome large subunit (LSU) biogenesis. Ribosome biogenesis is a spatially and temporally regulated cellular process essential for growth and development[9,10]. Dysregulated ribosome biogenesis has been linked to human disorders including cancer[11], Alzheimer’s disease[12], and congenital ribosomopathies[13]. In humans,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ribosome biogenesis starts with nucleolar transcription of the 47S precursor rRNA (pre-rRNA), which is modified and complexed with assembly factors and ribosomal proteins to form the 90S pre-ribosomal particle. Endonucleolytic cleavage of the pre-rRNA subsequently generates the rRNA components of the pre-40S and pre-60S particles. While the large and small subunit rRNAs are produced stoichiometrically from cleavage of a single pre-rRNA, the pre-60S and pre-40S subunit undergo distinct maturation pathways prior to export to the cytoplasm via independent transporters. Pulse-chase radiolabeling experiments in *Tetrahymena* demonstrated that newly synthesized pre-60S appears in the cytoplasm later than pre-40S[14], suggesting that pre-60S matures more slowly. Consistent with this observation cryo-electron microscopy (cryo-EM)[15,16] has revealed quality control checkpoints for 5S ribonucleoprotein particle incorporation and rotation, active site formation, and removal of internal transcribed spacer 2 (ITS2) prior to pre-60S export. Once in the cytoplasm, the final steps of maturation occur to produce large 60S and small 40S subunits, which associate to form translation-competent ribosomes.

Here, we developed a chemoproteomic pipeline to identify nascent alt-proteins, which improves the sensitivity of metabolic unnatural amino acid incorporation technology[17] for microprotein enrichment. We identified 22 actively translated human alt-proteins or N-terminal extensions of canonical proteins[18]. We confirmed the translation of six selected alt-proteins, and functionally defined an alt-ORF nested in the human RBM10 CDS. We name this nucleolar alt-protein MINAS-60 (*MI*croprotein that *N*egatively regulates *AS*sembly of the pre-*60*S ribosomal subunit), which associates with LSU assembly factors GTPBP4 and MRTO4. Loss of MINAS-60 increases global protein synthesis and cell proliferation by increasing the rate of late-stage pre-60S assembly and export into the cytoplasm. Over-expression of MINAS-60 slightly decreases cytoplasmic 60S levels and significantly decreases 80S levels, supporting a role in inspecting correctly assembled, functional LSU. These results implicate MINAS-60 as a checkpoint inhibitor regulating late nuclear steps in human LSU biogenesis. More broadly, our results demonstrate that chemoproteomics can prioritize alt-ORFs for functional study.

Results

Chemoproteomic profiling of nascent alt-proteins

Motivated by the hypothesis that alt-proteins undergoing regulated synthesis could play important cellular roles, we developed a chemoproteomic approach to identify them (Fig. 1a). We leveraged bio-orthogonal non-canonical amino acid tagging (BONCAT)[17], in which the methionine analogue azidohomoalanine (AHA) is metabolically incorporated into all newly synthesized proteins by the endogenous protein translation machinery. Previous BONCAT workflows derivatized AHA-labeled proteins with biotin-alkyne, requiring column-based removal of excess biotin-alkyne prior to streptavidin enrichment – a step that may eliminate small proteins. We therefore designed a protocol to enrich small proteins from whole proteomes after AHA-labeling with a C8 column[19], followed by copper-free Click chemistry capture directly on dibenzocyclooctyne magnetic beads. On-bead digest was followed by liquid chromatography/tandem mass spectrometry-based alt-protein identification[20]. Our C8 column/on-bead BONCAT workflow detected more than

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

twice as many canonical small proteins as biotin-streptavidin/peptide gel-based BONCAT capture (Supplementary Table 1). Consistent with a previous report[19], C8 columns afford less stringent small protein enrichment than peptide gels (Extended Data Fig. 1a). The expression level and size distribution of gene products detected in AHA-labeled samples is comparable to unlabeled samples, suggesting that our method is unbiased with respect to protein length and abundance (Extended Data Fig. 1a-b).

We then profiled nascent small proteins in HEK 293T cells under control conditions (DMSO treatment), oxidative stress (sodium arsenite treatment), DNA damage stress (etoposide treatment) and unfolded protein response stress (DTT treatment) (Supplementary Table 2 and Extended Data Fig. 1c). Searching our data against the annotated human proteome database followed by GO analysis revealed gene expression signatures consistent with the stress conditions applied (Extended Data Fig. 1d-e). We then searched the data against a 3-frame RNA-seq translation database with stringent peptide-spectral match parameters as previously reported[7], leading to identification of a total of 22 unannotated alt-proteins or N-terminal extensions of canonical proteins (Extended Data Fig. 1f, Supplementary Data S1). Nine were only detected under stress, including alt-CNPY2, which was observed after etoposide treatment (Extended Data Fig. 1g, Extended Data Fig. 2a), and 16 were detected with mass spectrometry in this study for the first time[7,19,21-23] (Supplementary Data S1).

We validated translation of six alt-proteins detected under various conditions in our screen (Extended Data Fig. 2b-h, Supplementary Table 3). The cDNA sequence comprising the 5'UTR of the encoding transcript through the stop codon of the putative alt-ORF was over-expressed from a mammalian expression vector with a FLAG-HA epitope tag on the C-terminus of the alt-protein, followed with immunostaining. As shown in Figure 1b and Extended Data Figure 1h, over-expressed alt-DRAP1 (previously validated and included as a positive control[7]), alt-PRR3, alt-PRH1, alt-CNPY2 and alt-CACTIN exhibited nucleocytoplasmic immunofluorescence. Over-expressed MINAS-60 co-localized with a nucleolar protein, fibrillarin (Fig. 1c). Immunoblotting confirmed the sizes of the six translated alt-proteins (Fig. 1d). MINAS-60 and alt-CNPY2 produced two immunoreactive bands, which is due to multiple in-frame start codon for MINAS-60 (vide infra), and may be caused by multiple start sites or post-translational modifications for alt-CNPY2.

To determine whether expression of alt-CNPY2 is induced by DNA damage, we treated HEK 293T over-expressing alt-CNPY2 with etoposide, followed by immunoblotting. As shown in Figure 1e, alt-CNPY2 exhibited dose- and time-dependent upregulation upon etoposide treatment of up to two fold. And the mRNA level of *CNPY2* did not change (Fig. 1f), indicating that upregulation of alt-CNPY2 is likely post-transcriptional. Taken together, these results suggest that our chemoproteomic pipeline can detect alt-proteins undergoing active synthesis during normal cell growth and cellular stress.

MINAS-60 is cell-cycle regulated and conserved in mouse

We selected MINAS-60, which was detected in normally growing cells, for further study, because it localizes to the nucleolus, the site of ribosome biogenesis. MINAS-60 is nested within the CDS of human RBM10, so probing its biological role could shed light on the under-explored class of overlapping alt-proteins (Fig. 2a, Supplementary Data S2-4). A

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

different MINAS-60 tryptic peptide was previously detected in colorectal cancer, supporting expression of MINAS-60 in human tissue[22]. To validate expression of MINAS-60 from the *RBM10* genomic locus, we generated two independent Cas9-directed knock-in (KI) HEK 293T cell lines with a 3×GFP11-FLAG-HA tag appended to the 3' end of MINAS-60 alt-ORF[24], followed with immunostaining. As shown in Figure 2b, endogenously expressed MINAS-60 co-localizes with fibrillarin, suggesting MINAS-60 functions in the nucleolus.

The nucleolus is the site of ribosome biogenesis in eukaryotic cells. Ribosome biogenesis starts with the transcription of rDNA, which oscillates during the cell cycle, nearly ceasing during M phase, increasing during G1, and maximizing during S and G2 in human cells[25-27]. We hypothesized that cell cycle expression of MINAS-60 might be correlated with ribosome biogenesis activity if it functions in this process. Immunostaining of synchronized MINAS-60 KI cells revealed that MINAS-60 increased at early S phase, peaking at the 2 h time point. MINAS-60 decreased by late S phase, and was very low during G2/M. At G1, MINAS-60 intensity again increased (Extended Data Fig. 3a). These results were confirmed with immunoblotting (Fig. 2c). MINAS-60 expression is therefore coordinated with ribosome biogenesis activity during the cell cycle.

We then identified the start codon(s) that initiate MINAS-60 translation. We tested two upstream in-frame non-AUG start codons[6,7,28], A_{383}TC and A_{386}GG (numbered relative to the first nucleotide of the cDNA), as well as seven internal AUG start codons[5,29] (numbered ATG1 – ATG7) relative to the RBM10 CDS (Extended Data Fig. 3b). The ~20 kDa MINAS-60 isoform initiates at ATG1 and the ~15 kDa MINAS-60 isoform at ATG6 or 7, because the indicated species is abrogated only when the corresponding start codon is deleted or mutated (Extended Data Fig. 3c-d). These results were confirmed by over-expressing truncated MINAS-60 coding sequences starting from ATG1 or ATG6 and comparing their products' sizes with the wild-type construct (Extended Data Fig. 3e). Both the long (~20 kDa) and short (~15 kDa) isoforms of MINAS-60 localize to the nucleolus when over-expressed, suggesting that both may be functional (Extended Data Fig. 3f). The MINAS-60 smORF is therefore entirely contained within the RBM10 coding sequence (Fig. 2a). We speculate that translation of these two MINAS-60 isoforms from *RBM10* transcript variant 1 may occur via leaky scanning, and that the additional MINAS-60 isoform diversity observed in KI cells could be generated by alternative splicing, analogous to PTBP3[29].

ClustalW alignment of hypothetical murine, bovine, and primate MINAS-60 homologs revealed significant sequence similarity, suggesting MINAS-60 is conserved in mammals (Fig. 2d). To determine whether mouse MINAS-60 can be translated, an expression vector containing the 5'UTR of mouse *RBM10* transcript variant 1 through the stop codon of mouse MINAS-60 was transfected into 3T3 cells, followed with immunostaining. As shown in Figure 2e, over-expressed mouse MINAS-60 appeared in the nucleolus. Together, these results indicate that MINAS-60 is endogenously expressed, cell-cycle regulated and conserved from humans to mouse.

MINAS-60 is associated with nucleolar LSU assembly factors

Because many microproteins and alt-proteins bind to and regulate other proteins[30], we performed a co-immunoprecipitation (co-IP) with nuclear lysates from MINAS-60-FLAG

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KI cells, and HEK 293T cells as a control for non-specific bead-binding proteins[31]. Two specific SDS-PAGE bands were observed in the KI co-IP that, upon label-free quantitative proteomics[32], yielded 17 proteins enriched >30-fold over control. GO analysis showed that the top 2 enriched biological processes were ribosome biogenesis and ribosomal large subunit (LSU) biogenesis (Extended Data Fig. 4a-c, and Supplementary Data S5-6). To obtain a more comprehensive picture of MINAS-60-associated proteins, we performed co-IPs and analyzed the entire molecular weight range with quantitative proteomics; while abundant ribosomal proteins limited detection sensitivity for other proteins, we observed enrichment of four LSU assembly factors: GTPBP4, MRTO4, BRIX1 and NOP2 (Fig. 3a and Supplementary Data S7-8). Co-IP followed by immunoblotting confirmed the association of these four factors with MINAS-60 (Fig. 3b). Interaction of MINAS-60 with GTPBP4 and MRTO4 did not depend on RNA, because these associations largely survived RNaseA treatment. In contrast, co-purification of MINAS-60 with BRIX1 and NOP2 was diminished after treatment with RNaseA, suggesting that their association with MINAS-60 was likely indirect. We therefore hypothesized that MINAS-60 interacts with nucleolar, late pre-60S particles containing GTPBP4 and MRTO4 to regulate ribosome biogenesis[15,33]. Additional negative controls, including HEK 293T cells stably transformed with empty vector, could be useful in further refining the remaining MINAS-60-enriched proteins listed in Supplementary Data S6 and S8.

To determine whether MINAS-60 associates with pre-60S particles, we performed sucrose gradient fractionation of nuclear extracts of HEK 293T cells stably expressing epitope-tagged MINAS-60, followed by immunoblotting. A subpopulation of over-expressed MINAS-60 co-sedimented with ribosome assembly factors in high-molecular weight fractions coincident with pre-ribosomal particles (Fig. 3c), consistent with a role for MINAS-60 in pre-60S biogenesis.

MINAS-60 inhibits protein synthesis and cell proliferation

We hypothesized that MINAS-60 could regulate ribosome biogenesis via its association with GTPBP4 and MRTO4. Ribosome biogenesis is required for protein synthesis, which promotes cell growth and proliferation, and thus is commonly upregulated in cancer cells[11,34]. We reasoned that, if MINAS-60 regulates ribosome biogenesis, its absence should result in changes to protein synthesis and cell proliferation. We required a system to query the function of MINAS-60 independent of RBM10, despite their co-expression from the same transcript. To this end, we knocked down (KD) *RBM10* in HEK 293T cells with two different shRNAs, which silence the entire mRNA and both proteins. To deconvolute phenotypic effects specific to MINAS-60, as well as to exclude off-target effects of shRNA, we generated rescue cell lines stably expressing MINAS-60 (Rescue_MINAS-60) or RBM10 (Rescue_RBM10) on the KD background. qRT-PCR and immunoblotting revealed that *RBM10* mRNA is efficiently silenced by both shRNAs, and the KD cells express MINAS-60 or RBM10 after rescue (Extended Data Fig. 5a-b). Noting that the RBM10 rescue construct could be subject to leaky translation to produce both RBM10 and MINAS-60, we also rescued the KD with an RBM10 construct bearing an $A_{398}TG$ to TAA mutation that eliminates the first MINAS-60 start codon while preserving RBM10 translation (Extended Data Figure 5a-b).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To test the effect of MINAS-60 on cellular protein synthesis, we labeled nascent peptides in the control, *RBM10* KD, Rescue_MINAS-60, Rescue_RBM10 and Rescue_RBM10(A_{398}TG-TAA) cell lines with puromycin followed by anti-puromycin immunoblotting[35]. As shown in Figure 3d and Extended Data Figure 5c, KD of *RBM10* led to a significant increase in global protein synthesis, which was rescued by reintroduction of MINAS-60. Partial rescue by RBM10 was also observed, but was likely due to leaky MINAS-60 expression, because this effect was not present in the RBM10(A_{398}TG-TAA) rescue cells. Similar results were observed for a second shRNA targeting *RBM10* (Extended Data Fig. 5d-e). MINAS-60, and not RBM10, may therefore play a negative regulatory role in ribosome biogenesis.

We next sought to determine which regions of MINAS-60 are required for its protein-protein interactions and inhibitory effect on protein translation. We transfected deletion mutants of MINAS-60 into HEK 293T cells, followed by FLAG-IP and immunoblotting. As shown in Extended Data Figure 6a, deletion of the N-terminal or middle region of MINAS-60 maintained association with GTPBP4 and MRTO4, while deletion of the C-terminal region abolished the association. We then examined the effect of the deletion constructs on global protein synthesis in HEK 293T. As shown in Extended Data Figures 6b-c, over-expression of MINAS-60 full-length, N-terminal deletion or middle region deletion reduced puromycylation, while over-expression of MINAS-60 C-terminal deletion did not, compared with empty vector control. The C-terminus of MINAS-60 is therefore required for association with GTPBP4 and MRTO4 and negative regulation of ribosome biogenesis.

Because protein synthesis, cell growth and proliferation are linked[10,11], we asked whether MINAS-60 regulates cell proliferation. As shown in Figure 3e, *RBM10* depletion led to a significant increase in cell proliferation, consistent with a published report[36]. Remarkably, this increase was rescued by reintroduction of MINAS-60 alone, and partially rescued by RBM10. Similar results were observed for a second shRNA targeting *RBM10* (Extended Data Fig. 5f). These results suggest that MINAS-60 negatively regulates ribosome biogenesis, subsequently decreasing cellular protein synthesis and cell proliferation.

MINAS-60 downregulates cytoplasmic LSU export and function

Eukaryotic ribosome biogenesis can be divided into sequential processes[37]: pre-rRNA transcription; chemical modification and processing of the pre-rRNA in the nucleolus; folding, assembly and maturation of the pre-ribosomal subunits in the nucleolus and nucleus; and export and final maturation of ribosome subunits in the cytoplasm. Because MINAS-60 co-purified with proteins involved in various steps of LSU biogenesis, we wished to determine the step in this process that it regulates.

To determine whether MINAS-60 controls pre-rRNA transcription, we performed qRT-PCR targeting the primary pre-rRNA (47S/45S/30S) using previously published primers[35] in lysates from control, *RBM10* KD, Rescue_MINAS-60 and Rescue_RBM10 cells. No significant differences were observed (Extended Data Fig. 7a). To determine whether MINAS-60 regulates the processing of LSU pre-rRNA, we performed northern blotting analysis with these cell lines using a probe complementary to ITS2, which detects all LSU pre-rRNA processing products[35]. Again, no significant differences were observed (Extended

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Data Fig. 7b). Finally, no significant differences were observed in the ratio of the mature 28S/18S rRNAs in whole-cell lysates (Extended Data Fig. 7c). MINAS-60 therefore does not regulate transcription or processing of the pre-rRNA.

To determine whether MINAS-60 regulates nucleocytoplasmic export of the pre-LSU, we quantified the ratio of nuclear vs. cytoplasmic RPL29-GFP[38] stably expressed in the control, KD, and rescue cell lines. As shown in Figure 4a and Extended Data Figure 8a, KD of *RBM10* decreased the ratio of nuclear to cytoplasmic RPL29-GFP, which can be rescued by reintroduction of MINAS-60, and partially rescued by RBM10 reintroduction, likely due to leaky MINAS-60 translation. Similar results were observed for a second shRNA targeting *RBM10* (Fig. 4a, right, and Extended Data Fig. 8b-c). As a control, to determine whether MINAS-60 regulates nucleocytoplasmic export of the pre-40S ribosomal subunit, we quantified the ratio of nuclear vs. cytoplasmic RPS2-GFP[38] stably expressed in these cell lines. No significant changes were observed, though RPS2-GFP signal was predominantly cytoplasmic, consistent with prior reports of faster pre-40S maturation and export[14,38] (Extended Data Fig. 9). MINAS-60 therefore may specifically inhibit a step in the LSU biogenesis pathway between rRNA processing and cytoplasmic export.

The observation of faster pre-60S export in the absence of MINAS-60 predicted that the same conditions should increase mature cytoplasmic 60S ribosomal subunits. To test this, we performed cytoplasmic polysome profiling. As shown in Figures 4b-c, knockdown of *RBM10* resulted in a small but significant increase in the ratio of cytoplasmic 60S/40S ribosome subunits. This increase was rescued by reintroduction of MINAS-60, and partially rescued by RBM10. Similar results were observed for a second shRNA targeting *RBM10* (Fig. 4c, right and Extended Data Fig. 10a). Conversely, over-expression of MINAS-60 led to a small but significant decrease in the ratio of cytoplasmic 60S/40S ribosome subunits as compared to empty vector (Extended Data Figure 10b-c). More significantly, over-expression of MINAS-60 dramatically decreased the levels of 80S monosomes (Extended Data Figure 10b-c), suggesting that fewer non-functional ribosomes are produced when MINAS-60 levels are high. Taken together, these results suggest that MINAS-60 negatively regulates the rate of assembly and nucleocytoplasmic export of pre-60S particles to ensure that they are correctly assembled.

MINAS-60 is a checkpoint inhibitor of pre-60S assembly

We speculated that MINAS-60 is a checkpoint inhibitor that ensures correct pre-60S assembly prior to export from the nucleus, which would suggest that the protein composition of pre-60S particles should change in cells lacking MINAS-60. BRIX1 and MRTO4 are markers of early-to-intermediate, or intermediate-to-late, pre-60S particles, respectively[15]. We examined their interactomes as a readout of changes in pre-60S protein composition in the presence or absence of MINAS-60 (Fig. 5a). Affinity purification of MRTO4-FLAG in control or *RBM10* KD HEK 293T cells, followed by quantitative proteomics, identified statistically significant increases in several intermediate- and late-stage LSU assembly factors co-purified with MRTO4 in *RBM10* KD cells (Fig. 5b, Supplementary Data S9-10). These increases were further confirmed by immunoblotting (Fig. 5c). Similar results were observed for BRIX1 co-IP, though fold changes were smaller (Extended Data Fig. 10d-f,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Supplementary Data S11-12). These results suggest that increased interaction between pre-60S assembly factors occurs in the absence of MINAS-60, consistent with its potential function as a checkpoint inhibitor of LSU assembly.

Discussion

In this work, we developed a bio-orthogonal strategy for detection of actively synthesized alt-proteins. BONCAT has previously been applied to examine nascent protein synthesis[17], but standard workflows include a column resolution step that de-enriches small proteins. We therefore developed a new method to interface azidohomoalanine labeling with small protein enrichment using a C8 column[19], coupled with direct on-bead capture and alt-protein proteomics. Using this approach, we identified 22 unannotated alt-proteins or N-terminal extensions of canonical proteins undergoing active synthesis in HEK 293T cells during normal growth or cellular stress. BONCAT detection of MINAS-60 in normally growing cells likely reflects active synthesis of this protein in the subpopulation of cells in S phase, consistent with its coordinated synthesis with the ribosome biogenesis machinery. These results show that our method is able to detect alt-proteins regulated by the cell cycle and cellular stress.

We then conducted a focused study of MINAS-60, which is a nucleolar alt-protein co-encoded with human RBM10. RBM10 is an RNA binding protein that regulates alternative pre-mRNA splicing[36]. Null mutations in the *RBM10* gene are found in patients with TARP syndrome[39], an X-linked inherited pathology associated with malformation of multiple organs and significant early-life mortality. The *RBM10* gene is frequently mutated in lung adenocarcinoma, and RBM10 inhibits cancer cell proliferation[40]. The finding that *RBM10* dually encodes MINAS-60, which also downregulates cell proliferation via repression of ribosome biogenesis, opens the question of whether this alt-protein contributes to *RBM10* mutation-associated disease phenotypes.

This work expands on recent findings that a single human transcript can encode overlapping, sequence-independent, yet functionally related proteins. Several studies have suggested that as many as 50% of multicistronic human genes encode proteins that directly interact to form complexes[3,41]. In contrast, alt-RPL36 regulates the same pathway (protein translation) as the canonical protein co-encoded with it (ribosomal protein L36), without directly interacting with each other[6]. Here, we show that MINAS-60 and RBM10 exhibit comparable antiproliferative effects but act via different pathways: ribosome biogenesis vs. alternative pre-mRNA splicing, respectively. We therefore speculate that direct interaction between alt-proteins and canonical proteins is only one mechanism by which the fitness effects of co-encoded proteins may be optimized. For example, *de novo* acquisition of an alt-ORF within a protein coding gene might lead to positive selection if both the alt-ORF and canonical coding sequence increase cellular fitness under the same conditions, without necessarily interacting or regulating on the same pathway or process, broadening our current understanding of multicistronic human gene function[3].

To our knowledge, MINAS-60 is only the second human microprotein or alt-protein shown to localize to the nucleolus to date, and the first reported to regulate ribosome biogenesis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

A previous study identified nucleolar C11orf98, which interacts with NPM1 and NCL, but no role for this microprotein in ribosome biogenesis has been investigated[42]. MINAS-60 is conserved from human to mouse, but the *RBM10* gene is not present in yeast. This suggests that MINAS-60 may have arisen in mammals, though we cannot yet rigorously conclude that a functional homolog of MINAS-60 is absent in lower eukaryotes. Interestingly, another putative mammal-specific LSU biogenesis inhibitor has recently been identified. A structural study revealed that an unidentified protein (protein X) is positioned to block the incorporation of nuclear export factor NMD3 into the late pre-60S particle, hypothetically suppressing pre-60S export into cytoplasm[33]. Protein X binds in the immature peptidyl transfer center and contacts GTPBP4 as well as helix 89 of the rRNA. Protein X was observed in human pre-60S particles and has not been detected in yeast LSU precursors[33]. Because 1) both MINAS-60 and protein X likely function as repressors of LSU biogenesis at the late assembly and export stage; and 2) both MINAS-60 and protein X associate with GTPBP4, we speculate that MINAS-60 functions by a similar mechanism to protein X. It is possible that protein X and MINAS-60 are the same protein, and that protein X could not be identified due to its absence from the UniProt database; however, our efforts to model MINAS-60 sequences into the reported structures remain inconclusive to date. It is also possible that multiple previously unknown mammalian proteins slow the rate of LSU biogenesis, suggesting that regulation of 60S biogenesis may be more complex in human cells than in yeast.

Why, then, might MINAS-60 have arisen as part of the ribosome biogenesis machinery in mammalian cells? In order to ensure that improperly assembled ribosome subunits do not lead to mistranslation[23], ribosome biogenesis needs to be precisely monitored in cells. However, while hundreds of proteins required for ribosome biogenesis in human cells have been identified[43], few inhibitors or checkpoints for this process have been reported[44-47]. For example, the Rio1-Nob1-Pno1 network establishes a checkpoint to safeguard against the release of immature 40S subunits into translating ribosomes[48]. LSU biogenesis involves multiple composition and conformational changes during generation of the peptidyl transferase center (PTC). Our data are most consistent with a model in which MINAS-60 may be a core component of the ribosome biogenesis machinery that acts as a checkpoint for LSU biogenesis to safeguard faithful pre-60S assembly. This conclusion is supported by the observation that MINAS-60 depletion promotes the late-stage assembly and export of pre-60S particles into cytoplasm, leading to increases in cytoplasmic 60S ribosomal subunits, global protein synthesis, and cell proliferation (Fig. 6). Conversely, over-expression of MINAS-60 leads to a small decrease in cytoplasmic 60S levels, and a much larger decrease in 80S levels. This suggests that over-expression of MINAS-60 promotes assembly of functional LSU by slowing its production, and that MINAS-60 may be one factor that contributes to previously observed slower maturation and export of the LSU relative to the SSU[14], consistent with checkpoint inhibition. However, we do not detect a statistically significant change in 80S levels in our stable silencing cell lines. This could be a result of compensatory effects in the stable cell lines, or could suggest that the observed decrease in 80S levels is an artifact of transient over-expression. In the future, our hypothesis that MINAS-60 is a checkpoint inhibitor for LSU assembly could be tested by examining

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the effects of MINAS-60 levels on the fidelity of global protein translation, and on the composition and structure of the LSU.

In conclusion, our study, combined with previous literature[8], demonstrates the power of chemoproteomics to reveal alt-proteins that are regulated in important cellular processes including the cell cycle, and reveals an entirely new regulatory node of the mammalian ribosome biogenesis pathway, meriting further development of chemical tools to identify functional alt-proteins.

Methods

Data analysis.

Two-tailed *t*-test was performed using Microsoft Excel (version 2013) or GraphPad Prism (version 8.0.0), one-way or two-way ANOVA (Dunnett's test) was performed using GraphPad Prism (version 8.0.0), and *F*-tests were performed to evaluate equal variance between samples.

Antibodies.

Primary antibodies for immunoblotting include: anti-FLAG (1:500, Sigma, F3165 or 1:1500, Cell Signaling, 14793); anti-HA (1:3000, Invitrogen, 71-5500); anti-β-actin (1:3000, Invitrogen, MA5-15739); anti-GTPBP4 (1:1000, Abclonal, A4565); anti-MRTO4 (1:1000, ThermoFisher, 20194-1-AP); anti-NSA2 (1:1000, Abclonal, A14475); anti-NOP2 (1:1000, Cell Signaling, 25017); anti-BRIX1 (1:1000, Abclonal, A14481); anti-RPF2 (1:1000, Abclonal, A17224); anti-GNL2 (1:1000, Abclonal, A13191); anti-RPL11 (1:2000, Cell Signaling, 18163); anti-RPL36 (1:1000, Bethyl Laboratories, A305065A-M); anti-RPS2 (1:1000, Invitrogen, PA5-30160); anti-RBM10 (1:1000, Abcam, ab72423); anti-puromycin (1:1000, Kerafast, EQ0001); anti-GFP (1:1000, Abcam, ab183734); anti-cyclin B1 (1:1000, Cell Signaling, 4138). Immunoprecipitation was performed with anti-FLAG M2 affinity gel (Sigma, A2220). Secondary antibodies for immunoblotting were: goat anti-rabbit IgG horseradish peroxidase conjugate (1:8000, Rockland, 611-1302) and goat anti-mouse IgG horseradish peroxidase conjugate (1:4000, Rockland, 610-1319-0500). Primary antibodies for immunostaining were: rabbit anti-HA (1:500, Invitrogen, 71-5500) and mouse anti-Fibrillarin (1:500, Abcam, ab4566). Secondary antibodies for immunostaining were: goat anti-rabbit IgG Alexa fluor 568 (1:500, Invitrogen, A11011) and goat anti-mouse IgG Alexa fluor 647 (1:500, Invitrogen, A21235).

Cloning and genetic constructs.

The 5'UTR of human *DRAP1*, *PRR3* transcript variant (tv) 2, *PRH1* tv 1, *CNPY2* tv 1, *CACTIN* tv 1, or *RBM10* tv 1 through the stop codon of each alt-ORF was PCR amplified with a 3' FLAG-HA epitope tag from reverse-transcribed HEK 293T cDNAs, then cloned into pcDNA3. MINAS-60 deletions or mutations bearing a FLAG-HA tag were generated by ligating PCR products into BamHI and EcoRI sites in the pcDNA3 vector. For generation of HEK 293T cells stably expressing MINAS-60, the FLAG-HA tagged coding sequence was PCR amplified and cloned into pLJM1. The RPS2 cDNA clone was purchased from Addgene (a gift from Thomas Tuschl), and the coding sequences of MRTO4,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

BRIX1 and RPL29 were amplified by PCR from HEK 293T cDNAs, then subcloned into pJLM1. RNAi constructs were made by synthesizing oligonucleotides encoding a 21 bp short hairpin RNA that targets RBM10 (shRNA1, CTTCGCCTTCGTCGAGTTTAG; shRNA2, TCCAACGTGCGCGTCATAAAG), then cloned into pLKO.1. Empty pLKO.1 vector control was purchased from Sigma (SHC001). qPCR primer sequences are provided in Supplementary Table 4.

Cell culture, lentivirus production and stable cell line generation.

HEK 293T and 3T3 cells were purchased from ATCC and early-passage stocks were established. Cells cultured as previously described[6] up to only 10 passages. To produce lentivirus and generate stable cell lines, HEK 293T cells were co-transfected using polyethyleneimine (Polysciences, 23966) with expression construct in pLJM1, along with pMD2.G and psPAX2, and growth media were replaced after 7-8 h. 48 h post-transfection, media containing viruses was harvested, filtered through a 0.45-μm filter, and infection was performed by mixing with two volumes of fresh media containing suspended HEK 293T cells. 24 h post-infection, the growth media was replaced. 48 h post-infection, stable cells were selected with 4 μg/mL puromycin for 2 days. Early stocks of stable cell lines were established after selection. Stable cell lines were released from puromycin for 2 days prior to use in experiments.

Immunostaining and live-cell imaging.

HEK 293T cells were plated on glass coverslips and, when necessary, transfected the next day. Forty-eight hours later, the cells were fixed in 10% formalin for 15 min at room temperature (RT), permeabilized with PBS containing 0.2% (v/v) Triton X-100, then incubated with primary antibodies for 18 h at 4°C. After washing with PBS, the cells were incubated with secondary antibodies and DAPI for 1 h at RT, washed with PBS and mounted with Mowiol (Sigma, 81381) before viewing.

HEK 293T cells stably expressing RPL29-GFP or RPS2-GFP were grown to 70% confluency on coverslips in 12-well plates. Coverslips were inverted and imaged in pre-warmed DMEM with 10% FBS, 1% penicillin-streptomycin in MatTek imaging dishes. Confocal imaging was performed on a Leica SP8 LS confocal microscope with 63× oil immersion objective with atmosphere-controlled stage at 37°C. Nuclear/cytoplasmic ratios of RPL29 and RPS2 were measured using the ImageJ (version ImageJ2) Intensity Ratio Nuclei Cytoplasm Tool (RRID:SCR_018573; https://github.com/MontpellierRessourcesImagerie/imagej_macros_and_scripts/wiki/Intensity-Ratio-Nuclei-Cytoplasm-Tool).

Immunoprecipitation and proteomics.

Control or MINAS-60 KI HEK 293T cells were grown to 80-90% confluency in 15 cm dishes. Cells were harvested and suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$ with 0.5% NP40 and Roche Complete protease inhibitor cocktail tablets (Roche, 11873580001)), and incubated on ice for 10 min, followed by centrifugation at 3,000 g, 4°C, 3 min. Nuclear pellets were suspended in 1 mL lysis buffer [Tris-buffered saline (TBS) with 1% Triton X-100 and Roche Complete

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

protease inhibitor cocktail tablets], followed with sonication and immunoprecipitation as previously described[6]. Elution was in 40 μL of 3× FLAG peptide (Sigma, F4799), at a final concentration of 100 μg/mL in lysis buffer at 4°C for 1 h. Eluted proteins were subjected to SDS-PAGE prior to LC-MS/MS.

Gel slices (protein bands or entire lanes) were digested with trypsin at 37°C for 14-16 h. Resulting peptide mixtures were extracted from the gel, dried, and residual detergent was removed with ethyl acetate followed by de-salting with peptide cleanup C18 spin column (Agilent Technologies, 5188-2750). Peptides were re-suspended in 35 μL 0.1% formic acid (FA), followed by centrifugation at 21,130 g, 4°C, 30 min. 5 μL of each sample was injected on a pre-packed column attached to a nanoAcquity UPLC (Waters) in-line with a Thermo Scientific™ Q Exactive™ Plus Hybrid QuadrupoleOrbitrap™ mass spectrometer and a 130-min gradient was used to further separate the peptide mixtures as follows (solvent A: 0.1% FA; solvent B: acetonitrile (ACN) with 0.1% FA): Isocratic flow was maintained at 0.1 μL/min at 1% B for 40 min, followed by linear gradients from 1% B to 6% B over 2 min, 6% B to 24% B over 48 min, 24% B to 48% B over 5 min, 48% B to 80% B over 5 min. Isocratic flow at 80% B was maintained for 5 min, followed by a gradient from 80% B to 1% B over 5 min, and isocratic flow at 1% B was maintained for 10 min. The full MS was collected over the mass range of 300-1,700 m/z with a resolution of 70,000 and the automatic gain control (AGC) target was set as 3×10^6. MS/MS data was collected using a top 10 high-collisional energy dissociation method in data-dependent mode with a normalized collision energy of 27.0 eV and a 1.6 m/z isolation window. MS/MS resolution was 17,500 and dynamic exclusion was 90 seconds.

For alt-protein and microprotein identification, ProteoWizard MS Convert was used for peak picking and files were analyzed using Mascot (version 2.5.1). Oxidation of methionine and N-terminal acetylation were set as variable modifications, and a previously reported three-frame translation of assembled transcripts from HEK 293T mRNA-seq was used as the database exactly as previously described[7]. Identified peptides were searched against human UniProt protein database (version 2019) using a string-matching algorithm. Peptides found to be identical to annotated proteins were discarded. The remaining peptides were searched against non-redundant human protein sequences using on-line BLAST (Basic Local Alignment Search Tool). Any peptides found to be less than two amino acids different from the nearest protein match were discarded. For co-IP proteomics searches and quantitative analysis, files were analyzed using MaxQuant (version 1.6.8.0), oxidation of methionine and N-terminal acetylation were set as variable modifications, and human UniProt plus MINAS-60 was used as the search database. For all analyses, a mass deviation of 20 p.p.m. was set for MS1 peaks, MS/MS tolerance was 0.6 Da, and a maximum of two missed cleavages were permitted. Maximum false discovery rates (FDR) were set to 1% both on peptide and protein levels. Minimum required peptide length was five amino acids. Protein quantitation was calculated from the LFQ intensity ratio of pulldown to negative control samples using MaxQuant with standard parameters. *P*-values (*t*-test) were calculated using Perseus (version 1.5.8.5) with standard parameters.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Dibenzocyclooctyne (DBCO) bead construction.

NHS-activated beads (Pierce, 88826) were conjugated to dibenzocyclooctyne-amine (Sigma, 761540) by 90 min incubation at room temperature with rotation in a saturated 100 mM sodium bicarbonate (pH 8.0) solution, then washed and blocked for 2 h with 3 M ethanolamine. After blocking, the beads were re-suspended in PBS prior to immediate protein conjugation.

BONCAT (bio-orthogonal non-canonical amino acid tagging) and streptavidin pulldown.

HEK 293T cells were grown to 80-90% confluency in 15 cm dishes, treated with methionine aminopeptidase inhibitor TNP470 (50 nM) for 2 h, then immersed in methionine-free DMEM (Corning, 17-204-CI) for 30 min before addition of 4 mM AHA (Click Chemistry tools, 1066-1000) in methionine-free DMEM with 10% FBS. For stress conditions, cells were treated with 4 mM AHA with simultaneous exposure to 500 μM sodium arsenite, 20 μM etoposide (Sigma, E1383), or 1 mM DTT. After a 2 h incubation, the cells were washed twice with cold PBS, harvested, and flash frozen.

For gel-based selection of AHA-labeled small proteins, cells from one 15 cm dish were lysed in 2 mL 1% (w/v) SDS in PBS by boiling for 10 min and then brought up to 0.1% SDS by adding 18 mL PBS with 0.2% TritonX-100. In a separate tube, 200 μL of 70 mM $CuSO_4$ was combined with 20 μL of 100 mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA, Click Chemistry tools, 1061-100) and 70 μL of 500 mM Tris(2-carboxyethyl)phosphine hydrochloride, then incubated for 10 min at RT. 20 μL of 25 mM biotin-alkyne (BALK, Click Chemistry tools, 1266-5) was added to the mixture and then combined with 20 mL cell lysate, followed with rotation at RT for 6 h. Labeled lysates were then passed through a 3 kDa Amicon Ultra 15 mL column, and diluted 1000 times with RIPA buffer (10 mM Tris-HCl (pH=7.4), 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl) to remove excess BALK, followed by concentration to 1 mL. A 50 μL aliquot of streptavidin magnetic beads (NEB) was washed twice with 1 mL RIPA before gentle rotation with 1 mL of biotinylated proteins for 1 h at RT. After binding, beads were washed twice with 1 mL RIPA, once with 1 M KCl, once with 0.1 M sodium carbonate, once with 2 M urea, and twice with RIPA. Bound proteins were eluted by boiling in 1×SDS loading buffer containing 20 mM DTT and 2 mM biotin for 10 min. After elution, proteins were electrophoresed on a 16% tricine gel, and gel bands corresponding in the 2-15 kDa size range were excised for proteomic analysis. As a control, unlabeled cell lysates were size-selected on a tricine gel as described above.

DBCO bead enrichment of AHA-labeled proteins and ERLIC (electrostatic repulsion hydrophilic interaction chromatography) fractionation.

AHA-labeled cells from one 15 cm dish were lysed by boiling in 2 mL 50 mM HCl with 0.01% 2-mercaptoethanol and 0.05% TritonX-100 for 10 min. Lysates were centrifuged at 21,100 g, 4°C for 30 min, and passed through a 5 μm filter before size selection with a C8 column (Agilent Technologies, 12102100) essentially as previously reported[20]. The C8 column was pre-conditioned with 1 bed volume of methanol and 2 bed volumes of 0.25 M triethylammonium formate (TEAF, pH 3.0). Up to 2 bed volumes of cell lysate were loaded on the column, followed by two washes with 2 bed volumes of TEAF and elution

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with 2 bed volumes of ACN:TEAF (3:1). The sample was dried and reconstituted in 300 μL PBS. Size-selected lysates were then incubated with 50 μL DBCO beads for 1 h at RT. The beads were washed twice with 1 mL RIPA, once with 1 M KCl, once with 0.1 M sodium carbonate, once with 2 M urea, twice with RIPA, and 6 times with PBS. Proteins conjugated to the beads were subjected to reduction, alkylation and on-bead trypsin digestion according to standard protocols[50].

For method validation, digested peptides were directly analyzed via 1D LC-MS/MS as described above. For subsequent analysis of control and stress conditions, digested peptides were pre-fractionated using ERLIC on an Agilent 1100 HPLC. Peptides were re-suspended in 55 μL of 85% ACN/0.1% FA and 50 μL was loaded onto a polyWAX LP column (150×1.0 mm; 5 μm 300 Å; PolyLC). Samples were run on an 80 min gradient protocol as follows (Solvent A: 80% ACN 0.1% FA; Solvent B: 30% ACN 0.1% FA): Isocratic flow was maintained at 100% A at a flow rate of 0.3 mL/min for 5 min, followed by a 17 min linear gradient to 8% B, and a 25 min linear gradient to 45% B. Finally, a 10 min gradient to 100% B was followed by a 5 min hold at 100% B before a 10 min linear gradient back to 100% A, followed by an 8 min hold at 100% A. The digested peptides were separated into 15-18 fractions which were dried and re-suspended in 7 μL of 3:8 70% FA : 0.1% TFA before LC-MS/MS analysis.

Generation of knock-in cell lines.

MINAS-60 3xGFP11-FLAG-HA KI HEK 293T cells were generated using CRISPR-Cas9. Guide RNAs (gRNAs) were designed with the guide design tool from the Zhang lab (crispr.mit.edu) to target the RBM10 genomic region: gRNA1, 5'- TGTCGGCCAGGATTCCTACG-3'; gRNA2, 5'- CCCGATAGTCGCCGTCTCGG-3'. Double-stranded DNA oligonucleotides corresponding to the gRNAs were inserted into pSpCas9(BB)-2A-GFP vector (Addgene, as a gift from F. Zhang, MIT, Cambridge, MA). A donor plasmid containing 300 bp left-arm homology and 300 bp right-arm homology around the stop codon of MINAS-60, which are separated by 3xGFP11-FLAG-HA tag and BamHI / NotI restriction sites, was synthesized by GenScript, and a DNA sequence containing pGK promoter and hygromycin B resistance gene were subcloned into the donor plasmid using the BamH1 and NotI restriction sites. An equal mixture of the gRNA and donor plasmids were transfected into HEK 293T cells using polyethyleneimine, and hygromycin B selection was performed 2 days post-transfection. MINAS-60-3xGFP11-FLAG-HA KI cells were validated by genomic PCR and sequencing.

Puromycin incorporation.

SUnSET was used to measure protein synthesis[51]. HEK 293T cells were grown to 80-90% confluency in 6 well plates, then growth media was replaced with media containing 1 μM puromycin and cells were incubated at 37°C for 1 h according to a previously published protocol[43]. Cells were then washed once with PBS, harvested and analyzed with immunoblotting.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Sucrose gradient profiling.

HEK 293T cells were seeded in 15 cm dishes at 2.5×10^7 cells per dish and cultured 24 h. Cells were then treated with 100 μg/mL cycloheximide for 5 min, washed with cold PBS containing 100 μg/mL cycloheximide twice, harvested and flash frozen.

For cytoplasmic polysome profiling, after thawing on ice, cells were lysed in polysome lysis buffer on ice for 10 min (20 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 100 μg/mL cycloheximide, 1 mM DTT with 1% TritonX-100, Roche Complete protease inhibitor cocktail tablets and ribonuclease inhibitors (Promega N2511)), followed by centrifugation at 21,130 g, 4°C, 10 min. Supernatants were normalized according to absorbance (A260) and layered onto 12 mL 10-50% sucrose gradients (20 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 100 μg/mL cycloheximide, 1 mM DTT with Roche Complete protease inhibitor cocktail tablets and ribonuclease Inhibitors), followed by centrifugation in an SW-41Ti rotor at 252,878 g, 4°C, 3 h, then sampled using a Biocomp Gradient profiler (model 251) with constant monitoring of optical density at 254 nm using standard parameters. Data were analyzed with Excel (version 2013) and ImageJ2.

For sucrose gradient sedimentation analysis of nuclear lysates, thawed cells were suspended in 1 mL nuclear isolation buffer (10 mM HEPES-KOH pH 7.4, 100 mM KCl, 5 mM MgCl2 with 0.5% NP40, Roche Complete protease inhibitor cocktail tablets and ribonuclease inhibitors), incubated on ice for 10 min, followed by centrifugation at 3,000 g, 4°C, 3 min. Nuclear pellets were suspended in 1 mL nuclear lysis buffer (20 mM HEPES-KOH pH 7.4, 300 mM KCl, 5 mM $MgCl_2$, 1 mM DTT with 1% TritonX-100, Roche Complete protease inhibitor cocktail tablets and ribonuclease inhibitors), followed with sonication (50% intensity, 5 s pulse with 25 s rest, 5×, MICROSON XL 2000), and centrifugation at 21,130 g, 4°C, 10 min. The supernatants were normalized to A260 and layered onto 12 mL 10-50% sucrose gradients, centrifuged and analyzed as described above.

Crystal violet staining.

4×10^4 HEK 293T cells were seeded in 12-well plates, and fixed in 10% formalin for 15 min at RT every 24 h. After washing with ddH_2O, cells were stained with 0.1% crystal violet in methanol for 30 min at RT in the dark, followed with three ddH_2O washes and dried. Cells were then immersed in 1 mL 10% acetic acid with shaking for 20 min. 20 μL solution was combined with 80 μL ddH_2O in a 96-well plate, and optical density at 590 nm was monitored with Synergy™ HT.

Cell cycle synchronization.

Following a protocol for cell cycle synchronization[52], 1.5×10^5 HEK 293T cells were seeded in 12-well plates and cultured overnight. Cells were then treated with 2 mM thymidine for 16 h, followed by two PBS washes, and incubated with fresh media for 9 h before the second 2 mM thymidine block for 14 h. Cells were washed with PBS, then incubated with fresh media to release from the G1/S boundary, and fixed in 10% formalin for immunostaining or harvested for immunoblotting.

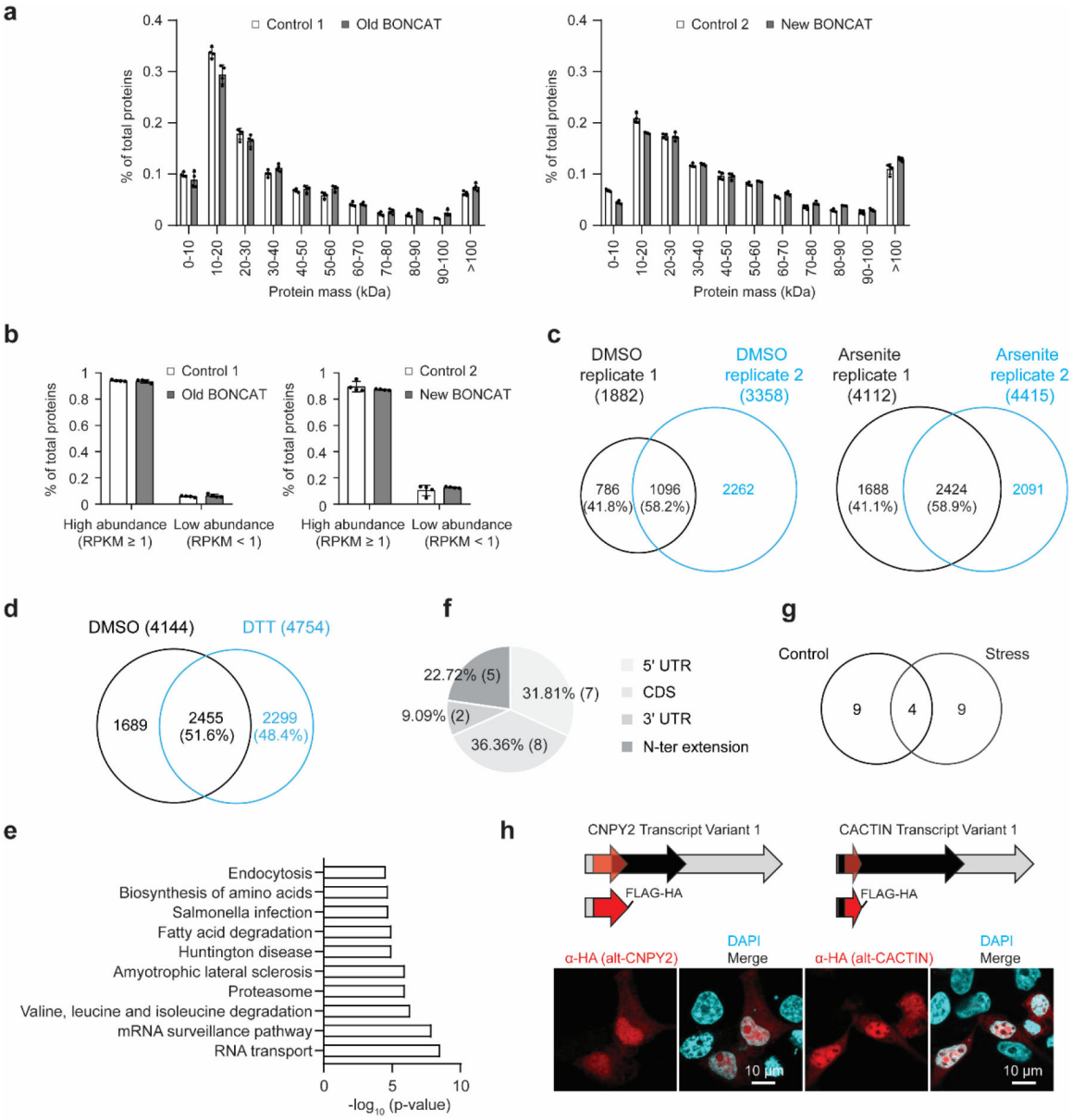
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Northern blot.

Total RNA was extracted from HEK 293T cells using TRIzol. To determine changes in levels of LSU pre-rRNA intermediates, 3 μg of total RNA was run on a 1% agarose/1.25% formaldehyde gel in a 1.5 M tricine/ 1.5 M triethanolamine buffer, followed by overnight transfer to a Hybond XL nylon membrane (GE Healthcare, RPN 303S) in 10× saline-sodium citrate transfer buffer after a 15 min soak in 0.5 M sodium hydroxide solution. Membranes were exposed to UV (254 nm) to immobilize the RNA, followed by incubation with denatured yeast tRNA for 1 h at 42°C, and hybridized overnight at 37°C with 5' end radiolabeled oligonucleotide probe (P4 5'-CGGGAACTCGGCCCGAGCCGGCTCTCTCTTTCCCTCTCCG-3') in a solution of 7.5× Denhardt's solution, 5× sodium chloride-sodium phosphate-EDTA buffer with 0.1% SDS, as previously reported[46]. Membranes were also hybridized with a 7SL probe (7SL 5'-TGCTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3') as a loading control.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Extended Data

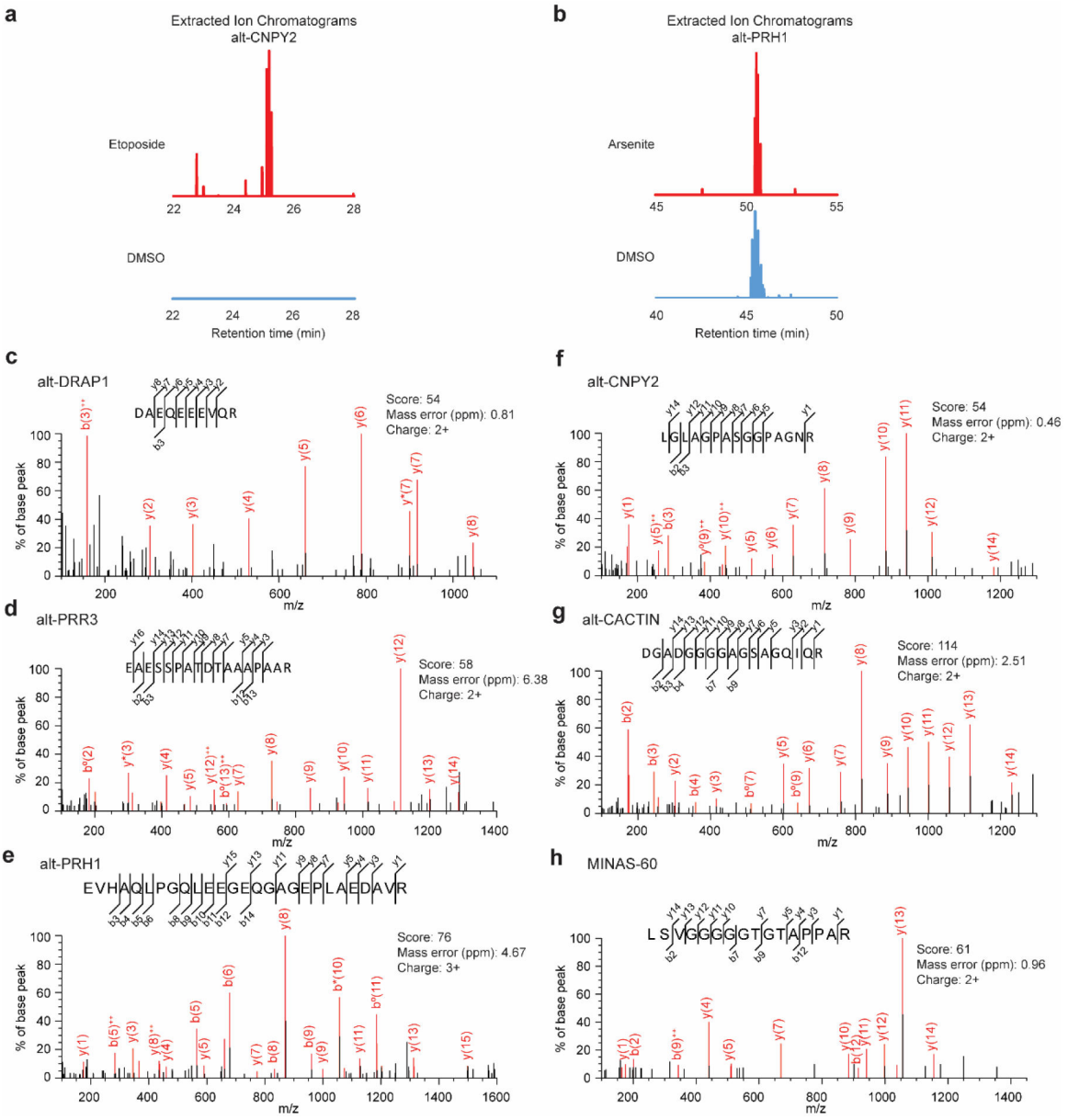


Extended Data Fig. 1 | Overview of BONCAT-based chemoproteomics method optimization.
a, Size distribution of canonical, annotated proteins detected from vehicle HEK 293T lysate subjected to gel-based microprotein isolation (control 1), AHA-labeled HEK 293T lysate subjected to Click chemistry with biotin, 3 kDa MWCO membrane removal of excess biotin, streptavidin enrichment and gel-based microprotein isolation (old BONCAT), vehicle HEK 293T lysate subjected to C8 column-based microprotein isolation (control 2), or AHA-labeled HEK 293T lysate subjected to on-bead direct capture and digest after C8 column-based size selection (new BONCAT). Error bars, standard error of the mean (s.e.m.), $N = 4$ (two biologically independent experiments, each with two technical replicates).
b, Expression level distribution of canonical, annotated proteins detected from the four

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

strategies described above. Error bars, standard error of the mean (s.e.m.), $N = 4$ (two biologically independent experiments, each with two technical replicates). **c**, Venn diagram of the reproducibility of control samples (**c**, left, DMSO) and AHA-labeled samples (**c**, right, Arsenite). **d**, Venn diagram of detected annotated proteins from combined DMSO or DTT treatment samples. **e**, GO (KEGG) analysis of the 2,299 DTT-induced genes with g:Profiler identified processes associated with endoplasmic reticulum stress and proteostasis (e.g., proteasome, mRNA surveillance pathway). **f**, Distribution of locations of identified alt-ORFs relative to annotated coding sequences (CDS). **g**, Venn diagram of alt-proteins identified under control or combined stress conditions. **h**, Top: Schematic representation of human *CNPY2* transcript variant (tv) 1 or *CACTIN* tv 1; light gray arrow, 5' and 3' untranslated regions; red, alt-ORF; black, annotated protein CDS. Middle: Schematic representation of the expression construct containing complete 5'UTR and alt-ORF of the transcript indicated above, with a FLAG-HA appended to the C-terminus of the alt-protein. Bottom: HEK 293T cells transfected with the expression construct were immunostained with anti-HA (red) and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates.

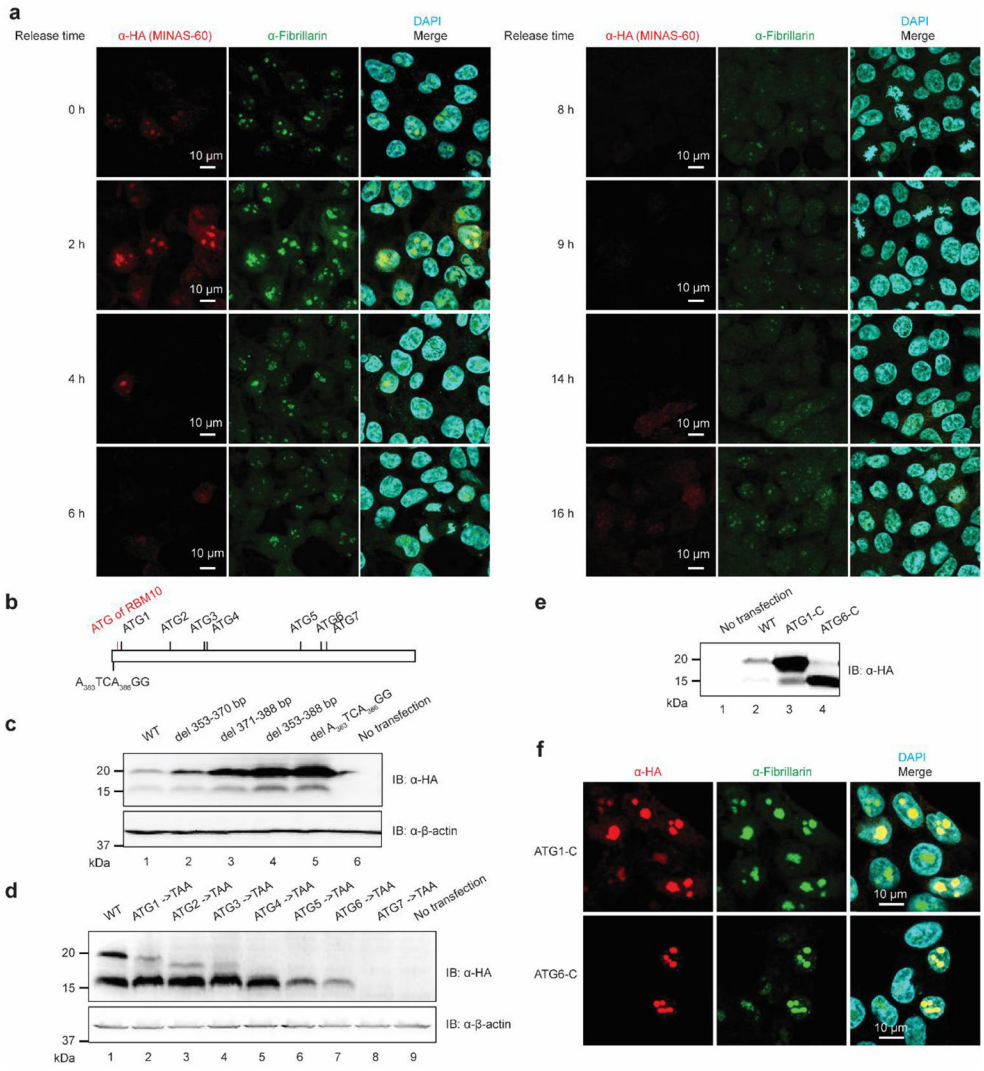
Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript



Extended Data Fig. 2 |. BONCAT-based chemoproteomic identification of newly synthesized alt-proteins.

a,**b**, Extracted ion chromatograms (EICs) from MS1 spectra corresponding to tryptic peptides from alt-CNPY2 (**a**) or alt-PRH1 (**b**) under control (DMSO) vs. stress (DNA damage vs. oxidative damage, respectively) conditions. **c**-**h**, MS/MS spectra of tryptic peptides assigned to alt-DRAP1 (**c**), alt-PRR3 (**d**), alt-PRH1 (**e**), alt-CNPY2 (**f**), alt-CACTIN (**g**), and MINAS-60 (**h**) using BONCAT in HEK 293T cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

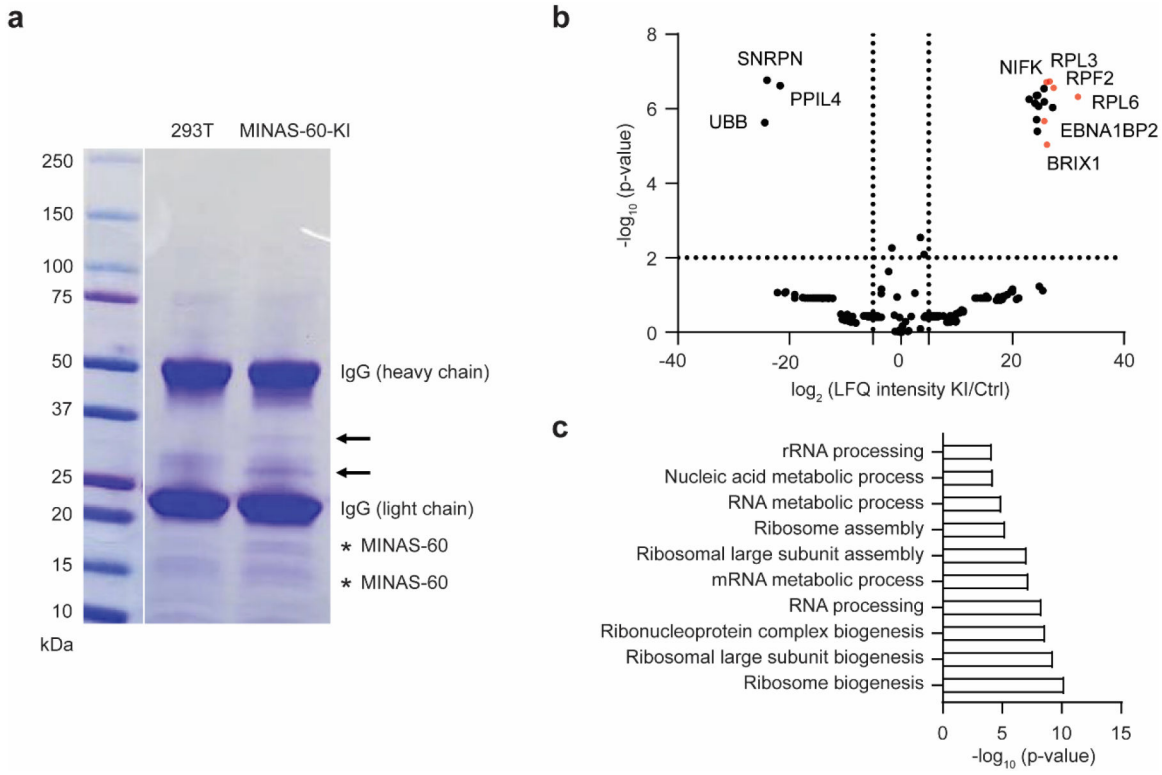


Extended Data Fig. 3 |. MINAS-60 is cell-cycle regulated, and translated from multiple start codons.

a, Immunostaining of synchronized MINAS-60 KI cells released from the G1/S boundary by a double thymidine block at the indicated time points with anti-HA (red), anti-fibrillarin (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates. **b**, Schematic representation of the region of human *RBM10* tv 1 with candidate MINAS-60 start codons indicated. The ATG start codon of RBM10 is highlighted in red, and the two upstream non-ATG and seven internal ATG start codons that are upstream of and in-frame with the detected tryptic fragment of MINAS-60 are indicated in black. **c**,**d**, Expression of a construct containing the full 5'UTR and wild-type MINAS-60 coding sequence (WT, lane 1), or MINAS-60 coding sequence with each candidate start codon deletion (del, lanes 2-5, **c**), or mutation (lane 2-8, **d**), indicated on the top derived from *RBM10* tv 1, with an HA tag appended to the C-terminus of MINAS-60, in HEK 293T cells was followed by immunoblotting with the antibodies indicated to the right. Untransfected (no transfection) HEK 293T cells served as a control. Data are representative of three biological replicates. **e**, Expression of a construct containing the full 5'UTR and MINAS-60

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

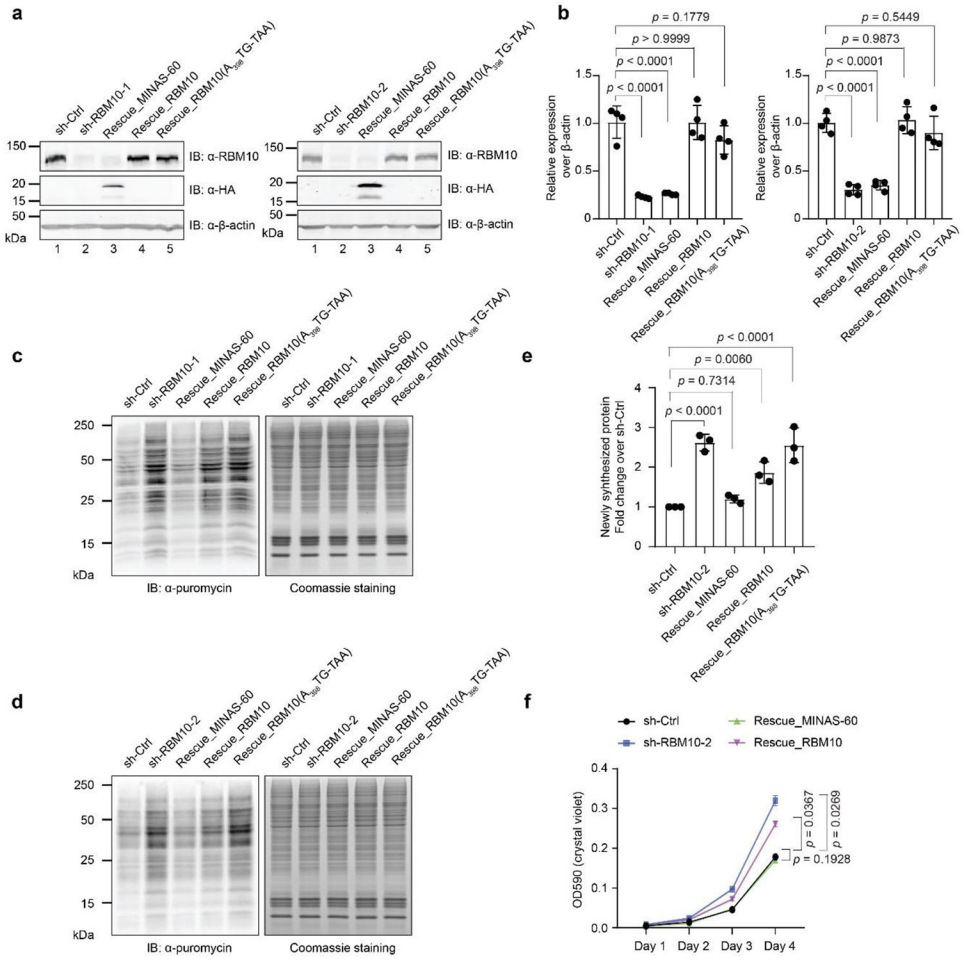
coding sequence (WT, lane 2), or ATG1 to the C-terminus of MINAS-60 (ATG1-C, lane 3), or ATG6 to the C-terminus of MINAS-60 (ATG6-C, lane 4) derived from *RBM10* tv 1, with an HA tag appended to the C-terminus of MINAS-60, in HEK 293T cells was followed by immunoblotting with the antibodies indicated to the right. Untransfected HEK 293T cells served as a control. Data are representative of three biological replicates. **f**, HEK 293T cells were transfected with ATG1-C or ATG6-C of MINAS-60, followed by immunostaining with anti-HA (red), anti-fibrillarin (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates.



Extended Data Fig. 4 |. MINAS-60 associates with ribosomal large subunit (LSU) biogenesis factors.

a, Coomassie blue staining of FLAG-IP from control (293T) or MINAS-60 KI (MINAS-60-KI) HEK 293T cells. Two major enriched bands from the KI FLAG-IP are indicated with black arrows, and MINAS-60 bands are indicated with stars. **b**, Volcano plot of quantitative proteomics ($N = 3$ biologically independent experiments) of gel bands (25-45 kDa) excised from MINAS-60 KI (KI) or control (Ctrl) HEK 293T nuclear lysate FLAG-IP samples after SDS-PAGE. LSU biogenesis factors are indicated in red and the gene names are labeled. For complete quantitative proteomics results, see Supplementary Data S5-6. **c**, GO (biological processes) analysis of genes enriched (fold change ≥ 30) in MINAS-60 KI FLAG-IP over control with g:Profiler.

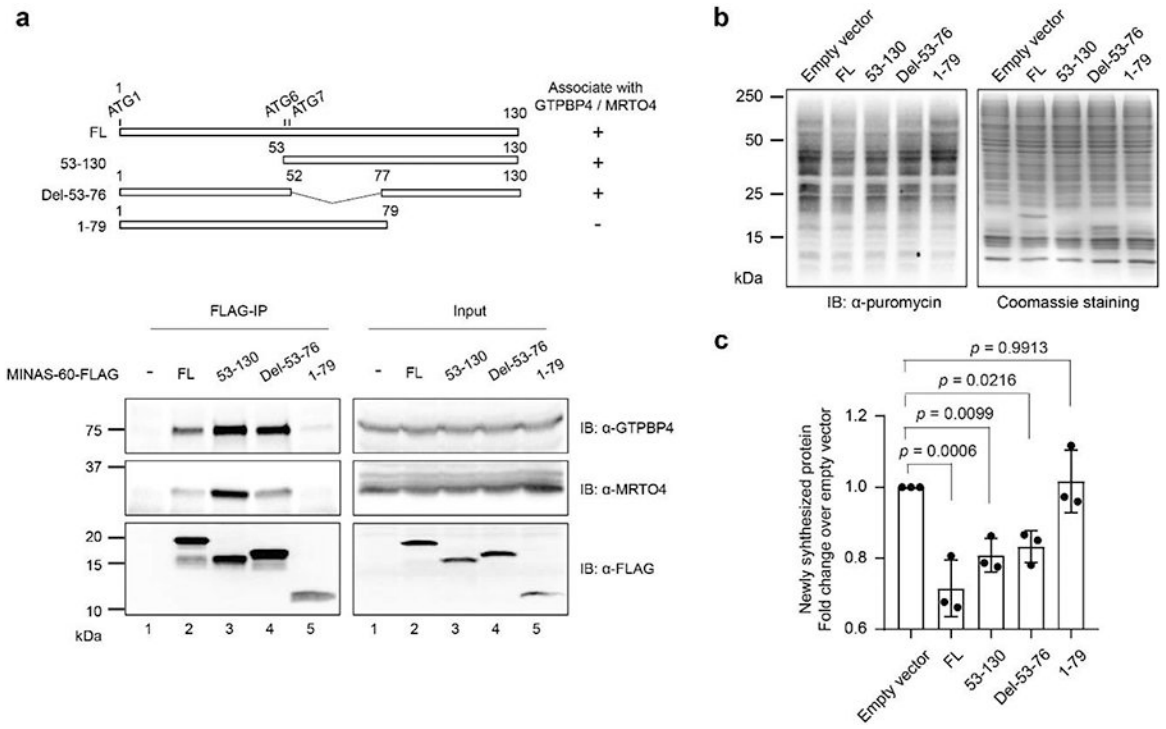
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 5 |. MINAS-60 downregulates global protein synthesis and cell proliferation. **a**, Western blot analysis of HEK 293T stably expressing empty pLKO.1 vector control (lane 1, sh-ctrl), one of the two *RBM10* shRNAs (lane 2, sh-RBM10-1 (**a**, left), sh-RBM10-2 (**a**, right), rescue with MINAS-60 (lane 3, Rescue_MINAS-60), rescue with RBM10 (lane 4, Rescue_RBM10), or rescue with RBM10 bearing an A_{398}TG to TAA mutation (lane 5, Rescue_RBM10(A_{398}TG-TAA). Data are representative of three biological replicates. **b**, Quantitative RT-PCR of the cell lines described above with primers specific to *RBM10*. Error bars, standard error of the mean (s.e.m.), $N = 4$ biologically independent samples, one-way ANOVA (Dunnett's test). **c**,**d**, HEK 293T cells stably expressing empty pLKO.1 vector control (sh-Ctrl), one of the two *RBM10* shRNA (sh-RBM10-1 (**c**) sh-RBM10-2 (**d**), rescue with MINAS-60 (Rescue_MINAS-60), rescue with RBM10 (Rescue_RBM10), or rescue with RBM10 bearing an A_{398}TG to TAA mutation [Rescue_RBM10(A_{398}TG-TAA)] were treated with 1 μM puromycin for 1 hour at 37°C before harvesting and western blotting with anti-puromycin antibody. Coomassie staining served as a loading control. Data are representative of three biological replicates. **e**, ImageJ was used to quantify the relative puromycin incorporation for cells indicated at the bottom relative to sh-Ctrl from three biological replicates. Data represent mean values ± s.e.m., and significance was evaluated with one-way ANOVA (Dunnett's test). **f**, Growth curve of control (sh-Ctrl), *RBM10* knockdown with a second shRNA (sh-RBM10-2), rescue with MINAS-60

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

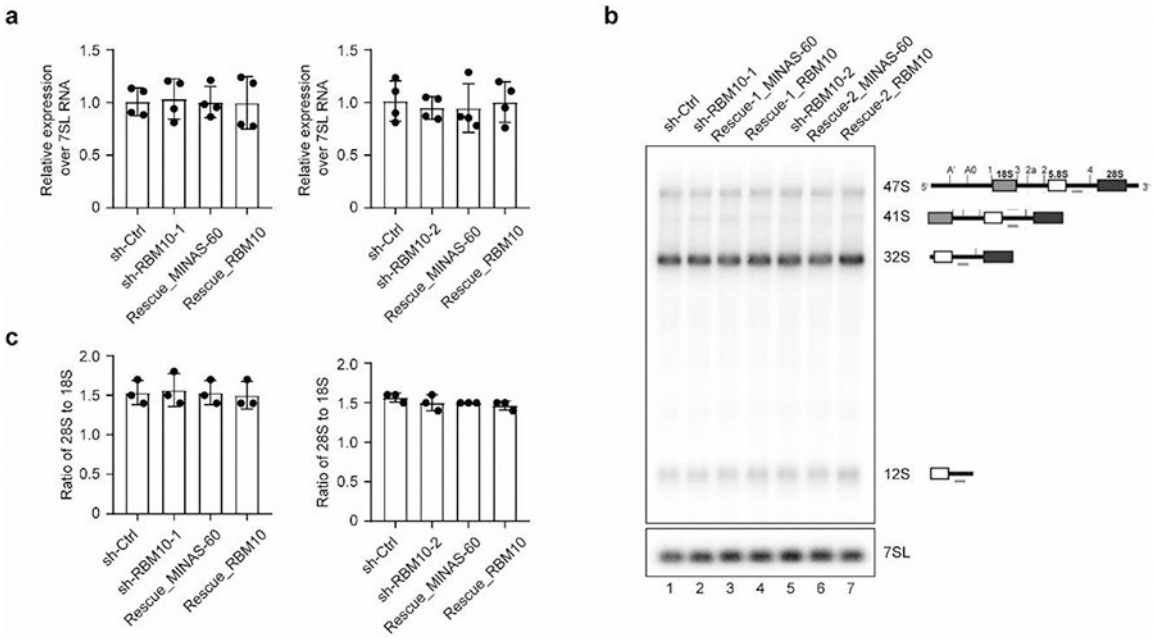
(Rescue_MINAS-60) and rescue with RBM10 (Rescue_RBM10) HEK 293T cells at the indicated number of days ($N = 3$). Data represent mean values ± s.e.m., and significance was evaluated with two-way ANOVA (Dunnett's test).



Extended Data Fig. 6 |. The C-terminal region of MINAS-60 is required for its interactions and function.

a, Top: Schematic representation of the domain structures of the wild-type (full-length, FL) and deletion mutants of MINAS-60, with amino acid residue numbers above. GTPBP4 and MRTO4 association status of each construct are listed on the right. Bottom: HEK 293T cells were transfected with MINAS-60 wild-type or mutants (listed at top), and immunoprecipitations (IP) were performed with anti-FLAG antibody followed by immunoblotting (IB) with the antibodies indicated on the right. Untransfected HEK 293T cells served as a control. Data are representative of three biological replicates. **b**, HEK 293T cells transfected with pcDNA 3.1 empty vector (Empty vector), wild-type (FL) or deletion mutants of MINAS-60 listed at top were treated with 1 μM puromycin for 1 hour at 37°C before harvesting and IB with anti-puromycin antibody. Coomassie staining served as a loading control. Data are representative of three biological replicates. **c**, ImageJ was used to quantify the relative puromycin incorporation for cells indicated at the bottom relative to empty vector from three biological replicates. Data represent mean values ± s.e.m., and significance was evaluated with one-way ANOVA (Dunnett's test).

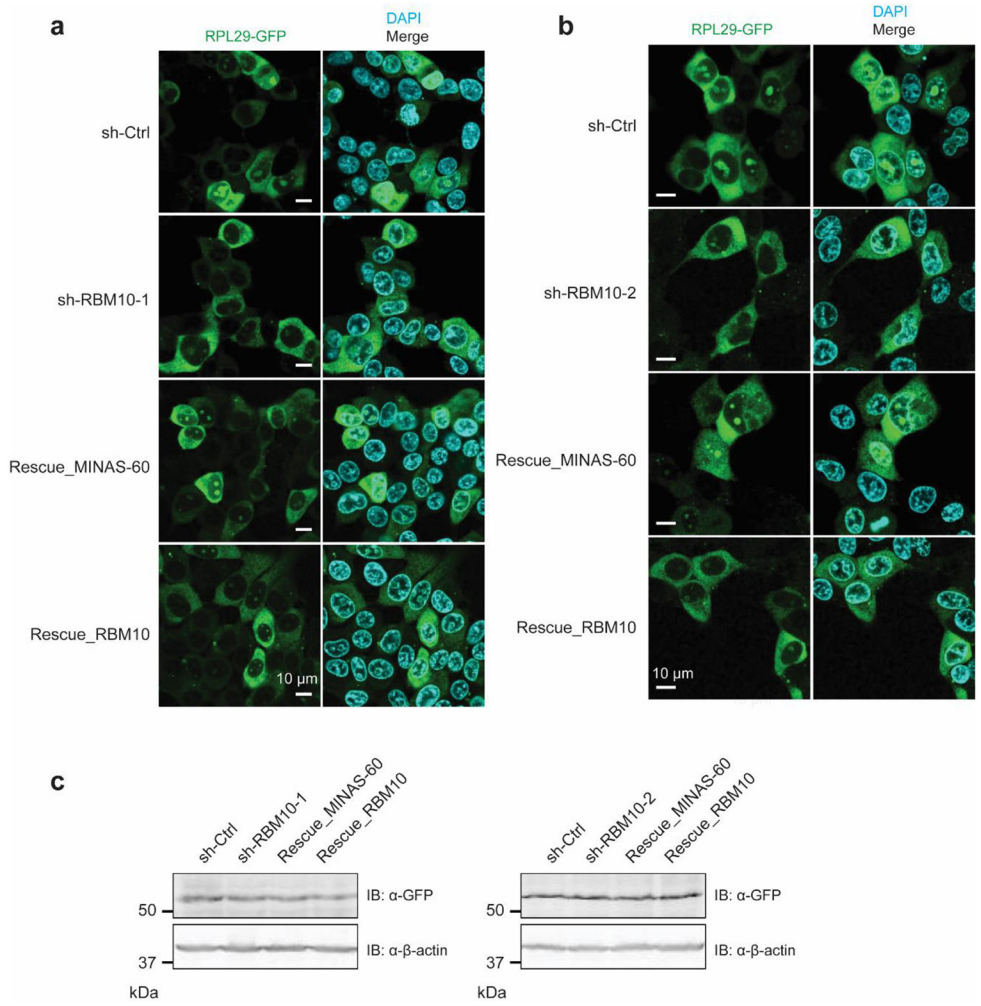
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 7 |. MINAS-60 does not regulate transcription of pre-rRNA, nor LSU pre-rRNA processing.

a, Quantitative RT-PCR with primers specific to the primary pre-rRNA (47S/45S/30S) of HEK 293T cells stably expressing empty pLKO.1 vector control (sh-ctrl), one of the two *RBM10* shRNAs (sh-RBM10-1 (**a**, left), sh-RBM10-2 (**a**, right), rescue with MINAS-60 (Rescue_MINAS-60), or rescue with RBM10 (Rescue_RBM10). Error bars, standard error of the mean (s.e.m.), $N = 4$ biologically independent samples. **b**, Total RNA from HEK 293T cells stably expressing empty pLKO.1 vector control (lane 1, sh-ctrl), one of the two *RBM10* shRNAs [sh-RBM10-1 (lane 2), sh-RBM10-2 (lane 5)], rescue with MINAS-60 [Rescue-1_MINAS-60 (rescue on sh-RBM10-1 background, lane 3), Rescue-2_MINAS-60 (rescue on sh-RBM10-2 background, lane 6)], or rescue with RBM10 [Rescue-1_RBM10 (rescue on sh-RBM10-1 background, lane 4), Rescue-2_RBM10 (rescue on sh-RBM10-2 background, lane 7)] were isolated with TRIzol, and pre-rRNAs were separated by gel electrophoresis, followed with northern blotting using radioactively labeled P4 probe (gray lines, diagram at right). Northern blotting with a probe against the 7SL RNA was used as a loading control. Illustrations of the pre-rRNAs detected by P4 are indicated to the right of their respective bands. Data are representative of three biological replicates. **c**, Total cellular RNA from the cell lines described above was isolated with TRIzol, and the ratio of 28S/18S was calculated using an Agilent 2100 Bioanalyzer. Error bars, standard error of the mean (s.e.m.), $N = 3$ biologically independent samples.

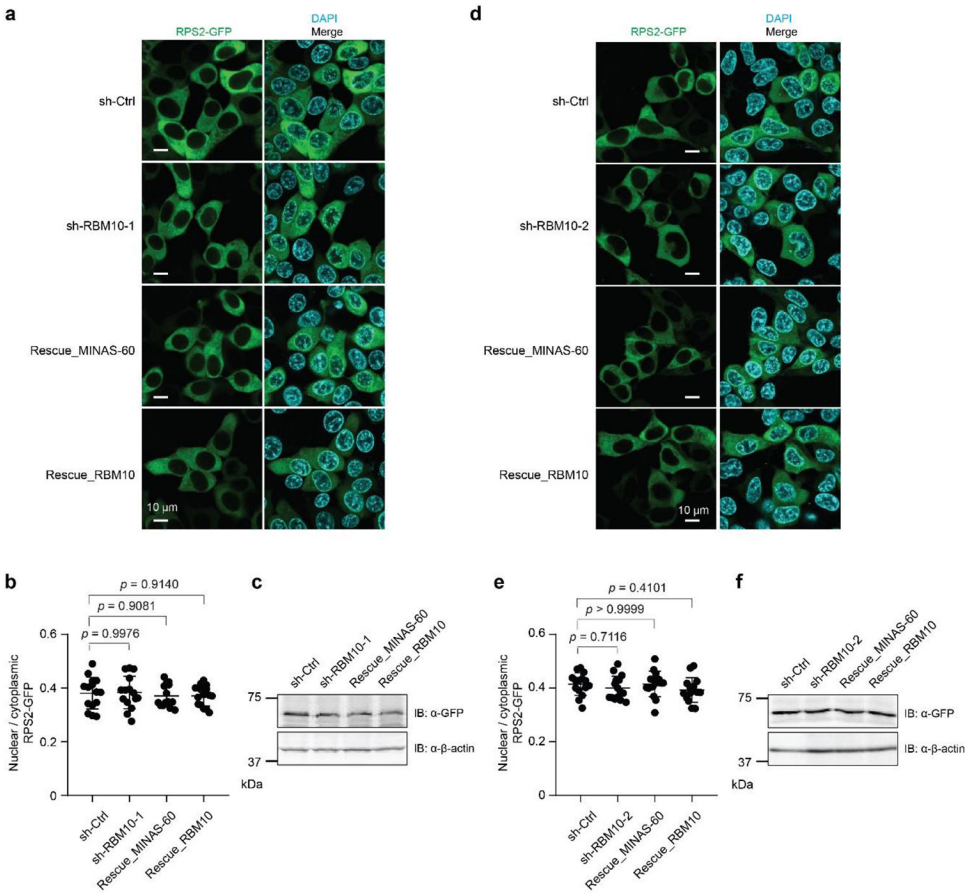
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 8 |. MINAS-60 downregulates LSU export.

a,**b**, Confocal live-cell imaging of control (sh-Ctrl), *RBM10* knockdown with two independent shRNAs [sh-RBM10-1 (**a**), sh-RBM10-2 (**b**)], rescue with MINAS-60 (Rescue_MINAS-60) or rescue with RBM10 (Rescue_RBM10) HEK 293T cells stably expressing RPL29-GFP. Scale bar, 10 μm. Data are representative of three biological replicates. **c**, Western blot of the cell lines described above with antibodies indicated on the right for comparison of RPL29-GFP expression. Data are representative of three biological replicates.

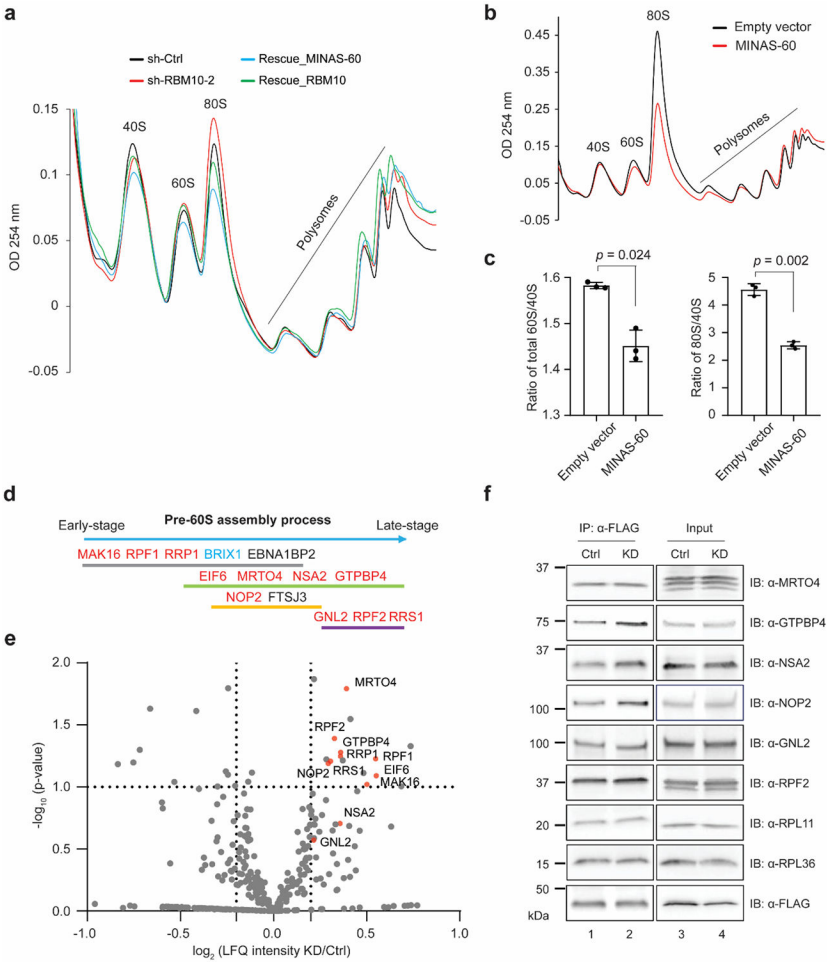
Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 9 |. MINAS-60 does not regulate 40S ribosomal subunit export.

a,**d**, Confocal live-cell imaging of control (sh-Ctrl), *RBM10* knockdown with two independent shRNAs [sh-RBM10-1 (**a**) and sh-RBM10-2 (**d**)], rescue with MINAS-60 (Rescue_MINAS-60) or rescue with RBM10 (Rescue_RBM10) HEK 293T cells stably expressing RPS2-GFP. Scale bar, 10 μm. Data are representative of two biological replicates. **b**,**e**, Quantitation of the RPS2-GFP signals in the cell lines described above. At least 13 fields of view were analyzed, totaling > 350 cells for each measurement. Data represent mean values ± s.e.m., and significance was evaluated with one-way ANOVA (Dunnett's test). **c**,**f**, Western blot of the cell lines described above with antibodies indicated on the right for comparison of RPS2-GFP expression. Data are representative of two biological replicates.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



Extended Data Fig. 10 |. MINAS-60 downregulates cytoplasmic LSU export and pre-60S assembly.

a, Sucrose gradient sedimentation analysis of polysome fractions of cytoplasmic lysates from control (sh-Ctrl), *RBM10* knockdown with a second shRNA (sh-RBM10-2), rescue with MINAS-60 (Rescue_MINAS-60) or rescue with RBM10 (Rescue_RBM10) HEK 293T cells. Data are representative of three biological replicates. **b**, Transient over-expression of MINAS-60 reduces 60S and 80S levels. Sucrose gradient sedimentation analysis of polysome fractions of cytoplasmic lysates from HEK 293T cells transfected with pcDNA 3.1 empty vector (Empty vector) or pcDNA 3.1-MINAS-60 (MINAS-60). Data are representative of three biological replicates. **c**, Quantitation of the ratio of cytoplasmic total 60S to total 40S subunits (**c**, left) or 80S to free 40S subunits (**c**, right) in "**b**". The area under each peak was measured using ImageJ. Data represent mean values ± s.e.m., and significance was evaluated with two-tailed *t*-test. **d**, Schematic representation of pre-60S assembly factors associated with different pre-60S states. The bait protein BRIX1 is indicated in blue. LSU assembly factors exhibiting increased co-purification with BRIX1 in *RBM10* KD cells are indicated in red. **e**, Volcano plot of quantitative proteomics (N = 5 biologically independent experiments) of BRIX1-FLAG pulldown from HEK 293T cells stably expressing BRIX1 and control shRNA (Ctrl), or BRIX1 and *RBM10* shRNA (KD), to quantify changes in BRIX1 interactome in *RBM10* KD over control HEK 293T cells.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Increased association with ribosome assembly factors are indicated in red and gene names are labeled. For complete quantitative proteomics results, see Supplementary Data S11-12. **f**, BRIX1-FLAG-IP and western blotting with antibodies indicated on the right using the two cell lines described above. Cell lysates (4%) before IP (input) were used as the loading control. Data are representative of three biological replicates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Franziska Bleichert and all members of the Slavoff and Baserga labs for helpful conversations. We thank Susan Gerbi at Brown University for discussion of the literature on ribosome subunit transport. We thank the Yale West Campus Imaging Core for the support and assistance in this work. This work was supported by a Searle Scholars Program Award, an Odyssey Award from the Richard and Susan Smith Family Foundation, and start-up funds from Yale University West Campus (to S. A. S.). X.C. was supported in part by a Rudolph J. Anderson postdoctoral fellowship from Yale University. A.K. was in part supported by an NIH Predoctoral Training Grant (5T32GM06754 3-12). S.J.B, C.J.B. and C.M.H. were supported by R35 GM131687. C.M.H. was supported by an NSF GFRP.

Data availability

Proteomic data are publicly available in ProteomeXchange under accession PXD026880. All data supporting the findings of this work are provided within the manuscript and its related source data.

References

1. Orr MW, Mao Y, Storz G & Qian SB Alternative ORFs and small ORFs: shedding light on the dark proteome. Nucleic Acids Res 48, 1029–1042 (2020). [PubMed: 31504789]
2. Brunet MA, Levesque SA, Hunting DJ, Cohen AA & Roucou X Recognition of the polycistronic nature of human genes is critical to understanding the genotype-phenotype relationship. Genome Res 28, 609–624 (2018). [PubMed: 29626081]
3. Chen J et al. Pervasive functional translation of noncanonical human open reading frames. Science 367, 1140–1146 (2020). [PubMed: 32139545]
4. Prensner JR et al. Noncanonical open reading frames encode functional proteins essential for cancer cell survival. Nat Biotechnol (2021).
5. Brunet MA et al. The FUS gene is dual-coding with both proteins contributing to FUS-mediated toxicity. EMBO Rep 22, e50640 (2021). [PubMed: 33226175]
6. Cao X et al. Alt-RPL36 downregulates the PI3K-AKT-mTOR signaling pathway by interacting with TMEM24. Nat Commun 12, 508 (2021). [PubMed: 33479206]
7. Slavoff SA et al. Peptidomic discovery of short open reading frame-encoded peptides in human cells. Nat Chem Biol 9, 59–64 (2013). [PubMed: 23160002]
8. Schwaid AG et al. Chemoproteomic discovery of cysteine-containing human short open reading frames. J Am Chem Soc 135, 16750–3 (2013). [PubMed: 24152191]
9. Pena C, Hurt E & Panse VG Eukaryotic ribosome assembly, transport and quality control. Nat Struct Mol Biol 24, 689–699 (2017). [PubMed: 28880863]
10. De Keersmaecker K, Sulima SO & Dinman JD Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. Blood 125, 1377–82 (2015). [PubMed: 25575543]
11. Pelletier J, Thomas G & Volarevic S Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat Rev Cancer 18, 51–63 (2018). [PubMed: 29192214]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

12. Nyhus C, Pihl M, Hyttel P & Hall VJ Evidence for nucleolar dysfunction in Alzheimer's disease. Rev Neurosci 30, 685–700 (2019). [PubMed: 30849050]
13. Farley-Barnes KI, Ogawa LM & Baserga SJ Ribosomopathies: Old Concepts, New Controversies. Trends Genet 35, 754–767 (2019). [PubMed: 31376929]
14. Leick V & Andersen SB Pools and turnover rates of nuclear ribosomal RNA in Tetrahymena pyriformis. Eur J Biochem 14, 460–4 (1970). [PubMed: 5479377]
15. Kater L et al. Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. Cell 171, 1599–1610 e14 (2017). [PubMed: 29245012]
16. Sanghai ZA et al. Modular assembly of the nucleolar pre-60S ribosomal subunit. Nature 556, 126–129 (2018). [PubMed: 29512650]
17. Dieterich DC, Link AJ, Graumann J, Tirrell DA & Schuman EM Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Natl Acad Sci U S A 103, 9482–7 (2006). [PubMed: 16769897]
18. Koch A et al. A proteogenomics approach integrating proteomics and ribosome profiling increases the efficiency of protein identification and enables the discovery of alternative translation start sites. Proteomics 14, 2688–98 (2014). [PubMed: 25156699]
19. Ma J et al. Improved Identification and Analysis of Small Open Reading Frame Encoded Polypeptides. Anal Chem 88, 3967–75 (2016). [PubMed: 27010111]
20. Khitun A & Slavoff SA Proteomic Detection and Validation of Translated Small Open Reading Frames. Curr Protoc Chem Biol 11, e77 (2019). [PubMed: 31750990]
21. Cao X et al. Comparative Proteomic Profiling of Unannotated Microproteins and Alternative Proteins in Human Cell Lines. J Proteome Res 19, 3418–3426 (2020). [PubMed: 32449352]
22. Wang T, Liu Y, Liu Q, Cummins S & Zhao M Integrative proteomic analysis reveals potential high-frequency alternative open reading frame-encoded peptides in human colorectal cancer. Life Sci 215, 182–189 (2018). [PubMed: 30419281]
23. Vanderperre B et al. Direct detection of alternative open reading frames translation products in human significantly expands the proteome. PLoS One 8, e70698 (2013). [PubMed: 23950983]
24. Kamiyama D et al. Versatile protein tagging in cells with split fluorescent protein. Nat Commun 7, 11046 (2016). [PubMed: 26988139]
25. Iyer-Bierhoff A & Grummt I Stop-and-Go: Dynamics of Nucleolar Transcription During the Cell Cycle. Epigenet Insights 12, 2516865719849090 (2019). [PubMed: 31206100]
26. Gorski SA, Snyder SK, John S, Grummt I & Misteli T Modulation of RNA polymerase assembly dynamics in transcriptional regulation. Mol Cell 30, 486–97 (2008). [PubMed: 18498750]
27. Hernandez-Verdun D Assembly and disassembly of the nucleolus during the cell cycle. Nucleus 2, 189–94 (2011). [PubMed: 21818412]
28. Cao X & Slavoff SA Non-AUG start codons: Expanding and regulating the small and alternative ORFeome. Exp Cell Res 391, 111973 (2020). [PubMed: 32209305]
29. Tan LY et al. Generation of functionally distinct isoforms of PTBP3 by alternative splicing and translation initiation. Nucleic Acids Res 43, 5586–600 (2015). [PubMed: 25940628]
30. Saghatelian A & Couso JP Discovery and characterization of smORF-encoded bioactive polypeptides. Nat Chem Biol 11, 909–16 (2015). [PubMed: 26575237]
31. Mellacheruvu D et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods 10, 730–6 (2013). [PubMed: 23921808]
32. Cox J et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics 13, 2513–26 (2014). [PubMed: 24942700]
33. Liang X et al. Structural snapshots of human pre-60S ribosomal particles before and after nuclear export. Nat Commun 11, 3542 (2020). [PubMed: 32669547]
34. Liu Y, Deisenroth C & Zhang Y RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis and Tumor Surveillance. Trends Cancer 2, 191–204 (2016). [PubMed: 28741571]
35. Sondalle SB, Longerich S, Ogawa LM, Sung P & Baserga SJ Fanconi anemia protein FANCI functions in ribosome biogenesis. Proc Natl Acad Sci U S A 116, 2561–2570 (2019). [PubMed: 30692263]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

36. Bechara EG, Sebestyen E, Bernardis I, Eyras E & Valcarcel J RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Mol Cell 52, 720–33 (2013). [PubMed: 24332178]
37. Bassler J & Hurt E Eukaryotic Ribosome Assembly. Annu Rev Biochem 88, 281–306 (2019). [PubMed: 30566372]
38. 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). [PubMed: 21048991]
39. Niceta M et al. TARP syndrome: Long-term survival, anatomic patterns of congenital heart defects, differential diagnosis and pathogenetic considerations. Eur J Med Genet 62, 103534 (2019). [PubMed: 30189253]
40. Imielinski M et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150, 1107–20 (2012). [PubMed: 22980975]
41. Gagnon M et al. Potentiation of B2 receptor signaling by AltB2R, a newly identified alternative protein encoded in the human bradykinin B2 receptor gene. J Biol Chem 296, 100329 (2021). [PubMed: 33497625]
42. Chu Q et al. Identification of Microprotein-Protein Interactions via APEX Tagging. Biochemistry 56, 3299–3306 (2017). [PubMed: 28589727]
43. Farley-Barnes KI et al. Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. Cell Rep 22, 1923–1934 (2018). [PubMed: 29444442]
44. Nguyen le XT, Raval A, Garcia JS & Mitchell BS Regulation of ribosomal gene expression in cancer. J Cell Physiol 230, 1181–8 (2015). [PubMed: 25336383]
45. Xing YH et al. SLERT Regulates DDX21 Rings Associated with Pol I Transcription. Cell 169, 664–678 e16 (2017). [PubMed: 28475895]
46. Ogawa LM 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). [PubMed: 33689394]
47. Bryant CJ, McCool MA, Abriola L, Surovtseva YV & Baserga SJ A high-throughput assay for directly monitoring nucleolar rRNA biogenesis. Open Biol 12, 210305 (2022). [PubMed: 35078352]
48. Parker MD, Collins JC, Korona B, Ghalei H & Karbstein K A kinase-dependent checkpoint prevents escape of immature ribosomes into the translating pool. PLoS Biol 17, e3000329 (2019). [PubMed: 31834877]
49. Gregory B et al. The small and large ribosomal subunits depend on each other for stability and accumulation. Life Sci Alliance 2(2019).
50. Weerapana E, Speers AE & Cravatt BF Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP)--a general method for mapping sites of probe modification in proteomes. Nat Protoc 2, 1414–25 (2007). [PubMed: 17545978]
51. Schmidt EK, Clavarino G, Ceppi M & Pierre P SUnSET, a nonradioactive method to monitor protein synthesis. Nat Methods 6, 275–7 (2009). [PubMed: 19305406]
52. Cao X et al. Histone H4K20 Demethylation by Two hHR23 Proteins. Cell Rep 30, 4152–4164 e6 (2020). [PubMed: 32209475]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

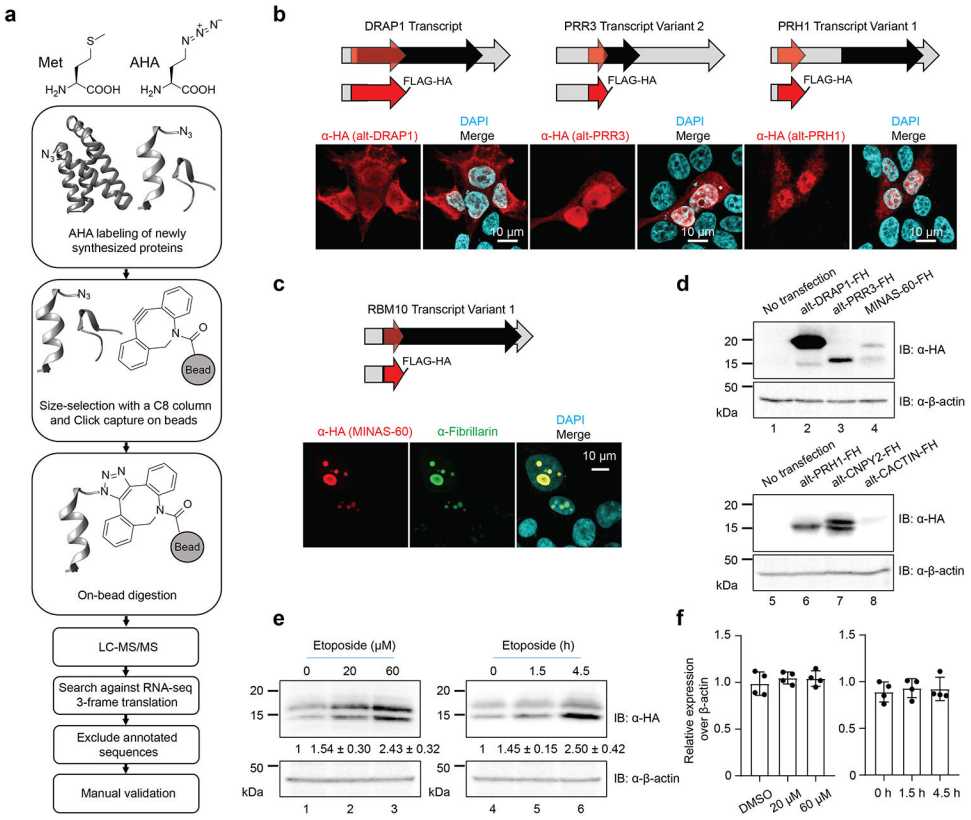


Fig. 1 | BONCAT chemoproteomic profiling of nascent alt-proteins.

a, Schematic workflow. **b**,**c**, Top: Schematic representation of human *DRAP1* transcript, *PRR3* transcript variant (tv) 2, *PRH1* tv1, or *RBM10* tv1 (**c**); light gray arrow, 5' and 3' untranslated regions; red, alt-ORF; black, annotated protein CDS. Middle: Schematic representation of the expression construct containing complete 5'UTR and alt-ORF of the transcript indicated above, with a FLAG-HA appended to the C-terminus of the alt-protein. Bottom: HEK 293T cells transfected with the expression construct were immunostained with anti-HA (red), DAPI (cyan), and anti-fibrillarin (green, **c**). Scale bar, 10 μm. Data are representative of three biological replicates. **d**, HEK 293T cells transfected with the expression construct were lysed and immunoblotted (IB), with untransfected HEK 293T cells as a control. Data are representative of three biological replicates. **e**, **f**, HEK 293T cells transfected with the alt-CNPY2 expression construct were treated with increasing amounts of etoposide or vehicle for 2 h (left), or with 60 μM etoposide for different times or vehicle (right), followed by quantitative analysis ($N = 3$ biologically independent experiments) of the alt-CNPY2-FLAG-HA immunoblotting signal (**e**); or followed by quantitative RT-PCR with primers specific to *CNPY2* tv1 (**f**). Data represent mean values ± s.e.m., $N = 4$ biologically independent samples.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

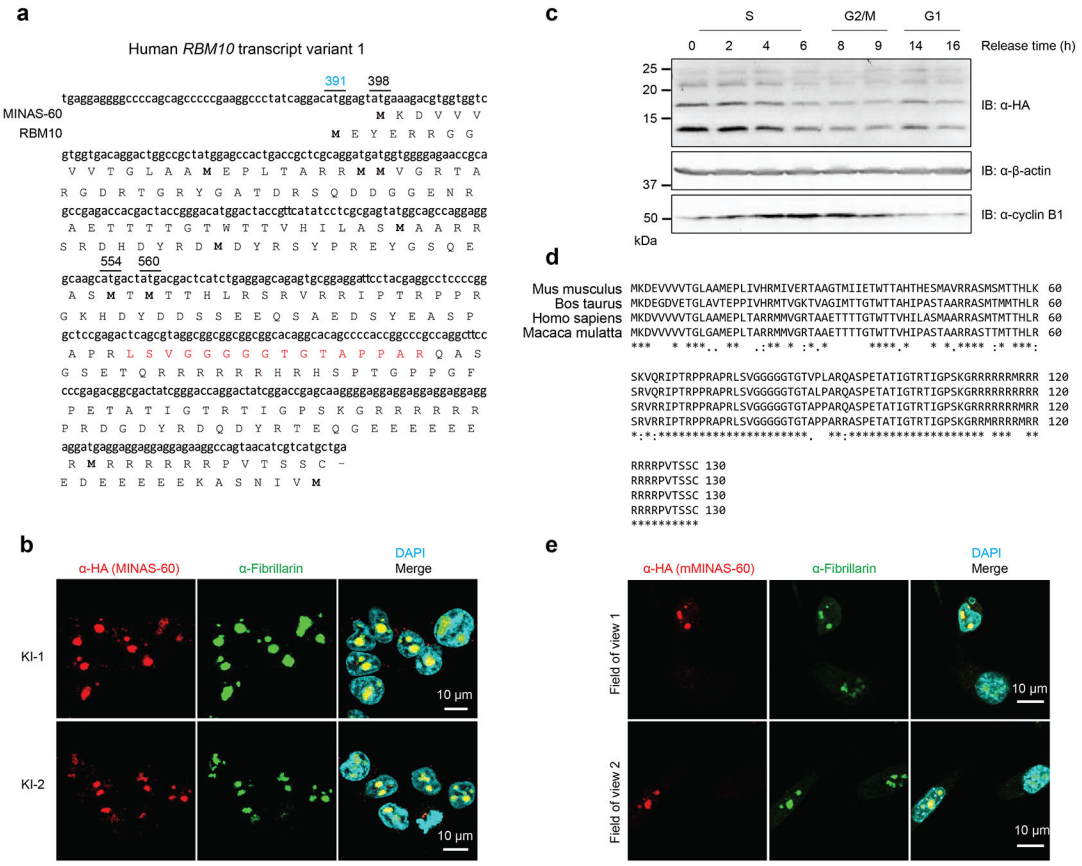


Fig. 2 | MINAS-60 is endogenously expressed, cell-cycle regulated and conserved in mouse. **a**, The cDNA sequence of human *RBM10* tv1 is shown with the protein sequences of MINAS-60 and RBM10 indicated below. The start codons of RBM10 (391) and MINAS-60 (398, 554 and 560) are numbered above the cDNA sequence. The tryptic peptide of MINAS-60 detected by LC-MS/MS in this study is highlighted in red. **b**, Immunostaining of two MINAS-60 epitope tag knock-in (KI) HEK 293T cell lines with anti-HA (red), anti-fibrillarin (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates. **c**, Western blotting analysis of synchronized MINAS-60 KI cells released from the G1/S boundary at the indicated time points with antibodies indicated on the right. Data are representative of three biological replicates. **d**, *RBM10* mRNAs from *Mus musculus* (NM_145627.3), *Bos taurus* (NM_001206693.1), and *Macaca mulatta* (XM_015127275.2) were obtained from NCBI nucleotide database, then translated in the +1, +2 and +3 frames using the Expasy translate tool. The full-length sequences of hypothetical MINAS-60 homologs were predicted in each transcript and aligned with ClustalW2. **e**, 3T3 cells transfected with mouse MINAS-60 (mMINAS-60) expression construct were immunostained with anti-HA (red), anti-fibrillarin (green), and DAPI (cyan). Scale bar, 10 μm. Data are representative of three biological replicates.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

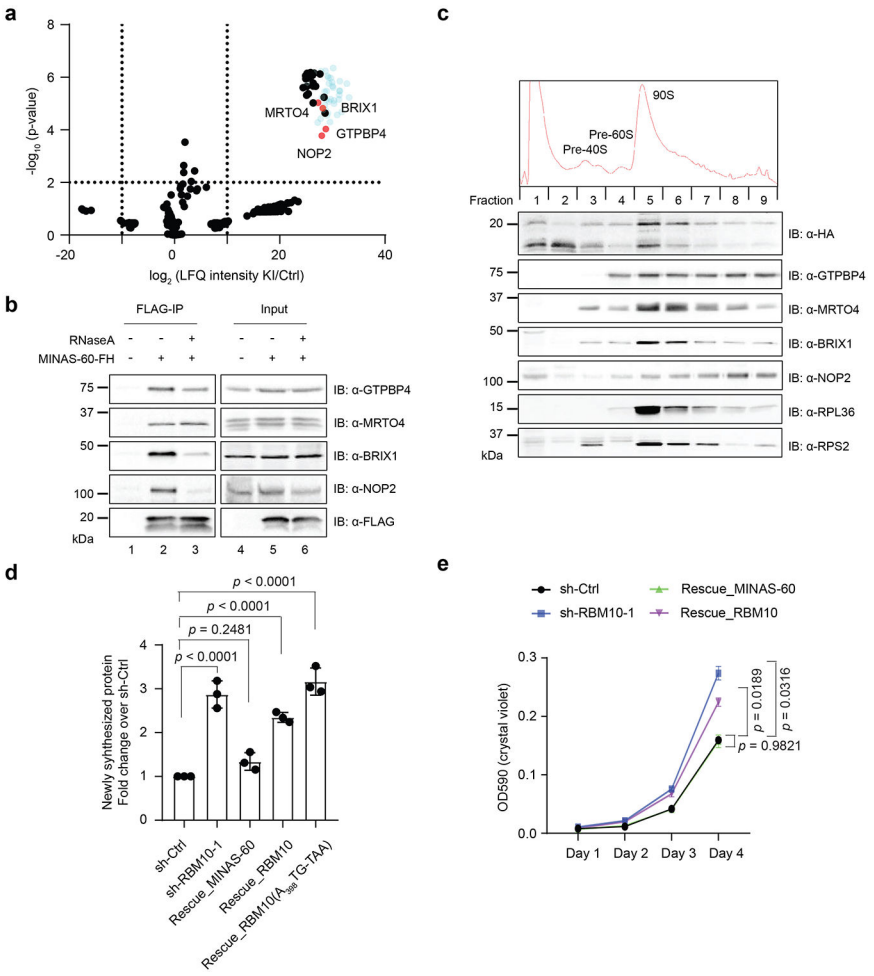


Fig. 3 |. MINAS-60 associates with ribosomal large subunit (LSU) assembly factors to downregulate protein synthesis and cell proliferation.

a, Volcano plot of quantitative proteomics ($N = 3$ biologically independent experiments) of anti-FLAG pulldown from MINAS-60 KI (KI) or control (Ctrl) HEK 293T nuclear lysates. Ribosomal proteins are indicated in blue. Enriched LSU assembly factors are indicated in red and gene names are labeled. For complete quantitative proteomics results, see Supplementary Data S7-8. **b**, HEK 293T cells were transfected with MINAS-60-FLAG-HA (MINAS-60-FH, lanes 2 and 3) or vehicle (lane 1), and immunoprecipitation (FLAG-IP) was performed in the absence (lanes 1 and 2) or presence (lane 3) of RNaseA, followed by immunoblotting with antibodies indicated on the right. Cell lysates (4%) before IP (input, lanes 4-6) were used as loading controls. Data are representative of three biological replicates. **c**, Top: Sucrose-gradient sedimentation analysis of nuclear lysates containing ribosome precursor complexes from HEK 293T cells stably expressing MINAS-60-FLAG-HA. Bottom: Immunoblot analysis of fractions numbered at the top with antibodies indicated on the right. Data are representative of three biological replicates. **d**, ImageJ was used to quantify puromycin incorporation for cells indicated at the bottom relative to sh-Ctrl from three biological replicates. Data represent mean values ± s.e.m., one-way ANOVA (Dunnett's test). **e**, Growth curve of control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1), rescue with MINAS-60 (Rescue_MINAS-60) and rescue with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RBM10 (Rescue_RBM10) HEK 293T cells at the indicated number of days ($N = 3$ biological independent experiments). Data represent mean values ± s.e.m., two-way ANOVA (Dunnett's test).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

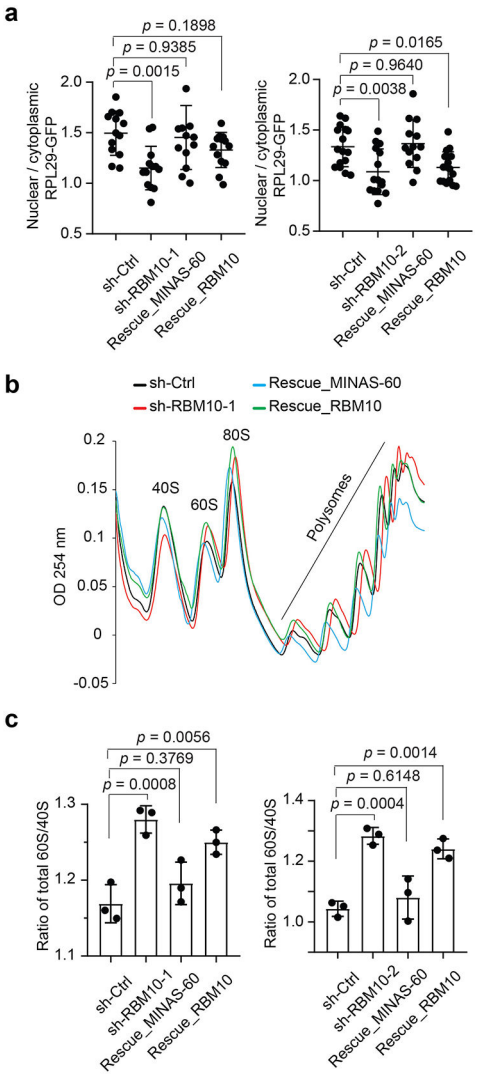


Fig. 4 |. MINAS-60 downregulates cytoplasmic LSU export.

a, Quantitation of RPL29-GFP intensity in the nucleus vs. cytoplasm in control (sh-Ctrl), *RBM10* knockdown with shRNA (sh-RBM10-1 (left), sh-RBM10-2 (right)), rescue with MINAS-60 (Rescue_MINAS-60), or rescue with RBM10 (Rescue_RBM10) HEK 293T cells stably expressing RPL29-GFP. At least 13 fields of view were analyzed, totaling > 350 cells for each measurement. Data represent mean values ± s.e.m., and significance was evaluated with one-way ANOVA (Dunnett's test). **b**, Sucrose-gradient sedimentation analysis of ribosomal fractions (40S, 60S, 80S and polysomes) of cytoplasmic lysates from control (sh-Ctrl), *RBM10* knockdown (sh-RBM10-1), rescue with MINAS-60 (Rescue_MINAS-60) or rescue with RBM10 (Rescue_RBM10) HEK 293T cells. Data are representative of three biological replicates. **c**, Quantitation of the ratio of cytoplasmic 60S to 40S subunits in the cell lines indicated below after sucrose gradient fractionation. The area under each peak was measured using ImageJ[49]. Data represent mean values ± s.e.m., and significance was evaluated with one-way ANOVA (Dunnett's test).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

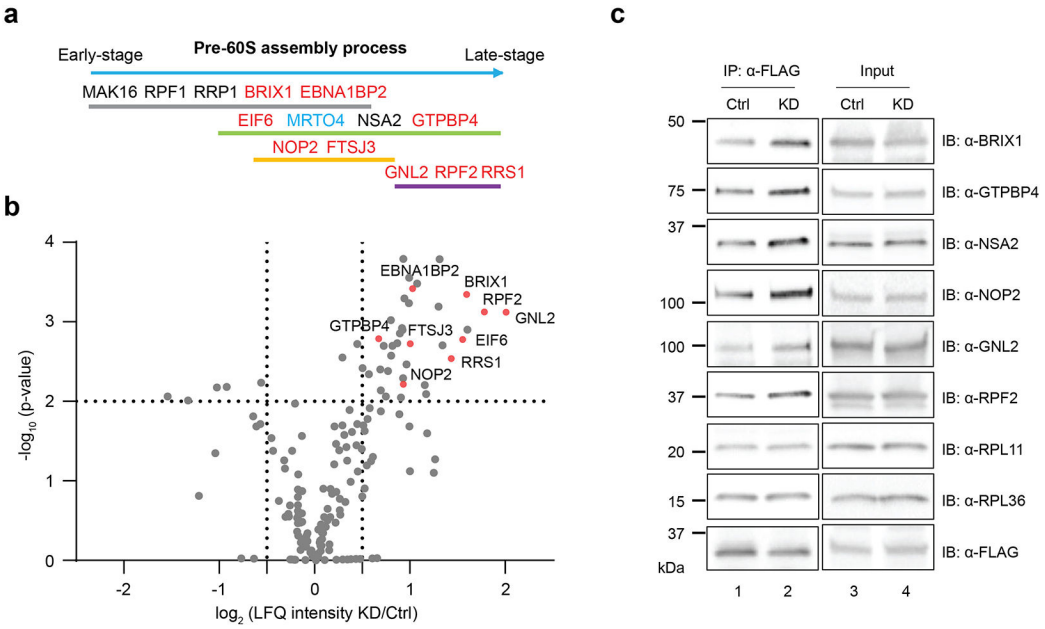


Fig. 5 | *RBM10* silencing promotes pre-60S assembly.

a, Schematic representation of pre-60S assembly factors associated with different pre-60S states[15]. The bait protein MRTO4 is indicated in blue. LSU assembly factors exhibiting increased co-purification with MRTO4 in *RBM10* KD cells are indicated in red. **b**, Volcano plot of quantitative proteomics ($N = 4$ biologically independent experiments) of MRTO4-FLAG pulldown from HEK 293T cells stably expressing MRTO4 and control shRNA (Ctrl), or MRTO4 and *RBM10* shRNA (KD), to quantify changes in MRTO4 interactome in *RBM10* KD over control HEK 293T cells. Increased association with ribosome assembly factors are indicated in red and gene names are labeled. For complete quantitative proteomics results, see Supplementary Data S9-10. **c**, MRTO4 FLAG-IP and western blotting with antibodies indicated on the right using the two cell lines described above. Cell lysates (4%) before IP (input) were used as the loading control. Data are representative of three biological replicates.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

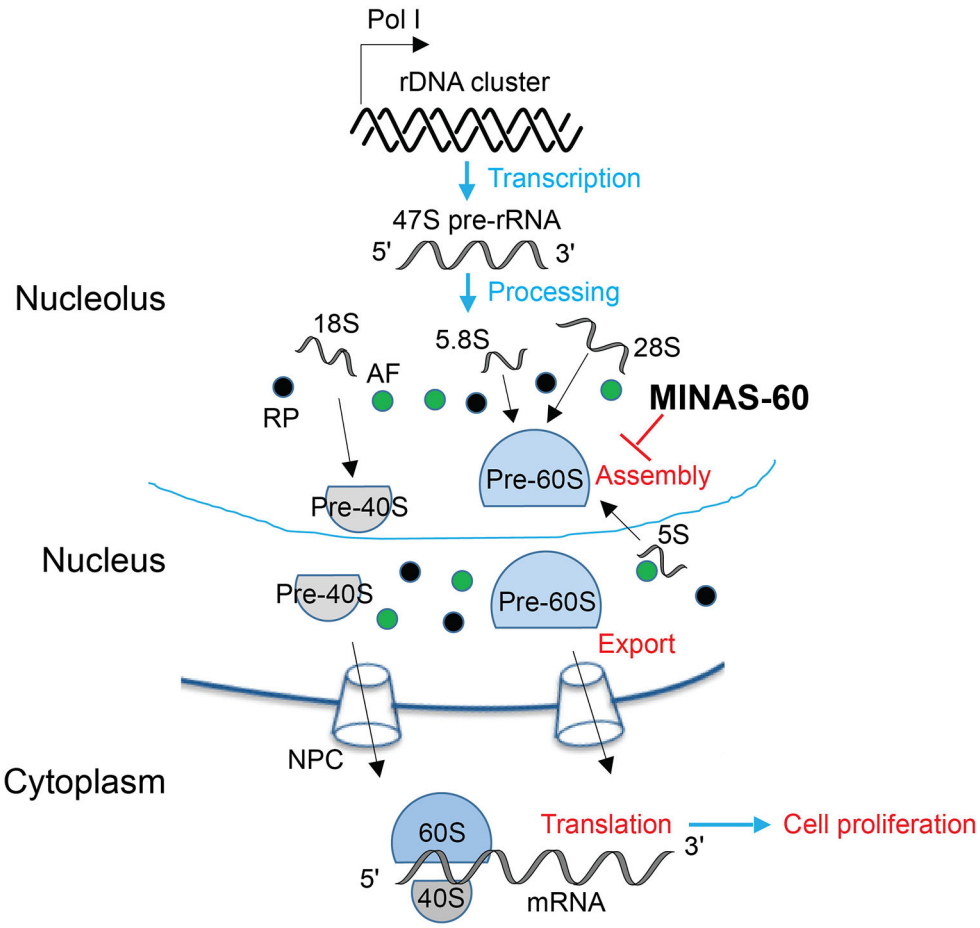


Fig. 6 |. Model of MINAS-60 regulatory pathway.

MINAS-60 localizes to the nucleolus, where it may acts as a checkpoint inhibitor of late-stage pre-60S assembly, consequently slowing cytoplasmic LSU export, global protein synthesis and cell proliferation.